The “bilingual advantage”
How does speaking two languages affect cognitive control and language processing abilities?

Susan E. Teubner-Rhodes, BA, Alan Mishler, BA, Ryan Corbett, BA, Llorenç Andreu, PhD, Monica Sanz-Torrent, PhD, John C. Trueswell, PhD, Jared M. Novick, PhD

We report the results of a study designed to investigate whether having a high degree of proficiency in two languages impacts individuals' cognitive and linguistic abilities in ways that are relevant to Language and Intelligence Analysts (see attached article). In this study, monolinguals were compared to “balanced” bilinguals—native speakers of two languages who use both on a regular basis. The balanced bilinguals were shown to have better cognitive control and linguistic ambiguity resolution abilities, a result that may be attributed to their frequent switching between languages and their concomitant need to suppress the non-active language. The pattern of the bilinguals’ results suggests that this cognitive advantage arises from an enhanced ability to detect conflict or interference in the environment, and, consequently, yields greater cognitive flexibility to deal with such conflict. These findings further strengthen the link between domain-general cognitive control and the ability to understand complex, degraded, or ambiguous language materials, and they highlight again the potential value of executive function training for analysts. Finally, the results underscore the importance of achieving high proficiency in a second language because of the broad-based cognitive benefits.

RESEARCH GOALS AND BACKGROUND

This study was designed to achieve the following goals: (1) replicate and extend previous results showing that balanced bilinguals have superior cognitive control abilities compared to monolinguals; (2) characterize the nature of this so-called “bilingual advantage” by identifying the specific cognitive functions that are enhanced in bilinguals; and (3) determine whether this advantage extends to analyst-relevant tasks involving linguistic ambiguity resolution.

Evidence from previous research has suggested that bilinguals enjoy certain cognitive advantages relative to their monolingual peers, but the nature of the bilingual advantage has been in dispute. Bilinguals typically outperform monolinguals on cognitive control tasks such as the Stroop task, which requires participants to name the color of the ink that a word appears in while ignoring (i.e., not reading) the word itself. For example, participants might have to respond “blue” to the word yellow printed in blue ink, which requires resolving the conflict between the color and word representations. This is a particularly challenging cognitive control task because one must override the dominant bias to read the word in order to perform an atypical color-naming (perceptual) task instead. One account of the bilingual advantage, the Inhibitory Control account, holds that bilinguals have an enhanced ability to suppress the interference that arises when there are competing sources of information (e.g., the word itself versus the color that it is printed in). Another account, the Conflict
Monitoring account, holds that bilinguals have a greater ability to detect the presence of multiple, competing sources of information, but that they are not necessarily better than monolinguals at resolving the conflict that arises once it is detected.

In addition, previous evidence of enhanced cognitive abilities in bilinguals has come from non-syntactic cognitive tasks (e.g., Stroop), leaving unanswered the question of whether the bilingual advantage manifests in situations involving complex, ambiguous, or degraded linguistic input of the kind that Language Analysts frequently encounter. Recent CASL research has suggested that in monolinguals, domain-general cognitive abilities are involved in the processing of ambiguous sentences. We predicted, therefore, that the bilingual advantage would extend to this domain as well.

STUDY DESIGN

This study compared native monolingual Spanish speakers to native bilingual speakers of Spanish and Catalan. Participants completed the following three tasks over the course of approximately one hour: (1) a pretest sentence processing task that included both complex/ambiguous sentences as well as simple/unambiguous sentences (20 minutes); (2) either a high- or a low-interference version of the N-back task (20 minutes), a non-linguistic working memory task that engages the same processes involved in the resolution of temporary linguistic ambiguities; and finally, (3) a posttest version of the initial sentence processing task (20 minutes). During N-back, subjects viewed single words sequentially and indicated whether an item appeared 3 trials previously. Only the high-interference version contained “lures”—words that appeared 2, 4, or 5 items before, compelling subjects to override a familiarity bias to correctly indicate that the item was not a 3-back target. All tasks were conducted in Spanish.

RESULTS

Bilinguals were more accurate than monolinguals on the high- but not the low-interference version of the N-Back task, and they exhibited measurably better sentence comprehension than monolinguals across all sentence types. These results thus replicate and extend the bilingual advantage. Notably, in the high-interference version of the N-back task, bilinguals outperformed monolinguals both on trials that involved cognitive conflict (lures) and on non-conflict trials (targets and fillers); similarly, in the sentence processing task, bilinguals were more accurate than monolinguals on simple/unambiguous sentences as well as on complex/ambiguous sentences. These results suggest that in contexts that required participants to continually monitor for the presence of conflict, including conflict arising from temporary syntactic ambiguities, bilinguals were able to more rapidly assess whether conflict was present and react accordingly. In other words, bilinguals’ performance appeared to be due to superior conflict detection skills (consistent with the Conflict Monitoring account) rather than to a selective ability to resolve conflict per se (inconsistent with the Inhibitory Control account). This indicates that the bilingual advantage may be more global than previously thought; in particular, when conflict or irrelevant information is present in the environment and must be ignored, bilinguals readily detect its presence and can more rapidly and accurately respond to all trial types, presumably due to a generally heightened state of awareness. Interestingly, when conflict trials were removed in the low-interference version of N-back, monolinguals performed just as well as bilinguals, suggesting that short-term memory abilities are equivalent between the groups when there is no conflict or interference to detect. Finally, improvement on N-Back lure trials selectively predicted improvement in syntactic ambiguity resolution from pretest to posttest in both monolinguals and bilinguals, suggesting that the two tasks draw on shared cognitive control resources. This finding is especially important because the N-back intervention task lasted for just 20 minutes, suggesting that even relatively brief training may bestow temporary benefits on untrained language processing tasks.

RELEVANCE

The results of this study further substantiate the link between domain-general cognitive control and linguistic ambiguity resolution. Bilinguals’ greater ability to detect conflict manifested not only on a non-
linguistic cognitive task (N-Back), but also on a reading task involving temporary syntactic ambiguities that often lead to misinterpretation. This result is counterintuitive, considering that both monolinguals and bilinguals were native speakers of Spanish, the language in which the reading task was presented. Given how frequently Language Analysts encounter language materials that are ambiguous, complex, or degraded, these results further underscore the potential value of training regimens designed to improve domain-general cognitive control. The findings also suggest that a high proficiency in a second language may serve as an important natural form of cognitive control training; namely, the frequent usage of, and switching between, multiple language systems requires frequent conflict detection and may therefore enhance cognitive control abilities across a range of linguistic and non-linguistic tasks where attention demands are high. The refined understanding of the effect of bilingualism on cognitive control that this study has yielded will help us as we continue to develop such language and cognitive interventions for analysts.

Finally, the results of this research could also ultimately argue for a bilingual workforce in the USG, not merely because of the obvious benefits that linguistic and cultural knowledge bestows, but also because of the general-purpose cognitive benefits enjoyed by bilingual speakers. The present findings also suggest that the effects of the bilingual advantage might be promoted through various forms of intensive training—both cognitive and linguistic—and might further place a high value on a bilingual staff (perhaps even a subset of which is a native-bilingual staff) of the future. The results may eventually inform job-selection criteria and training of current government employees.

Corresponding Author and Reprints:
Jared M. Novick, PhD, Associate Research Scientist, University of Maryland Center for Advanced Study of Language, (301) 226-8841, jnovick@casl.umd.edu, www.casl.umd.edu.

Funding/Support: This material is based upon work supported, in whole or in part, with funding from the United States Government. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the University of Maryland, College Park and/or any agency or entity of the United States Government. Nothing in this report is intended to be and shall not be treated or construed as an endorsement or recommendation by the University of Maryland, United States Government, or the authors of the product, process, or service that is the subject of this report. No one may use any information contained or based on this report in advertisements or promotional materials related to any company product, process, or service or in support of other commercial purposes. This report is not Releasable to the Defense Technical Information Center per DoD Directive 3200.12. The Contracting Officer’s Representative for this project is John Walker, Government Technical Director at CASL, (301) 226-8912, jwalker@casl.umd.edu.
The “bilingual advantage”

How does speaking two languages affect cognitive control and language processing abilities?

