Published in final edited form as:

Linguist Approaches Biling. 2016 ; 6(1-2): 86–118. doi:10.1075/lab.15003.iva.

Does Bilingual Language Control Decline in Older Age?

Iva Ivanova^{a,*}, Mayra Murillo^b, Rosa I. Montoya^b, and Tamar H. Gollan^b

^aDepartments of Psychiatry and Psychology, University of California, San Diego

^bDepartment of Psychiatry, University of California, San Diego

Abstract

We investigated age-related decline of bilingual language control. Thirteen older and 13 younger bilinguals performed a verbal fluency task (completing the same letter and semantic categories in each language and switching languages after every category), and a non-linguistic flanker task. In letter fluency, bilinguals produced fewer correct responses after switching languages, suggesting inhibition of the previously-used language. However, this testing-order effect did not differ between groups and older bilinguals produced few wrong-language intrusions, implying intact ability to apply inhibition in older age. In contrast, age-related deficits in the flanker task were robust, implying dissociations between language control and domain-general executive control. In semantic fluency, there were no testing-order effects but older bilinguals produced more intrusions than younger bilinguals, and more intrusions than in letter fluency. Thus, bilinguals may flexibly modulate the degree of inhibition when they can benefit from semantic priming between languages, but less efficiently so in older age.

Keywords

bilingualism; language control; domain-general executive control; age-related decline; verbal fluency task; letter fluency; semantic fluency

1. Introduction

Being a bilingual is in some ways similar to being a railroad switch operator: It is necessary to control which language stops and which language goes. If bilinguals let the wrong language go, monolingual conversation partners will not understand them, and communication will break down. In general, bilinguals are very good at this task – unintended languages switches are rare (1%; Gollan, Schotter, Gomez, Murillo, & Rayner, 2014; Poulissee, 1999), but seem to increase in older age (Gollan, Sandoval, & Salmon, 2011), although experimental investigations of this topic have been scarce.

Researchers generally agree that bilinguals avoid wrong-language intrusions with the help of *language control mechanisms*, which regulate the activation of the two languages and may inhibit the wrong language (e.g., Green, 1998). Language control is assumed to function

*Corresponding author: 9500 Gilman Drive, 0948, Department of Psychiatry, University of California, San Diego, La Jolla, CA 92093-0948, USA, Tel.: +1 858 246 1278, Fax: +1 858 246 1284, iva.m.ivanova@gmail.com.

through the same mechanisms which regulate any type of behavior – broadly known as (*domain-general*) *executive control mechanisms*. These include planning and execution of tasks, task-switching, conflict monitoring and conflict resolution and involve inhibitory control, working memory, decision-making and problem-solving. Domain-general executive control mechanisms that might be particularly relevant for bilingual language control are those that enable engagement of the appropriate task, avoidance of irrelevant or inappropriate actions (for example by applying inhibition) and detection of conflicting stimuli or situations (e.g., Norman & Shallice, 1986). Here, we investigate to what extent language control mechanisms are impaired or preserved in older age, and explore the possibility that some aspects of language control might be independent of domain-general executive control (Gollan et al, 2014).

1.2 Language Control in Older Age

Language control might decline in older age, because it may rely on domain-general executive control, which declines in older age. Consistent with this possibility, bilinguals sometimes outperform monolinguals on non-linguistic measures of executive control (Bialystok, Craik, & Luk, 2008; Bialystok, Craik, Klein, & Viswanathan, 2004; Costa, Hernández, & Sebastián-Gallés, 2008; but see Paap & Greenberg, 2013). These processing advantages in bilinguals may reflect their daily practice with controlling their two languages. For example, bilinguals who reported switching languages often in daily life exhibited smaller language-switching costs, and also smaller non-linguistic task-switching costs, than bilinguals who switch languages less often (Prior & Gollan, 2011; Tao, Taft, & Gollan, 2015; but see Hernández, Martin, Barceló, & Costa, 2013; see also Calabria, Hernández, Branzi, & Costa, 2011). Often-switching bilinguals also outperformed rarely-switching bilinguals on two other non-linguistic executive control tasks (flanker and Simon tasks, see below: Verreyt, Woumans, Vandelandotte, Szmalec, & Duyck, in press). These findings suggest a relationship between the control bilinguals exercise over their two languages on a daily basis and their performance on domain-general executive control tasks. Such a relationship implies that bilinguals recruit domain-general executive control mechanisms to achieve language control.

Furthermore, there is an emerging consensus among researchers of cognitive aging that at least some aspects of domain-general executive control decline in older age (see Verhaeghen, 2011) independently of age-related general slowing (Myerson, Hale, Wagstaff, Poon, & Smith, 1990; Salthouse, 1985). One example for such decline is performance on the Simon task (Simon & Ruddell, 1967). In this task, participants are instructed to respond to a stimulus dimension (e.g., press a button on the left when they see color X and a button on the right when they see color Y) independent of the stimulus position on the computer screen. However, participants tend to respond more slowly when the stimulus appears on the opposite side of the screen from the response hand (e.g., when color X appears on the right), than when it is on the same side (i.e., there is a *conflict effect*). The Simon task thus measures inhibition abilities because, to perform well, participants have to inhibit the automatic tendency to give a response congruent with the stimulus location. In this task, older participants exhibit a larger conflict effect than younger participants even after adjustments are made to account for age-related general slowing, suggesting that inhibitory

control declines in older age (van der Lubbe & Verleger, 2002; Bialystok et al., 2008; but see Verhaeghen, 2011, for counterevidence). Furthermore, older individuals also exhibit larger task-mixing (and to some degree also task-switching) costs than younger individuals (Wasylyshyn, Verhaeghen, & Sliwinski, 2011; for aging effects in bilinguals on related domain-general control tasks, see Goral, Campanelli, & Spiro III, 2015).

Despite expectations based on the evidence reviewed above, language control does not always seem to follow exactly the same decline trajectory as domain-general executive control, suggesting at least some dissociation between the two. In picture naming with cued language switching, older bilinguals responded more slowly, and also failed to switch languages more often when cued to do so, than younger bilinguals (Weissberger, Wierenga, Bondi, & Gollan, 2012; Hernandez & Kohnert, 1999). However, when allowed to switch freely (i.e., when instructed to use “whichever language comes to mind” to name pictures), older bilinguals switched languages as often as younger bilinguals, and exhibited little to no increase in switching and mixing costs (Gollan & Ferreira, 2009). Furthermore, an explicit comparison of language switching with non-linguistic task switching (switching between making color or shape judgments) revealed different patterns of age-related decline on the two tasks suggesting some dissociation in the control mechanisms involved (Weissberger et al., 2012). Specifically, the color-shape task showed greater decline than the language task. Additionally, a subset of older bilinguals who could not do the color-shape task were easily able to do the language-switching task. Further evidence for a dissociation between language control and domain-general executive control in older age comes from the study of Calabria, Branzi, Marne, Hernández, and Costa (2013), who tested younger, middle-aged and older highly-proficient bilinguals on a language-switching and a non-linguistic-switching task. They found that both tasks exhibited age-related general slowing. However, there was a relationship between switching costs and age only for the non-linguistic but not the linguistic task, and there was no relationship between the switching costs in the two tasks.

Most relevant to the current investigation are studies on language control in older bilinguals using the verbal fluency task. In this task, participants are required to generate as many words as possible belonging to a specific letter category (words starting with a specific letter, e.g., *F*) or semantic category (e.g., *animals*) in a given time period (typically 1 minute). To complete the task, speakers must search for, access, phonologically encode, and then produce task-appropriate lexical items, while also monitoring prior responses to avoid repetitions and inhibit non-category members (Henry, Crawford, & Phillips, 2004). For bilinguals, these include responses that fit the category perfectly except for being in the wrong-language – a particularly challenging test of monitoring mechanisms (Sandoval, Gollan, Ferreira, & Salmon, 2010).

