Examining bilingual language switching across the lifespan in cued and voluntary switching contexts

Angela de Bruin[1,2], Arthur G. Samuel[2,3,4], & Jon Andoni Duñabeitia[5,6]

[1]Department of Psychology, University of York, York, UK

[2] Basque Center on Cognition, Brain and Language (BCBL), Donostia-San Sebastián, Spain

[3] Ikerbasque, Basque Foundation for Science, Bilbao, Spain

[4] Department of Psychology, Stony Brook University, Stony Brook, NY, USA

[5] Centro de Ciencia Cognitiva – C3, Universidad Nebrija, Madrid, Spain

[6] Department of Language and Culture, The Arctic University of Norway, Tromsø, Norway

Corresponding author:

Angela de Bruin

Department of Psychology

University of York

York YO10 5DD

UK

Phone: +44 (0)1904 322868

Email: angela.debruin@york.ac.uk

Word count: 17409

Abstract

How bilinguals control their languages and switch between them may change across the lifespan. Furthermore, bilingual language control may depend on the demands imposed by the context. Across two experiments, we examined how Spanish-Basque children, teenagers, younger, and older adults switch between languages in voluntary and cued picture-naming tasks. In the voluntary task, bilinguals could freely choose a language while the cued task required them to use a pre-specified language. In the cued task, youths and older adults showed larger language mixing costs than young adults, suggesting that using two languages in response to cues was more effortful. Cued switching costs, especially when the switching sequence was predictable, were also greater for youths and older adults. The voluntary switching task showed limited age effects. Older adults, but not youths, showed larger switching costs than younger adults. A voluntary mixing benefit was found in all ages, implying that voluntarily using two languages was less effortful than using one language across the lifespan. Thus, while youths and older adults experience greater difficulties using multiple languages in response to external cues, they are affected less when they can freely use their languages. This shows that age effects on bilingual language control are context-dependent.

Keywords: Cued language switching; voluntary language switching; cognitive ageing; cognitive development

Public significance of the study: This study shows that age effects on bilingual language control depend on the context. In response to external cues, youths and older adults have greater difficulty using two languages and switching between them. However, when bilinguals can use two languages freely, this is less effortful than using one language for all age groups.

Introduction

A bilingual's two languages may always be active, even when only one is used (e.g., Spivey & Marian, 1999). Nevertheless, bilinguals usually manage to use the language that is contextually appropriate and to switch between languages when needed or wanted. This ability to control two languages and switch between them has been studied extensively in young bilingual adults. However, it remains unclear if and how bilingual language control and switching change across the lifespan. There is much evidence that at least some aspects of cognition are affected by age (e.g., Cepeda, Kramer, & Gonzalez de Sather, 2001; Hasher & Zacks, 1988). The current study examined whether bilingual language control is modulated by age in terms of development during childhood as well as decline during later adulthood.

Typically, language switching is tested by asking bilinguals to name pictures or digits in response to a cue indicating which language to use. This can be done in a *blocked* single-language condition in which all stimuli have to be named in the same language or in a *mixed* dual-language condition in which two languages need to be used, with the choice of language determined by the cue. This mixed condition contains both switch trials (a response in a different language than on the previous trial) and non-switch trials (using the same language as on the previous trial). With young adults, two main findings have been observed in these cued tasks. First, when blocked trials are compared to mixed non-switch trials, responses are typically slower to the non-switch than blocked trials (a 'mixing cost'). This cost is argued to reflect the proactive control mechanisms that are needed to maintain and control two languages or tasks (cf. Rubin & Meiran, 2005) and suggests that using two languages can be more effortful than using one language. The second main finding concerns the 'switching cost': bilinguals respond more slowly on switch than non-switch trials in the mixed condition (e.g., Costa & Santesteban, 2004; Meuter & Allport, 1999). This effect reflects the reactive cost associated with trial-by-trial switching that requires reconfiguration of the task or language set. The reconfiguration includes activating the target language, but has also been linked to inhibition of the non-target language (Green, 1998).

Using cued paradigms, language switching and mixing have been found to be costly and cognitively demanding (cf. Declerck & Philipp, 2015, for a review of (mainly cued) language switching). However, bilinguals do not always need to use their languages in response to external cues. When talking with a bilingual interlocutor who speaks the same languages, a bilingual is free to use a language of choice and to switch whenever convenient, and this is a common practice in bilingual societies. The amount of language control has been proposed to depend on the interactional context in which the languages are used. In the Adaptive Control Hypothesis (Green & Abutalebi, 2013), three daily-life interactional contexts are described. In the first, single-language context (most comparable to blocked single-language conditions used in language-switching paradigms), a bilingual uses each language in a separate context (e.g., one language at work and another language at home). Considering that the other language may still be active (cf. Spivey & Marian, 1999), control is needed to select the appropriate target language and to avoid interference from the non-target language (possibly by proactively suppressing that language). In the second, strict dual-language context, the two languages are used in the same context but with different people (e.g., two languages are used at work but with different colleagues). In this scenario, the bilingual not only needs to monitor the environment to use the appropriate target language, avoiding interference from the other language, but also needs to switch between languages at the appropriate moment. This is argued to recruit processes such as conflict monitoring and task (dis-)engagement in addition to interference suppression and goal maintenance. This interactional context is analogous to cued language-switching tasks. Lastly, when bilinguals are surrounded by other bilinguals who speak the same languages, language switching can take place freely. In this so-called dense code-switching context, less language control may be needed. Consequently, there may be little or no need for conflict monitoring or suppression of one of the languages. Bilinguals can take advantage of the constant availability of two languages and can use an opportunistic approach to use the words that come to mind first, regardless of the language.

Indeed, studies comparing different interactional contexts have suggested that language control may be more effortful in a strict dual-language context than in a free, voluntary switching context. Voluntary switching contexts have produced faster responses, smaller switching costs, and smaller mixing effects than cued switching contexts (e.g., Gollan, Kleinman, & Wierenga, 2014; Jevtović, Duñabeitia, & de Bruin, in press). Furthermore, comparing voluntary language mixing to blocked single-language conditions has revealed mixing *benefits* rather than the costs typically observed during cued language mixing (de Bruin, Samuel, & Duñabeitia, 2018; see Gollan & Ferreira, 2009, for similar findings in the non-dominant language of unbalanced bilinguals). That is, in these free circumstances, using two languages may be less effortful than having to stay in one language, supporting Green and Abutalebi's hypothesis (2013).

At the same time, even voluntary switching may come with a *switching* cost (e.g., de Bruin et al., 2018; Gollan & Ferreira, 2009; but cf. Blanco-Elorrieta & Pylkkänen, 2017, for cost-free switching in voluntary production). It has furthermore been proposed that the cognitive mechanisms underlying voluntary language use and switching include both bottom-up mechanisms driven by lexical access as well as top-down cognitive control mechanisms (e.g., de Bruin et al., 2018; Gollan & Ferreira, 2009). Evidence for bottom-up mechanisms driving voluntary language use comes from studies showing that language choice is related to lexical access. For instance, in unbalanced bilinguals, easier-to-name pictures (e.g., picture names with a higher frequency or shorter word length) are more often named in the non-dominant language (Gollan & Ferreira, 2009). In balanced bilinguals, language choice was found to be affected by *individual* lexical access. That is, participants who were faster to name a specific picture in language A were more likely to name it in that language when given a choice of languages to use (de Bruin et al., 2018). At the same time, language choice and switching are not entirely driven by bottom-up processes. For example, bilinguals do not switch between languages in daily life on every word (Fricke & Kootstra, 2016) and often prefer to use one language as the base language. Furthermore, the often observed switching costs suggest that voluntary language switching is not entirely cost-free. Switching costs have not only been observed in experimental settings, but also in corpus analyses of daily-life switching (Fricke & Kootstra, 2016). This cost suggests that language switches, even when made voluntarily, require some form of (reactive) language control.

While cued language switching, and to some extent voluntary language switching, have been studied frequently in young adults, relatively little is known about if and how language switching and control change across the lifespan. Work comparing older adults to younger adults has shown mixed results. Some studies find overall slower reaction times (RTs) for older adults but no age effects on the mixing or switching costs (Calabria, Branzi, Marne, Hernández, & Costa, 2015) while others find larger mixing and/or switching costs for older adults (e.g., Hernandez & Kohnert, 1999; Weissberger, Wierenga, Bondi, & Gollan, 2012). In addition, considering that voluntary language switching may be less demanding than cued language switching, age effects may depend on the task that is used. Indeed, Gollan and Ferreira (2009) observed limited age effects on a voluntary language-switching task run with older and younger adults. However, a comparison of age effects across different types of switching contexts within the same bilingual population is missing. The current study therefore examined how younger and older adults performed in cued and voluntary language-switching contexts (Experiment 1). In Experiment 2, we examined the same tasks on the other end of the lifespan, by comparing children, teenagers, and young adults. Studying age effects on language mixing and switching in cued and voluntary contexts will improve our understanding of the cognitive demands involved in different switching contexts as well as our knowledge of age-related developments in language control.