Susan E. Teubner-Rhodes, BA, Alan Mishler, BA, Ryan Corbett, BA, Llorenç Andreu, PhD, Monica Sanz-Torrent, PhD, John C. Trueswell, PhD, Jared M. Novick, PhD

ABSTRACT

Bilinguals typically outperform monolinguals on non-linguistic tasks requiring cognitive control, the ability to regulate attention and resolve competing alternatives. This “bilingual advantage” has been attributed to bilinguals’ extensive practice using inhibitory control when selecting between their languages. Recent evidence suggests that, in monolinguals, resolving conflict (i.e., ambiguity) during language processing engages domain-general cognitive control resources, yet it is unknown if the bilingual advantage extends to sentence processing. The present study compared bilinguals and monolinguals on a parsing task involving syntactic ambiguity both before and after brief training on N-back, a non-linguistic cognitive control task. Bilinguals were more accurate on a high-interference but not a low-interference N-back task and exhibited measurably better sentence comprehension than monolinguals; however, the advantages were observed for both conflict and non-conflict trial types across the high-interference N-back and parsing tasks. Additionally, improvement on N-back trials requiring interference resolution selectively predicted improvement in syntactic ambiguity resolution. These findings provide additional evidence for a bilingual advantage in cognitive control, show that the advantage extends to sentence processing, and substantiate a link between cognitive control and syntactic ambiguity resolution. Finally, because the advantage was not selective for conflict trials, our results suggest that the bilingual advantage may be better characterized by increased cognitive flexibility rather than inhibitory control.
INTRODUCTION

Recent research has demonstrated that balanced bilinguals, individuals who are equally proficient in two languages, experience a host of cognitive advantages over monolinguals. In particular, relative to their monolingual peers, bilinguals seem to have better cognitive control, the ability to regulate mental activity in order to resolve among competing alternatives. This advantage is evident across the lifespan: for example, young bilingual children outperform monolinguals on executive function tasks requiring inhibitory and attentional control (Bialystok, 1999; Bialystok & Martin, 2004; Kovács & Mehler, 2009; Martin-Rhee & Bialystok, 2008); healthy adult bilinguals are faster than monolinguals on tasks requiring cognitive control (Bialystok, 2006; Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009); and older adult bilinguals are more resilient than monolinguals against the effects of cognitive decline due to natural aging and Alzheimer’s disease (Bialystok, Craik, Klein, & Viswnathan, 2004; Schweizer, Ware, Fischer, Craik, & Bialystok, 2012).

Though there is a wealth of evidence supporting the existence of cognitive advantages in bilinguals (for review, see Bialystok, Craik, Green, and Gollan, 2009), there are still a number of unanswered questions regarding the nature, specificity, and extent of the advantage. In particular, there is extensive debate about whether the bilingual advantage reflects enhanced inhibitory control (e.g., increased suppression of goal-irrelevant representations), or whether it reflects better conflict monitoring abilities (e.g., better conflict detection and increased flexibility in recruiting cognitive control resources). Additionally, while a general bilingual advantage in cognitive control is well-established, few studies have examined whether this advantage cascades into language processing. The absence of such investigations is surprising given that the proposed source of the bilingual cognitive advantage is the systematic control of two languages; that is, because the bilingual advantage presumably results from differential language experience, one might also expect benefits in language processing. Finally, it is unclear whether the bilingual advantage changes while performing a cognitively demanding task; during such tasks, monolinguals may “catch-up” to bilinguals in cognitive control abilities due to the effects of practice, or bilinguals could outstrip monolinguals further due to more flexible engagement of cognitive control mechanisms. The present study aims to address these issues by testing whether healthy young adult bilinguals outperform monolinguals on a reading task involving syntactic ambiguity, both before and after a brief cognitive intervention tapping cognitive control abilities.

Over the past decade-and-a-half, the dominant model of bilingual language processing has been the inhibitory control (IC) model (Green, 1998), which theorizes that, because semantic equivalents from different lexicons (e.g., native language (L1) versus second language (L2)) are both activated to some extent by their shared conceptual representation, lexical items from the irrelevant lexicon must be suppressed by a central inhibitory control mechanism in order to select the appropriate word. This model accounts quite nicely for many existing empirical findings regarding bilingual language use, particularly language-switching and translation. For instance, consider the counterintuitive asymmetry of language-switch costs, wherein bilinguals have more difficulty switching from their weaker L2 into their dominant L1, than they do switching from their dominant L1 into their weaker L2 (Meuter & Allport, 1999). Such an effect is easily explained by the IC model; namely, because participants must use stronger inhibitory control to suppress their L1 when outputting L2 than vice versa, L1 requires relatively greater reactivation following L2 use than L2 does following L1 use. Theoretically, the IC model could also explain a bilingual advantage in cognitive control: if switching between two languages involves inhibiting irrelevant lexical representations from the language not currently in use, then bilingual language experience essentially acts as cognitive training, strengthening domain-general inhibitory control mechanisms via extensive practice (Abutalebi & Green, 2008; Bialystok et al., 2009). These improved cognitive resources could be applied in non-verbal tasks requiring cognitive control, yielding the observed bilingual cognitive benefit.

However, it is uncertain whether the IC theory can fully account for the diverse empirical evidence demonstrating a bilingual advantage in cognitive control, partly because the advantage itself is not well characterized. Generally, it seems that bilinguals, especially balanced bilinguals who have been exposed to two languages from infancy or early childhood and are equally proficient in both, outperform monolinguals on a range of cognitive control tasks (Bialystok, 2010; Bialystok et al., 2004; Costa et al., 2009; Martin-Rhee & Bialystok, 2008). Particularly remarkable is that this advantage is observed on non-linguistic tasks: bilinguals
exhibit better performance (typically, faster response times) on 1) the Simon task (Bialystok et al., 2004), in which a visual stimulus appears on one side of a computer screen, and the correct response to a non-spatial attribute of the stimulus is located on the same (congruent) or opposite (incongruent) side; 2) the Flanker task (Costa et al., 2009), in which participants indicate the direction of an arrow that is flanked by task-irrelevant arrows pointing in the same (congruent) or opposite (incongruent) direction; and 3) the spatial Stroop task (Bialystok, 2006), in which participants indicate the direction of a single arrow that appears on the same (congruent) or opposite (incongruent) side as the correct response. These tasks all involve task-irrelevant stimulus attributes that occasionally conflict with information from task-relevant attributes; thus, they all require cognitive control to resolve between competing sources of information. However, in some instances (Bialystok et al., 2004; Kovács & Mehler, 2009), the bilingual advantage appears to reflect superior inhibitory control (i.e., more efficient conflict resolution indexed by performance on incongruent compared with congruent trials, as predicted by the IC model), whereas in others (Abutalebi et al., 2012; Costa et al., 2009), it seems to reveal better conflict monitoring (e.g., enhanced detection of conflict in the environment and thus more flexible adjustments between conflict and non-conflict trials, yielding better performance generally; for review, see Hilchey & Klein, 2011).

Support for the IC account comes from findings that, as compared to monolinguals, bilinguals exhibit less interference (e.g., selective performance decrements) on conflict trials—those requiring inhibition of irrelevant information—relative to baseline non-conflict trials. For example, Kovács and Mehler (2009) found that 7-month-old infants being raised in bilingual environments successfully inhibited looks to a previously rewarded, but now incorrect, location, whereas infants being raised in monolingual environments did not. Additionally, on the Simon task, middle-aged and older bilinguals are faster on incongruent relative to congruent trials than their monolingual peers (Bialystok et al., 2004), although this effect only reached significance in older adults. Thus, if there is a bilingual advantage in inhibitory control, it appears more robust at the poles of the developmental spectrum, i.e., in older adults and young children whose inhibitory control abilities are naturally reduced. Notably, in the middle-aged adults, bilinguals were faster than monolinguals on both congruent and incongruent trial types; while this population of bilinguals did not demonstrate a significant (or selective) inhibitory control advantage, they still exhibited an overall advantage on the Simon task. Because responding to congruent trials does not involve inhibition, this overall RT advantage is not easily explained by the IC account, suggesting an alternative account of the bilingual advantage may be necessary.