Verbal fluency performance declines in older compared to younger speakers (Delis, Kaplan, & Kramer, 2001; but see Axelrod & Henry, 1992; Boone, 1999), and is weaker in bilinguals compared to monolinguals (e.g., Gollan, Montoya, & Werner, 2002) and in the non-dominant relative to the dominant language (Sandoval et al., 2010). In both cases, group differences in performance are larger on semantic than on letter fluency (older relative to younger speakers: Brickman, Paul, Cohen, Williams, MacGregor, et al., 2005; Crossley, D’Arcy, & Rawson, 1997; Kozora & Cullum, 1995; Tombaugh, Kozak, & Rees, 1999;

Tomer & Levin, 1993; bilinguals relative to monolinguals: Gollan et al., 2002; Portocarrero, Burright, & Donovanick, 2007; Rosselli, Ardila, Araujo, Weekes, Caracciolo et al., 2000). Bilinguals have also been shown to differ from monolinguals with respect to the time course of retrieval. Bilinguals were slower to produce the first response and had delayed retrieval of items over the course of the minute-long trial relative to monolinguals (Sandoval et al., 2010, Luo, Luk, & Bialystok 2010; but see Ljungberg, Hansson, Andrés, Josefsson, & Nilsson, 2013). The delayed time-course of retrieval of category members is consistent with the assumption that bilinguals face interference from the non-target language during fluency generation – particularly at the beginning of the fluency trial, and particularly in semantic fluency (Sandoval et al., 2010).

Of particular relevance here, Gollan et al. (2011) employed the verbal fluency task to explore how language control declines in older age. They focused on unintended language switches (i.e., production of words in the non-target language), or *wrong-language intrusion errors*, which evidence language control difficulties. In this study, the oldest bilinguals, and particularly bilinguals who also showed impairment in the flanker task (a non-linguistic conflict-resolution task), produced a greater proportion of intrusion errors. The latter result implies an explicit link between domain-general executive control and language control. However, another finding from the same study suggests a very different conclusion. Despite serious deficits in domain-general executive control for some of the older bilinguals, their overall rate of wrong-language intrusion errors was extremely low (between 0 and 3%), even for those individuals who had pronounced difficulty performing the flanker task. This result implies substantial independence between language control and domain-general executive control mechanisms. Specifically, it suggests the presence of mechanisms forming part of language control but not of domain-general executive control, which are relatively spared in older age.

Taken together, most results reviewed so far suggest that non-linguistic executive control and language control might be differentially affected by older age, with language control being relatively more preserved. However, given the small number of studies on how older age may alter language control, we aimed to extend this investigation by examining performance in a modified verbal fluency task designed to be especially demanding for language control mechanisms.

1.2. The Present Study

We investigated age-related decline of language control by comparing older and younger bilinguals' performance on the letter and semantic fluency tasks. To obtain a measure of how older age affects aspects of domain-general executive control which could overlap with language control, we administered the flanker task. We additionally administered the digit span task as a measure of how older age affects other aspects of executive control such as working memory.

In an attempt to maximize language control requirements in the fluency task, we asked speakers to produce exemplars in the same category first in one language, henceforth *first naming language*, and immediately thereafter in the other language, henceforth *second naming language*; bilinguals thus switched languages after every fluency trial. Order effects

(reduced fluency in a language when it was the second naming language relative to when it was the first naming language) were our first measure of the efficiency of language control. We assumed that, during production in the first naming language, the second naming language needs to be inhibited. When the second naming language is then required on the subsequent trial, it takes time and effort to release the inhibition, and fluency in this language is reduced.

Furthermore, order effects tend to show an asymmetry between the dominant and the non-dominant language (Green, 1998). Specifically, the dominant language, being stronger, must be inhibited more strongly to allow production in the non-dominant language than vice versa. For this reason, language switches back to the dominant language are sometimes more effortful than switches to the non-dominant language (Meuter & Allport, 1999).

Accordingly, in a similar paradigm to the one used here (using the same letters in both languages but without language switches after every trial), Van Assche, Duyck, and Gollan (2013) found asymmetrical testing-order effects for college-aged bilinguals. Specifically, dominant language fluency was significantly reduced by prior production of category members in the non-dominant language, whereas the non-dominant language was unaffected by prior production in the dominant language.¹

In addition to order effects, we also examined wrong-language intrusion errors as an index of language control abilities (e.g., Gollan et al., 2011), insofar as the production of responses in the wrong language arguably provides the most direct evidence for a failure of language control.

Our predictions were as follows. In the letter task, we first expected to observe language order effects for both older and younger bilinguals, possibly limited to the dominant language, based on Van Assche et al.'s (2013) results. Second, inhibitory control might decline in older age (Hasher & Zacks, 1988; Zacks & Hasher, 1994). In our task, such decline might have several different manifestations in our two indexes of language control, order effects and wrong-language intrusion errors. One possibility is that older bilinguals are less efficient at applying inhibition on the previous trial than younger bilinguals. Recovery from that smaller amount of inhibition should be relatively easy, and hence we should observe smaller order effects (greater fluency in the second naming language) for older than for younger bilinguals. In addition, if older bilinguals are less efficient in applying inhibition, they should also produce more wrong-language intrusions than younger bilinguals. Another possibility is that older age does not reduce the ability to apply inhibition, but instead reduces the ability to *release* inhibition once it has been applied. If so, we should observe larger language order effects (reduced fluency in the second naming language) for older than for younger bilinguals, but a similar number of intrusions for older and younger bilinguals, since efficient application of inhibition should minimize the number of intrusions.

¹Note that such a testing order effect appeared also when Van Assche et al. (2013) used *different* letters in the two languages. The effect, however, only held for Chinese-English but not Dutch-English bilinguals.

Our predictions for the semantic task were less clear. As noted above, between-group differences in performance are sometimes greater for semantic than for letter fluency (including differences between older and younger participants as well as between bilinguals and monolinguals). In the current study, the semantic task differed from the letter task in that it required bilinguals to produce words with the same meanings first in one and then immediately afterwards in the other language. One possibility is that such frequent switching between words with the same meanings increases language control demands relative to the letter task. This would be because the second naming language requires the production of the exact same words which were just inhibited as they were the most likely source of interference (they have the same meaning) during production in the first naming language. For example, when producing exemplars in the category Vegetables in Spanish (e.g., *zanahorias* [carrots], *espinacas* [spinach], *lechuga* [lettuce]), their translations need to be inhibited; but, on the very next trial, Vegetables in English, those same translations (*carrots*, *spinach*, *lettuce*) are the appropriate responses. In this scenario, we would expect a similar pattern to the one for the letter task, with possibly larger order effects, and larger differences between older and younger bilinguals.

Another possibility is that the second naming language might benefit from semantic priming, leading to smaller order effects than in the letter task (or absent order effects) for both groups. This is because concepts activated for production in the first naming language would boost the activation of translation equivalents in the second naming language, and such priming could partially or completely offset inhibition effects (and possibly age-related decline). However, while having a positive effect on the production of correct responses, such a scenario may increase the number of wrong-language intrusions, specifically for older bilinguals and maybe in the second naming language, because the conceptual boost might lead to less efficient inhibition of the first naming language.

Further predictions concern the extent to which language control (measured here by the fluency task) and domain-general executive control (measured by the flanker task) use overlapping mechanisms. If language control relies completely on domain-general executive control, we should observe age-related decline in both the fluency and flanker tasks. Specifically, in the flanker task, older bilinguals should show a larger conflict effect than younger bilinguals (disproportionately slower and less accurate responses on incongruent relative to congruent trials). In the fluency task, they should produce fewer correct responses, different order effects (see above) and a higher proportion of language intrusions than younger bilinguals. Conversely, if language control is independent (or partially independent) of domain-general executive control, we should observe age-related decline in the flanker task but in no, or only some but not all, aspects of the fluency task.

Finally, we predicted that older bilinguals would perform worse than younger bilinguals on the digit span task (Bopp & Verhaeghen, 2005).

2. Method

2.1 Participants

Thirteen older Spanish-English bilinguals participated in the study. Six were recruited from the Alzheimer's Disease Research Center at the University of California, San Diego (UCSD). These participants were diagnosed as cognitively intact by two neurologists on the basis of medical, neurological, and neuropsychological exams, in the same year as they participated in our study. The remaining 7 participants were recruited during the Health Fair at the Kimball Park Senior Center in National City (CA), and by means of flyers placed in various San Diego County Libraries. They were classified as cognitively intact on the basis of reported daily activities, scores on the Dementia Rating Scale (DRS; Mattis, 1988), and scores on the Mini Mental Status Examination (MMSE; Folstein, Folstein, & McHugh, 1975).