Experiment 1. Language switching in older and younger adults

Introduction

Cued language switching in older and younger adults

In an ageing society, effects of age on cognitive functioning have been studied extensively and age effects have been observed on a wide range of cognitive tasks. A frequently studied, but still open question, is whether this age-related decline is due to an overall slowing in processing speed (Salthouse, 1991) or whether age also affects (components of) executive control. According to the former account, complex cognitive tasks require fast processing speed, which is diminished in older adults, leading to diminished performance. Support for this theory comes from reduced or absent age-related effects in cognitive tasks after accounting for decreased processing speed (Salthouse, 1991). Beyond processing speed, age may also affect subcomponents such as inhibitory control (Hasher & Zacks, 1988) or task switching (e.g., Kray, Li, & Lindenberger, 2002) and age-related decline on cognitive tasks may be explained by both slower processing speed and diminished executive control (e.g., Bugg, Zook, DeLosh, Davalos, & Davis, 2006). Task-switching paradigms allow for the examination of two different cognitive processes and consequently can examine age effects on different components of executive control within the same paradigm. As described above, cued task- and language-switching paradigms typically show mixing and switching costs. Mixing costs reflect the effort associated with maintaining and using two tasks or languages and may be related to working memory processes due to the need to keep multiple tasks/languages in mind. Switching costs reflect the effort associated with the actual switch between tasks/languages, including the deactivation or inhibition of the previously used task/language, selection and activation of the new target task/language, and release of any inhibition previously applied to that task/language (cf. also Wasylyshyn, Verhaeghen, & Sliwinski, 2011). The two effects also reflect more proactive (mixing costs) versus more reactive (switching costs) types of cognitive control.

Several studies of non-linguistic task switching have shown that age affects switching costs (e.g., Cepeda et al., 2001; Kray et al., 2002) and/or mixing costs (e.g., Kray & Lindenberger, 2000; Mayr, 2001). Results from meta-analyses, however, suggest that these age effects only occur consistently on mixing costs, but not on switching costs (e.g., Wasylyshyn et al., 2011). These findings are in line with other meta-analyses showing age effects in dual-task paradigms (e.g., Verhaeghen, Steitz, Sliwinski, & Cerella, 2003), suggesting that older adults have more difficulty completing two tasks in parallel. These findings have been interpreted in light of age-related differences in working memory, considering that the different task sets need to be coordinated and maintained in working memory (e.g., Wasylyshyn et al., 2011). In contrast, the absence of age effects on switching costs suggests that more reactive control mechanisms are more resistant to ageing, in line with meta-analyses showing no age effects on selective attention or inhibition tasks (Verhaeghen, 2011). At the same time, there is a large literature suggesting that age affects at least some types of reactive attention and inhibitory control (e.g., Castel, Balota, Hutchison, Logan, & Yap, 2007; de Bruin & Della Sala, 2018; Proctor, Pick, Vu, & Anderson, 2005; Rey-Mermet & Gade, 2018). With inhibition having been argued to play an important role in (cued) language switching (e.g., Green, 1998), it remains an open question whether age affects language mixing costs only (in line with the non-linguistic switching literature) or language *switching* too.

A few studies have examined age effects in (cued) language switching tasks. Some of these observed larger switching and/or mixing costs for older than younger adults. For instance, Hernandez and Kohnert (1999) found differences between the mixed and blocked conditions (in terms of accuracy and response times) to be larger for older than younger adults. In addition, a comparison of RTs on switch trials versus trials following a switch trial (i.e., non-switch trials) showed larger differences for older than younger adults (suggesting larger switching costs). Weissberger et al. (2012) also found that age effects can go beyond overall RTs. In this study, older adults showed larger switching costs in terms of RTs and errors and a larger mixing cost in terms of RTs. Weissberger and colleagues furthermore suggested that the age effects were somewhat stronger for the switching than mixing costs (in particular for accuracy). In contrast, Hernandez and Kohnert (2015) observed RT age effects on mixing costs, but not on overall switching costs. Lastly, Calabria et al. (2015) reported that middle-aged and older adults were slower overall than younger adults, but showed comparable language switching costs (mixing costs could not be examined as there were no single-language blocks). Thus, from the few studies that have compared younger to older adults on language-switching tasks, it remains unclear whether age only affects overall response times, mixing costs, switching costs, or all of these.

Predictable versus unpredictable task sequences

Part of these inconsistencies may stem from methodological differences across studies. One potentially modulating factor is the predictability of the switching sequence. Hernandez and Kohnert (1999) used a predictable switching pattern, in which participants had to switch on every trial, and an unpredictable condition in which participants randomly had to alternate between switch and non-switch trials. Age effects were larger in the predictable condition. Age effects on mixing costs were also found in the predictable switching paradigm employed by Hernandez and Kohnert (2015). In contrast, of the two recent studies using unpredictable switching sequences, only one observed age effects on mixing and switching costs (Weissberger et al., 2012). The other study (Calabria et al., 2015) showed no age effects beyond overall slowing.

It is possible that the degree of predictability modulates age effects on mixing/switching costs. However, in the one direct comparison between predictable and unpredictable switching (Hernandez & Kohnert, 1999), participants had to switch on every trial in the predictable sequence, thus leading to a major confound between predictable and unpredictable conditions (i.e., age effects may have been larger in the predictable condition because constant switching was required). In addition, the larger age effects were found in a comparison between the mixed condition as a whole and the blocked condition, potentially conflating proactive (mixing) and reactive (switching) control processes. Furthermore, from the non-linguistic switching literature, the (direction of a) predictability effect is unclear. Some task-switching studies have suggested that age effects on switching costs are more likely when the switch is unpredictable (Kray et al., 2002; Van Asselen & Ridderinkhof, 2000). For instance, Kray and Lindenberger (2000) used a predictable switching pattern and observed age effects on the mixing cost, but not on the switching cost. In contrast, using an unpredictable switching pattern in a later study, Kray et al. (2002) showed age effects on the switching but not on the mixing costs. However, a direct comparison of the results found in the predictable (Kray & Lindenberger, 2000) versus unpredictable (Kray et al., 2002) task is difficult as cues were only present in the unpredictable task and as such predictability and cue presence were confounded.

The presence or absence of cues could be particularly important considering that age effects on mixing costs may partly stem from working memory deficits. The need to memorise trial sequences in a predictable task without cues could explain why older adults showed larger mixing costs in this condition but not in the unpredictable task. In addition, cue processing may have been more effortful for older adults, leading to larger switching costs in the unpredictable task. These confounds thus hinder an evaluation of the role of *predictability* as such. Furthermore, while there is some evidence for a modulating role of predictability in individual studies, the meta-analysis by Wasylyshyn et al. (2011) did not find systematic evidence for predictable versus unpredictable switching sequences modulating age effects. Based on the existing language- and task-switching literature, it is thus unclear whether age effects on (language) switching should be larger in predictable sequences (as in Hernandez & Kohnert, 1999), larger in unpredictable sequences (as in Kray et al., 2002 vs Kray & Lindenberger, 2000), or should be comparable for predictable and unpredictable switching sequences (as suggested by Wasylyshyn et al., 2011). The current study advances the work on language switching and ageing by assessing the role of predictability more systematically within-subject and while minimising differences in task demands. In both predictable and unpredictable contexts, participants had to switch on half of the trials and they were always presented with cues prior to the picture that had to be named.

Voluntary language switching in older and younger adults

Age effects have been observed on cued language switching, but previous studies with young adults have suggested that this type of language mixing and switching may come with higher cognitive demands than voluntary language use. Not all older bilinguals need to switch languages in response to external cues in daily life. Instead, some bilinguals can switch between languages more freely. It is therefore important to understand whether age effects on switching and mixing generalise across different interactional contexts. Furthermore, assessing age effects on mixing effects in voluntary tasks may help to better understand the age effects that have been observed on cued mixing costs. These age-related effects have been linked to deficits in working memory capacity. Cued paradigms, however, oblige participants to use the tasks in a pre-specified way, thus making it unclear whether age affects the ability to maintain and use two tasks *in general* or rather to maintain two tasks and use them in a specific, externally cued way. By studying voluntary language mixing, these two processes can be teased apart.

Considering that voluntary language use is less effortful than cued language use in young adults (e.g., de Bruin et al., 2018; Gollan et al., 2014; Jevtović et al., in press), age effects may be more limited on voluntary than cued tasks. This is in line with one previous study comparing older and younger adults on a voluntary switching task (Gollan & Ferreira, 2009) that showed comparable mixing effects for younger and older adults. In terms of the switching cost, older adults showed a larger cost than younger adults, but only in the non-dominant language. Thus, this study suggests that during voluntary language switching, age effects may be limited to switching costs. However, considering that not all cued language-switching studies have observed age effects on mixing or switching costs either, it is important to compare the two tasks within the same group of participants.

Current experiment

The first experiment therefore examined how age affects language switching in different contexts. Older and younger highly proficient Spanish-Basque bilinguals completed two picture-naming tasks. In the **cued task**, pictures had to be named in Spanish or Basque in response to a cue. In the **voluntary task**, bilinguals were free to choose in which language they wanted to name each picture.

