Individual differences in cognitive control advantages of elderly late Dutch-English bilinguals

Merel C. J. Keijzer and Monika S. Schmid
University of Groningen / University of Essex

This study addresses a gap in the literature on executive function advantages among bilingual speakers by investigating a group of elderly, long-term, immersed bilinguals. Our participants are native Dutch speakers who emigrated to Australia as adults and have spent many years in that country. They are compared on a range of cognitive and linguistic measures to native Dutch and native English control groups. We argue that, due to the massive differences in the bilingual experience, group analyses may fall short of capturing the full picture. We argue instead for a more qualitative approach, which takes into account as detailed a picture of bilingual development, daily language habits and, in particular, code-switching habits as possible.

Keywords: language and aging, executive function, language attrition, late bilinguals, bilingual advantage

1. Introduction

Recent years have seen a host of studies on healthy aging, reflecting the increasing proportion of the elderly in developed countries (see Alho, 2008). It is well documented that cognitive resources tend to decline with age: processing speed, working memory, attention span and inhibition mechanisms are all reported to suffer due to changes in the neural substrate (Wingfield & Grossman, 2006). Equally well known is that such cognitive decline impacts on language. Previous work has shown that elderly language users experience deficits such as increased difficulty comprehending speeded speech, show an impaired ability to comprehend and produce complex syntax, and experience word finding and selecting difficulties (Burke & Shafto, 2008).

Linguistic Approaches to Bilingualism 6:1/2 (2016), **64–85**. DOI 10.1075/lab.14032.kei
ISSN 1879–9264 / E-ISSN 1879–9272

Intriguing in this respect is that – due to increased international mobility – many older adults reside in environments where their mother tongue is not spoken. In other words, their ecological reality is that of multilingual aging. Evidence from cognitive (and language) control in bilinguals suggests that early bilinguals (those having acquired their two languages from birth or a very young age onwards) may possess some protection against age-related cognitive decline, even to the extent that the cognitive reserve they have built up over the years can delay the onset of dementia (Bialystok, Craik, Klein & Viswanathan, 2004; Bialystok, Craik & Freedman, 2007; Craik, Bialystok & Freedman, 2010). To date, It has not been clearly established what factors linked to bilingual development may modulate this bilingual advantage, and whether the effect will also hold for late bilinguals. Such late bilinguals (who acquired their second language beyond puberty) are the focus of the present investigation. Not only are studies investigating language and cognitive control in late bilinguals scarce, but those studies that do exist (typically involving college-aged late bilinguals) are far from uniform in their outcomes (Luk, Bialystok, Craik & Grady, 2011; Tao, Marzecová, Taft, Asanowicz, & Wodniecka, 2011, Festman & Münte, 2012): while some work has found cognitive advantages for late bilinguals in their twenties, other studies have not or only partially revealed such an advantage. The question of whether the bilingual cognitive advantage applies across the full range of ages of acquisition thus remains controversial. In a recent keynote article by Valian (2015), the mixed findings that characterize the field are translated to two possibilities: 1) there is a cognitive benefit of bilingualism, which is sometimes obscured by other benefits prevalent in both bilinguals and monolinguals, such as musical training, education, etc.; 2) there is no cognitive benefit of bilingualism. The author herself favors the first hypothesis, and claims that the advantage has most potential to be uncovered in elderly bilinguals. The present collection of papers speaks to this gap in research.

We examine a group of older (71+) late Dutch-English bilinguals (n = 29), all L1 Dutch long-term immigrants in Australia. As opposed to life-long bilinguals, our late bilinguals – despite some secondary school basic courses that some of our subjects took – all started using their second language English at a post-puberty age and are therefore expected to be more variable in their proficiency levels but also contexts of use of their L2 (see also *2* below for more studies on late bilinguals). Our late Dutch-English bilinguals took part in a battery of language tests (language use and history questionnaire, vocabulary, and grammar measures, among others) as well as a number of cognitive measures (consisting of several working memory tasks, inhibition tasks, processing speed measures, and set shifting tests). In addition, detailed questionnaires were administered, tapping the extent of and contexts in which subjects used their two languages, but also containing questions about participants' experiences other than bilingualism: education, active lifestyle, etc. The late bilinguals' data were set off against SES-matched Dutch and English monolingual controls. The data analysis focused on group scores, but – as a follow-up – qualitative analyses more closely inspected the best and worst performing bilinguals and augmented this with correlation analyses to see in which way any cognitive advantage was related to their language use patterns. No two bilingual experiences are the same, so this mixed methods design was employed to uncover more about the consistency of the bilingual cognitive advantage and the individual predictor variables that may underlie it.

2. The bilingual advantage

It is widely recognized that experience shapes cognition and may even alter neurological organization: testaments of that can be found, for instance, in London cab drivers exhibiting a more voluminous posterior hippocampal area (Maguire et al., 2000) or Canadian postal workers responding to letters and numbers as a single category due to their experience with postal codes (Polk & Farah, 1995). In the past decade, bilingualism studies have added another dimension to this, by showing that the bilingual experience impacts on cognition beyond the language domain. Work on this phenomenon, dubbed the *bilingual advantage*, has shown that using two languages on a daily basis leaves its mark on domain-general cognitive skills, with bilinguals outperforming monolinguals on non-linguistic cognitive tasks that tap constructs such as inhibition and working memory (Bialystok, 2011).

The bilingual experience is different from the settings mentioned above in that it is highly variable. The age at which a bilingual acquires his or her languages, the extent to and contexts in which (s)he uses them, the nature of the input on which acquisition was based, but also differences in personal factors like education and *a priori* variability in cognitive ability result in no two bilingual experiences being the same. Partly based on those differences, in recent years the debate has turned to the question on what constitutes the bilingual advantage and whether it is even possible to talk about just one such advantage. Just as the bilingual experience is multidimensional, the components that make up the bilingual advantage are multi-faceted and difficult to break down. In the past, the bilingual advantage has largely been ascribed to enhanced inhibitory control, but in recent years the term mental flexibility has instead come to be preferred, to reflect that the bilingual advantage is not made up of a single underlying cognitive mechanism (Kroll & Bialystok, 2013).