Nine of the 13 participants were followed longitudinally by the ADRC for two years after their participation in this study (three joined the ADRC after their participation in our study), and seven of these nine retained a "normal" diagnosis. Two were diagnosed with Mild Cognitive Impairment (MCI) in the following year.² Of the remaining participants, three had DRS scores within the normal range upon participation in another non-ADRC bilingual study two years later, while, around the same time, the last participant was "showing cognitive decline" according to family members. This participant's data were nevertheless included in the present study based on DRS scores above the group average at the time of testing, and which had not declined (in fact, had increased slightly) on the testing date relative to her performance three years prior to participation. One additional participant was excluded because she was diagnosed with MCI 7 days after participating in this study, and the following year her diagnosis was changed to probable Alzheimer's Disease.

In addition, 64 Spanish-English, college-aged bilinguals participated in the study for course credit, and 13 of these were matched to the older bilinguals for English and Spanish picture-naming scores using the Multilingual Naming Test (MINT; Gollan, Weissberger, Runnqvist, Montoya, & Cera, 2012). MINT scores were chosen as the matching criterion because picture naming provides an objective independent measure of productive vocabulary potentially accessible during the verbal fluency task, which was of primary interest here. Our focus on younger and older bilinguals with similarly-sized bilingual lexicons (as measured by the MINT) enabled us to isolate possible age differences in language control (that are independent of vocabulary size). On average, the older bilinguals had acquired English marginally later than the younger bilinguals² (see Table 1 for participant characteristics). Language dominance for both groups was determined on the basis of MINT score differences between English and Spanish.

²These diagnoses were made on the basis of executive control and memory impairments, respectively. We analyzed the data without these participants and their matched younger bilinguals and mention the two cases in which a significant effect became non-significant; in three other cases, significant effects became marginal [$.05 < p < .09$]. We think that the most likely explanation for these discrepancies was loss of power, and none of these effects was related to our main conclusions. Also note that, without these participants, the between-group difference in English Age of Acquisition was non-significant [$p = .22$], suggesting that the main pattern of results did not change with a better matched English AoA between groups.

2.2 Materials and procedure

Participants performed a letter fluency task (ten letters, each in both languages), a semantic fluency task (eight categories, each in both languages), a non-linguistic flanker task, a digit-span task, and the MINT (Gollan et al., 2012). Task and language order for the letter and semantic fluency tasks, and language order for the MINT, were counterbalanced between participants. Task order for the first 4 participants in each group is shown in Table 2. Flanker task stimuli were presented using PsyScope 1.2.5 software (Cohen, MacWhinney, Flatt, & Provost, 1993) on a MacBook Pro laptop with a 15-inch display.³

2.2.1 Fluency task—Participants were given 60 seconds to produce as many exemplars as they could from letter categories (e.g., words that start with the letter S) or semantic categories (e.g., animals). They were instructed not to say proper names, numbers, or morphological variants of the same word (e.g., *run* and *running*). After the first trial, participants were informed that they would be performing this task multiple times, that they would be switching between English and Spanish, and that they would be informed which language to use before the start of each trial. Participants performed each letter and semantic category first in one language and then in the other language immediately afterwards, thus switching languages after every minute-long trial (e.g., words starting with the letter A in English, on Trial 1; words starting with A in Spanish, on Trial 2, etc.). Instructions in English preceded English trials, and instructions in Spanish preceded Spanish trials. Letter categories were words that start with *A, D, F, L, M, N, P, R, S* and *T*. The average frequency of these letters as word-initial letters in English and Spanish did not significantly differ between the two languages [English: $M = 5.68$, $SD = 2.59$; Spanish: $M = 5.67$, $SD = 2.64$; $p = .99$].⁴ Semantic categories were *clothing, countries, fruit, insects, musical instruments, sea animals, sports* and *vegetables*. Letter and category order was varied across participants. Responses were recorded manually, and digitally for later verification of scoring.

2.2.2 Flanker task—The task was adapted from Fan et al. (2002). Stimuli included a central arrow with two flanking arrows on each side, and appeared in three conditions: congruent (central arrow pointing in the same direction as the flankers), incongruent (central arrow pointing in the opposite direction as the flankers), and neutral (central arrow flanked by lines without arrowheads). There was an equal number of left-facing and right-facing arrows in each condition. There were 48 trials in each condition. To increase difficulty (and

³Older bilinguals also performed another version of the flanker task, which differed from the one described below in that all stimuli were presented in the center of the screen. Younger participants performed the trail making test (Reitan, 1958) in its place (Table 2). We do not report either of these tasks as we had data for only one participant group.

⁴Word-initial letter frequencies for each letter (in the form of a percentage of the total number of words in a dictionary) are provided below. Frequencies for English were obtained from the tenth edition of the Concise Oxford English Dictionary, and, for Spanish, from the third edition of the Gran Diccionario Oxford (reproduced from http://lavengro.typepad.com/peter_harvey_linguist/2012/03/frequency-of-initial-letters-in-english-and-spanish.html).

	A	D	F	L	M	N	P	R	S	T
English	5.64	5.39	4.26	3.71	5.71	2.29	8.13	4.49	11.65	5.52
Spanish	9.86	6.10	3.51	3.10	6.57	1.79	9.81	4.82	5.53	5.57

possibly sensitivity to bilingual effects) relative to the original flanker task, an even number of stimuli per trial type were presented on the center, left, or right sides of the screen (creating double incongruence on some trials, for example, when a stimulus consisting of a left-pointing center arrow flanked by right-pointing arrows appeared on the right side of the screen; see Gollan et al., 2011). Thus, the flanker task incorporated some aspects of the Simon task described above. Each trial was preceded by a location cue (a procedure that produced the largest speed advantage for bilinguals compared to monolinguals in Costa et al., 2008). Trials of each type (congruent, neutral, and incongruent displays each presented on the left, right, and center of the screen) were presented at least once in a random order before any type was repeated. Trials were structured as follows: 1) a central fixation point appeared for a total of 800 ms; 2) after 400 ms, the fixation point was accompanied by a location cue for 100 ms; 3) the target stimulus appeared for 1,700 ms or until a response was made; 4) there was an inter-trial interval of 1,700 ms. Participants pressed a button on the left or right of a button box to indicate the direction of the central arrow. Practice blocks began with six neutral trials, followed by six congruent trials, then six incongruent trials, and ended with nine trials with equal numbers of the different trial types presented in a random order.

2.2.3 Digit span task—On the forward digit span task (Wechsler, 1997), the experimenter produced increasingly longer strings of numbers and participants were asked to repeat them in the order in which they were presented. On the backward digit span task (Wechsler, 1997), participants repeated different, increasingly longer number strings in reversed order from which they were presented. One point was given for every correctly-repeated sequence. The task was discontinued after participants failed to repeat correctly two strings of the same length. The digit span task was administered in the preferred language for each participant (English for 10 older and 12 younger participants, and Spanish for 3 older and 1 younger participants).

2.2.4 Multilingual Naming Test (MINT)—The MINT (Gollan et al., 2012) consists of 68 black-and-white line drawings, administered in order of progressing difficulty (e.g., item #1 is *hand*, and item # 68 is *axle*). This test was designed to assess picture-naming ability in four languages (English, Spanish, Mandarin, Hebrew), and also both dominant and non-dominant language proficiency. Bilinguals were asked to name half of the MINT items (odd or even) in only one and then the other of their two languages (“Only” condition), and the other half, in both languages in whatever order came to mind (“Both” condition; see Table 2 for counterbalancing). A $2 \times 2 \times 2$ ANOVA with the factors Age group (Older, Younger), Language (English, Spanish) and Condition (Only, Both) revealed no differences between naming conditions (main effect of Condition: $p = .40$; interactions involving Condition: all $ps > .2$). Thus, to obtain a more complete measure of participants’ naming abilities, MINT scores were calculated by adding the responses in the “Only” and “Both” conditions for each bilingual in each language.

The study procedures conformed to Federal guidelines for the protection of human subjects and were approved by the UCSD Institutional Review Board. Informed consent was

obtained from all participants prior to testing and after the procedures of the study had been fully explained.

3. Data analyses

We analyzed separately the number of correct responses, within-language errors, and wrong-language intrusions. In the letter task, one trial (letter A in the dominant language as second naming language) was eliminated for one older bilingual due to distractions during the trial. To preserve an equal number of trials per language, order and group, a few further trials were removed for this older bilingual (A in the non-dominant language as first, and S in the dominant language as first and the non-dominant language as second), and for her matched younger bilingual (A and S in both languages). Likewise, in the semantic task, four trials (*vegetables* and *musical instruments* in the dominant language as second, and non-dominant language as first naming language) were not administered to one older bilingual due to experimenter error. Following the same strategy as in the letter task, two further trials (*clothing* and *insects*) were removed for this bilingual, and all four categories (*vegetables*, *musical instruments*, *clothing* and *insects* in both languages), for his matched younger bilingual. The subsequently eliminated trials (letter S and semantic categories *clothing* and *insects*) had the same position within a run (e.g., category #5 in a Spanish-English run) as the originally eliminated (letter A) or mistakenly excluded trials (*vegetables* and *musical instruments*; e.g., category #5 in an English-Spanish run; see Table 2).