Based on the few studies examining age effects on cued language switching, age effects might be found only on overall RTs, or on the switching and/or mixing costs too. Considering the task-switching literature, age may be more likely to affect the proactive processes associated with language mixing than the reactive trial-by-trial adjustments related to language switching. Within the cued task, we furthermore included predictable and unpredictable switching conditions to assess whether age effects on mixing and switching costs differed depending on the predictability of the switching pattern. Two important changes were made compared to previous studies (e.g., Hernandez & Kohnert, 1999; Kray et al., 2002) to make the predictable and unpredictable tasks more comparable: Cues were presented even in the predictable condition and both tasks required a language switch on half of the trials. Given these changes, it is difficult to predict based on previous studies if and how predictability in language switching may alter age effects.

Voluntary language use may impose fewer cognitive demands than cued use, and younger Spanish-Basque bilinguals were previously found to show a language mixing *benefit.* If voluntary language mixing is relatively effortless, age effects may be limited. If older adults show larger cued but not voluntary mixing effects, this would suggest that it is not the use of two languages as such but rather controlling them in response to external demands that deteriorates with age. In contrast, given that previous studies have observed voluntary switching costs, older adults may show larger voluntary switching costs than younger adults, in line with Gollan and Ferreira (2009).

Methods

Participants

Experiment 1 was completed by thirty older adults. Five of the older adults did not meet the criteria for the study: Four participants acquired Basque after the age of 17, thus not meeting the requirement of early bilingualism, and one participant reported several health issues (including dyslexia and hearing loss). All other participants did not report neurological, reading, or language problems. All but one completed the Mini Mental State Examination (MMSE, a short dementia screening test; Folstein, Folstein, & McHugh, 1975) and scored 27 or higher out of the maximum score of 30 points (a score of 24 or higher indicates normal cognitive functioning). The participants provided written informed consent. The study was approved by the BCBL Ethics Review Board and complied with the guidelines of the Helsinki Declaration.

The 25 older adults (15 female) who met the selection criteria were on average 67.7 (*SD* = 4.2, range 59-82) years old. They were early Basque-Spanish bilinguals with a high and relatively balanced proficiency in both languages. To compare the performance of the older adults to a group of younger adults with a similar language profile, from the data reported in Experiment 2 in de Bruin et al. (2018), we selected the 25 young adults (14 female, *M* age = 26.8, *SD* = 4.3, range 19-35) with the highest Basque proficiency. We therefore selected participants who scored between 61-65 points on the BEST picture-naming proficiency measurement (maximum score of 65; de Bruin, Carreiras, & Duñabeitia, 2017). In this way, we ensured that younger and older adults were matched in terms of objectively measured proficiency (respectively *M* = 63.3, range 61-65, and *M* = 63.7, range 59-65). In addition, the younger and older participants were comparable in the amount of Basque versus Spanish language use (see Table 1). Twenty-five participants per age group furthermore ensured over 80% power to detect the voluntary mixing benefit (fixed-effect size = 0.066) and switching cost (fixed-effect size = 0.042) reported in Experiment 1 in de Bruin et al. (2018), with comparable power expected for the cued switching tasks. Based on the previous literature, it is very difficult to estimate the size of a potential age effect on mixing/switching costs. Simulations based on a fixed-effect size of 0.05 showed that 25 participants per age group would yield over 80% power to detect an age effect of this size on mixing/switching costs. This power is also driven by the large number of trials per condition (480 trials per task) that are taken into account in the mixed-effect models based on individual trials employed here.

Table 1 shows the results from the different language proficiency and use measurements that are part of the BEST. We used a repeated measures ANOVA with age group as a between-subject variable and language as a within-subject variable to assess whether the older and younger adults differed on any aspects of their language profile (see Supplementary Table 1 for the statistics). On most measurements, there were no age effects or interactions between language and age. The age of acquisition analysis showed that older adults reported acquiring Spanish later than Basque, but the average Spanish age of acquisition was still below three years old. On some self-ratings, older adults reported a larger difference between Spanish and Basque, but there were no interactions on the objective proficiency task or the exposure to or use of the languages. With respect to overall language effects, Basque and Spanish were similar in terms of exposure and use and self-rated speaking, understanding, and writing proficiency. However, Spanish self-rated reading and self-rated general proficiency, as well as the objective proficiency score, were higher than in Basque.

In addition to the measurements described in Table 1, we asked participants to indicate on a scale from 1 ('never') to 4 ('all the time') how often they switch between languages on a daily basis, within a conversation, and within a sentence (data from two older adults are missing). Younger and older adults switched similarly often on a daily basis (younger *M*: 3.4, *SD*: 0.8; older *M*: 3.2, *SD* = 0.9; $t(46) = -.60$, $p = .56$) and within a conversation (younger *M*: 2.5, *SD*: 0.9; older *M*: 2.3, *SD* = 1.0; $t(46) = -.95$, $p = .35$), but younger adults switched significantly more often within a sentence (younger *M*: 2.0, *SD*: 0.8; older *M*: 1.4, *SD* = 0.7; $t(46) = -2.62$, $p = .012$)

The age groups were comparable in their years of education (younger *M*: 16.9, *SD*: 1.9; older *M*: 16.0, *SD* = 5.3; $t(45) = -.78$, $p = .44$; data from three older adults missing). We also assessed the participants' lifestyle through a questionnaire (Scarmeas et al., 2003, scale ranging from 18-54) asking them about their daily intellectual, physical, and social activities (e.g., reading, exercising, visiting friends). There was no significant difference between younger (*M*: 35.3, *SD*: 3.5) and older adults (*M*: 37.2, *SD* = 3.7; $t(46) = 1.86$, $p = .069$; data from two older adults missing).

Table 1. *Summary of the language profiles of the 25 younger and 25 older adults. Means and (standard deviations) are provided for the different measurements of language proficiency and use for Spanish and Basque per age group. A single asterisk indicates a significant difference between older and younger adults across languages; a double asterisk indicates a significant interaction between language and age group.*

	Spanish Younger adults	**Spanish** Older adults	**Basque** Younger adults	**Basque** Older adults
AoA*,**	0.9 (1.3)	2.9 (2.7)	0.6 (1.5)	0.08 (0.3)
Picture naming[1] (0-65)	64.8 (0.4)	64.7 (0.7)	63.3 (1.4)	63.7 (1.7)
Self-rated proficiency[2] (0-10)				
Speaking	9.4 (0.8)	9.2 (1.2)	9.5 (0.8)	8.7 (1.3)

*Understanding***	9.5 (0.8)	9.6 (0.6)	9.5 (0.8)	9.1 (1.0)
*Writing***	8.8 (1.1)	9.1 (1.3)	9.2 (0.9)	7.9 (2.1)
Reading	9.5 (0.8)	9.2 (1.5)	9.3 (1.2)	8.4 (1.8)
General,***	9.5 (0.6)	9.4 (0.8)	9.4 (0.8)	8.6 (1.2)
%exposure[3] (0-100)*	44.4 (17.8)	50.2 (22.0)	41.6 (19.1)	45.6 (23.0)
%speaking[3] (0-100)*	41.6 (18.6)	47.8 (24.0)	48.8 (21.3)	49.2 (25.2)

[1]Data from 1 older adult missing
[2] Data from 1 younger adult missing
[3]Young adults reported exposure to and use of a third language (English), leading to lower exposure/speaking percentages across Basque and Spanish compared to older adults.

Materials

Thirty coloured drawings that were selected from the MultiPic database (Duñabeitia et al., 2018) were used for the voluntary picture-naming task. These pictures depicted easy to name objects, with high naming agreement, that were non-cognates in Spanish, Basque, and English. Picture names were matched between Spanish and Basque on word length (syllables and phonemes) and log frequency (see de Bruin et al., 2018, for more details about the stimulus selection and the list of stimuli). A different set of thirty pictures was used in the cued picture-naming task. Like the voluntary pictures, they were matched between Basque and Spanish in word length and frequency (see de Bruin et al., 2018, for the stimuli). In the cued task, each picture was preceded by a country flag indicating which language had to be used. Two versions of each flag were used. The use of two flags per language ensured that there was a cue switch on every trial, even when there was no language switch. When only one cue is used per language, the switching cost not only reflects a switch between languages but also a switch between cues that is absent on non-switch trials (cf. Logan & Bundesen, 2003; Mayr & Kliegl, 2003).

Procedure

The procedure was similar to the one described in de Bruin et al. (2018). All participants completed a voluntary and cued picture-naming task, on two separate days. Of the 25 younger adults, 11 completed the cued task on day 1 and the voluntary task on day 2. Of the 25 older adults, 13 completed the cued task on day 1 and the voluntary task on day 2. The voluntary picture-naming task was followed by a series of background measurements (described in the Supplementary Materials) examining inhibitory control and working memory span, as well as some questionnaires. In addition, participants completed the proficiency measurement if these data were not already in the database, and the older adults completed the MMSE.