Based on evidence that bilinguals outperform monolinguals on both congruent and incongruent trials, without exhibiting reduced interference effects (e.g., performance decrements on incongruent trials relative to congruent trials), Costa et al. (2009) have proposed that the source of the bilingual advantage is improved conflict monitoring, the ability to detect the occurrence of conflict and reactively increase cognitive control recruitment (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Indeed, Costa et al. (2009) demonstrated that the magnitude of the bilingual advantage was influenced by how much switching was necessary between congruent and incongruent trials on a Flanker task. When switching occurred frequently, imposing heavy demands to monitor for conflict and adjust cognitive control accordingly, bilinguals were significantly faster at both trial types. But, when very little switching occurred, bilinguals performed no differently from monolinguals, even when the vast majority of the trials (92%) were incongruent and required inhibitory control (Costa et al., 2009). Recently, fMRI research has found that overlapping voxels in the anterior cingulate cortex (ACC), the medial-frontal region thought to be responsible for detecting conflict and signaling adjustments in control, are activated by language-switching during a picture-naming task and by incongruent Flanker trials (Abutalebi et al., 2012). This suggests that the unique bilingual experience of language-switching may engage the domain-general conflict monitoring system, strengthening these resources and contributing to the observed bilingual advantage in cognitive control.

Although many studies have observed a general bilingual advantage in cognitive control, the role of this advantage in language processing has not been examined. Mounting evidence suggests that some types of language processing, specifically cases involving linguistic ambiguity or selection between competing lexical representations, involve the same domain-general cognitive control resources that have been implicated in the bilingual advantage (January, Trueswell, & Thompson-Schill, 2009; Novick, Kan, Trueswell, & Thompson-Schill, 2009; Novick, Trueswell, & Thompson-Schill, 2005; Ye & Zhou, 2009). For instance, a case-study of a patient with focal damage to the left inferior frontal gyrus (LIFG), thought to play a major role in resolving
among competing alternatives (after the ACC has signaled that adjustments are necessary), revealed concomitant deficits in resisting proactive interference on an item-recognition task and recovering from misinterpretation during processing of temporarily ambiguous sentences (Novick et al., 2009), suggesting that the LIFG was involved in resolving both syntactic and non-syntactic conflict.

If bilingualism enhances the same cognitive control resources engaged by syntactic ambiguity resolution, then it should also improve performance on monolingual sentence processing tasks (i.e., sentence processing within a single language) involving syntactic ambiguity. However, exactly how such improvement should manifest depends on whether the bilingual advantage reflects better inhibitory control or better conflict monitoring. If it involves better inhibitory control, then bilinguals should exhibit better syntactic ambiguity resolution than monolinguals, as measured by faster reading times and higher comprehension accuracy selectively on temporarily ambiguous sentences relative to unambiguous sentences. In contrast, if it involves better conflict monitoring, then bilinguals should be more flexible in adjusting cognitive control recruitment, outperforming monolinguals on both temporarily ambiguous and unambiguous sentences in situations that require the frequent detection of linguistic ambiguities. By comparing how bilinguals and monolinguals perform on a sentence processing task involving occasional temporary syntactic ambiguities, we can thus assess 1) whether the advantage conferred by bilingualism taps the same cognitive control resources as the resolution of linguistic conflict and 2) whether bilinguals enjoy enhanced inhibitory control or enhanced conflict monitoring abilities.

Some evidence suggests that despite bilinguals’ initial advantage on cognitive control tasks, monolinguals may improve their performance with a small amount of practice, effectively “catching up” to bilinguals and eliminating their advantage. For instance, Bialystok et al. (2004) compared the interference effects of bilinguals and monolinguals (the difference in RTs between incongruent and congruent trials) over several blocks of the Simon task, presented within a single session. They found that bilinguals initially exhibited less interference than monolinguals, but that only monolinguals gradually improved, eventually equaling bilingual performance. If advantages endowed by a lifetime of bilingual language experience can be achieved by monolinguals via 240 trials of practice on the Simon task, then the effect of bilingualism on cognitive control seems rather weak. Moreover, the lack of improvement in bilinguals suggests that they may already operate at a capacity limit for cognitive control. Thus, whether bilinguals can continue to improve their cognitive control abilities via practice is of theoretical importance for the question of whether cognitive capacity is fixed or flexible.

Relatedly, research on cognitive training suggests that, with extensive practice at the threshold of their ability level, healthy young monolingual adults are able to increase their working memory and cognitive control; moreover, these improvements transfer to non-trained tasks relying on the same cognitive resources, suggesting that the effect of training extends beyond task-specific practice effects (Chein & Morrison, 2010; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008). Recent findings also demonstrate that domain-general cognitive control training can improve syntactic ambiguity resolution as indexed by increased sentence comprehension and decreased re-reading during revision (Novick, Hussey, Teubner-Rhodes, Harbison, Dougherty, & Bunting, submitted). This body of work suggests that, even in populations who are at their developmental peak with respect to executive function, abilities like working memory and cognitive control, which were once thought to be fixed in capacity (Engle & Bukstel, 1978), are amenable to improvement.

However, evidence from cognitive training also reveals that the extent of training gains varies considerably among individuals (Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Novick et al., submitted). The reasons for this have not yet been elucidated, but it is conceivable that some individuals may hit an endogenous capacity limit, preventing further improvement. Since bilinguals have already benefitted from extensive practice with language control, they may not be able to improve further through training. Yet accuracy on tasks that have typically been used to assess the bilingual advantage (e.g., Simon, Flanker) is relatively high, so that effects are primarily observed in reaction times, which are constrained by a physical lower bound. Therefore, bilinguals

---

1 It is worth mentioning that some studies (Gollan, Montoya, Cera, & Sandoval, 2008; Ivanova & Costa, 2008; Sandoval, Gollan, Ferreira, & Salmon, 2010) have observed a bilingual disadvantage in speed of lexical access in each of their languages. We believe that bilinguals should enjoy a sentence processing advantage in cases of ambiguity despite these apparent lexical disadvantages, because of their improved cognitive control resources. We return to this issue in the Discussion.
may be performing at ceiling on these tasks, without being at their cognitive control limit, making it impossible to observe continued improvements. To determine whether monolinguals and bilinguals truly benefit differentially from cognitive control practice, experiments must employ alternate tasks allowing greater room for improvement.

The present study therefore sought to integrate research on the bilingual advantage, the role of cognitive control in language processing, and the function of cognitive control training, in order to address three open questions in the bilingualism literature: 1) Does the bilingual advantage stem from enhanced inhibitory control or conflict monitoring abilities?; 2) Does the bilingual advantage extend to language processing when ambiguity is present?; and 3) Do bilinguals and monolinguals benefit differentially from brief cognitive control training, when given equal opportunity to improve?

**Experimental Preliminaries and Predictions**

We used a pretest/posttest training design to test Spanish-Catalan bilinguals and Spanish monolinguals on a sentence processing task involving temporary syntactic ambiguity before and after training on either a high- or low-interference version of an N-back memory task (where N=3). The high-interference version of N-back is a highly demanding cognitive control task that requires resistance to proactive interference (Burgess, Gray, Conway, & Braver, 2011) and has been used successfully in various forms as a training paradigm to improve general fluid intelligence (Jaeggi et al., 2008) and garden-path recovery (Novick et al., submitted). In contrast, the low-interference version omits proactive interference trials such that only recognition memory is required for successful performance; thus, the low-interference N-back serves as a baseline measure of attention and memory, where differences between bilinguals and monolinguals are not expected to emerge because the cognitive control demands are minimal.

Overall, the pretest/posttest training design allowed us to test (i) whether bilinguals exceed monolinguals in their baseline sentence processing abilities, (ii) whether bilinguals and monolinguals improve differentially during brief cognitive control training, and (iii) whether training-related improvements transfer to syntactic ambiguity resolution and if such transfer is dependent on language history. Previous work has found a causal relationship between cognitive control training and syntactic ambiguity resolution (Novick et al., submitted), but the present study extends prior investigations by examining the effect of brief (20-minute) rather than long-term (20 hours over 5 weeks) training and by comparing the effects of training with and without interference resolution. Moreover, because the sentence processing and the high-interference N-back tasks contained both conflict and non-conflict trials, we could assess the specificity of the bilingual advantage to inform whether it is better characterized as an advantage in inhibitory control or in conflict monitoring more generally.