Because we tested a relatively small number of bilinguals, we adopted the following approach in our interpretation of the findings. We interpreted significant effects, and also interpreted marginal effects if they replicated previous findings (e.g., Van Assche et al., 2013). We interpreted null effects with caution. Correct responses and errors were analyzed in a $2 \times 2 \times 2$ ANOVA with the factors Age group (older, younger), Language order (first naming language, second naming language), and Language dominance (dominant language, non-dominant language). We additionally report simple effects of Language order for each age group and each language, if any of them was significant.

4. Results

4.1 Letter fluency task

Correct responses—Means are plotted on Figure 1a. Older bilinguals produced a similar number of correct responses as younger bilinguals [no main effect of Age group, $F < 1$]. Participants produced more correct responses in the first than in the second naming language [significant main effect of Language order, $F(1,24) = 17.68$, $MSE = 33.46$, $p < .001$, $\eta_p^2 = .42$], indexing inhibition of the non-target language. The language order effect was similarly sized in the two age groups [no interaction between Age group and Language order, $F < 1$], implying intact application of inhibitory control in older bilinguals.

Bilinguals also produced more correct responses in the dominant than in the non-dominant language [significant main effect of Language dominance, $F(1,24) = 26.94$, $MSE = 122.57$, $p < .001$, $\eta_p^2 = .53$], and simple effects further indicated that, for the dominant language, there was a significant language order effect for older, and marginal language order effect

for younger bilinguals [older: $F(1,24) = 8.03$, $p = .009$, $\eta_p^2 = .25$; younger: $F(1,24) = 3.03$, $p = .09$, $\eta_p^2 = .12$]. However, language order effects for the non-dominant language were not significant for either group [older: $F(1,24) = 1.48$, $p = .24$, $\eta_p^2 = .06$; younger: $F(1,24) = 2.00$, $p = .17$, $\eta_p^2 = .08$]. The Language order by Language dominance interaction in this analysis was not significant [$F < 1$]. Separate ANOVA analyses on the correct responses in each language confirmed that there was a significant language order effect in the dominant language [$F(1,24) = 10.47$, $MSE = 44.13$, $p = .004$, $\eta_p^2 = .30$] but only a marginal order effect in the non-dominant language [$F(1,24) = 3.46$, $MSE = 48.09$, $p = .08$, $\eta_p^2 = .13$], and this was similar for both groups [for the interaction between Age group and Language order in the separate ANOVAs, both $ps > .45$]. The absence of a significant interaction between Language order and Language dominance in the main analysis could reflect the marginal trend towards an order effect on the non-dominant language (not observed in Van Assche et al., 2013, possibly because bilinguals in the present study were relatively more balanced bilinguals). Setting this slight difference between studies aside, in the current study the language order effect was highly robust for both young and older bilinguals suggesting that, during testing in the first naming language, inhibition was applied by both groups to the non-target language that then needed to be released to enable production in the second naming language after the language switch.

Within-language errors—Classified as within-language errors were (a) non-words, including incorrectly pronounced words (e.g., *obuse* instead of *abuse*), or production of just parts of words; (b) perseverations, including repetitions of the same word; (c) morphological variants (e.g., *running* after the response *run*); and (d) other instruction violations (e.g., producing proper names or numbers).

Means are plotted on Figure 1b. Older bilinguals produced significantly more within-language errors than younger bilinguals [main effect of Age group, $F(1,24) = 21.18$, $MSE = 15.54$, $p < .001$, $\eta_p^2 = .47$], and bilinguals produced more within-language errors in the first than in the second naming language [main effect of Language order, $F(1,24) = 6.85$, $MSE = 3.65$, $p = .02$, $\eta_p^2 = .22$]. This effect, however, was driven by the greater number of total responses in the first than in the second naming language and did not differ across age groups: Analyses conducted with proportion within-language errors per language order condition produced a non-significant effect of Language order, and no interaction [both $ps > .20$]. Also, participants produced more within-language errors in the non-dominant than in the dominant language [main effect of Language dominance, $F(1,24) = 5.23$, $MSE = 5.56$, $p = .04$, $\eta_p^2 = .18$].

Wrong-language intrusion errors—Intrusion rates (saying a Spanish word on an English trial or vice versa) were very low for both older (0.3%) and younger bilinguals (0). Means are plotted on Figure 1c. Because the variance of the younger bilingual group was 0 (and hence differed from the variance of the older bilingual group), Welch's t-tests for samples with unequal variances were performed comparing each pair of means across groups and conditions. There were no significant effects [all $ps > .15$].

In sum, the letter fluency task revealed no evidence for an age-related decline of language control: Relative to younger bilinguals, older bilinguals produced more within-language

errors, but a similar number of correct responses. Older bilinguals also produced almost no wrong-language intrusions, and the number produced was statistically indistinguishable from the zero intrusions produced by younger bilinguals. Furthermore, bilinguals produced fewer correct responses in the second than in the first naming language, and this effect did not differ between age groups. This effect was present in the dominant language, but was weaker or absent in the non-dominant language (replicating Van Assche et al., 2013).

4.2 Semantic fluency task

Number correct responses—Means are shown on Figure 2a. Bilinguals produced more correct responses in the dominant than in the non-dominant language [main effect of Language dominance, $F(1,24) = 34.82$, $MSE = 56.93$, $p < .001$, $\eta_p^2 = .59$]. No other effects were significant [all $ps > .20$].

Within-language errors—Older bilinguals produced more within-language errors than younger bilinguals [main effect of Age group, $F(1,24) = 14.39$, $MSE = 12.00$, $p = .001$, $\eta_p^2 = .36$], in both the first and the second naming language [no interaction between Age group and Language order, $F < 1$] (see Figure 2b). It appeared that bilinguals made marginally more within-language errors in the first than in the second naming language [main effect of Language order, $F(1,24) = 3.11$, $MSE = 1.78$, $p = .09$, $\eta_p^2 = .12$], which was driven by older bilinguals making marginally more within-language errors in the first than in the second naming language when it was the dominant [$F(1,24) = 3.23$, $p = .09$, $\eta_p^2 = .12$] but not the non-dominant language [$F(1,24) = 1.08$, $p = .31$, $\eta_p^2 = .04$]; such an effect was not present for younger bilinguals in either language [both $ps > .50$]. No other effects were significant [all $ps > .12$].

Wrong-language intrusion errors—Older bilinguals produced significantly more wrong-language intrusions (5.3%) than younger bilinguals (0.4%) [main effect of Age group, $F(1,24) = 9.92$, $MSE = 10.69$, $p = .004$, $\eta_p^2 = .29$]. Also, older bilinguals produced more intrusions in the first than in the second naming language but only when this was the dominant language [$F(1,24) = 5.05$, $p = .03$, $\eta_p^2 = .17$]⁵ and not when this was the non-dominant language [$F(1,24) = 1.29$, $p = .27$, $\eta_p^2 = .05$]; for younger bilinguals, effects in both languages were not significant (both $ps > .75$) (see Figure 2c). There was a marginal effect of Language order [$F(1,24) = 3.12$, $MSE = 13.02$, $p = .09$, $\eta_p^2 = .12$]⁵, driven by the just-reported Language order effect for older bilinguals in the dominant language. No other effects were significant [all $ps > .13$]. Note too that bilinguals produced significantly more wrong-language intrusions in the semantic than in the letter task (see below for a direct comparison between the tasks).

In sum, the semantic fluency task revealed evidence of age-related decline only in error production: Relative to younger bilinguals, older bilinguals produced more within-language errors, and more wrong-language intrusions. There was no age-related decline evident in the

⁵This effect was not significant after excluding from analysis the two older bilinguals diagnosed with MCI in the subsequent year, and the two matched younger bilinguals [main effect of Language order, $p = .21$; simple effect of Language order for older bilinguals in the dominant language, $p = .11$].

production of correct responses. Also, unlike in the letter task, bilinguals produced a similar number of correct responses in the second as in the first naming language.