With all this in mind, it is unsurprising that past studies examining bilingual advantage phenomena have produced mixed results. This is most true for late bilinguals, who – compared to early bilinguals – are more heterogeneous as to the age at which they first came into contact with their L2, the attainment levels they achieved and the contexts in and extent to which they use their languages. Research into the cognitive advantage of late bilinguals is greatly outnumbered by studies focusing on early bilinguals. The work that does exist is highly variable in its outcome. Tao et al. (2011) looked at three groups of young adults in Australia: a monolingual English group as well as two bilingual groups (one consisting of early and the other of late Chinese-English bilinguals). Both bilingual groups were found to have denser and more efficient executive function networks than the monolinguals, but there were intergroup differences between the bilinguals: whereas the early bilinguals appeared better at monitoring tasks, late bilinguals' superior performance lay especially in conflict monitoring, both assessed by means of a lateralized attention networks test. By contrast, no cognitive advantage effect for late bilinguals was found in Luk et al. (2011), who compared monolingual English speakers with two groups of bilinguals (early vs. late; this binary distinction is complicated by the discretely imposed cut-off point of 10.0 years old being imposed in this study, rather than treating age of acquisition as a continuous variable). All speakers were in their early 20s when tested. Whereas the early bilinguals as a group did outperform the others on a Flanker test, the late bilinguals performed on a par with the monolinguals. Festman and Münte (2012), finally, explored a group of highly proficient late bilinguals in their 20s, and found a differential cognitive advantage effect for those who regularly produced switching errors on a bilingual naming task versus those who did not. The non-switchers (those who did not switch as part of the bilingual naming task) obtained significantly higher scores on tasks such as Go/No-Go, and the Wisconsin Card Sorting Task.

That context very much determines the magnitude of the bilingual advantage has received abundant attention. Most notably, in 2011, Green formulated his behavioral ecology of bilingualism framework, later elaborated in Green and Abutalebi (2013) and labeled the Adaptive Control Hypothesis. Here, the interactional context in which a bilingual speaker is situated (single language, dual language or dense code-switching environment), leads to different adaptive changes in the neural circuits and regions that are associated with specific control processes. In other words, habitual (dense) code-switchers can freely exploit the joint activation of their two languages, while dual language speakers, who need to keep their two language systems strictly separate (at home vs. at work, for instance), have become more skilled at resolving language conflict. Similarly, no great effects are expected for single language bilinguals who effectively only use one of their two languages, as they do not have to resolve conflict on a regular basis. As such, more of a cognitive control advantage is expected for non-switchers who still find themselves in dual language environment (rather than bilinguals who effectively move in monolingual settings). Control processes themselves are then broken up into constructs like control monitoring, suppression, opportunistic planning, etc. As both the interactional contexts and control processes needed for those contexts tend to be even more variable in late than early bilinguals, the behavioral ecology of bilingualism plays a more deterministic role in the former group.

The Adaptive Control Hypothesis has found its way into several recent studies, which have used it to add a more fine-grained method to uncover bilingualism effects on cognitive control. One such study is the one by Verreyt, Woumans, Vandelanotte, Szmalec & Duyck (2016), in which three groups of bilinguals were compared on their performance on two cognitive control tasks (a Simon and Flanker procedure): unbalanced bilinguals, balanced yet non-switching balanced bilinguals, and switching balanced bilinguals. Whereas the unbalanced bilinguals and non-switching bilinguals were not found to differ in their performance, the switching bilinguals did produce significantly better results than both other groups. The authors conclude that, more than language proficiency, language use patterns and most notably code-switching practices are at the basis of the bilingual cognitive advantage. While this is an important finding, the Adaptive Control Hypothesis does not predict that switching *per se* leads to enhanced cognitive control; rather, it is being faced with two languages that have to be kept apart in particular contexts. Habitual, dense code-switchers are not predicted to show any advantage. The language setting (dual or dense code-switching) is hard to assess in the Verreyt et al. study as code-switching (as well as language proficiency) was evaluated through self-assessment methods (on a scale of 0 to 7 (0 being never and 7 being very often) how often do you code-switch?).

In another recent commentary paper by de Bruin and Della Sala (2015), language use patterns are related to the neuroanatomical work that has tried to detail the basis for bilingualism, both functionally and structurally. Here too mixed results characterize the findings, leading de Bruin and Della Sala to claim that neuroanatomical work needs to align with behavioral work on the bilingual advantage, especially taking language use patterns of individual bilinguals into account. In this paper, a further suggestion is put forward – in line with Hartsuiker's (2015) earlier claim – that we have now arrived at a situation where uncovering the bilingual advantage resembles a treasure hunt, without being guided by clear predictions that are in turn firmly embedded in and build on a theoretical framework.

Valian (2015), in reviewing the mixed findings of past work, summarized the current state-of-the-art as only showing the tip of the iceberg of the complex cognitive construct that is the bilingual advantage. Postulating that the bilingual advantage does not exist (cf. Kirk, Fiala, Scott-Brown, & Kempe, 2014) does not do justice to what we know about how experience in general shapes cognition.

Rather, Valian described how future work should focus on teasing apart those advantages which can be ascribed to the bilingual experience itself and those that are at work in both bilinguals and monolinguals alike, stemming for instance from musical training, education, (active) lifestyle, etc. Perhaps now more than ever is the time to move away from the group analyses that have dominated the search for the cognitive advantage and to, instead, acknowledge the individual differences in the bilingual advantage and relate these to predictor variables, both those pertaining to the bilingual experience domain and beyond. In other words, to uncover the nature of the cognitive advantage, it is informative to move away from putting bilinguals' performance in categories such as good, bad or indifferent vis-à-vis the monolingual control group (cf. Bialystok, 2009) and carry out carefully controlled case studies.

The main aim of this study is to examine the bilingual cognitive advantage in a group of late Dutch-English bilinguals. As opposed to most available work that has looked at late bilinguals now in their 20s, however, this study targets older adult speakers – aged 71 years or older. While previous work has argued that such participants are likely to show the greatest effect, elderly late bilinguals still remain a greatly under-researched population. By investigating the participants' scores on a variety of cognitive tests through both group analyses and follow-up qualitative case-by-case analyses, a more detailed picture of the bilingual advantage can be obtained. This will be underscored by carefully and objectively detailing the participants language backgrounds and L1 and L2 proficiency levels as well as observing their code-switching behavior in free speech to look more closely at the language use pattern of each individual subject. Such a mixed method design is likely to shed more light on the cause of mixed findings in earlier studies. In particular, based on the theoretical discussion above, the following predictions guide the present study:

1. Bilingualism being a highly variable experience and *late* bilingualism even more so, the bilinguals' cognitive task scoring range is likely to be greater than that shown by the monolinguals. The best and worst performing subjects overall will be bilinguals. In other words, the *late* bilingual experience presents an advantage for some and an overloading effect for others, leveling off in group analyses. The direction of the effect (positive or negative) will be related to the individual bilingual's language use patterns:
2. Those bilingual subjects in dual language situations (rather than single language or dense code-switching settings) will show the greatest cognitive effect of bilingualism.
3. The effect of language context (single vs. dual vs. dense code-switching settings) will override any individual factors like age of acquisition, active lifestyle

or educational background in explaining any cognitive advantage and will also have more explanatory power than language proficiency backgrounds of individual participants.