We hypothesized that bilinguals would outperform monolinguals on both the sentence processing and the high-interference N-back task, reflecting bilinguals’ better cognitive control abilities. However, the precise pattern of expected results differs for the IC versus the conflict monitoring accounts of the bilingual advantage. The IC model predicts that bilinguals should be faster and more accurate than monolinguals selectively on conflict trials, i.e. on syntactically ambiguous sentences and on N-back trials requiring interference resolution. In contrast, the conflict monitoring account predicts that bilinguals should be faster and more accurate than monolinguals on both conflict (ambiguous sentences and interference N-back trials) and non-conflict (unambiguous sentences and non-interference N-back trials) trial types, so long as the task demands occasional conflict resolution. Crucially, however, under both accounts, bilinguals should perform equivalently to monolinguals on the low-interference N-back task, because without any conflict, neither conflict monitoring nor conflict resolution should take place. Finally, we hypothesized that—when given equal opportunity to improve—both bilinguals and monolinguals should benefit from brief practice with a cognitive control task (e.g., the high-interference N-back), and that these benefits should transfer selectively to syntactic ambiguity resolution, replicating previous training results (Novick et al., submitted). Moreover, if bilinguals’ cognitive advantage stems from conflict monitoring, then they may achieve greater gains than monolinguals, due to more flexible recruitment of cognitive control following the detection of conflict.
METHOD

Participants
Participants included healthy adult balanced Spanish-Catalan bilinguals (N=59; 7 males; M_age = 20.78, SD_age = 3.38) and Spanish monolinguals (N=51; 12 males; M_age = 26.51, SD_age = 5.94) recruited from the University of Barcelona community. Participants in each language group were randomly assigned to either the high- or low-interference N-back condition. The final distribution included 32 high-interference bilinguals (4 males; M_age=20.53, SD_age=3.15), 27 low-interference bilinguals (3 males; M_age=21.07, SD_age=3.67), 26 high-interference monolinguals (6 males; M_age = 25.54, SD_age = 5.39) and 25 low-interference bilinguals (6 males; M_age = 27.52, SD_age = 6.42). All subjects were given the option of receiving payment (12 Euros) or psychology course credit for their participation; because a choice was given to all subjects, it is unlikely that any observed group differences could be ascribed to motivational factors related to compensation. Also, despite the gender imbalance in the experiment, females accounted for the same high distribution of participants across the two language groups and across the high- and low-interference versions of the N-back task.

Language status was verified using language use and background questionnaires borrowed from Costa et al. (2009; see Appendix B). Bilinguals were included if: their first language was Spanish, Catalan, or both; they had some exposure to both Spanish and Catalan before or during primary school; they continued using both languages through adulthood; they used both languages approximately equally during either childhood or adolescence; they reported at least sufficient proficiency in speaking, writing, listening and reading in both languages; and they were not fluent in a third language. Monolinguals were included if: their first language was Spanish; they had little or no exposure to any language other than Spanish before or during primary school; they did not use Catalan more than a quarter of the time in adolescence; and they were not fluent in speaking or listening comprehension in any language other than Spanish. An additional 25 subjects participated, but were dropped from analyses because they did not meet the requirements for either language group (N=19), because they were less than 75% accurate on filler sentences or non-target N-back trials (N=5; 2 bilinguals), or because of computer error (N=1; monolingual).

Materials and Procedure

We administered a sentence processing task before and after brief training on either a high- or low-interference version of the N-back working memory task. For consistency, the entire experiment was conducted in Spanish for both language groups. This pretest/posttest design allowed us to assess if (1) bilinguals and monolinguals differed in baseline sentence processing abilities, (2) brief training on a working memory task improves sentence processing abilities and if training with interference is necessary for such improvement, and (3) bilinguals and monolinguals respond differentially to brief working memory training.

Sentence processing assessment. Participants completed a moving window self-paced reading task (Just et al., 1992) at pre- and posttest. Two initial lists of Spanish sentences were created, consisting of 32 critical items and 64 fillers each. The critical items were eleven words long and structured to be interpretable as either subject-first or object-first cleft sentences until the seventh, disambiguating word (Betancort, Carreiras, & Sturt, 2009; del Río et al., 2011); however, the subject-first interpretation is strongly preferred. For example:

(1) Este es el general que vigilaba al espía desde la ventana. (Subject-first)
   (This is the general who watched the spy from the window.)

(2) Este es el general que vigilaba el espía desde la ventana. (Object-first)
   (This is the general who the spy watched from the window.)

In Spanish, the subject-first construction is much more frequent and the al/el manipulation results in large ambiguity effects for object-first constructions (Betancort et al., 2009). For instance, relative to subject-first sentences, object-first sentences elicit significant processing difficulty in the disambiguating region of the sentence (e.g., el espía), as indexed by increased first-pass and total reading times (Betancort et al., 2009). Moreover, evidence from magnetoencephalography (MEG) studies demonstrates that this processing difficulty is associated with increased activation in regions implicated in the cognitive control network, namely the LIFG and dorsal lateral prefrontal cortex (del Río et al., 2011), suggesting that readers engage cognitive control to revise their misinterpretation in object-first sentences. Also, on average, participants in del Río et al.’s study
failed to assign the correct meaning to object-first sentences more than 20% of the time, compared with only 5-10% misinterpretation in subject-first sentences. Thus, object-first sentences, but not subject-first sentences, require interference resolution for successful interpretation.

Half the critical items in each list contained “al” (marking subject-first), and half contained “el” (object-first). Additionally, complementary versions of the two lists were created in which the “al” and “el” conditions were swapped for each sentence, such that subject-first sentences from the initial two lists became object-first sentences, and vice versa. Filler sentences were seven to fourteen words long and varied in terms of syntactic structure and complexity. None of the fillers contained temporary syntactic ambiguities, but sixteen of the 64 fillers in each list were designed to be relatively difficult to process compared to the other fillers. The ‘hard’ fillers contained a variety of harder-to-process structures, including multiple embedded prepositional phrases, passive verbal constructions, and fronted direct objects. Each sentence was followed by a True-False comprehension probe (e.g., El general vigilaba al espía (The general watched the spy)). Responding correctly to a critical item probe required correctly processing its syntactic structure. Because comprehending an object-relative clause requires overcoming a strong “subject-first” parsing bias in Spanish, the majority of the critical-item probes (75%) were designed to be false, so that in the case of the object-relative items, subjects would have to successfully reanalyze the sentences in order to respond correctly (syntactic reanalysis is hypothesized to involve cognitive control; see Novick et al., 2005). Filler probes were balanced so that overall, each list contained half True and half False probes. True and False probes occurred with the hard fillers in the same proportions as with the rest of the fillers.

Subjects saw one list before and another list containing different sentences after the N-back task. Sentences were presented in pseudorandom order such that critical items were never adjacent. Presentation of the lists was counterbalanced across subjects.

N-back training task. Within each language group, participants were randomly assigned to a high- or low-interference version of the N-back task, such that 32 bilinguals and 26 monolinguals completed the high-interference version, and 27 bilinguals and 25 monolinguals completed the low-interference version. For this task, 144 four- to eight-letter Spanish nouns and adjectives were selected from the subset of the LEXESP database contained in the BuscaPalabras software tool, using the following criteria: frequency between 20 and 30, familiarity rating between 5 and 7, concreteness rating between 1 and 3.9, and imageability rating between 3.5 and 7 (Davis & Perea, 2005; Sebastián-Gallés et al., 2000).

During the 20-minute training period, all participants completed three separate blocks of N-back, each of which contained 96 trials, lasted about 6.5 minutes, and was followed by a 1-minute break. A different set of words was used in each of the three blocks. In both versions of the task, a series of Spanish words appeared one at a time for 2-seconds each, with a 2-second inter-stimulus interval. Participants indicated as quickly and accurately as possible via button-press whether the current item matched or mismatched the item presented three trials previously, or “3-back.” Importantly, participants were required to make a response on every trial, indicating “yes” for targets (e.g., 3-back matches) and “no” for non-matches. In each block, 3-back targets comprised 50% of the trials. However, in the low-interference version, all non-match trials were non-target words that had not appeared before, whereas in the high-interference version, 75% of the non-match trials (i.e., 36 out of 48 trials) were lure items that had appeared recently, but two, four, or five trials previously, instead of the target three. Thus, while both versions involved maintenance of attention and memory, the high-interference version additionally required conflict resolution on lure trials, where participants had to override their familiarity for the item to correctly reject it as a non-match trial.