4.3 Comparison of the letter and semantic tasks

Since we administered 10 letter but only 8 semantic categories, for these analyses we divided responses by the respective number of trials (10 for the letter task and 8 for the semantic task; for those bilinguals for whom trials were missing or excluded, we divided by the respective number of valid trials). We conducted a 2 (Age group) $\times$ 2 (Task) $\times$ 2 (Language order) ANOVA on the transformed responses. There was a Language order effect on correct responses (fewer correct in second than in first naming language) [$F(1,24) = 4.49$, $MSE = .94$, $p = .05$, $\eta_p^2 = .16$]⁶, but this effect was more robust for the letter than for the semantic task (in which the means show no order effects) [significant Task $\times$ Language order interaction, $F(1,24) = 6.43$, $MSE = 1.21$, $p = .02$, $\eta_p^2 = .21$]. ANOVAs conducted separately on the dominant and non-dominant languages further indicated that order effects on the correct responses were carried by the dominant language [significant main effect of Language order, $F(1,24) = 11.88$, $MSE = .43$, $p = .002$, $\eta_p^2 = .33$] but not by the non-dominant language [$F < 1$].

Additionally, bilinguals produced marginally fewer within-language errors in the letter than in the semantic task [marginal main effect of Task, $F(1,24) = 3.68$, $MSE = .33$, $p = .07$, $\eta_p^2 = .13$]. This effect was likewise driven by the dominant language [significant main effect of Task in the separate ANOVA for the dominant language, $F(1,24) = 5.30$, $MSE = .15$, $p = .03$, $\eta_p^2 = .18$] and was not significant in the separate ANOVA for the non-dominant language [$F < 1$].

Bilinguals also produced fewer wrong-language intrusions in the letter than in the semantic task [significant main effect of Task, $F(1,24) = 12.58$, $MSE = .17$, $p = .002$, $\eta_p^2 = .34$]. This effect was driven mainly by the older bilinguals [significant Age group $\times$ Task interaction, $F(1,24) = 9.43$, $MSE = .17$, $p = .005$, $\eta_p^2 = .28$]. T-tests assuming unequal variances between groups (because younger bilinguals produced no intrusions in the letter task) confirmed that a) in a between-group comparison, older bilinguals produced more intrusions than younger bilinguals in the semantic task [$t(12.32) = 3.31$, $p = .006$] but a similar number of intrusions in the letter task [$t(12) = 1.34$, $p = .21$], and b) in a between-task comparison, older bilinguals produced significantly more intrusions in the semantic than in the letter task [$t(12.39) = 3.37$, $p = .005$] while younger bilinguals produced marginally more intrusions in the semantic task relative to the zero intrusions they produced in the letter task [$t(12) = 2.13$, $p = .06$]. The separate ANOVAs on the dominant and non-dominant languages further showed that older bilinguals produced more intrusions in the semantic than the letter task independently of language dominance [simple effects of Task for older bilinguals; dominant language: $F(1,24) = 6.73$, $p = .02$, $\eta_p^2 = .22$; non-dominant language: $F(1,24) = 10.92$, $p = .003$, $\eta_p^2 = .31$] (see Figure 3).

⁶This effect was not significant [$p = .14$] after excluding from analysis the two older bilinguals diagnosed with MCI the following year, and the two matched younger bilinguals.

4.4 Flanker task

Results are plotted on Figures 4a and 4b. Response times (RTs) longer than three standard deviations above each participant's mean and shorter than 250ms were excluded as outliers (0.44% for older bilinguals and 1.95% for younger bilinguals). Both response times and errors analyses indicated a robust age-related decline, such that older bilinguals responded more slowly [$F(1,24) = 71.09$, $MSE = 17061.53$, $p < .001$, $\eta_p^2 = .75$], produced more errors [$F(1,24) = 6.34$, $MSE = .004$, $p = .02$, $\eta_p^2 = .21$], and exhibited greater interference than younger bilinguals comparing incongruent to congruent trials [significant Age group x Condition interaction; RTs: $F(1,24) = 16.73$, $MSE = 9579.50$, $p < .001$, $\eta_p^2 = .41$; Errors: $F(1,24) = 4.95$, $MSE = .01$, $p = .04$, $\eta_p^2 = .17$], as well as comparing incongruent to neutral trials [RTs: $F(1,24) = 12.79$, $MSE = 13878.68$, $p = .002$, $\eta_p^2 = .35$; Errors: $F(1,24) = 5.44$, $MSE = .01$, $p = .03$, $\eta_p^2 = .19$].⁷ The pattern of results remained similar when RTs were ln-transformed [main effect of Age group: $F(1,24) = 73.79$, $MSE = .03$, $p < .001$, $\eta_p^2 = .76$; Age group x Condition interaction, incongruent vs. congruent: $F(1,24) = 5.53$, $MSE = .01$, $p = .03$, $\eta_p^2 = .19$; Age group x Condition interaction, incongruent vs. neutral: $F(1,24) = 3.85$, $MSE = .02$, $p = .06$, $\eta_p^2 = .14$], and thus cannot be accounted for by general age-related slowing. These results suggest robust age-related decline in domain-general executive control.⁸

4.5 Digit span task

Unlike in the flanker task, older bilinguals did not significantly differ from younger bilinguals in digit span measures (see Table 1). The digit span task typically does show significant age-related decline (see Bopp & Verhaeghen, 2005, for a meta-analysis). Indeed, older bilinguals in the present study had numerically lower scores than younger bilinguals, and these differences were larger (1.2 for both the forward and backward tasks) than the age-related decline reported across other studies (.53 for forward and .54 for backward: Bopp & Verhaeghen, 2005). The lack of a significant difference may have been due to a lack of statistical power, but also suggests that age-related decline on this measure is less robust than it is on the measures above that did exhibit age effects. However, due to the inconclusiveness of the digit span results in the current study, we do not discuss them further.

5. Discussion

In this study, we investigated age-related decline of language control with the letter and semantic verbal fluency tasks. Bilinguals completed the same category in both languages, and switched languages after every minute-long fluency trial – a manipulation we

⁷We also analyzed separately time-out responses (in which participants failed to respond before the three-second response deadline). These produced similar findings, except that the Group x Condition interaction in the planned comparison between the Incongruent and Neutral conditions was just marginally significant [$p = .07$].

⁸To investigate any relationships between language control as measured by the fluency task, and domain-general executive control as measured by the flanker task, Gollan et al. (2011) conducted correlations between fluency and flanker measures. For consistency, we conducted such analyses here too. However, note that the fluency task versions differed between the two studies, and the small sample available for the present study makes interpreting such correlations questionable. Indeed, for older bilinguals, the three significant correlations we found became non-significant with the exclusion of only one participant, and for younger bilinguals, the only two significant correlations each involved a variable with at least 7 data points equal to 0. In short, we did not find any conclusive evidence for a relationship between language control and domain-general executive control as measured with these two tasks, but, due to the sparseness of the data, we do not discuss these correlations further.

hypothesized would be especially taxing for language control mechanisms. In the letter task, we observed a signature effect of language control. Bilinguals produced fewer correct responses after having just produced members of the same category in the other language than when they produced members of a category for the first time. Additionally, these order effects tended to be stronger in the dominant than in the non-dominant language (as reported in Van Assche et al., 2013). The replication of significant order effects in the current study supports prior conclusions that the letter fluency task measures inhibitory control processes in bilingual language production. When starting off with a letter category in one language, the other language (and specifically members of this letter category in the other language) is inhibited to avoid interference (Green, 1998). The effects of such inhibition are evident on the next trial, calling for production of items *from the same category* in the just-inhibited language. To release the inhibition takes effort, and performance in the just-inhibited language suffers, relative to when it has not been preceded by production in the other language. This effect might be asymmetrical (particularly in bilinguals with one clearly dominant language as in Van Assche et al., 2013) because the dominant language is likely to interfere more in non-dominant language production than the reverse, and thus the dominant language should be inhibited more strongly than the non-dominant language (Green, 1998; see also Meuter & Allport, 1999).

Contrary to our predictions, the language order effects that we observed on correct responses in the letter task, and which we suggest index inhibitory control abilities, did not differ significantly between older and younger bilinguals. Coupled with the almost complete lack of wrong-language intrusions by older bilinguals in this task, this result suggests that the ability to *apply* inhibition did not decline in older age. We did observe a trend implying that older bilinguals might have difficulty with another aspect of inhibitory control relative to younger bilinguals: the ability to *release* inhibition after it is applied. Such a possibility is suggested by a numerical trend towards a bigger order effect for older ($\eta_p^2 = .32$) than for younger bilinguals ($\eta_p^2 = .22$; but note that a significant interaction between Age group and Language order would be needed to support this interpretation, and conclusions drawn in light of the present data must be tempered by the relatively small sample size). In any case, we consider it noteworthy that we did not find any evidence for, or trends towards, an age-related decline in one aspect of language control – the application of inhibition.