Voluntary picture-naming task

In the voluntary picture-naming task, participants were asked to name pictures in one language only (blocked condition) or in a mixed condition in which they were free to decide which language to use for each picture. Depending on the language choice, trials in the mixed condition were classified as switch or non-switch trials. The main aim of the study focused on effects of age on mixing and switching effects, respectively defined as the difference between blocked and mixed non-switch RTs and the difference between switch and non-switch RTs.

The structure of the session was as follows: Familiarisation – Blocked1 – Mixed – Blocked2. In the familiarisation phase, participants saw each picture on the screen with the corresponding Basque and Spanish words. Each of the blocked conditions consisted of two parts: One part in which all pictures had to be named in Basque and one part in which they had to be named in Spanish. The order of languages within the blocks was counterbalanced (13 older adults and 14 younger adults started with Spanish) and participants who started with one language in the Blocked1 condition started with the other language in the Blocked2 part. Instructions for each blocked condition were given in the language of that block. In total, the two blocked conditions consisted of 120 items (60 per language). In the Mixed condition, participants were free to name each picture in Basque or Spanish, following these instructions:

'In the following part, you can name the pictures in Spanish or Basque. You are free to switch between languages whenever you want. Try to use the word that comes to mind first, but don't use the same language throughout the whole task'. The instructions were given in both languages, with the order on the screen (top or bottom half) counterbalanced across participants. Each picture was repeated 12 times, leading to a total of 360 mixed trials, with breaks presented after every 60 trials. The first blocked and the mixed conditions were preceded by practice trials (4 per language in the blocked condition and 8 in the mixed condition).

A trial started with a fixation cross presented for 500 ms. Next, the picture was presented and stayed on the screen for 2500 ms, during which time the participants' responses were recorded. Pictures stayed on the screen for 2500 ms, regardless of when a response was made. This was done to avoid a voice key being triggered inadequately by hesitations, background noise, or non-speech sounds (which was especially important for the children tested in Experiment 2, cf. Gross & Kaushanskaya, 2015, 2018, for fixed presentation times in language switching studies with children).

Cued picture-naming task

The cued picture-naming task followed the same structure as the voluntary task. The main difference was that participants always saw a Spanish or Basque flag prior to the picture, asking them to name the picture in the language corresponding to the flag. The mixed condition consisted of a predictable (180 trials) and unpredictable (180 trials) part. In the predictable part, language cues followed the predictable order of Basque – Basque – Spanish – Spanish. For the unpredictable part, thirty different lists were created in which the language cues were presented pseudo-randomly (with the restriction that the same language was not used more than four times in a row). Prior to the predictable condition, participants were informed about the fixed order of language cues. Prior to the unpredictable condition, participants were told that the language cues would appear in a random order.

In both parts, the items were pseudo-randomly distributed across conditions, ensuring that the same item did not occur twice in a row. Half of the predictable and half of the unpredictable trials were switch trials; the other half were non-switch trials. In each part, half of the trials had to be named in each language. Due to participant exclusion, the counterbalancing of predictable/unpredictable first was off by a few participants: 9 older adults and 11 younger adults started with the predictable task.

The trial structure was similar to the one used in the voluntary task, but the language cue (flag) was presented for 500 ms prior to the picture. Next, the picture appeared with a smaller version of the flag above it. Each part started with eight practice trials.

Both the cued and the voluntary picture-naming tasks were presented with Psychopy 1.83.04 (Peirce, 2007). The stimuli were presented on a Viewsonic E90f monitor using a 90Hz refresh rate, screen resolution of 1024x768, and a white background screen. The stimuli were 200x200 pixels in size while the cues were 200x100 pixels when presented alone and 100x50 pixels when presented above the picture. The participants' responses to the pictures were recorded using a Sennheiser PC 151 headset with microphone.

Data analysis

The data are available at: osf.io/qmxk5/

Responses were scored during the tasks by the research assistants running the experiment. For both tasks, accuracy was scored. In addition, the naming language was reported for the voluntary task; trial type in the voluntary mixed condition (switch or non-switch) was coded afterwards. For both tasks, accuracy was coded as in de Bruin et al. (2018), using the following options: A) no response given; B) correct answer; C) incorrect word in correct language; D) wrong language (for blocked and cued conditions only); E) hesitation; F) combination of two languages; G) correct items in the cued and voluntary mixed condition that could not be classified as a switch or non-switch trial (namely, the first trial after a break and mixed trials preceded by responses of type A, D, or F); H) recording failures (4 trials for the younger adults). The responses were recorded and response onset was determined with Chronset software (Roux, Armstrong, & Carreiras, 2017) and checked manually with CheckVocal (Protopapas, 2007). Responses from one older adult were not recorded in the voluntary task and from another older adult in the cued task and were not included in the analyses.

To calculate switching frequency in the voluntary task, we divided the number of switch trials by the total number of mixed trials (only including trials of type B, C, E, H, as others could not be classified as switch or non-switch trials). For all RT analyses, we excluded inaccurate responses (A, C, D, F) and trials of type G and H.

The general procedure for data-analysis was as follows. Reaction time data were analysed with linear mixed-effect models, using the lme4 package (version 1.1-21) in R (Bates, Maechler, Bolker, & Walker, 2014). The RTs were skewed and were therefore log transformed, which improved the normality of the residual distribution. Outliers were removed with the help of the trimr package (Grange, 2015). All two-level categorical factors (e.g., language and age group) were contrast coded as -0.5 and 0.5.

In all models, we included participants and items as random effects and always started with a maximal structure (cf. Barr, Levy, Scheepers, & Tily, 2013). Thus, the initial model always included random intercepts and random slopes for all within-item and within-subject predictors. In the case of non-convergence, correlations between the random slopes and intercepts were removed, followed by removal of the item slopes that explained the lowest amount of variance. We tried to keep the maximal slope-structure for subjects, but if the model did not converge we proceeded with the removal of subject slopes that explained the least variance. The results from the models with all fixed effects (including the significant and non-significant predictors) are reported. All models were checked for collinearity and all VIFs were below 2.5 (Franck, 2011). We interpreted *t* and *z* values > 2 to represent significant findings (Gelman & Hill, 2007). In addition, model comparisons were conducted to determine the best fitting model. To do this, we first compared the simple model with significant predictors only to the model including all possible predictors. If likelihood-ratio chi-square tests did not show a significant effect favouring the model with all predictors, the simpler model with fewer predictors was preferred (Baayen, 2008). We then removed each significant predictor from the simple model and examined whether exclusion of that predictor worsened the model fit. In all analyses, the best model was the one only including the significant predictors while removal of significant predictors worsened the fit. In addition to the models presented in the main text below, we also assessed whether the age-related findings were related to task order (cued/voluntary first), order of predictability (predictable/unpredictable first), and language order (Basque/Spanish first). As described in the Supplementary Materials, order did not interact with age-related findings and all results presented in the main text below were observed when order was included in the analysis.

While RT analyses were performed on log RTs, all figures, tables, and in-text means represent untransformed RTs (i.e., the actual means).

Voluntary task

To examine age effects on the RTs, two models were constructed: One examined the mixing effect (only including blocked and non-switch trials) and one examined the switching effect (only including non-switch and switch trials). Trial type, age, and language were included as fixed effects. Age was coded as -0.5 for younger adults and 0.5 for older adults. For language, Basque trials were coded as -0.5 and Spanish as 0.5. In the mixing-effect analysis, blocked trials were scored as -0.5 and non-switch trials as 0.5. This model converged with all subject slopes and item slopes for trial type, language, age, and language x age. In the switching-effect analysis, non-switch trials were coded as -0.5 and switch trials as 0.5. This model converged with all subject and item slopes.

Cued task

For the cued task, we also constructed two main models to examine age effects on mixing and switching costs. First, to examine the switching costs we included mixed switch and non-switch trials and constructed a model with the predictors language (Basque coded as -0.5; Spanish as 0.5), trial type (non-switch coded as -0.5; switch as 0.5), predictability (predictable coded as -0.5; unpredictable as 0.5), and age (younger adults coded as -0.5; older adults as 0.5), as well as all interactions. This model converged with all subject slopes and item slopes for trial type, language, age, predictability, language x age, trial type x language x age, trial type x language x predictability, language x age x predictability, and trial type x language x age x predictability. To examine age effects on mixing costs, we constructed a model only including mixed non-switch trials and blocked trials. This model included language and age (coded as described above) as well as condition (blocked / predictable non-switch / unpredictable non-switch trials). Condition was coded as follows. To assess the mixing effect ('trial type'), the levels were coded as -2/3 (blocked), 1/3 (predictable non-switch), and 1/3 (unpredictable non-switch), thus comparing blocked to non-switch trials. To compare predictable and unpredictable non-switch trials ('predictability'), we coded these levels respectively as 0, -0.5, and 0.5. This model converged after removing all item slopes (keeping the subject slopes).