3. The study

3.1 Participants

Participants in this study ($n = 29$, 16 males) were L1 Dutch speakers who were born and raised in the Netherlands. None had been raised bilingually. At a mean age of 27.23 (range 13–61 years) they moved to an English-speaking (Australian) environment, where they subsequently learned English as their L2 in a naturalistic setting. A number of participants indicated that they had had basic English classes at school prior to immigration. All participants were above the age of 71 (mean age 77.93, with a range between 71 and 86).

These bilingual older adults were compared to 17 Dutch monolinguals (8 males) with a mean age of 78.35 (age range: 71–85) and 16 (Australian) English monolinguals (5 males) with a mean age of 76.31 (age range: 71–87), none of whom had ever lived abroad for a substantial period of time (i.e. longer than two months). The English native speakers were true monolinguals, whereas the Dutch native speakers invariably had at least a rudimentary command of one or several foreign languages, most typically English, through formal schooling but also informal exposure (e.g. undubbed television shows with Dutch subtitles). They were no different in this respect from the bilingual group prior to moving abroad. In other words, this design effectively created a three-way bilingual hierarchy.

The three groups were compared by means of one-way ANOVA tests of variance on the independent variables of age in years and years of formal education. The groups were not found to be statistically different in the age at which they were tested ($F(2,58) = 1.02$, $p = .37$), but an a priori difference in educational level was attested, operationalized as number of years of formal education: $F(2,57) = 3.92$, $p < .05$. Subsequent Tukey post-hoc procedures located this effect as the Australian monolinguals having received significantly more years of formal education than the Dutch controls ($p < .05$). The bilinguals as a group did not diverge significantly from either of the monolingual groups. In other words, the premise of this study was not affected by the difference in educational level between the two control groups. In addition, the three groups were matched for fluid IQ level, assessed by means of Raven Advanced Progressive Matrices and their mental states (all groups being composed of older adults) through the mini-mental state examination (Folstein, Folstein, & McHugh, 1975). On none of these measures were the three groups found to differ statistically: $F(2,51) = .95$, $p = .39$ and $F(2.58) = .01$, $p = .91$, respectively. Finally, older adults' processing times are known be longer than those of younger subjects, but none of the groups were statistically different in terms of their overall reaction times, which was assessed by means of an auditory reaction time experiment (participants heard a loud beep and were asked to press a button as soon as possible upon hearing this cue): $F(2,54) = .42$, $p = .66$. All subjects had normal or corrected to normal sensory acuity.

3.2 Materials

Participants were given cognitive as well as language (proficiency) tasks. The test battery was administered individually for each participant. For the bilingual population, many of the tasks were administered in both their L1 (Dutch) and their L2 (English), so that the bilingual sessions lasted approximately 2.5 hours, while the monolinguals took around 1.5 hours. The participants were first presented with the Dutch stimuli, followed by the English stimuli. A 15 to 30 minute break marked the transition from one language to the next. All instruction took place in the language of administration (Dutch for the first part of the test battery, and English for the second). All bilingual testing took place in a quiet room at Monash University in Melbourne, Australia. There was one case where a participant did not have any means of transportation, and he was tested in a quiet room in his home. Data collection for the Australian control group likewise took place at Monash University, whereas the Dutch controls were tested in a quiet room at Utrecht University in the Netherlands. The test sessions were mainly computer-based, but also included an orally administered language and social background questionnaire, which was recorded using Audacity software. Participants received a 30 Australian-dollar reimbursement for their participation (20 Australian dollars and 20 euros for both monolingual groups, respectively). Table 1 presents an overview of the language (proficiency) tasks as well as cognitive tests as well as language tests included in the test battery. The tasks marked with an asterisk were administered in both Dutch and English. The same is true for the semi-structured interview (where questions about pre-emigration life in the Netherlands were asked in Dutch, and those questions about life in Australia were formulated in English). The other – cognitive – tasks (i.e. the backward digit span, modified Wisconsin card sorting test and Simon task) were all administered in the preferred language of the subject (indicated at the start of the test battery). The semi-structured interview always marked the start of the test session, the remaining tests were counterbalanced across subjects.

Table 1. The test battery

Cognitive tasks	Language (proficiency) tasks
– Reading span task* (Daneman & Carpenter, 1980) – Backward digit span task – Modified Wisconsin Card Sorting Test (Grant & Berg, 1948) – Simon task (Simon & Wolf, 1963) – Stroop task* (Stroop, 1935)	– Semi-structured interview to elicit free speech – C-test* (Keijzer, 2007) – Grammaticality judgment task* – Peabody Picture Vocabulary Test (Dunn & Dunn, 2007)* – Category and letter verbal fluency task* [1]

*administered in both L1 Dutch and L2 English for the bilinguals.

Cognitive tasks
To provide more information about the administrative (including timing) procedures of the cognitive tasks, the standardized computerized Reading Span Tasks (L1 Dutch and L2 English) used in this study were short forms (60 rather than 100 sentences) of those developed by van den Noort, Bosch, Haverkort and Hugdahl (2008). All sentences were presented for a maximum duration of 6 seconds, unless the subject hit the space bar before. Also in the working memory range, the backward digit span test that was administered presented subjects with an ever increasing list of digits they subsequently had to type – in reverse order – in a separate window. The longest list of digits correctly recalled formed the score of this test, with a ceiling score of 10. Each digit series was projected onto the screen for 2 seconds before disappearing.