RESULTS

N-back Performance

We first examined performance in terms of accuracy and reaction time (RT) on the N-back training task to determine if bilinguals demonstrated better domain-general cognitive control than monolinguals (i.e., whether we replicated the bilingual advantage). There were four participants (1 low-interference monolingual, 1 low-interference bilingual, and 2 high-interference bilinguals) who initially misunderstood the task instructions, and
thus had abnormally low accuracy on Block 1. Consequently, Block 1 was removed from all analyses for these subjects.

**Accuracy.** To correct for boundedness issues, statistics were calculated both on raw proportions and on elogit-transformed proportions for all accuracy data. For clarity, we report only the data from the raw proportion analysis, but both analyses yielded the same pattern of results unless otherwise noted.

Because the high- and low-interference N-back tasks contained different trial types and were thus not directly comparable, we conducted mixed-ANOVAs separately for each interference condition using language group as a between-subjects factor and block and trial type as within-subject factors. In the high-interference condition, we observed a main effect of language group ($F(1, 54) = 5.97, p = .018$), a main effect of block ($F(2, 108) = 12.52, p < .001$), a main effect of trial type ($F(2, 108) = 71.16, p < .001$), and an interaction between block and trial type ($F(4, 216) = 3.59, p = .007$). No other effects reached significance ($ps > .23$). In the low-interference group, there were also significant effects of block ($F(2, 96) = 7.44, p = .001$), trial type ($F(1, 48) = 110.49, p < .001$), and a block-by-trial type interaction ($F(2, 96) = 8.58, p < .001$). But, unlike the high-interference condition, there was no main effect of language group ($F(1, 48) = .034, p = .85$). No other effects reached significance ($ps > .31$). Together, these results suggest that bilingualism influences memory and attention performance only when the task also involves interference resolution, and thus, cognitive control.

![Figure 1](image-url)

**Figure 1.** Overall performance on N-back by trial type and language group for (A) the high-interference task and (B) the low-interference task. (A) In the high-interference task, there was a significant main effect of language group ($p < .05$) because bilinguals were more accurate than monolinguals. There was also a significant main effect of trial type, such that participants were more accurate on non-targets than on lures ($p<.001$) or targets ($p<.001$). (B) In the low-interference task, there was a significant main effect of trial type, such that participants were less accurate on targets than non-targets ($p<.001$), but no effect of language group ($p=.85$).
**High Interference.** In the high-interference task, the main effect of language group was due to higher overall accuracy ($t(56) = 2.36, p = .02$) in bilinguals ($M = .74, SD = .12$) than in monolinguals ($M = .67, SD = .11$; see Figure 1A). This reflected a general bilingual advantage, as opposed to a conflict-specific advantage, since we did not observe a group-by-trial type interaction. From block 1 to block 3, participants demonstrated significant improvement on target trials ($t(55) = -4.46, p < .001$) and marginal improvement on lure trials ($t(55) = -1.92, p = .06$), but not on non-target filler trials ($t(55) = -1.31, p = .20$), although the absence of significant improvement on non-target filler trials may be because participants were near ceiling performance for these trials at block 1 ($M = .94, SD = .10$). Indeed, participants were significantly more accurate on non-target filler trials ($M = .94, SD = .07$) than on lure ($M = .67, SD = .19$; $t(57) = 10.64, p < .001$) or target trials ($M = .67, SD = .16$; $t(57) = 11.56, p < .001$; see Figure 1A). Because target trials required a “yes” response and lure trials required a “no” response, improvement on both trial types suggests that participants did not merely shift their response criterion, but actually improved their sensitivity in distinguishing targets from non-matches. Additionally, the absence of a language group-by-block interaction suggests that participants improved on N-back by the same magnitude regardless of their language status (see Figure 2).

![Figure 2. Accuracy on high-interference N-back by block.](image)

Both language groups demonstrated significant gains ($p < .001$) over the course of the task.

**Low Interference.** In the low-interference task, the source of the block-by-trial type interaction was that participants demonstrated significant improvement from Block 1 to Block 3 on target trials ($M_{\text{improvement}} = .08, SD_{\text{improvement}} = .16$; $t(49) = 3.64, p = .001$) but not on non-target trials ($M_{\text{improvement}} = -.003, SD_{\text{improvement}} = .04$; $t(49) = -0.43, p = .67$), although again, this result might be attributable to near-ceiling performance on non-target trials at block 1 ($M = .98, SD = .04$). Overall, participants were significantly more accurate ($t(51) = 10.64, p < .001$) on non-target ($M = .97, SD = .04$) than on target trials ($M = .73, p < .17$; see Figure 1B), reflecting increased attention and memory demands for target detection.

**RT.** We also analyzed the effect of language experience on RT during the high- and low-interference versions of N-back. Incorrect trials were excluded from this analysis because they may reflect fundamentally different underlying cognitive processes than correct trials (this affected 22% of the data). To reduce the effect of outliers, we replaced all values that were more than 2.5 standard deviations beyond a participant’s mean RT with the 2.5 standard-deviation threshold value.

**High Interference.** In the high-interference condition, a 3 (block) x 3 (trial type) x 2 (language group) mixed-ANOVA revealed significant main effects of block ($F(2, 53) = 28.4, p < .001$) and trial type ($F(2, 53) = 83.70, p < .001$), a significant interaction between block and language group ($F(2, 53) = 3.81, p = .03$), and a significant block-by-trial type interaction ($F(4, 51) = 2.70, p = .04$). There was also a marginal main effect of language group ($F(1, 54) = 2.95, p = .09$), which was due to faster RTs in the bilingual group ($M = 995.50$, ...
$SD=179.53$) than the monolingual group ($M=1098.12, SD=258.78$; see Figure 3). Although this difference did not reach significance, it suggests that the higher accuracy observed for the bilingual group was not merely the result of slower responding (a speed-accuracy tradeoff), but instead appears to reflect more efficient cognitive control processes.

![Figure 3. Reaction time (in ms) on N-back for (A) the high-interference task and (B) the low-interference task.](image)

(A) Bilingual participants were marginally faster than monolinguals in the high-interference condition ($p=.09$). Overall, participants were slower on lures than on non-targets ($p<.001$) or targets ($p<.001$). (B) Bilinguals and monolinguals performed equivalently in the low-interference condition. Participants were slower on targets than on non-targets.

Participants in the high-interference group became faster with practice on N-back ($t(55)=8.76$, $p<.001$), exhibiting significantly slower RTs in block 1 ($M=1156.89, SD=259.59$) than in block 3 ($M=966.12, SD=219.22$). They were significantly slower on lure trials than on non-target ($t(57)=9.49$, $p<.001$) and target trials ($t(57)=11.80$, $p<.001$), demonstrating the increased effort required for interference resolution on lure trials. The overall difference between target and non-lure trials was not significant ($t(57)=-.51$, $p<.61$); however, participants were significantly slower on targets than on non-targets in block 1 ($t(55)=2.72$, $p=.009$), but by block 3, they were numerically faster on targets than on non-targets ($t(57)=-1.64$, $p=.11$). Apparently, although participants became significantly faster from block 1 to block 3 on all three trial types (lures: $M_{diff}=153.79$, $SD_{diff}=207.11$, t(55) = 5.56, $p<.001$; non-lures: $M_{diff}=99.55$, $SD_{diff}=213.28$, t(55) = 3.49, $p=.001$; and targets:
$M_{\text{diff}} = 223.66$, $SD_{\text{diff}} = 233.47$, $t(55)=7.17$, $p<.001$, participants improved more on targets than on non-targets over the course of the N-back task.

An investigation of the language group-by-block interaction for the high-interference condition showed that bilinguals were marginally faster than monolinguals in blocks 1 ($t(54)=-1.97$, $p=.05$) and 3 ($t(56)=-1.80$, $p=.08$), but not in block 2 ($t(56)=-.99$, $p=.33$). However, both language groups improved significantly throughout the course of the task (see Figure 4), with bilinguals improving, on average, 84.86 ms from block 1 to block 2 ($t(29)=3.75$, $p=.001$) and 92.27 ms from block 2 to block 3 ($t(31)=4.73$, $p<.001$) and monolinguals improving, on average, 163.92 ms from block 1 to block 2 ($t(25)=5.42$, $p<.001$) and 49.19 ms from block 2 to block 3 ($t(25)=2.24$, $p=.04$). This parallels the results from the accuracy analyses, where participants improved on the cognitive control task regardless of language status.