Though younger and older bilinguals exhibited similar order effects in the fluency task, we found robust age-related decline in the non-linguistic flanker task. In this task, older bilinguals showed a larger conflict effect (difference between congruent and incongruent trials) than younger bilinguals, which remained significant even when the RTs were corrected for general slowing. The fact that linguistic and non-linguistic control appeared to be differentially affected by older age in the present study suggests that language control is relatively less impaired in older age, and only partially overlaps with, domain-general executive control (in line with the conclusions reached by Calabria et al., 2013; Gollan et al., 2011; Gollan, Kleinman, & Wierenga, 2014; Prior & Gollan, 2013; Weissberger et al., 2013).

In contrast, in the semantic task, we did not observe strong signatures of inhibitory control in either age group (order effects were not significant or relatively weak), while at the same

time, in this version of the fluency task, we obtained some evidence for age-related decline in language control. This absence of significant order effects in semantic fluency could suggest conceptual priming effects that facilitated production in the second naming language, perhaps undoing inhibition effects, as follows. Upon production of items in the first naming language, the non-target (second naming) language would be inhibited to avoid interference. On the following trial, producing items in the just-inhibited second naming language would require verbalizing the same, very recently activated concepts, and this activation may allow the control system to more easily and quickly overcome the effects of inhibition at the lexical level. In contrast, the letter task would produce no conceptual priming (because translation equivalents generally do not begin with the same letter or sound across languages). Indeed, repeated items (items produced in the second naming language which were translation equivalents of items produced in the same trial in the first naming language) constituted 43% of the correct responses for the second naming language in the semantic task, but only 12% in the letter task [$F(1,24) = 576.34$, $MSE = .004$, $p < .001$, $\eta_p^2 = .96$]. These rates did not differ across groups, suggesting that the two groups benefitted to a similar extent from semantic priming [Group x Task interaction: $p = .19$; simple effects of Task: both $ps > .38$].

However, both older and younger bilinguals produced more intrusions in the semantic than in the letter task, which might suggest that the same conceptual boost that primed production of the second naming language might also have made it more difficult to inhibit this language on the following trial, and particularly so for older bilinguals. On this view, the application of inhibition may be flexible to a degree depending on task demands. When bilinguals stand to benefit in some way from maintaining both languages active, the control system will exercise this option despite the possible concomitant interference costs that might be observed (e.g., an increased rate of wrong-language intrusion errors). Note that, when conversing with bilinguals, the cost of producing unintended language switches would either be relatively small (a small cost in time), or non-existent (Kleinman & Gollan, submitted). Thus, it would be advantageous for bilinguals to develop flexible control settings that modulate the amount of inhibition based on different task demands or speaking contexts.

The fact that older bilinguals produced more wrong-language intrusions than younger bilinguals in the semantic task could stem from an age-related decline in domain-general executive control (found here in the non-linguistic flanker task), or a decreased ability to toggle back and forth between control settings within the language system. Another possibility is that the increased rate of intrusions produced by older bilinguals in the semantic fluency task was driven by a failure to remember the required response language for a given trial (wrong-language perseverance). However, such a failure arguably also indexes a language control deficit, and in particular it seems telling that this occurred only in the semantic but not in the letter task (a simple ‘failure to remember’ mechanism should have applied equally to the two types of fluency task). When removing all responses we suspected might have been caused by a failure to remember which language was supposed to be produced (consecutive wrong-language responses, typically country names or sports), older bilinguals produced more intrusions than younger bilinguals in the non-dominant language [$p = .02$] (although the means were identical in the dominant language). Most

importantly for our conclusions, both groups still produced more intrusions in semantic than in letter fluency [$p = .001$].

An unexpected result we reported was that older bilinguals produced more intrusions in the dominant language when it was *first* than when it was second, only in the semantic fluency task. Our explanation of this result is necessarily speculative, but might also reflect the conceptual priming boost and its effects after category switches (which are also language switches). Specifically, language switches occurred in the current paradigm both within and between categories (e.g., a speaker would complete English-animals, Spanish-animals, and then English-fruits followed by Spanish-fruits). In this example, dual language activation (i.e., a conceptual boost) would occur on Spanish-animals and Spanish-fruits (because words with the same meanings were named in English on the immediately previous trial).

However, it might be more difficult to inhibit a language that has just benefitted from a conceptual boost. Thus, when switching categories on the following trial (e.g., from Spanish-animals to English-fruits), Spanish would have to be inhibited, but at this point the activation of Spanish-animals has been boosted (and activation might have spread to other Spanish words), which makes Spanish overall harder to inhibit and leads to more Spanish intrusions on the English-fruits trial. Such a scenario might make the non-dominant language (Spanish in our example) particularly harder to inhibit on a new category trial (English-fruit) because this language, due to being weaker, might be especially benefitting from the conceptual boost (see Van Hell & Dijkstra, 2002). Such a difficulty might also be especially pronounced for older bilinguals. This post-hoc explanation raises some questions: For example, it is not clear why we did not observe analogous effects in the number of correct responses. Replication of this unexpected effect with a larger number of participants is warranted before drawing any major conclusions.

We found no differences between older and younger bilinguals in the number of correct responses in either the semantic or the letter task, despite well-documented age-related decline especially in semantic fluency (e.g., Brickman et al., 2005; Kozora & Cullum, 1995; but see Sandoval, 2010). It is possible that such decline is relatively small compared to the other effects we observed (see Tomer & Levin, 1993) and we did not detect it due to our small sample size. It is also possible that the relatively high education level of our older bilinguals (13.38 years on average indicating at least some college education) counteracted the age-related decline in semantic fluency; education seems to influence verbal fluency performance (see Bolla, Gray, Resnick, Galante, & Kawas, 1998; Tombaugh et al., 1999).

We concluded that there must be at least a partial dissociation between language control and domain-general executive control in bilinguals. But which mechanisms are separate, and which are shared between the two? One possibility is that there are separate mechanisms for applying inhibition on linguistic and non-linguistic representations, but shared mechanisms for determining the need to apply inhibition on either kind of representation (conflict monitoring). In older age, conflict monitoring and the application of inhibition on non-linguistic representations would decline, but the application of inhibition on linguistic representations would be relatively spared. There is a hint for such a possibility in our data. In the letter task, the non-response language could be consistently inhibited, since keeping it active would not bring any advantage. Accordingly, in this task, we did not find evidence for

age-related decline in the application of inhibition of the non-target language (different from non-linguistic inhibition). Conversely, in the semantic task, there were advantages of keeping both languages active on certain trials (naming in the second naming language) but not others (naming in the first naming language). It was in this variable-control-demand task that older bilinguals in our study made more intrusion errors than younger bilinguals. Following this logic, we could expect that age-related decline on language control would be observed specifically in a linguistic task requiring flexibility in applying or releasing inhibition, rather than in tasks requiring the consistent application of inhibition (note that the ability to release inhibition after it has been applied might also be a domain-general mechanism, but our data do not speak to this issue.)

Such a scenario seems consistent with several lines of evidence. First, as mentioned above, older bilinguals perform more poorly than younger bilinguals in cued language switching (Weissberger et al., 2012), but, when allowed to switch freely, they perform equally well (Gollan & Ferreira, 2009). Notably, cued language switching is a task that requires constant shifting between applying and releasing inhibition (bilinguals were cued to switch 50% of the time in Weissberger et al.). In contrast, free language switching allows for inhibition to be applied for longer periods without the need to release it as often (older bilinguals switched 25–33% of the time in Gollan & Ferreira). Second, (college-aged) bilinguals outperform monolinguals only in situations of high conflict monitoring but not low conflict monitoring demands (Costa, Hernandez, Costa-Faidella, & Sebastian-Galles, 2009). Finally, Calabria et al. (2013) proposed that balanced bilinguals (who are equally proficient in their two languages) exhibit the strongest dissociation between language control and domain-general executive control (specifically, non-linguistic inhibitory control; see also Goral et al., 2015). This is because such bilinguals might not rely on inhibiting the non-response language but use other mechanisms instead to prevent interference (Costa & Santesteban, 2004). Further research is needed, however, for conclusive evidence on this issue.