In the executive functioning domain, the Modified Wisconsin Card Sorting test first of all followed the guidelines posited in the test manual (Heaton, Chelune, Talley, Kay & Curtiss, 1993). The Simon task mainly followed the design explicated in Bialystok et al. (2004), study 2: subjects were asked to respond to four-colored square-shaped stimuli that were either presented congruently or incongruently in relation to the response box. Thirdly, the Stroop test employed three colors: yellow, black and pink. These colors were selected because they are among the few color words that are not phonologically similar in Dutch and English (the Dutch color words are *geel, zwart*, and *roze*, respectively) (see Tzelgov, Henik & Leiser, 1990). Following a 1000 ms fixation point, the color words were projected one-by-one in the middle of the screen. Subjects were asked to press a response box button that had been taped with yellow, black, and pink to correspond to the color words. All color words were displayed in a set order until subjects pressed the key, with

1. Two category fluency tasks were administered (animal exemplars, and fruit & vegetable exemplars). Similarly, two letter fluency tasks were administered (words starting with A and words starting with F). The mean of these two letter and category fluency tasks is reported in the results section below.

a maximum duration of 4000 ms, following Gass & Lee (2010). In the two blocks (one for Dutch and one for English), 12 trials for each color were presented (six congruent ones and six incongruent ones). Because of the three colors that were used in this design, this resulted in a total of 36 items. In addition to the 36 items, 14 neutral trials were added. These neutral trials were words presented in an unfamiliar language for all participants, namely Georgian (familiarity with foreign languages other than L2 English had been checked by means of a screening questionnaire).

Language tasks

The semi-structured interview was not just a way of gathering important individual background information per subject (including that relating to their language use patterns), but also resulted in free speech to be analyzed, notably in relation to code-switching occurrences. To this end, the first part of the interview was conducted in Dutch before a (functional) switch to English was made (i.e. a switch from questions relating to the subjects' lives in the Netherlands to their situations in Australia). The task was untimed.

A C-test formed the main measure of general language proficiency (in both the L1 and L2) in this study. Being a cloze format, the C-test builds on the concept of reduced redundancy and internalized pragmatic expectancy grammar developed by Oller and Streiff (1975). Each C-test consisted of five texts that each contained 20 gapped items. Both texts had been standardized prior to testing according to the guidelines set out in Grotjahn (1987). Participants were given a maximum of 5 min per text, the rationale being that time pressure can distinguish between individual levels of expectancy grammars. All blanks and unacceptable words (with respect to grammar and/or text content) were considered incorrect (and awarded a score of 0), while all intended words or acceptable alternatives were scored as correct (a score of 1). Acceptable alternatives belonged to the same word class as the original word and also semantically fitted the context. Correct alternatives or original words containing spelling errors were also considered correct. With a total of 20 gaps per text, the maximally obtainable score on both the L1 Dutch and L2 English C-test was 100, allowing for easy comparison between the two languages. Also in the general language proficiency realm, a receptive vocabulary test was presented in the form of a Peabody Picture Vocabulary Test, administered in line with the test manual (cf. Dunn & Dunn, 2007 for English, and Dunn, Dunn & Schlichting, 2005 for Dutch). Both PPVTs (both languages) were untimed.

The grammaticality task that was used in this study was a 5-point Likert scale format. This methodology is statistically stronger than a binary grammaticality judgment task (see Sorace, 2010 for a discussion on different GJT formats). The areas under investigation were clausal embeddings, V2, discontinuous word order, auxiliary selection/do support and reflexive constructions. These features were chosen as they are present in both the subjects' L1 Dutch and L2 English, but have different surface forms, allowing a closer inspection into cross-linguistic influence. The GJT was untimed.

The verbal fluency measure, finally, presented both a phonological and semantic component, which were themselves subdivided into two one-minute recall tests. In the semantic verbal fluency task, and for both Dutch and English, the score was a composite between the categories *animals* and *fruits and vegetables*. The phonological fluency score was also a composite one: this time between words starting with an *F* and *A*. These word-initial segments were kept constant for both Dutch and English. Frequency-wise, there are roughly as many words starting with an *A* in both languages, but there are relatively more words starting with an *F* in English. This was not taken into account as much as warranted, and may have colored the results.

3.3 Data analysis

Data analysis proceeded in two steps and followed a mixed methods design. The first step consisted of group analyses, where the three groups' performance (per cognitive or language task) was compared by means of either ANOVA analyses of variance (in case all three groups had completed the same cognitive test) or independent samples t-tests (for the language-specific components), carried out in SPSS.

Secondly, the quantitative, group-based analyses were followed up by qualitative analyses. During this subsequent step, an attempt was made to further break down the variability in scores for the three groups at hand. More specifically, to test the first prediction that the bilinguals would be more extreme in their performance at both ends of the spectrum, the standard deviations and range of scores for all three groups separately were inspected per test. Moreover, to test the second prediction, the language use pattern of the bilingual subjects was related to their performance on the cognitive measures; their self-reported daily language use and code-switching patterns were examined but their actual code-switching behavior in free speech was also correlated with their scores on cognitive tasks such as the Simon and Stroop procedures. To do this, a code-switching index was devised based on the number of switches per minute produced as part of the semi-structured interview for each bilingual speaker. As the semi-structured interview was administered in both Dutch (questions pertaining to the participants' life in the Netherlands) and English (for those questions relating to the participants' time in Australia) switches could occur in either language direction (i.e. L1 to L2 or vice versa). The code-switching index was used rather than the absolute number of switches to control for different interview lengths among the participants. Finally, the third prediction was tested by correlating individual factors such as language proficiency (in both L1 and L2) as well as factors such as active lifestyle and educational level to the cognitive test battery outcomes. This was done because 1) active lifestyle and educational level have been named as factors that might be important in shaping cognitive test outcomes (cf. Valian, 2015) and 2) the earlier study by Verreyt et al. (2016) had shown that language use patterns had more explanatory power in relation to outcomes on a Simon and Flanker test than language proficiency levels. They established this on the basis of self-report data of college-aged participants. This study uses objective language measures and targets older adults.

4. Results

Step 1: quantitative group comparison

Table 2 below lists all the cognitive tasks that made up this study's language proficiency test battery and details each of the three groups' mean scores, as well as range of scores, on these tasks. Table 3 does the same for the cognitive tasks.