**Figure 4. Reaction time (in ms) on the high-interference N-back by block.** Bilinguals were marginally faster than monolinguals on block 1 ($p=.05$) and block 3 ($p=.08$), but not on block 2 ($p=.33$). However, both bilinguals and monolinguals became significantly faster over the course of the task ($ps <.001$).

**Low Interference.** We also conducted a mixed 3 (block) x 2 (trial type) x 2 (language group) ANOVA for the low interference condition, revealing significant main effects of block ($F(2, 96)=44.59$, $p<.001$) and trial type ($F(1, 48)=21.86$, $p<.001$), a significant interaction between block and language group ($F(2, 96)=3.77$, $p=.03$), and a significant block-by-condition interaction ($F(2, 96)=19.09$, $p<.001$). As in the accuracy data, there was no effect of language group in the low-interference condition ($F(1, 48)=.49$, $p=.49$).

Participants in the low-interference condition became significantly faster from block 1 ($M=993.75$, $SD=362.80$) to block 3 ($M=870.83$, $SD=368.33$) on the N-back task ($t(49)=6.88$, $p<.001$). They exhibited significantly slower RTs ($t(51)=4.83$, $p<.001$) for target trials ($M=1022.22$, $SD=401.53$) than for non-target trials ($M=861.57$, $SD=361.81$). In conjunction with the accuracy results, this suggests that even on the low-interference task, participants had greater difficulty identifying 3-back target matches than they did identifying non-target items. Although participants decreased their RTs on both trial types during the low-interference N-back task, for targets they improved significantly from block 1 to block 2 ($t(49)=6.37$, $p<.001$) and block 2 to block 3 ($t(51)=3.57$, $p=.001$), whereas for non-targets they improved only from block 1 to block 2 ($t(49)=2.95$, $p=.005$; block 2 vs. block 3: $t(51)=1.68$, $p=.10$). This suggests that there was a greater benefit of practice for target trials, perhaps indicating improved recognition processes, or a strategic shift to ignore the position information (which was irrelevant for this version of the task). The observed block-by-language interaction resulted from different rates of improvement in the bilingual and monolingual groups, with the monolinguals getting faster from block 1 to block 2 ($t(23)=5.69$, $p<.001$) and from block 2 to block 3 ($t(24)=4.86$, $p<.001$), but the bilinguals demonstrating significant improvements only from block 1 to block 2 ($t(25)=3.45$, $p=.002$; block...
2 vs. block 3: \( t(26)=1.53, p=.14 \). However, the language groups were not significantly different from each other in the low-interference condition at any block (\( ps>.38 \)).

**Sentence Processing Performance**

We examined the effects of bilingual language experience on baseline sentence processing abilities at pretest, as measured by accuracy on sentence comprehension probes and by reading times. If the bilingual advantage is the result of improved inhibitory control specifically, then bilinguals should outperform monolinguals only on temporarily ambiguous sentences requiring re-interpretation. In contrast, if the bilingual advantage is the result of superior conflict monitoring abilities, then bilinguals should outperform monolinguals on all sentence types, including ambiguous items, unambiguous items, and fillers. We explicitly included fillers in our analyses because, regardless of their status, all trials that occur within a task containing conflict require conflict monitoring and should therefore yield a bilingual advantage, at least according to the conflict monitoring account.

We also examined whether cognitive training (i.e., the intervening 3-back task) mediated the relationship between language experience and sentence processing. We expected that participants in the high-interference N-back group would show greater improvements in syntactic ambiguity resolution from pretest to posttest than those in the low-interference group, because only the high-interference group practiced implementing cognitive control.

**Sentence Comprehension.** For sentence comprehension accuracy, statistics were calculated both on raw proportions and on elogit-transformed proportions to correct for effects of boundedness. Unless otherwise reported, both analyses produced the same general pattern of results. Because they are assumed to be more reliable against boundedness issues, we report the statistics for the elogit-transformed data.

To assess baseline differences in comprehension accuracy between bilinguals and monolinguals, we conducted a 2 (language group) x 3 (sentence type) mixed ANOVA at pretest, that is, before participants received training. We did not include interference group as a factor because the interference variable was not introduced until training and so could not influence pretest performance. We observed a significant main effect of trial type (\( F(2, 216) = 247.63, p<.001 \)) and a marginal main effect of language group (\( F(1, 108) = 2.87, p=.09 \)), but no interaction (\( F(2, 216) = .34, p=.71 \)). Accuracy was significantly lower for object-first (\( M= -.48, SD=1.61 \)) than for subject-first sentences (\( M=2.13, SD=1.03 \); \( t(109)=-14.87, p<.001 \)), and for subject-first than for filler sentences (\( M=2.48, SD=.79; t(109)=-3.52, p<.001 \)). The marginal effect of language group was due to higher comprehension accuracy in bilinguals (\( M=1.72, SD=.62 \)) than monolinguals (\( M=1.57, SD=.74 \)).

To assess the influence of training on the relationship between bilingualism and sentence comprehension, we conducted a 2 (language group) x 2 (N-back interference condition) x 2 (assessment) x 3 (sentence type) mixed ANOVA. There were significant main effects of block (\( F(1, 106)=7.32, p=.008 \)) and sentence type (\( F(2, 212)=231.18, p<.001 \)), as well as a significant block-by-sentence type interaction (\( F(2, 212)=8.25, p<.001 \)). No other effects reached significance (\( ps>.10 \)). The interaction occurred because participants improved significantly from pretest to posttest on object-first sentences (\( t(109)=-4.35, p<.001 \)), but not on subject-first (\( t(109)=-.48, p=.63 \)) or filler sentences (\( t(109)=.365, p=.72 \)). Overall, this pattern of results suggests that all participants improved selectively on object-first sentences irrespective of language group, but that the type of N-back training (high versus low) did not influence improvement selectively on object-first sentences from pretest to posttest (as the 3-way interaction between interference, block, and sentence type was not significant; \( F(2, 212)=2.28, p=.11 \)).

However, individual differences in training gains may be critical to the relationship between training and transfer. Previous training research has shown that those individuals who achieve the greatest gains during training also demonstrate the largest improvements on transfer tasks (Jaeggi et al., 2011; Novick et al., submitted). We therefore examined the continuous relationship between improvement on N-back and improvement on sentence comprehension. We predicted that only improvements on lure trials should correlate with improvement on object-first sentences, since only lure trials tapped the underlying cognitive control...
resources hypothesized to implement ambiguity resolution in object-first sentences. Indeed, we found that performance gains on lure trials from block 1 to block 3 of N-back were significantly positively correlated with the amount of improvement on object-first sentence comprehension accuracy from prettest to posttest (r=.28, p=.04; see Figure 5), whereas improvements on other N-back trial types (targets and non-targets for either interference group) did not predict object-first accuracy gains (ps>.18), even though both groups exhibited significant accuracy gains in target performance. The one exception was that non-target gains in the high-interference group were significantly negatively correlated with improvements on object-first sentences (r=-.27, p=.04), perhaps because those individuals who gained substantially in non-target accuracy in the high-interference group were training recognition processes instead of interference resolution processes. Importantly, improvement on lure trials was not related to gains on the other sentence types that did not involve syntactic ambiguity resolution (i.e., subject-first and filler sentences; ps>.30), suggesting that lure gains selectively influenced performance on object-first sentences, as one might predict under a domain-general cognitive control account (Botvinick et al., 2001; Novick et al., 2005).

**Figure 5.** Correlation between lure accuracy gains and object-first sentence comprehension improvements. For each subject, lure accuracy gains were computed by subtracting percent correct on lure trials at block 1 from block 3. Similarly, object-first accuracy gains were calculated by subtracting percent correct on object-first comprehension probes at pretest from posttest.

**Reading Times.** Only experimental items were analyzed, and only trials with correct responses were included in the analysis, since incorrect responses may reflect the absence of the cognitive processes of interest. The final word of each sentence was also excluded, since it was presumed that wrap-up effects and other noise in the region would be likely to either obscure the effects of interest or to create spurious effects.