Results[1]

Voluntary task

Switching frequency and accuracy

On average, the younger adults switched on 39.1% of the trials (*SD* = 10.5, see Figure 1). One older participant did not switch at all while another older participant only switched on 4% of

[1] Studies comparing older and younger adults typically perform additional analyses to correct for age-related differences in overall response speed. We therefore also calculated ANOVAs on z-scored RT costs. These analyses showed similar result patterns as the reported mixed-effect models.

the trials and did not produce any Spanish non-switch trials. These two participants were excluded from all further analyses. Without these two participants, mean switching frequency in the group of older adults was 32.8% (*SD* = 9.8). Overall, older adults switched less often than younger adults (see Appendix A for the statistics). Of the trials that could be classified as switch or non-switch trials, younger adults named 57.5% (*SD* = 7.8) in Basque, similar to the older adults (*M* = 55.8%, *SD* = 10.3). For both age groups, the percentage of switch trials was higher in Spanish than Basque. Of the Basque trials, 35.1% (*SD* = 11.2) were switch trials for the younger adults and 30.2% (*SD* = 10.4) for the older adults. Of the Spanish trials, 46.6% (*SD* =12.1) were switches for the younger adults and 39.9% (*SD* =16.8) for the older adults.

Accuracy was high in all conditions for both age groups and was not analysed further. In the blocked Basque condition, mean accuracy was 98.4% (*SD* = 1.6) for younger adults and 94.5% (*SD* = 4.4) for older adults. For the Spanish blocked condition, mean accuracy was 97.9% (*SD* = 3.3) for younger adults and 96.6% (*SD* = 3.1) for older adults. In the voluntary mixed condition, mean accuracy was 99.1% (*SD* = 1.2) for younger adults and 97.8% (*SD* = 1.8) for older adults.

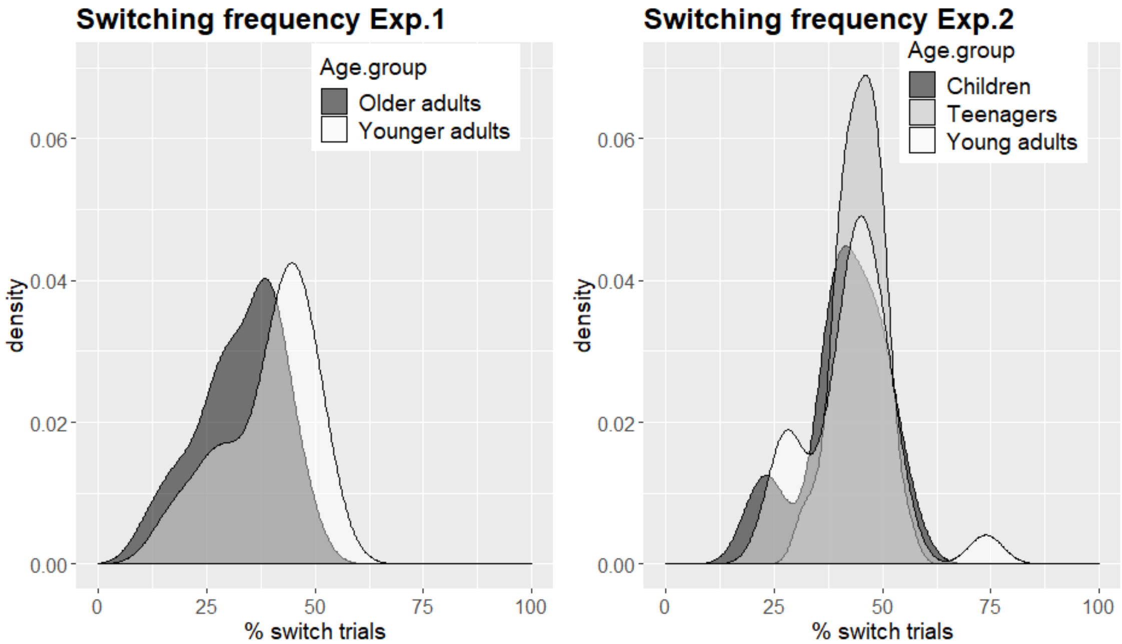

Figure 1. Density plots showing the switching frequency per age group in Experiment 1 (left) and Experiment 2 (right). The density plot for young adults is presented in white, the plot for teenagers is in light grey, and the plots for the older adults and children are dark grey. Only participants who were included in the analyses (i.e., those that produced switch and non-switch trials in both languages) are shown here.

Reaction times

Prior to the RT analysis, RTs more than 2.5 *SD* above or below the mean (of the log RTs per participant, trial type, and language; 2.0% of trials) were removed. Figure 2 shows the RT mixing and switching effects per age group and language. Mean RTs per trial type and language are provided in Table 2. To foreshadow the voluntary RT results, older adults showed larger switching costs than younger adults but similar mixing benefits.

Switching effects

Table 3 shows the full results from the switching analysis. Not surprisingly, older adults responded more slowly than younger adults (see Table 2). Basque responses were faster than Spanish responses. There was a main effect of trial type as responses were faster to non-switch than switch trials. Given the clear switching cost, a central question in Experiment 1 is whether this cost is greater for older adults than for younger adults. The answer is Yes: The switching cost interacted with age, with older adults producing larger switching costs than younger adults (see Figure 2). There were no further interactions between age, language, or switching cost.

Mixing effects

Table 4 shows the full results from the mixing analysis. The model examining mixing costs/benefits also showed faster responses for younger than older adults and faster Basque than Spanish responses. There was a main effect of mixing, indicating an overall mixing benefit with faster responses to non-switch than blocked trials (see Table 2 and Figure 2). Of particular interest here, the mixing benefit did not interact significantly with age, showing that it was comparable for younger and older adults (see Figure 2). The mixing effect was furthermore comparable for both languages in both age groups.

To sum up the findings, older adults showed larger voluntary switching costs than younger adults but comparable mixing benefits.

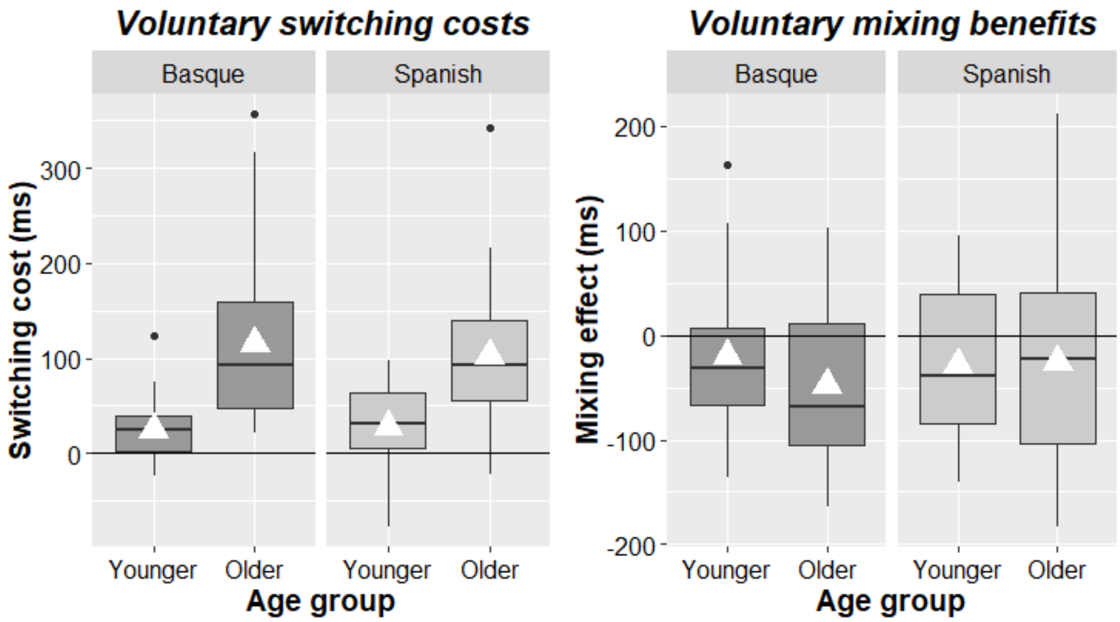

Figure 2. Boxplots showing the voluntary switching costs (left; RT difference between switch and non-switch trials) and mixing effects (right; RT difference between mixed non-switch and blocked trials) for younger and older adults in Basque (left panel) and Spanish (right panel). The boxplot shows the interquartile range with the black dots representing the outliers falling outside 1.5*interquartile range. The median is indicated by the horizontal black line and the centres of the white triangles show the means.