Examining the mean scores in group-wise comparisons, only four significant results were found and these all belonged to the language proficiency realm: on the C-test that taps general language proficiency, the Dutch-English bilinguals' score lay significantly lower than that produced by the English monolinguals ($t(36) = -3{,}27$, $p < .005$). The same was true for English receptive vocabulary as measured through the Peabody Picture Vocabulary Test ($t(40) = -2{,}55$, $p < .05$) and the number of English category exemplars produced as part of the verbal fluency procedure: $t(40) = -2{,}28$, $p < .05$. These three results are attributable to the Dutch-English bilinguals having smaller vocabulary sizes in their L2 (receptively and productively as well as lower general English skills compared to native speakers). The Dutch-English bilinguals were also found to retrieve significantly fewer items on the Dutch category fluency task in relation to the Dutch monolinguals: $t(40) = -3.55$, $p < .005$. This is most likely not due to a smaller Dutch vocabulary size (there was no difference on the other, untimed (lexical) language tests, but more to a lower activation threshold, making the items harder to access (see Schmid & Jarvis, 2014). Crucially, no group differences were attested on any of the cognitive tests that made up the test battery: the bilinguals at no point overperformed or underperformed in relation to the two monolingual groups. In other words, on the basis of the group analyses, no bilingual advantage was attested for our group of late Dutch-English older bilingual adults.

Table 2. Language proficiency test battery results, split per group

	Dutch-English bilinguals	Dutch monolinguals	English monolinguals
C-test Dutch (max = 100)	*N* = 24[2] M: 73.04 (SD: 23.91) Range: 21–97	*N* = 16 M: 84.63 (SD: 18.32) Range: 34–100	–
C-test English (max = 100)	*N* = 25 M: 68.76 (SD: 47.46) Range: 15–95	–	*N* = 13 M: 86.46 (SD: 12.02) Range: 50–97
Grammaticality judgment task Dutch (max = 26)	*N* = 22 M: 19.23 (SD: 3.92) Range: 11–26	*N* = 15 M: 18.60 (SD: 3.85) Range: 9–23	–
Grammaticality judgment task English (max = 28)	*N* = 25 M: 22.36 (SD: 3.86) Range: 13–28	–	*N* = 13 M: 19.69 (SD: 4.77) Range: 11–24
Peabody Picture Vocabulary Test Dutch Standardized score	*N* = 26 M: 101.38 (SD: 10.97) Range: 75–120	*N* = 17 M: 97.59 (SD: 12.56) Range: 78–116	–
Peabody Picture Vocabulary Test English Standardized score	*N* = 26 M: 105.77 (SD: 10.23) Range: 88–129	–	*N* = 16 M: 114.19 (SD: 10.68) Range: 94–134
Mean number of exemplars category verbal fluency task Dutch	*N* = 26 M: 15.11 (SD: 3.84) Range: 8–24.5	*N* = 16 M: 19.31 (SD: 3.52) Range: 12–25	–
Mean number of exemplars category verbal fluency task English	*N* = 26 M: 16.48 (SD: 4.01) Range: 9.5–27	–	*N* = 16 M: 20.16 (SD: 6.48) Range: 5.5–34.5
Mean number of exemplars letter verbal fluency task Dutch	*N* = 26 M: 7.5 (SD: 2.57) Range: 4–15	*N* = 16 M: 8.19 (SD: 3.07) Range: 1.5–12.5	–
Mean number of exemplars letter verbal fluency task English	*N* = 25 M: 10.82 (SD: 3.32) Range: 3–17.5	–	*N* = 16 M: 12.88 (SD: 5.33) Range: 3.5–22.5

2. The varied sample sizes in Tables 2 and 3 are due to data cleansing procedures (outlined in the footnotes below) as well as missing data, resulting in different sample sizes per test.

Table 3. Cognitive test battery results, split per group

	Dutch-English bilinguals	Dutch monolinguals	English monolinguals
Reading Span Task Dutch (max = 60)	*N* = 28 M: 31.00 (SD: 7.41) Range: 14–51	*N* = 17 M: 31.41 (SD: 6.5) Range: 18–41	–
Reading Span Task English (max = 60)	*N* = 28 M: 30.21 (SD: 6.42) Range: 18–46	–	*N* = 16 M: 33.88 (SD: 8.11) Range: 18–49
Backward Digit Span Test (max = 10)	*N* = 27 M: 6.30 (SD: 1.90) Range: 3–10	*N* = 17 M: 5.12 (SD: 1.58) Range: 2–7	*N* = 16 M: 6.13 (SD: 2.03) Range: 3–10
Modified Wisconsin Card Sorting Test – raw composite score[3]	*N* = 28 M: 110.96 (SD: 16.48) Range: 72–145	*N* = 17 M: 107.65 (SD: 18.947) Range: 73–132	*N* = 16 M: 111.50 (SD: 16.387) Range: 67–133
Simon Task accuracy score (max = 24)[4]	*N* = 22 M: 19.91 (SD: 2.65) Range: 12–24	*N* = 15 M: 18.00 (SD: 3.76) Range: 12–22	*N* = 15 M: 18.33 (SD: 3.13) Range: 15–24
Simon Task: Simon effect (in ms.)[5]	*N* = 22 M: 82 (SD: 75) Range: −55–272	*N* = 15 M: 69 (SD: 53) Range: −34–137	*N* = 15 M: 61 (SD: 59) Range: −18–179
Stroop Task Dutch accuracy score (max = 50)[6]	*N* = 26 M: 44.96 (SD: 6.85) Range: 21–50	*N* = 12 46.50 (SD: 4.48) Range: 34–50	–
Stroop Task Dutch: Stroop effect (in ms.)	*N* = 21 M: 59 (SD: 179) Range: −514–410	*N* = 9 M: 149 (SD: 94) Range: −34–318	–
Stroop Task English: accuracy score (max = 50)	*N* = 26 M: 47.08 (SD: 5.11) Range: 27–50	–	*N* = 13 M: 47.92 (SD: 2.57) Range: 40–50
Stroop Task English: Stroop effect (in ms.)	*N* = 23 M: 160 (SD: 143) Range: −203–392	–	*N* = 9 M: 180 (SD: 140) Range: 30–401

3. The raw composite scores were used here rather than the standardized scores, because the latter already take background variables such as age, gender and education into account, which we fed as covariates into our analyses instead.

4. With a total number of 24 items, participants who scored lower than chance level (i.e. 12 items correct) were excluded from further analysis. For both the Simon and Stroop tasks, only correct responses were fed into the Simon effect and Stroop effect calculations.