In order to reduce the potential effect of outliers, the data was Winsorized: reading times that fell more than 2.5 standard deviations from the mean for each subject were set to the 2.5 standard deviation threshold value. This procedure affected 11.78% of the data. The data was then residualized for each participant in order to factor out the effect of word length on reading time duration as well as individual differences in reading times (Ferreira & Clifton, 1986; Trueswell, Tanenhaus & Garnsey, 1994).

The pretest data was initially examined separately in order to assess baseline differences between monolinguals and bilinguals. The residualized reading times for each of the first ten words in the sentence were analyzed using linear mixed effects models with fixed effects for Group (monolingual/bilingual), Trial Type (subject-/object-first), and the interaction between those factors; and crossed random effects for Subject and Item. In order to assess the effect of the N-back training intervention, the residualized reading times for pre- and

---

2 All of the correlations reported were calculated using exclusively elogit-transformed data to be consistent with the sentence comprehension analyses.
posttest combined were then examined in the same fashion with additional fixed effects for Block (pre/posttest), the Interference level of the intervening N-Back task (high/low interference), and all possible interactions between the fixed effects. The models were built using the lmer function from the package “lme4” in R. P-values were obtained through Markov chain Monte Carlo sampling using the pvals.fnc function from the package “languageR” in R. Mean raw and residualized reading times for the combined pre- and posttest data are presented in Table 1.

Table 1. Mean raw (uncorrected) and residual reading times for the disambiguating regions of the subject- and object-first items, pooled across pre- and posttest and across monolinguals and bilinguals.

<table>
<thead>
<tr>
<th>Sentence Type</th>
<th>Word7</th>
<th>Word8</th>
<th>Word9</th>
<th>Word10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>492.63</td>
<td>723.11</td>
<td>493.81</td>
<td>418.91</td>
</tr>
<tr>
<td>Object</td>
<td>528.09</td>
<td>992.80</td>
<td>640.69</td>
<td>507.98</td>
</tr>
<tr>
<td>Difference</td>
<td>35.46</td>
<td>269.70</td>
<td>146.88*</td>
<td>89.07*</td>
</tr>
</tbody>
</table>

Note. *pMCMC≤.05. Negative residual values reflect faster reading times than predicted given word length; positive residuals reflect slower reading times than predicted given word length.

Since word 7 (el/al) was the disambiguating region, and since the subject and object-first items were identical up to this region, the primary regions of interest were words 7-10. At pretest, significant effects for Trial Type at words 8, 9, and 10 (p=.0001, p=.0001, p=.028) reflect the fact that participants were significantly slower on object-first items at these regions, thus replicating the well-attested garden-path effect (Table 2). There were no other significant effects at pretest, indicating that monolinguals did not differ from bilinguals at baseline.

Table 2. Fixed effects with significant p-values for residualized reading times at pretest.

<table>
<thead>
<tr>
<th>Region</th>
<th>Significant Model Parameters</th>
<th>Beta Estimate</th>
<th>SE</th>
<th>pMCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word1</td>
<td>Intercept</td>
<td>-66.37</td>
<td>10.01</td>
<td>.0001</td>
</tr>
<tr>
<td>Word6</td>
<td>Intercept</td>
<td>34.12</td>
<td>15.14</td>
<td>.0238</td>
</tr>
<tr>
<td>Word7</td>
<td>Intercept</td>
<td>119.44</td>
<td>13.67</td>
<td>.0001</td>
</tr>
<tr>
<td>Word8</td>
<td>Intercept</td>
<td>211.34</td>
<td>18.95</td>
<td>.0001</td>
</tr>
<tr>
<td>TrialTypeSubject</td>
<td>-91.98</td>
<td>20.82</td>
<td>.0001</td>
<td></td>
</tr>
<tr>
<td>Word9</td>
<td>Intercept</td>
<td>114.70</td>
<td>13.98</td>
<td>.0001</td>
</tr>
<tr>
<td>TrialTypeSubject</td>
<td>-61.04</td>
<td>15.75</td>
<td>.0001</td>
<td></td>
</tr>
<tr>
<td>Word10</td>
<td>Intercept</td>
<td>40.43</td>
<td>10.93</td>
<td>.0002</td>
</tr>
<tr>
<td>TrialTypeSubject</td>
<td>-26.19</td>
<td>11.93</td>
<td>.0280</td>
<td></td>
</tr>
</tbody>
</table>

Note. Parameters are listed only if they significantly improved the fit of the model (pMCMC<.05). Negative parameter estimates indicate faster reading times, while positive parameter estimates indicate slower reading times.
Table 3 reports all significant effects for the combined pre- and posttest data. The garden-path effect persisted in the combined data, as indicated by significant effects for Trial Type in words 8, 9, and 10 (\(p<.05\)). Participants exhibited a predictable practice effect: they were faster on the posttest block than in the pretest block, as indicated by significant effects for Block at every word (\(p<.001\)). There were also significant main effects of Group at word 1 (\(p=.0001\)) and Interference at words 1 and 6 (\(p<.05\)). Additionally, the following interactions were significant: Group*Interference (word 1, \(p<.01\)); Group*Block (words 4 and 7, \(p<.05\)); Group*Trial Type*Interference (word 10, \(p<.05\)); Group*Interference*Block (word 10, \(p<.05\)); and Group*Trial Type*Interference*Block (word 10, \(p<.01\)).

To confirm that the garden path effect observed at words 8, 9, and 10 was not due to the pretest data alone, the posttest data was analyzed separately using the same model specifications used for the pretest data above plus an additional fixed effect for Interference (high/low) and the corresponding additional possible interactions. Significant effects of Trial Type were found in the same regions (\(p<.05\)), confirming that the garden-path effect was present at posttest (Table 4). Significant main effects were also found for Group (word 1, \(p=.0001\); word 6, \(p<.05\)) and Interference (words 1 and 6, \(p<.001\)). Significant interaction effects were found for Group*Interference (word 1, \(p<.001\); word 6, \(p<.05\)) and Group*Trial Type*Interference (word 10, \(p<.05\)).

Table 3. Fixed effects with significant \(p\)-values for residualized reading times for combined pre- and posttest data.

<table>
<thead>
<tr>
<th>Region</th>
<th>Significant Model Parameters</th>
<th>Beta Estimate</th>
<th>SE</th>
<th>pMCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word1</td>
<td>Intercept</td>
<td>-136.16</td>
<td>10.23</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>GroupMonolingual</td>
<td>64.55</td>
<td>14.91</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>InterferenceLow</td>
<td>36.92</td>
<td>14.94</td>
<td>0.0128</td>
</tr>
<tr>
<td></td>
<td>BlockPost</td>
<td>-55.83</td>
<td>14.11</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>GroupMonolingual:InterferenceLow</td>
<td>-66.66</td>
<td>22.15</td>
<td>0.0024</td>
</tr>
<tr>
<td>Word2</td>
<td>Intercept</td>
<td>-71.41</td>
<td>8.96</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>BlockPost</td>
<td>-71.99</td>
<td>12.56</td>
<td>0.0001</td>
</tr>
<tr>
<td>Word3</td>
<td>Intercept</td>
<td>-66.81</td>
<td>8.64</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>BlockPost</td>
<td>-58.86</td>
<td>12.49</td>
<td>0.0001</td>
</tr>
<tr>
<td>Word4</td>
<td>Intercept</td>
<td>-116.99</td>
<td>15.07</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>BlockPost</td>
<td>-144.60</td>
<td>20.52</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>GroupMonolingual:BlockPost</td>
<td>67.32</td>
<td>32.26</td>
<td>0.0388</td>
</tr>
<tr>
<td>Word5</td>
<td>Intercept</td>
<td>-60.71</td>
<td>10.86</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>BlockPost</td>
<td>-69.70</td>
<td>15.62</td>
<td>0.0001</td>
</tr>
<tr>
<td>Word6</td>
<td>Intercept</td>
<td>-120.07</td>
<td>15.62</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>InterferenceLow</td>
<td>57.09</td>
<td>23.55</td>
<td>0.0148</td>
</tr>
<tr>
<td></td>
<td>BlockPost</td>
<td>-145.35</td>
<td>22.23</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
Note. Parameters are listed only if they significantly improved the fit of the model (pMCMC<.05). Negative parameter estimates indicate faster reading times, while positive parameter estimates indicate slower reading times.