Table 2. *Mean RTs (and standard deviations) for blocked, voluntary non-switch, and voluntary switch trials per language and age group.*

	Younger adults	**Older adults**
Blocked		
Spanish	861.3 (116.5)	950.0 (121.5)
Basque	814.8 (105.8)	908.9 (111.2)
Voluntary non-switch		
Spanish	832.9 (119.9)	924.9 (129.2)
Basque	795.6 (107.1)	861.0 (124.2)
Voluntary switch		
Spanish	860.9 (119.2)	1027.5 (156.6)
Basque	820.8 (109.3)	976.2 (161.6)

Table 3. *Results of the voluntary switching model (Experiment 1). For each predictor, the estimate, standard error, and t value are given with an asterisk indicating a significant effect (in bold). Basque trials were coded as -0.5; Spanish trials as 0.5. Non-switch trials were coded as -0.5; switch trials as 0.5. Younger adults were coded as -0.5; older adults as 0.5.*

Predictor	Estimate	SE	T value
Intercept	**6.753**	**0.021**	**318.46***
Trial type (switching)	**0.068**	**0.007**	**9.58***
Age	**0.124**	**0.039**	**3.15***
Language	**0.052**	**0.012**	**4.28***
Age x Trial type	**0.073**	**0.013**	**5.48***
Trial type x Language	-0.002	0.011	-0.22
Age x Language	0.011	0.020	0.56
Age x Trial type x Language	-0.014	0.023	-0.63

Table 4. *Results of the voluntary mixing model (Experiment 1). For each predictor, the estimate, standard error, and t value are given with an asterisk indicating a significant effect (in bold). Basque trials were coded as -0.5; Spanish trials as 0.5. Blocked trials were coded as -0.5; non-switch trials as 0.5. Younger adults were coded as -0.5; older adults as 0.5.*

Predictor	Estimate	SE	T value
Intercept	**6.736**	**0.020**	**335.39***
Trial type (mixing)	**-0.030**	**0.012**	**-2.38***
Age	**0.095**	**0.035**	**2.73***
Language	**0.051**	**0.014**	**3.69***
Age x Trial type	-0.010	0.024	-0.41
Trial type x Language	0.009	0.011	0.77
Age x Language	0.003	0.023	0.11
Age x Trial type x Language	0.038	0.023	1.70

Cued task

Accuracy

In the cued task, accuracy was high in all conditions for both age groups (Young adults blocked Basque: *M* = 98.3%, *SD* = 2.2; blocked Spanish: *M* = 97.4%, *SD* = 3.3; cued predictable: *M* = 98.3%, *SD* = 2.3; cued unpredictable: *M* = 98.6%, *SD* = 2.0; Older adults blocked Basque: *M* = 95.5%, *SD* = 5.1; blocked Spanish: *M* = 96.9%, *SD* = 2.5; cued predictable: *M* = 95.3%, *SD* = 4.1; cued unpredictable: *M* = 95.7%, *SD* = 3.7) and was therefore not analysed further.

Reaction times

RT outliers were removed per participant and condition (language, trial type, predictability), removing 1.9% of trials. Figure 3 (switching costs) and Figure 4 (mixing costs) show the effects of age per predictability condition and language. Table 5 shows the means and standard deviations per age group, trial type, condition, and language. The cued RTs showed larger mixing and switching costs for older than younger adults.

Switching effects

First, we examined age effects on the switching costs by only including the mixed (switch and non-switch) trials. Table 6 shows the full results from this model. To summarise the main findings, this model showed that older adults responded more slowly overall than younger adults (see Table 5). Overall, responses in Basque were faster than in Spanish. Response times were comparable in the predictable and unpredictable conditions. There was also a significant switching cost, with slower responses on switch than non-switch trials (see Figure 3).

With respect to the interactions, of main interest here was the interaction between age and trial type: Switching costs were larger for older than younger adults (see Figure 3). Switching costs were also larger in the predictable than unpredictable condition (see Figure 3). Numerically, age effects were larger in the predictable than unpredictable condition (see Figure 3), but this three-way interaction was not significant.

With respect to language effects, trial type interacted with language, with larger Basque than Spanish switching costs. Language did not interact with any other predictor, suggesting that language did not significantly modulate any of the age effects.

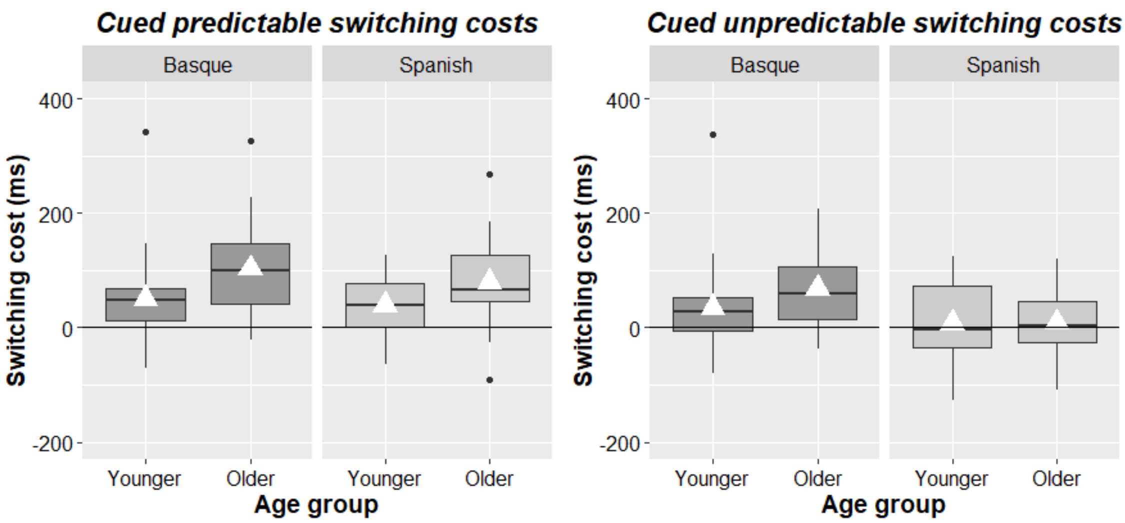

Figure 3. Boxplots showing the switching effects (RT difference between switch and non-switch trials) for younger and older adults in the cued predicable (left) and unpredictable condition (right). Per condition, the left panel shows the switching effects to Basque and the right panel to Spanish. The boxplot shows the interquartile range with the black dots representing the outliers falling outside 1.5*interquartile range. The median is indicated by the horizontal black line and the centres of the white triangles show the means.

Table 5. *Means (and standard deviations) for the younger and older adults in the blocked condition per language and in the mixed condition per trial type, language, and predictability condition.*

	Younger adults		Older adults	
Blocked				
Spanish	844.8 (108.0)		944.4 (119.6)	
Basque	782.9 (107.0)		917.8 (110.1)	
	Predictable	**Unpredictable**	**Predictable**	**Unpredictable**
Mixed non-switch				
Spanish	879.1 (143.2)	871.4 (158.0)	1030.5 (155.0)	1054.4 (167.3)
Basque	829.4 (110.5)	825.5 (117.6)	987.1 (129.5)	1010.5 (164.2)
Mixed switch				
Spanish	918.4 (149.3)	878.5 (149.2)	1109.4 (147.4)	1063.4 (189.5)
Basque	879.8 (150.9)	859.6 (172.9)	1090.0 (171.8)	1078.9 (197.6)

Table 6. *Results of the cued switching model (Experiment 1). For each predictor, the estimate, standard error, and t value are given with an asterisk indicating a significant effect (in bold). Basque trials were coded as -0.5; Spanish trials as 0.5. Non-switch trials were coded as -0.5; switch trials as 0.5. Predictable trials were coded as -0.5; unpredictable trials as 0.5. Younger adults were coded as -0.5; older adults as 0.5.*

Predictor	Estimate	SE	T value
Intercept	**6.831**	**0.022**	**311.09***
Trial type (switching)	**0.046**	**0.005**	**9.48***
Age	**0.191**	**0.040**	**4.76***
Language	**0.031**	**0.012**	**2.49***
Predictability	-0.012	0.011	-1.15
Age x Trial type	**0.022**	**0.010**	**2.28***
Trial type x Language	**-0.023**	**0.010**	**-2.23***
Age x Language	-0.023	0.020	-1.13
Trial type x Predictability	**-0.040**	**0.008**	**-5.00***
Language x Predictability	-0.014	0.008	-1.73
Age x Predictability	0.019	0.021	0.88
Age x Trial type x Language	-0.013	0.020	-0.65
Age x Trial type x Predictability	-0.016	0.016	-1.00
Age x Language x Predictability	-0.006	0.017	-0.34
Trial type x Language x Predictability	-0.018	0.017	-1.08
Age x Trial type x Language x Predictability	-0.024	0.032	-0.75

Mixing effects

Next, we examined whether older and younger adults differed in terms of mixing costs. The full results are presented in Table 7. In line with the previous analysis, there were main effects of age and language reflecting faster responses for younger than older adults and for Basque than Spanish. The main effect of trial type showed that there was a cued mixing cost, with slower responses to non-switch than blocked trials (see Table 5 and Figure 4). Of main interest, the mixing cost interacted with age, reflecting significantly higher mixing costs for older adults than younger adults (see Figure 4).

In summary, older adults showed larger cued switching and mixing costs than younger adults. In addition, switching costs were found to be larger when switching to Basque and in the predictable condition.