In sum, we employed the verbal fluency task to reveal possible changes in the mechanisms of language control in older age. Our results suggest that language control has a component that is independent from domain-general executive control, and remains relatively preserved in older age. Furthermore, language control mechanisms may operate differently when speakers face different task demands. Specifically, the language control system might allow more or less dual-language activation, and inhibitory control may be reduced (or absent) when speaking tasks might benefit from conceptual facilitation between languages. However, the ability to modulate the amount of inhibition in such tasks seems to decline in older age.

Acknowledgments

This research was supported by an R01 from NICHD (HD050287), an R01 from NIDCD (DC011492) and a P50 from NIA (AG05131). The authors declare no conflict of interest.

References

- Abutalebi J, Annoni JM, Zimine I, Pegna AJ, Seghier ML, Lee-Jahnke H, Lazeyras F, Cappa SF, Khateb A. Language control and lexical competition in bilinguals: An event-related fMRI study. *Cerebral Cortex*. 2008; 18:1496–1505. [PubMed: 17947346]
- Axelrod BN, Henry RR. Age-related performance on the Wisconsin card sorting, similarities, and controlled oral word association tests. *The Clinical Neuropsychologist*. 1992; 6:16–26.
- Bialystok E, Craik FIM, Luk G. Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology*. 2008; 34:859–873. [PubMed: 18605874]
- Bialystok E, Craik FIM, Klein R, Viswanathan M. Bilingualism, aging and cognitive control: Evidence from the Simon task. *Psychology and Aging*. 2004; 19:290–303. [PubMed: 15222822]
- Bolla KI, Gray S, Resnick SM, Galante R, Kawas C. Category and letter fluency in highly educated older adults. *The Clinical Neuropsychologist*. 1998; 12:330–338.
- Boone, KB. Neuropsychological assessment of executive functions: Impact of age, education, gender, intellectual level, and vascular status on executive test scores. In: Miller, BL.; Cummings, JL., editors. *The human frontal lobes: Functions and disorders. The science and practice of neuropsychology series*. New York, NY: Guilford Press; 1999. p. 247-260.
- Bopp KL, Verhaeghen P. Aging and verbal memory span: A meta-analysis. *Journal of Gerontology*. 2005; 60B:223–233.
- Brickman AM, Paul RH, Cohen RA, Williams LM, MacGregor KL, Jefferson AL, et al. Category and letter verbal fluency across the adult lifespan: Relationship to EEG theta power. *Archives of Clinical Neuropsychology*. 2005; 20:561–573. [PubMed: 15939182]
- Calabria M, Branzi FM, Marne P, Hernández M, Costa A. Age-related effects over bilingual language control and executive control. *Bilingualism: Language and Cognition*. 2015; 18:65–78.
- Calabria M, Hernández M, Branzi FM, Costa A. Qualitative differences between bilingual language control and executive control: Evidence from task-switching. *Frontiers in Psychology*. 2011; 2:399. [PubMed: 22275905]
- Cohen J, MacWhinney B, Flatt M, Provost J. PsyScope: An interactive graphical system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instrumentation and Computation*. 1993; 25:257–271.
- Costa A, Hernandez M, Sebastian-Galles N. Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*. 2008; 106:59–86. [PubMed: 17275801]
- Crossley M, D'Arcy C, Rawson NS. Letter and category fluency in community-dwelling Canadian seniors: A comparison of normal participants to those with dementia of the Alzheimer or vascular type. *Journal of Clinical and Experimental Neuropsychology*. 1997; 19:52–62. [PubMed: 9071641]
- Delis, DC.; Kaplan, E.; Kramer, JH. *Delis-Kaplan executive functioning system*. San Antonio, TX: Psychology Corporation; 2001.
- Fan J, McCandliss BD, Sommer T, Raz A, Posner MI. Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*. 2002; 14:340–347. [PubMed: 11970796]
- Folstein MF, Folstein SE, McHugh PR. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975; 12:189–198. [PubMed: 1202204]
- Gollan TH, Ferreira VS. Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2009; 35:640–665.
- Gollan TH, Kleinman D, Wierenga CE. What's easier: Doing what you want, or being told what to do? Cued versus voluntary language and task switching. *Journal of Experimental Psychology: General*. 2014; 143:2167–2195. [PubMed: 25313951]
- Gollan TH, Montoya RI, Werner G. Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology*. 2002; 16:562–576. [PubMed: 12382994]
- Gollan TH, Sandoval T, Salmon DP. Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological Science*. 2011; 22:1155–1164. [PubMed: 21775653]

- Gollan TH, Schotter ER, Gomez J, Murillo M, Rayner K. Multiple levels of bilingual language control: Evidence from language intrusions in reading aloud. *Psychological Science*. 2014; 25:585–595. [PubMed: 24367061]
- Gollan TH, Weissberger G, Runnqvist E, Montoya RI, Cera CM. Self-ratings of spoken language dominance: A multi-lingual naming test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. *Bilingualism: Language and Cognition*. 2012; 15:594–615.
- Goral M, Campanelli L, Spiro A III. Language dominance and inhibition abilities in bilingual older adults. *Bilingualism: Language and Cognition*. 2015; 18:79–89.
- Green DW. Mental control of the bilingual lexicosemantic system. *Bilingualism: Language and Cognition*. 1998; 1:67–81.
- Grogan A, Green DW, Ali N, Crinion JT, Price C. Structural correlates of semantic and phonemic fluency ability in first and second languages. *Cerebral Cortex*. 2009; 19:2690–2698. [PubMed: 19293396]
- Hasher L, Zacks RT. Working memory, comprehension, and aging: A review and a new view. *The Psychology of Learning and Motivation*. 1988; 22:193–225.
- Henry JD, Crawford JR, Phillips LH. Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia*. 2004; 42:1212–1222. [PubMed: 15178173]
- Hernandez AE, Kohnert KJ. Aging and language switching in bilinguals. *Aging, Neuropsychology, and Cognition*. 1999; 6:69–83.
- Hernández M, Martin CD, Barceló F, Costa A. Where is the bilingual advantage in task-switching? *Journal of Memory and Language*. 2013; 69:257–276.
- Kleinman D, Gollan TH. Speaking two languages for the price of one: Bypassing language control mechanisms via accessibility-driven switches. submitted.
- Kozora E, Cullum CM. Generative naming in normal aging: Total output and qualitative changes using phonemic and semantic constraints. *The Clinical Neuropsychologist*. 1995; 9:313–320.
- Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. *Behavioral and Brain Sciences*. 1999; 22:1–75. [PubMed: 11301520]
- Ljungberg JK, Hansson P, Andrés P, Josefsson M, Nilsson LG. A longitudinal study of memory advantages in bilinguals. *PLoS ONE*. 2013; 8:e73029. [PubMed: 24023803]
- Luo L, Luk G, Bialystok E. Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*. 2010; 114:29–41. [PubMed: 19793584]
- Martin A, Wiggs CL, Lalonde F, Mack C. Word retrieval to letter and semantic cues: A double dissociations in normal subjects using interference tasks. *Neuropsychologia*. 1994; 32:1487–1492. [PubMed: 7885578]
- Mattis, S. *Dementia rating scale: Professional manual*. Odessa, FL: Psychological Assessment Resources; 1988.
- Meuter RF, Allport A. Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*. 1999; 40:25–40.
- Myerson J, Hale S, Wagstaff D, Poon LW, Smith GA. The information-loss model: A mathematical theory of age-related cognitive slowing. *Psychological Review*. 1990; 97:475–487. [PubMed: 2247538]
- Paap KR, Greenberg ZI. There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*. 2013; 66:232–258. [PubMed: 23370226]
- Portocarrero JS, Burright RG, Donovick PJ. Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*. 2007; 22:415–422. [PubMed: 17336036]
- Poulisse, N. *Slips of the tongue: Speech errors in first and second language production*. Amsterdam, The Netherlands: John Benjamins; 1999.
- Prior A, Gollan TH. Good language-switchers are good task-switchers: Evidence from Spanish–English and Mandarin–English bilinguals. *Journal of the International Neuropsychological Society*. 2011; 17:682–691. [PubMed: 22882810]
- Prior A, Gollan TH. The elusive link between language control and executive control: A case of limited transfer. *The Journal of Cognitive Psychology*. 2013; 25:622–645. [PubMed: 24688756]