5. Following conventional methodology, all response latencies on the Simon task of shorter than 300 ms. or above 1500 ms. were excluded from further analysis. These same thresholds were also maintained for the Stroop procedure.

6. As part of the data cleansing procedure, all Stroop scores of 20 or lower were excluded from further analysis.

Step 2: qualitative individual differences examined

This section focuses exclusively and in more detail on the bilinguals in an attempt to look more at the individual profiles of these bilinguals and at how these relate to the cognitive test outcomes. In relation to the first prediction, and focusing on the cognitive task battery outcomes, the clearest trends emerged in relation to the tasks that have previously been described to tap inhibitory control: the Simon Task and the Stroop procedure. Here, the bilinguals were both markedly (though not significantly) faster in resolving conflict (i.e. the smallest Simon and Stroop effects were evident in the bilinguals), but also the slowest conflict resolution scores could be found within the bilingual condition. This is also evident in the larger standard deviations in the bilingual as opposed to the monolingual groups on these tasks. Interestingly, the tendency for the Stroop effect score to vary more in the bilinguals was only found when they were asked to do this task in their L1, i.e. Dutch and disappeared when the task was administered in their L2 English. Apart from these tendencies in the conflict resolution tasks, however, the other cognitive tests (the working memory capacity measures and Wisconsin card sorting test) did not reveal such a difference in scoring range between the bilinguals and the monolinguals. This is important, as despite discussions about what constitutes the bilingual advantage (cf. Kroll & Bialystok, 2013), conflict resolution tasks such as the Simon and Stroop procedures have most often been used to uncover it.

To more closely examine why the bilinguals' scores were so broadly distributed, particular on the conflict resolution tasks, the bilinguals' code-switching behavior was correlated with their Simon and Stroop scores, with code-switching behavior being operationalized as number of switches (in either language direction) produced as part of the oral semi-structured interview (see *3.3 Data analysis*). Table 4 details the mean code-switches indexes (i.e. mean switches per minute) for all the participants taken together, to get an idea of the extent of code-switches found within this group.

Table 4. Mean number of code-switches in the oral interviews

	From L1 Dutch to L2 English	From L2 English to L1 Dutch
CS index	$N = 19$[7] M: 0.982 (SD: 0.99) Range: 0.00–3.13	$N = 19$ M: 0.512 (SD: 0.82) Range: 0.00–3.46

Apart from the differences in extent of code-switching behavior among the bilinguals that is evident from this table and the more frequent switches from the

7. Oral interview data is missing from 10 bilingual participants due to technical failure of the recording equipment.

L1 to the L2 rather than vice versa, bivariate correlation analyses did not find the code-switching behavior to be related to either the Simon effect scores ($r = .33$, $p = .210$) or the Dutch Stroop task effect ($r = -.11$, $p = .688$), the conflict resolution tasks that showed the most diverse scoring range. One significant correlation that was attested, however, was between the Dutch Reading Span Task and the CS index from L1 to L2 ($r = -.47$, $p < .05$): the less a participant was found to switch to English when speaking Dutch, the better the score on the Dutch administration of the working memory capacity measure of the Reading Span.

In relation to language use patterns, the bilingual participants were also asked to indicate the percentage to which they used Dutch versus English in their daily lives, as well as what the main language used in the home was. Finally, they were asked to indicate how comfortable they felt using either language. These indicators were deemed important in classifying participants as moving in single, dual or dense code-switching settings, in addition to their behavioral code-switching data obtained through analyzing the oral interview. This led to one significant bivariate correlation: language in the home vs. Simon effect scores ($r = .53$, $p < .05$). Upon closer inspection, this result indicated that if participants continued to use Dutch in the home and used English mainly for outside-of-the-home encounters, they had smaller Simon effects, i.e. showed better conflict resolution.

To look into prediction 3, that the effect of language context (single vs. dual vs. dense code-switching settings) will override any individual factors like age of acquisition, active lifestyle or educational background in explaining any cognitive advantage, as well as yields more explanatory power than language proficiency (i.e. whether someone can be labeled a balanced vs. unbalanced bilingual), the scores on the cognitive measures were first of all correlated with the following predictor variables: 1) age at emigration (taken as an indicator of age of onset of bilingualism); 2) age at testing; 3) education (in years of formal education); 4) active lifestyle (dedifferentiated as staying mostly close to home, being mentally active, and being both mentally and physically active). This procedure yielded only one significant correlation: higher educated individuals showed better conflict resolution and subsequently produced smaller Simon effect scores ($r = .64$, $p < .01$). None of the other predictor variables that have been related to the magnitude of the bilingualism effect showed any predictive power in our sample.

Finally, the outcomes of the language proficiency test battery administered to the bilingual group were linked to the cognitive tests. This was done in order to assess Verreyt et al.'s (2016) claim that proficiency levels (or more specifically, whether bilinguals were balanced or not) were not as strong in predicting cognitive task performance as language use patterns. A substantial number of correlations was found, presented in Table 5 below.

Table 5. Significant correlation matrix between the cognitive and language proficiency measures

	Backward digit span	Reading Span Dutch	Reading Span English	Stroop task Dutch accuracy	Stroop task English accuracy
Dutch C-test		.42*			
Peabody English	.48*		.48*		
Category verbal fluency task – English		.62**	.62**	.59**	.44*
Letter verbal fluency task – English	.44*		.41*		
Category verbal fluency task – Dutch				.51*	.46*

*significant at the $p < .05$ level
**significant at the $p < .01$ level

As can be seen, the verbal fluency task, most notably the category naming task, correlated with several working memory tasks, explainable through the time constraints of the fluency task tapping working memory capacity. This was also represented in the Stroop accuracy scores. Crucially, language proficiency scores were in no way correlated with the conflict resolution aspect of the Stroop – or the Simon – task.