Table 7. *Results of the cued mixing model (Experiment 1). For each predictor, the estimate, standard error, and t value are given with an asterisk indicating a significant effect (in bold). Basque trials were coded as -0.5; Spanish trials as 0.5. Trial type was coded as -2/3 (blocked), 1/3 (predictable non-switch), and 1/3 (unpredictable non-switch); for predictability we coded these levels respectively as 0, -0.5, and 0.5. Younger adults were coded as -0.5; older adults as 0.5.*

Predictor	Estimate	SE	T value
Intercept	**6.785**	**0.021**	**327.33***
Trial type (mixing)	**0.067**	**0.010**	**6.65***
Age	**0.163**	**0.035**	**4.60***
Language	**0.045**	**0.010**	**4.76***
Predictability	0.006	0.011	0.53
Age x Trial type	**0.049**	**0.020**	**2.45***
Trial type x Language	-0.012	0.010	-1.14
Age x Language	-0.026	0.019	-1.34
Language x Predictability	-0.007	0.010	-0.75
Age x Predictability	0.027	0.022	1.19
Age x Trial type x Language	0.034	0.020	1.64
Age x Language x Predictability	0.006	0.019	0.31

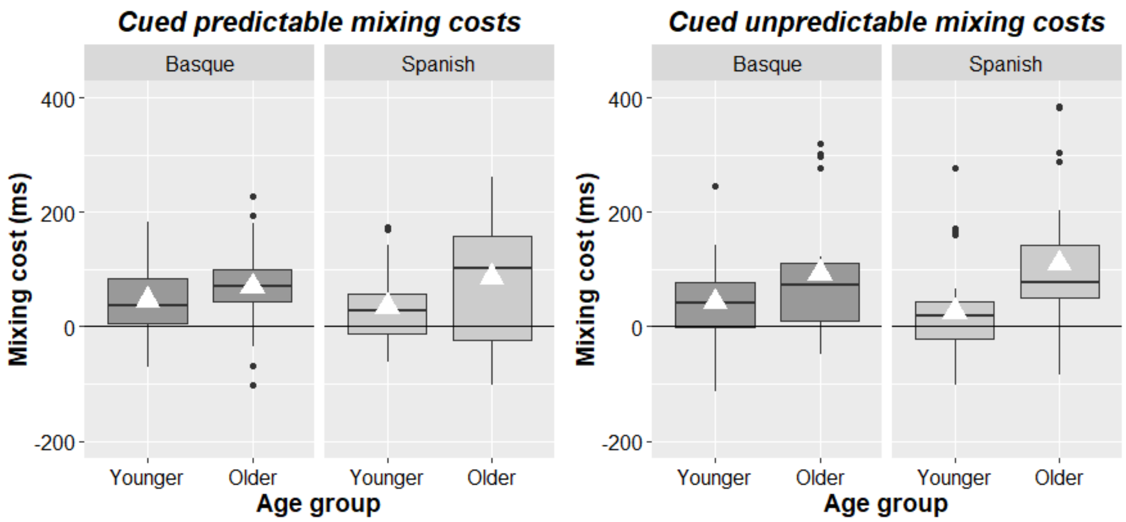

Figure 4. Boxplots showing the mixing effects (RT difference between mixed non-switch and blocked trials) for younger and older adults in the cued predicable (left) and cued unpredictable condition (right). Per condition, the left panel shows the mixing effects in Basque and the right panel in Spanish. The boxplot shows the interquartile range with the black dots representing the outliers falling outside 1.5*interquartile range. The median is indicated by the horizontal black line and the centres of the white triangles show the means.

Discussion

While accuracy was high for both older and younger adults, several age effects were observed in terms of response times, on both tasks. In addition to older adults being slower overall, they showed larger mixing and switching costs on the cued task than younger adults. On the voluntary task, they showed mixing benefits comparable to the younger adults, but larger switching costs. Together, these results suggest that for older adults, maintaining and using two languages as such is not more effortful than using one language when there are no external demands such as language cues. However, the more reactive processes believed to underlie switching costs may be more effortful regardless of whether the switch is executed voluntarily or in response to external cues.

The cued task thus showed an age effect on language mixing costs. This is consistent with several other studies that have shown larger mixing effects for older than younger adults, both in language-switching (e.g., Hernandez & Kohnert, 2015, Weissberger et al., 2012) and task-switching (e.g., Kray & Li, 2000) paradigms. In cued task-switching paradigms, this deficit has been linked to working memory capacity. However, from cued tasks, it is not clear whether older adults have greater difficulty using two tasks/languages as such or encounter difficulties when using two tasks/languages *in response to specific instructions*. The presence of age effects on cued mixing costs but not on voluntary mixing effects suggests that the latter is the case. Age effects do not appear to reflect overall deficits in maintaining and using two languages in memory, but rather the need to use the two languages in a specific, externally cued manner.

In addition, we observed an age effect on switching costs, showing larger switching costs for older than younger adults. This is in line with some previous language switching studies (e.g., Weissberger et al., 2012), while others have suggested that age effects are only found on mixing costs (Hernandez & Kohnert, 2015). This is in contrast to the task-switching literature showing reliable age effects on mixing but not on switching costs (Wasylyshyn et al., 2011). The current study suggests that age effects on language control go beyond proactive cognitive processes and can also be observed in the more reactive, transient processes related to language-set reconfiguration. While this may have implications for the task-switching literature, these findings may be specific to language control (cf. Calabria et al., 2015; Weissberger et al., 2012, showing diverging age effects on language versus task switching). We furthermore observed larger switching effects in the predictable condition. This was a surprising finding that might be related to unpredictable and more demanding conditions recruiting more controlled task processing, in particular when sufficient preparation time is given (cf. Tornay & Milán, 2001). We will elaborate on these findings in the General Discussion.

The few studies examining age effects on language switching have yielded inconsistent findings, with two studies showing age effects on both mixing and switching costs (Weissberger et al., 2012 and the current study), one study showing larger age effects when comparing the mixed condition as a whole to the blocked condition (Hernandez & Kohnert, 1999), one study showing age effects on mixing costs only (Hernandez & Kohnert, 2015), and one study not showing age effects beyond overall RTs (Calabria et al., 2015). Switching/mixing costs and age effects may be affected by methodological differences such as the timing of the task, including the interval between the cue and stimulus (CSI). Switching costs may decrease when participants are given sufficient time to process the cue (e.g., Mosca & Clahsen, 2016). In our study, we used a CSI of 500 ms to ensure that all age groups had sufficient time to process the cue without risking a reduction of switching costs by using a long CSI. The resulting CSI of 500 ms is somewhat longer than CSIs (around 250-300 ms) that have been used with younger adults (e.g., Gollan et al., 2014) but shorter than some other studies testing older adults (e.g., Weissberger et al., 2012 and Calabria et al., 2015 used CSIs of 1 second or longer). It is possible that the specific CSI modulates age effects on switching (and mixing) costs.

The same may be true for the response-to-stimulus interval (RSI). In the current study, all pictures stayed on the screen for 2500 ms regardless of when the responses were made. This resulted in a long RSI that was variable but on average around two seconds. Long RSIs may allow for more passive decay of the activation/interference of the previously used language. Both cued and voluntary switching costs were robust, but it is possible that they would have been larger with a shorter RSI (cf. Declerck, Koch, & Philipp, 2016; Horoufchin, Philipp, & Koch, 2011). The overall longer RTs for older adults led to somewhat smaller RSIs for this age group. While it is possible that this could explain part of the larger switching cost in older adults, the group differences in overall response times (and thus RSIs) were small, especially relative to the overall RSI (approximately 100 to 200 ms relative to an RSI of approximately two seconds).

The voluntary task of the current study showed age effects on the switching costs, with older adults showing a larger cost than younger adults. This is line with Gollan and Ferreira (2009), who also observed larger switching costs (to the non-dominant language) for older adults than younger adults and with a voluntary task-switching study that also showed larger switching costs for older adults (Butler & Weywadt, 2013). Even when made voluntarily, responses on switch trials are slower than non-switch trials. This shows that switching between languages is costly time-wise even when done voluntarily and that this may be especially demanding for older adults. Despite this cost, older adults voluntarily switched between their languages and did so frequently in both the current study and Gollan and Ferreira (2009). Even if voluntary language switching is costly, it is not so costly that older adults avoid switching between their languages. We will further discuss age effects on voluntary switching costs in the General Discussion.

In contrast to the switching cost, an overall mixing benefit was observed, suggesting that voluntarily using two languages was less costly than having to stay in one language. This mixing benefit was not affected by age, indicating that voluntarily using two languages was equally beneficial for older and younger adults. These findings are consistent with those of Gollan and Ferreira (2009), who also did not observe age effects on language mixing. Gollan and Ferreira (2009) reported a mixing cost for the dominant language, while the current study shows a mixing benefit for both languages. In addition to several potential methodological differences that might play a role (e.g., familiarisation task, exact task instructions, the duration of the stimulus on the screen, task order, and stimulus repetition), one of the most crucial difference is the type of bilinguals tested. Whereas we tested early balanced bilinguals, the participants in Gollan and Ferreira (2009) were more dominant in one language than the other, especially in terms of language use. Furthermore, the current study tested bilinguals living in a bilingual society. These bilinguals sometimes need to control their languages (e.g., when they are speaking to an interlocutor who only speaks one of their languages) but are often also surrounded by other bilinguals who speak the same languages, allowing free language use. Bilinguals living in a bilingual society may be more prone to show voluntary mixing benefits.