Table 4. Fixed effects with significant p-values for residualized reading times at post-test.
<table>
<thead>
<tr>
<th></th>
<th>GroupMonolingual</th>
<th>InterferenceLow</th>
<th>GroupMonolingual:InterferenceLow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>43.81</td>
<td>21.95</td>
<td>0.0480</td>
</tr>
<tr>
<td></td>
<td>59.08</td>
<td>22.00</td>
<td>0.0090</td>
</tr>
<tr>
<td></td>
<td>-63.16</td>
<td>32.62</td>
<td>0.0478</td>
</tr>
</tbody>
</table>

**Word7**

Intercept: 44.72

**Word8**

Intercept: 62.18

TrialTypeSubject: -53.83

**Word9**

Intercept: 29.73

TrialTypeSubject: -38.47

**Word10**

TrialTypeSubject: -34.24

GroupMonolingual:TrialTypeSubject:InterferenceLow: -61.44

Note: Parameters are listed only if they significantly improved the fit of the model (pMCMC<.05). Negative parameter estimates indicate faster reading times, while positive parameter estimates indicate slower reading times.

**DISCUSSION**

**N-back performance**

The results of both the accuracy and RT analyses of N-back performance indicated that bilinguals outperformed monolinguals, but only on the high-interference version, which involved conflict and thus cognitive control, manipulated via the presence of lure trials. No such advantage emerged on the low-interference N-back task, which involved the maintenance of attention and memory, but contained no lure-trial interference, and thus did not require cognitive control. This divergence of patterns across the two versions of N-back is critical; if we had also observed an advantage on the low-interference N-back, then our results would have suggested that bilinguals merely paid better attention than monolinguals, as cognitive control abilities should not deploy in the total absence of conflict. Instead, however, we found a bilingual advantage only on the task that involved occasional conflict, confirming that the advantage reflects improved cognitive control, rather than better attention or memory. Said another way, bilinguals do not appear to enjoy an advantage in the mnemonic aspects of working memory, when information must be temporarily processed and stored for ongoing use absent interfering representations; rather, their benefit emerges only when the demands for non-mnemonic control processes are relatively high, namely when interference must be detected and resolved throughout a particular task context.

Interestingly, although the bilingual advantage was selective for the task involving conflict, it was not specific to trials involving conflict; language group and trial type did not interact on the high-interference N-back task, where we observed better performance for the bilingual group. This lack of specificity challenges the inhibitory control account of the bilingual advantage (Green, 1998), which predicts that bilinguals should be faster and more accurate than monolinguals only on trials requiring conflict resolution—in this case, lure trials on the N-back task. However, our results are more compatible with the conflict monitoring account (Costa et al., 2009), wherein bilinguals should demonstrate a general advantage on any task requiring the detection of conflict and flexible adjustments in cognitive control. Under such a theory, the bilingual advantage emerges because the occasional presence of conflict heightens monitoring demands and therefore the readiness of cognitive control functions to deploy. Consequently, performance on even non-conflict trial types is better because of this increased state of awareness.
In contrast to previous studies, which may have been susceptible to artificial task-ceiling effects, we showed that both bilinguals and monolinguals demonstrate marked improvements during practice on a cognitive control task. Indeed, regardless of language group, participants in the high-interference condition increased their N-back accuracy by nearly 7%. In terms of reaction time, the group by block interaction suggested that bilinguals and monolinguals improved at different rates; however, bilinguals still became significantly faster with practice, and monolinguals never achieved bilingual-levels of performance. This novel finding is important because it suggests that despite the fact that bilinguals already possess better conflict monitoring and cognitive control abilities, they are still able to benefit from further training. Moreover, such a benefit can be realized within a relatively short timeframe (e.g., 20 minutes).

Sentence processing performance

Bilinguals exhibited a small, non-specific advantage over monolinguals in offline sentence processing prior to training, as evidenced by higher accuracy to comprehension probes following all trial types (ambiguous, unambiguous, and filler). However, bilinguals did not appear to be better than monolinguals at online sentence processing (i.e., moment-by-moment reading time). The pattern of a bilingual advantage in reading comprehension but not in online sentence processing suggests that the observed advantage may impact late-stage semantic-integration processes. However, it is worth noting that prior studies have observed slower lexical access in bilinguals relative to monolinguals (for review, see Bialystok et al., 2009), either because lexical frequency for individual words is reduced in bilinguals (Gollan, Montoya, Cera, & Sandoval, 2008) or because competition for word selection is increased in bilinguals due to interference from the irrelevant language (Sandoval, Gollan, Ferreira, & Salmon, 2010). It is therefore likely that bilinguals may suffer a measurable disadvantage at the early stages of sentence processing (e.g., lexical retrieval), but might be able to compensate for it in comprehension due to their increased cognitive control.

Crucially, the bilingual advantage in sentence comprehension was not selective for sentences requiring ambiguity resolution. These results are strikingly parallel to our findings on the N-back task, and further corroborate the idea that bilinguals are better at conflict detection and the flexible recruitment of cognitive control rather than conflict resolution per se. Again, however, we would not expect a global bilingual advantage in sentence comprehension in the complete absence of temporarily ambiguous sentences; indeed, the relatively low proportion of garden-path sentences (17%) in our task may be responsible for the small magnitude of the bilingual advantage in sentence comprehension, as monitoring demands are relatively low. The conflict monitoring theory predicts that the bilingual advantage should be largest when the need to monitor for conflict is high—that is, when switching between conflict and non-conflict trials is frequent. This prediction was confirmed by Costa et al. (2009), who found that when a high-proportion of trials on the Flanker task were the same type (either conflict or non-conflict), the bilingual advantage disappeared. Future studies should determine whether the bilingual advantage in sentence comprehension could be increased with a higher degree of switching between temporarily ambiguous and unambiguous sentences.

We found a selective relationship between training-related gains on lure trials and improvement from pretest to posttest on object-first comprehension accuracy. This finding provides supporting evidence for the notion that overcoming misinterpretation during language processing requires cognitive control abilities, and suggests that cognitive control can be enhanced via training. Remarkably, this training-related improvement occurred within a single session, after merely 20-minutes of practice. Thus, our study extends prior work on cognitive training to show that even short-term practice may benefit cognitive control abilities, and that both bilinguals and monolinguals can reliably improve depending on their performance increases on conflict trials during training. Though training studies typically enact long-term regimens over the course of several weeks, the current findings suggest the urgency of future research to test the effectiveness of shorter “booster” sessions, to determine the relative benefits of intervention types of varying length. Moreover, we provide a cross-linguistic replication of the finding that improvement over the course of cognitive control training predicts gains in garden-path recovery (Novick et al., submitted); to our knowledge, this is the first demonstration of the effect of training on sentence processing in Spanish.
The existence of a bilingual advantage in cognitive control is evident from our study, as well as from previous research. However, we are only beginning to understand the exact nature and extent of this advantage. If the bilingual advantage is better characterized as superior conflict monitoring than inhibitory control, then the mechanisms that would strengthen conflict monitoring in bilinguals need to be delineated. In other words, what are the processing demands associated with certain bilingual communities that may confer a conflict monitoring advantage? Although many studies have demonstrated the need for cognitive control in bilingual language production, few have investigated the role of cognitive control in bilingual language comprehension. Bilinguals in code-switching environments may have an especial need to monitor for conflict, because they may be charged with detecting unpredictable language switches and flexibly deactivating and reactivating lexical items. Unlike bilinguals living in single-language environments, code-switchers may not globally inhibit the language not currently in-use, but instead maintain activation of both languages (Green, 2011) in order to facilitate switching. Future studies should examine the extent to which code-switching comprehension requires conflict monitoring.

In conclusion, bilingual language experience apparently acts as a form of cognitive control training, bestowing measurable advantages in conflict monitoring—the ability to detect unpredictable conflict and flexibly adjust the recruitment of cognitive control resources. We demonstrate that this advantage applies not only to a general cognitive control task, but also to sentence processing involving occasional syntactic ambiguity resolution, suggesting that the conflict monitoring system acts across syntactic and non-syntactic domains. Moreover, this system continues to be amenable to improvement, as both bilinguals and monolinguals showed substantial gains over the course of practice. Taken together, our results support the notion of a flexible cognitive control system that underlies syntactic ambiguity resolution and resistance to proactive interference, and can be improved via cognitive training, including bilingual language experience.
REFERENCES