5. Discussion

At the group level, this study – like a number of studies that preceded it – failed to show a cognitive advantage for the late bilinguals under investigation: they did not outperform their monolingual peers (neither Dutch nor English). However, it would be too strong to claim on the basis of these findings that a bilingual advantage does not exist, as is claimed in some other studies (cf. Kirk et al., 2014). Bilingualism is an experience, which is likely to leave its mark on cognitive functioning, but since it is such a highly variable experience it may be difficult to establish its importance in group comparisons. To tap the variability of the bilingual experience, the mixed methods design that underlay the current study let go of the bilingual vs. monolingual group comparisons in the second part and instead focused entirely on the profiles of the bilinguals. In this, we acknowledge the basic methodological problem of making straightforward comparisons between bilingual and monolingual populations, since their life experiences are fundamentally different. That may be especially true for our sample, where immigration is at the foundation of our participants' bilingual experience. Although still under debate, it has been contended that immigration experience may well have an effect on brain structure, and in turn may influence behavioral patterns (de Bruin & Della Sala, 2015). It also needs to be stressed here that these findings are based on small sample sizes, having been set-up as an initial investigation into the question of individual differences in bilingual experiences underlying a cognitive advantage. To test the predictions formulated as part of this investigation more rigorously, however, future work is needed that encompasses larger and more controlled samples to pursue these important questions.

Solely focusing on the bilingual group, interesting trends did emerge that did not surface in the group analyses. These trends spoke to our predictions, but at the same time open interesting avenues for future work on this topic. Starting with prediction 1, the more variable range in scores that was predicted for the bilinguals was not uniformly borne out of their cognitive task data, except for the conflict resolution tasks of Simon and Stroop. Here, lowest and highest accuracy range but also the smallest and biggest Simon and Stroop effects respectively were witnessed for the bilingual group, albeit not being statistically different in comparison to the monolinguals. This in itself is indicative that bilingualism does have an effect on cognitive functioning, mostly notably on conflict resolution. However, this effect is not always a positive one: the lowest scores were also found in the bilingual group, although it does need to be considered that the bilingual group comprised most participants, which may have partly contributed to this skewedness. It appears that bilinguals may also present an overloading effect, and upon closer inspection, whether a negative or positive bilingualism effect was found seemed to strongly relate to the language use patterns of the bilinguals.

Indeed, prediction 2 was borne out of the data: language use patterns were found to significantly tie in with the bilinguals' scores on the conflict resolution tasks (Simon most notably). More specifically, those subjects who reported still using Dutch in the home (i.e. who found themselves in dual language contexts, where at home they spoke Dutch and in their other interactions in the community used solely English) were better at conflict resolution as measured through the Simon task. This finding is important in relation to the adaptive control hypothesis (Green & Abutalebi, 2013): dual language users show more of a cognitive advantage than non-switching bilinguals. The poorest performers were almost exclusively participants who indicated never to use Dutch or use Dutch less than 1% of their daily lives. Counterintuitively in this light, no effect was found between the number of code switches participants produced as part of the semi-structured interview (operationalized by means of the code-switching index, i.e. the number of switches per minute) and their scores on the conflict resolution tasks.

In relation to prediction 3, this study's outcomes largely replicated those found by Verreyt et al. (2016): the bilinguals' proficiency levels (general proficiency and productive and receptive vocabulary) were not found to correlate with the conflict resolution tasks. This is important because rather than self-reported proficiency (as was used in Verreyt et al.), this study employed several language proficiency tests, but arrived at the same result. In other words, language use patterns are a better predictor of cognitive advantages shown by bilinguals than their language proficiency (or dominance) patterns. At the same time, however, the language measures were found to significantly correlate with some of the working memory capacity measures (the reading span and backward digit span tasks). It seems most plausible to explain these effects as reflecting the verbal component of the reading span task as well as the time constraints inherent in the Stroop and verbal fluency measures tapping working memory. This differential outcome (no correlation between language measures and conflict resolution, but correlations between language measures and working memory and Stroop accuracy scores) does underline the importance of the current research trend to describe and break down the construct of mental flexibility (cf. Kroll & Bialystok, 2013).

Also in relation to the third prediction, a number of factors that have been coined in the past as possibly obscuring he bilingualism effect (age of acquisition, (active) lifestyle, education) were also related to the cognitive test results. Only one predictor was found to be related to the Simon effect: those bilingual participants with more formal years of education were better at conflict resolution that forms the basis the Simon task. It is difficult to interpret this finding in terms of magnitude of the effect: the correlation coefficient for education vs. Simon effect was bigger than for language use vs. Simon effect. In that sense education appears to be the strongest predictor of the two, but it is also likely that those who are more highly educated find themselves more often in dual language settings, if only as they may attach more importance to maintaining their mother tongue. This interaction is something that needs to be investigated in more detail in future studies.

6. Conclusion

This study has more closely examined the consistency of the bilingual advantage and has done so within the context of late, now elderly, bilinguals. Although the group effects did not reveal any bilingual advantage, subsequent analyses zooming in more closely on the profiles of the bilinguals and taking a dedifferentiation perspective did reveal interesting findings and trends. We found the bilinguals to show more diverse scoring ranges (both towards the upper and lower bounds of the performance spectrum) and these were correlated primarily with their language use patterns: participants in dual language contexts showed the smallest Simon effect scores. Another predictor variable with much explanatory power was education. However, language proficiency scores did not correlate with conflict resolution scores, although showing predictive power in the working memory domain.

We suggest that the findings presented here indicate the need for more in-depth, qualitative and, ideally, longitudinal analyses that will track and observe as many aspects of the bilingual experience as possible, in order to shed more light on what factors play into the development of cognitive reserve and enhanced executive function.

Acknowledgements

Preparation of this manuscript was made possible by the Netherlands Organisation for Scientific Research (NWO) Veni Innovation Scheme Grant 275-70-030 to Merel Keijzer, which is hereby gratefully acknowledged. We would furthermore like to thank the participants of the Multilingualism across the lifespan workshop (at UGhent) for their useful comments on an earlier presentation of this paper, most notably Wouter Duyck and the other experimental psychologists at UGhent for suggesting useful theoretical perspectives to include. Last but not least we would like to thank two anonymous reviewers and the editors of this special issue for the extremely useful comments they provided on an earlier version of this paper.

References

Alho, J.M. (2008). Migration, fertility, and aging in stable populations. *Demography*, 45(3), 641–650. doi: 10.1353/dem.0.0021

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12(1), 3–11. doi: 10.1017/S1366728908003477

Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology*, 65(4), 229–235. doi: 10.1037/a0025406

Bialystok, E., Craik, F. I., Klein, R. & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and aging*, 19(2), 290. doi: 10.1037/0882-7974.19.2.290

Bialystok, E., Craik, F. I. & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, 45(2), 459–464. doi: 10.1016/j.neuropsychologia.2006.10.009

Burke, D. M. & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (eds.), *The handbook of aging and cognition* (pp. 373–443). New York: Psychology Press.