- Rascovsky K, Salmon DP, Hansen LA, Thal LJ, Galasko D. Disparate letter and semantic category fluency deficits in autopsy-confirmed frontotemporal dementia and Alzheimer's disease. *Neuropsychology*. 2007; 21:20–30. [PubMed: 17201527]
- Reitan RM. Validity of the Trail Making Test to organic brain damage. *Perceptual Motor Skills*. 1958; 8:271–276.
- Rosselli M, Ardila A, Araujo K, Weekes VA, Caracciolo V, Padilla M, et al. Verbal fluency and repetition skills in healthy older Spanish–English bilinguals. *Applied Neuropsychology*. 2000; 7:1–24.
- Salthouse, TA. A theory of cognitive aging. Amsterdam: North Holland; 1985.
- Sandoval, T. Unpublished doctoral dissertation. 2010. The role of control in bilingual verbal fluency: Evidence from aging and Alzheimer's Disease.
- Sandoval TC, Gollan TH, Ferreira VS, Salmon DP. What causes the bilingual disadvantage in verbal fluency: The dual-task analogy. *Bilingualism: Language and Cognition*. 2010; 13:231–252.
- Simon JR, Ruddell AP. Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*. 1967; 51:300–304. [PubMed: 6045637]
- Strauss, E.; Sherman, EMS.; Spreen, O. A compendium of neuropsychological tests: Administration, norms, and commentary. 3. New York: Oxford University Press; 2006.
- Taler V, Johns BT, Young K, Sheppard C, Jones MN. A computational analysis of semantic structure in bilingual verbal fluency performance. *Journal of Memory and Language*. 2013; 69:607–618.
- Tao L, Taft M, Gollan TH. The bilingual switching advantage: Sometimes related to bilingual proficiency, sometimes not. *The Journal of the International Neuropsychological Society*. 2015; 21:531–544. [PubMed: 26527242]
- Tombaugh TN, Kozak J, Rees L. Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology*. 1999; 14:167–177. [PubMed: 14590600]
- Tomer R, Levin BE. Differential effects of aging on two verbal fluency tasks. *Perceptual and Motor Skills*. 1993; 76:465–466. [PubMed: 8483658]
- Van Assche E, Duyck W, Gollan TH. Whole-language and item-specific control in bilingual language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2013; 39:1781–1792.
- Van der Lubbe R, Verleger R. Aging and the Simon task. *Psychophysiology*. 2002; 39:100–110. [PubMed: 12206290]
- Van Hell JG, Dijkstra T. Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*. 2002; 9:780–789. [PubMed: 12613683]
- Verhaeghen P. Aging and executive control: Reports of a demise greatly exaggerated. *Psychological Science*. 2011; 20:174–180.
- Verreyt N, Woumans E, Vandelandotte D, Szmalec A, Duyck W. The influence of language switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition*. in press.
- Wasylshyn C, Verhaeghen P, Sliwinski M. Aging and task switching. A meta-analysis. *Psychology and Aging*. 2011; 26:15–20. [PubMed: 21261411]
- Wechsler, D. Wechsler adult intelligence scale. 3. San Antonio, TX: The Psychological Corporation; 1997.
- Weissberger GH, Wierenga CE, Bondi MW, Gollan TH. Partially over-lapping mechanisms of language and task control in young and older bilinguals. *Psychology and Aging*. 2012; 27:959–974. [PubMed: 22582883]
- Zacks, RT.; Hasher, L. Directed ignoring: Inhibitory regulation of working memory. In: Dagenbach, D.; Carr, TH., editors. *Inhibitory Processes in Attention, Memory, and Language*. San Diego, CA: Academic Press; 1994. p. 241–264.

Figure 1a

Figure 1b

Figure 1c

Figure 1.

Older and younger bilinguals' performance on the letter fluency task, in the dominant and the non-dominant language. Figure 1a: Number of correct responses. Figure 1b: Number of within-language errors. Figure 1c: Number of wrong-language intrusions. Error bars represent standard errors.

Figure 2a

Figure 2b

Figure 2c

Figure 2. Older and younger bilinguals' performance on the semantic fluency task, in the dominant and the non-dominant language. Figure 2a: Number of correct responses. Figure 2b: Number of within-language errors. Figure 2c: Number of wrong-language intrusions. Error bars represent standard errors.

Figure 3. Older and younger bilinguals' intrusions on the letter and semantic fluency tasks, divided by number of trials (Letter: 10; Semantic: 8). Error bars represent standard errors.

Figure 4a

Figure 4b

Figure 4.

Older and younger bilinguals' performance on the flanker task. Figure 3a: Response times. Figure 3b: Percentage errors. Error bars represent standard errors. Figure 1a

Table 1

Means, standard deviations and comparisons for all participants' characteristics

	Older bilinguals		Younger bilinguals		<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Age	77.00	9.28	21.23	2.35	21.00	<.001
Sex (% females)	62%	–	62%	–	–	–
Education	13.38	3.86	14.62	1.50	1.07	.30
English language						
AoA English	7.50	6.05	3.85	3.41	1.90	.07
% daily English use now	75.00	27.62	80.31	15.23	.61	.55
% daily English use growing up	46.62	39.87	46.15	18.95	.04	.97
English proficiency	6.00	1.20	6.42	.73	1.09	.29
English MINT	61.77	4.11	61.00	3.65	.51	.62
Spanish language						
AoA Spanish	.58	2.08	0	0	1.00	.34
Spanish proficiency	5.60	1.34	6.17	.71	1.37	.19
Spanish MINT	55.85	9.08	54.69	7.31	.36	.72
English-dominant	8	–	6	–	–	–
Spanish-dominant	3	–	2	–	–	–
Balanced	2	–	5	–	–	–
Digit span forward	7.0	1.78	8.2	2.08	1.52	.14
Digit span backward	5.2	2.55	6.4	1.89	1.31	.20

Note: MINT = Multilingual Naming Test (Gollan et al., 2012). Self-rated proficiency level was obtained by means of a language history questionnaire, and is averaged across self-ratings for four types of language use (speaking, comprehension of spoken speech, reading, writing). The ratings were made on a scale from 1 ("little to no knowledge") to 7 ("like a native speaker"). Dominance was determined on the basis of MINT score differences between the two languages. When these differences were 0 or +/-1, self-rated proficiency was used for classification (2 older bilinguals); when self-rated proficiency in the two languages also did not differ, the dominant language was classified as English because it is the dominant language of the environment (1 older bilingual).

Table 2a

Task order for four older participants

Participant number	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9	Task 10
1	Semantic Eng-Span 1-4	Flanker additional	Semantic Span-Eng 5-8	Digit span	Letter Eng-Span 1-5	Flanker	Letter Span-Eng 6-10	Eng MINT even items	Span MINT even items	Both MINT odd items
2	Semantic Eng-Span 1-4	Flanker additional	Semantic Span-Eng 5-8	Digit span	Letter Eng-Span 1-5	Flanker	Letter Span-Eng 6-10	Span MINT odd items	Eng MINT odd items	Both MINT even items
3	Letter Span-Eng 1-5	Flanker additional	Letter Eng-Span 6-10	Digit span	Semantic Span-Eng 1-4	Flanker	Semantic Eng-Span 5-8	Both MINT odd items	Eng MINT even items	Span MINT even items
4	Letter Span-Eng 1-5	Flanker additional	Letter Eng-Span 6-10	Digit span	Semantic Span-Eng 1-4	Flanker	Semantic Eng-Span 5-8	Both MINT even items	Span MINT odd items	Eng MINT odd items

Table 2b

Task order for four younger participants

Participant number	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9	Task 10
1	Semantic Eng-Span 1-4	Trails	Semantic Span-Eng 5-8	Digit span	Letter Eng-Span 1-5	Flanker	Letter Span-Eng 6-10	Eng MINT even items	Span MINT even items	Both MINT odd items
2	Semantic Eng-Span 1-4	Trails	Semantic Span-Eng 5-8	Digit span	Letter Eng-Span 1-5	Flanker	Letter Span-Eng 6-10	Span MINT odd items	Eng MINT odd items	Both MINT even items
3	Letter Span-Eng 1-5	Trails	Letter Eng-Span 6-10	Digit span	Semantic Span-Eng 1-4	Flanker	Semantic Eng-Span 5-8	Both MINT odd items	Eng MINT even items	Span MINT even items
4	Letter Span-Eng 1-5	Trails	Letter Eng-Span 6-10	Digit span	Semantic Span-Eng 1-4	Flanker	Semantic Eng-Span 5-8	Both MINT even items	Span MINT odd items	Eng MINT odd items