Craik, F. I., Bialystok, E. & Freedman, M. (2010). Delaying the onset of Alzheimer disease Bilingualism as a form of cognitive reserve. *Neurology*, 75(19), 1726–1729. doi: 10.1212/WNL.0b013e3181fc2a1c

Daneman, M. & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6 de Bruin, A. & Della Sala, S. (2015). The importance of language use when studying the neuroanatomical basis of bilingualism. *Language, Cognition and Neuroscience*. doi: 10.1080/23273798.2015.1082608

Dunn, L.M & Dunn, D.M. (2007). *The Peabody Picture Vocabulary Test – Fourth Edition*. Bloomington, MN: NCS Pearson, Inc.

Dunn, L. M., Dunn, L. M., Schlichting, L. (2005). *Peabody Picture Vocabulary Test-III-NL*. Amsterdam, The Netherlands: Pearson.

Festman, J. & Münte, T.F. (2012). Cognitive control in Russian-German bilinguals. *Frontiers in Psychology*, 3. doi: 10.3389/fpsyg.2012.00115

Folstein, M.F., Folstein, S.E. & McHugh, P.R. (1975). "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198. doi: 10.1016/0022-3956(75)90026-6

Gass, S. & Lee, J. (2010). Working memory capacity, inhibitory control, and proficiency in a second language. In M.S. Schmid & Lowie, W. (Eds.), *Modeling bilingualism: from structure to chaos* (pp. 59–84). Amsterdam: John Benjamins.

Grant, D.A. & Berg E.A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology*, 34, 404–411. doi: 10.1037/h0059831

Green, D.W. (2011). Language control in different contexts: the behavioral ecology of bilingual speakers. *Frontiers in Psychology 2011*(2). doi: doi: 10.3389/fpsyg.2011.00103

Green, D. W. & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25(5), 515–530. doi: 10.1080/20445911.2013.796377

Grotjahn, R. (1987). How to construct and evaluate a C-test: A discussion of some problems and some statistical analyses. *Quantitative Linguistics*, 23, 219–253.

Hartsuiker, R. (2015). Why it is pointless to ask under which specific circumstances the bilingual advantage occurs. *Cortex*. doi: 10.1016/j.cortex.2015.07.018

Heaton, R.K., Chelune, G.J., Talley, J.L., Kay, G.C. & Curtiss, G. (1993). Wisconsin card sorting test: manual. Odessa, FL: Psychological Assessment Resources.

Keijzer, M.. (2007). *Last in first out? An investigation of the regression hypothesis in Dutch emigrants in Anglophone Canada*. Vrije Universiteit Amsterdam, PhD dissertation.

Kirk, N. W., Fiala, L., Scott-Brown, K. C. & Kempe, V. (2014). No evidence for reduced Simon cost in elderly bilinguals and bidialectals. *Journal of Cognitive Psychology*, 26(6), 640–648. doi: 10.1080/20445911.2014.929580

Kroll, J. F. & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, 25, 497–514. doi: 10.1080/20445911.2013.799170

Luk, G., Bialystok, E., Craik, F. I. & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *The Journal of Neuroscience*, 31(46), 16808–16813. doi: 10.1523/JNEUROSCI.4563-11.2011

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences*, 97(8), 4398–4403. doi: 10.1073/pnas.070039597

Oller, J.W. & Streiff, V. (1975). Dictation: A test of grammar-based expectancies. In R.L. Jones & B. Spolsky (Eds.), *Testing language proficiency* (pp. 71–82). Arlington, VA: Center for Applied Linguistics.

Polk, T. A. & Farah, M. J. (1995). Late experience alters vision. *Nature*, 376, 648–649 doi: 10.1038/376648a0
Schmid, M. S. & Jarvis, S. (2014). Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition*, 17(4), 729–748. doi: 10.1017/S1366728913000771
Simon, J. R. & Wolf, J. D. (1963). Choice reaction time as a function of angular stimulus-response correspondence and age. *Ergonomics*, 6(1), 99–105. doi: 10.1080/00140136308930679
Sorace, A. (2010). Using magnitude estimation in developmental linguistic research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 57–72). Amsterdam: John Benjamins. doi: 10.1075/lllt.27.05sor
Stroop, J.R. (1935). "Studies of interference in serial verbal reactions." *Journal of Experimental Psychology*, 18, 643–662. doi: 10.1037/h0054651
Tao, L., Marzecová, A., Taft, M., Asanowicz, D. & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Frontiers in psychology*, 2. doi: 10.3389/fpsyg.2011.00123
Tzelgov, J., Henik, A. & Leiser, D. (1990). Controlling Stroop interference: Evidence from a bilingual task. *Journal of experimental psychology: Learning, memory, and cognition*, 16.5, 760–771.
Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, 18, 3–24. doi: 10.1017/S1366728914000522
Van den Noort, M.W.M.L., Bosch, P., Haverkort, M. & Hugdahl, K. (2008). "A standard computerized version of the Reading Span Test in different languages." *European Journal of Psychological Assessment*, 24.1: 35–42. doi: 10.1027/1015-5759.24.1.35
Verreyt, N., Woumans, E., Vandelanotte, D., Szmalec, A. & Duyck Verreyt, N., Woumans, E., Vandelanotte, D., Szmalec, A. & Duyck, W. (2016). Influence of language switching on the bilingual cognitive control advantage. *Bilingualism: Language and Cognition, 19(1), 181-190*. doi: 10.1017/S1366728914000352
Wingfield, A. & Grossman, M. (2006). Language and the aging brain: patterns of neural compensation revealed by functional brain imaging. *Journal of Neurophysiology*, 96(6), 2830–2839. doi: 10.1152/jn.00628.2006

Authors' addresses

Merel C. J. Keijzer
University of Groningen
Department of Applied Linguistics
Oude Kijk in 't Jatstraat 26
9712 EK Groningen
The Netherlands

m.c.j.keijzer@rug.nl

Monika S. Schmid
University of Essex
Department of Language and Linguistics
Wivenhoe Park
Colchester, CO4 3SQ
UNITED KINGDOM

mschmid@essex.ac.uk

Publication history

Date received: 25 November, 2014
Date accepted: 18 November, 2015
Published online: 10 March 2016