Our relatively balanced bilinguals tended to be somewhat more proficient in Spanish than Basque, but their responses were faster in Basque than Spanish. This result aligns with previous studies testing young adults, showing faster Basque responses in highly proficient Spanish-Basque bilinguals even when proficiency is somewhat higher in Spanish (e.g., de Bruin et al., 2018; Jevtović et al., in press). Most young adults were educated in Basque and Basque was the first acquired language for the older adults. Each factor could have contributed to faster Basque responses. In addition, both younger and older adults show an overall preference for Basque as a naming language, so this may be their preferred language to use even though this minority language may not always be the one they can use most often in daily life.

Experiment 2. Language switching in children, teenagers, and young adults

In the second experiment, we examined development of bilingual language control and switching on the other side of the lifespan by comparing children (7 to 9 years old), teenagers (12 years old), and young adults (18 to 35 years old). During childhood and adolescence, several aspects of cognition and executive control have been found to develop rapidly (e.g., Anderson, 2002; Davidson, Amso, Anderson, & Diamond, 2006). Different aspects of control might mature at different ages. Initial stages of attentional control, including the ability to hold information in mind, start to emerge during an infant's first year (cf. Diamond, 2006).
More complex functions develop during a child's first five years, with mechanisms related to inhibition argued to greatly improve between three and five years of age (Diamond, 2006). Children in this age range are able to inhibit responses and can hold information in mind to guide their behaviour. However, inhibitory control continues to mature throughout childhood and adolescence. Children's increased susceptibility to interfering information and their lower ability to suppress irrelevant responses have been linked to ongoing development of the prefrontal cortex during childhood and adolescence (e.g., Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002). Furthermore, it might not just be the application of inhibition but

also the ability to overcome previously applied inhibition that continues to develop (Davidson et al., 2006). During childhood, control processes are also argued to develop from being more reactive (i.e., responding to events as they happen) to becoming more proactive (i.e., keeping information in mind in anticipation of an event, e.g., Chatham, Frank, & Munakata, 2009). Related to this, cognitive flexibility (i.e., the ability to combine different demands, hold multiple pieces of information in mind, and continuously change between inhibiting and activating different pieces information) might be especially difficult for children and might mature at a later stage (e.g., Anderson, 2002; Davidson et al., 2006; Diamond, 2006). While cognitive flexibility might develop steeply between 5 and 11 years (e.g., Diamond, 2006), it does not appear to reach full maturation until early adulthood (Cepeda et al., 2001; Davidson et al., 2006).

In line with this pattern, studies assessing the effects of age on task-switching performance have shown an inverted U-shaped function across the lifespan. That is, performance increases (in terms of higher accuracy and faster responses) during childhood, reaching a peak in early adulthood, and then decreases during later adulthood. Cepeda et al. (2001), for example, showed larger RT switching costs and mixing costs for children and older adults compared to younger adults. Reimers and Maylor (2005) tested participants aged between 10 and 66 and found U-shaped age effects on response time mixing costs, but not on switching costs. Similar findings were observed by Kray, Eber, and Lindenberger (2004) testing 8-10 year old children, younger adults (20-25 years old), and older adults (61-72 years old). Thus, similar to older adults, age effects for children on task switching may be more likely to be observed on mixing costs than switching costs. That is, children may have increased difficulty maintaining two tasks in working memory and manipulating them in response to external task demands, in line with the argument that proactive control is developed later than reactive control (e.g., Chatham et al., 2009).

While there are several corpus analyses or case studies focusing on linguistic aspects of code switching in children (see e.g., Basnight-Brown & Altarriba, 2007, for an overview), there is little experimental work, especially studies with comparisons between children and adults. In a recent study (Gross & Kaushanskaya, 2018), 5 to 7 year old bilinguals named pictures in one language only or in two languages in response to cues. The children made more cross-language errors and produced slower responses in the mixed than single language condition, confirming that language mixing costs can be observed in children too. To examine possible age effects during childhood and adolescence, Jia, Kohnert, Collado, and Aquino-Garcia (2006) asked children of four different age groups (5-7 years, 8-10 years, 11-13 years, 14-16 years old) to name action pictures while using one language or two languages in a predictable switching pattern. The study focused on mixing costs and observed accuracy mixing costs for the three youngest age groups but not for the 14-16 year olds. Response time mixing costs were observed for all age groups and did not interact with age. Significant switching costs were observed in both accuracy and RTs, with some age-related changes with respect to the relative performance in each of the two languages. Similarly, Kohnert, Bates, and Hernandez (1999) studied single-language and mixed-language picture naming in children aged between 5 and 16 years old and in young adults. Age effects were observed in the accuracy scores and RTs, with mixing costs being smallest for the 14-16 year olds and young adults.

When it comes to voluntary switching, as in young adults (e.g., de Bruin et al., 2018; Gollan & Ferreira, 2009), 5 to 7 year old bilinguals have shown switching costs (Gross & Kaushanskaya, 2015). In terms of the mixing effect, Gross and Kaushanskaya (2015) found that the results depended on the language and measure. For the non-dominant language, either a benefit (accuracy) or no cost (RT) was found, while the dominant language showed a cost (RT) or no effect of mixing (accuracy). However, as this study only included children, it is unclear how these findings compare to the costs and benefits observed in young adults with a comparable language profile.

Thus, several previous studies have shown that children show cued mixing costs and cued and voluntary switching costs when using two languages in a picture-naming task. Cued mixing costs may decrease as children get older. However, few studies have compared children's performance to adults; furthermore, most focus on mixing costs and do not or only briefly discuss switching costs.

Based on the existing task- and language-switching literature, it appears that the ability to maintain two tasks and manipulate them in response to task demands continues to develop during childhood and adolescence. Given the age effect on language switching costs in older adults (which is not always found in non-linguistic switching tasks) and the scarcity of language-switching research in children, an open question is whether reactive control related to the process of switching is also still in development in children. Of course, while cognitive development and ageing may show comparable patterns in several non-linguistic domains (including task switching), this does not necessarily mean that bilingual language control is affected in the same way. Older bilinguals have a great deal of experience using and controlling two languages, while children have less experience with bilingual language control. Furthermore, while vocabulary continues to develop during childhood, it does not typically decline during later adulthood (e.g., Brysbaert, Stevens, Mandera, & Keuleers, 2016). As such, age effects may show different patterns than in Experiment 1.

Experiment 2 therefore examines how children, teenagers, and young adults perform on cued and voluntary picture-naming tasks. As in Experiment 1, we assessed whether age affected mixing as well as switching costs in cued and voluntary tasks. Based on task-switching studies (e.g., Cepeda et al., 2001), as well as the results reported with older adults in Experiment 1, we expected youths to show greater difficulty using two languages and switching between them (i.e., larger mixing and switching costs) in the cued task. Importantly, the children in the age group tested here (7-9 years old) should already be able to inhibit information and to hold multiple (simple) tasks in mind (unlike younger children, cf. Davidson et al., 2006). Thus, these age groups allow us to examine the development of more complex manipulations (i.e., language mixing and switching) that require inhibition and working memory. The children in this study were expected to be at an age at which they could apply inhibitory control, but we expected that they would require greater effort than young adults. Given that inhibition and cognitive flexibility have been argued to continue developing during childhood (e.g., Diamond, 2006), we also included a group of 11/12 year-olds (‘teenagers’). Including this age group allowed us to better monitor the ongoing development of language control by assessing whether language control continuous to develop after early childhood. Based on previous studies (e.g., Cepeda et al., 2001; Davidson et al., 2006), we expected ongoing development between early childhood and early adolescence.

Contrary to Experiment 1 (which tested bilinguals with a high and comparable proficiency level in both languages), Experiment 2 tested bilingual youths who were still in the process of acquiring their second language. Inhibitory control has been argued to play an important role in language switching, especially in unbalanced bilinguals who need to suppress their dominant language (Green 1998; Meuter & Allport, 1999). Previous (non-linguistic) studies have suggested that children not only have greater difficulty applying inhibition but also undoing inhibition (e.g., Davidson et al., 2006). In unbalanced bilinguals, more inhibition may be needed to suppress the stronger L1, resulting in longer RTs when switching back to the L1 due to the need to overcome the L1 inhibition applied during L2 production ('asymmetric switching cost', Green, 1998; Meuter & Allport, 1999). If this process of overcoming inhibition is more difficult for youths, age effects on switching/mixing costs may also be language-dependent (i.e., larger costs when using the L1).

With respect to the voluntary task, Experiment 1 suggested that age effects on mixing costs may only be found when the two languages need to be used in response to external cues but not when used freely. In line with this finding, we expected youths to show larger mixing costs than young adults on the cued task but not on the voluntary task. Considering the presence of voluntary switching costs that were larger for older adults, however, it is possible that youths will show both larger cued and voluntary switching costs than young adults.

Methods

Participants

Experiment 2 was completed by thirty children and thirty teenagers. The number of participants was based on the power analyses described for Experiment 1.To ensure that all participants had at least an intermediate proficiency level in both languages, we excluded all participants scoring less than 40 out of 65 points on the Basque objective proficiency measurement (six children); this proficiency cut-off was also used for the recruitment of young adults and was met by all participants in Experiment 1. In addition, we excluded participants who did not know the name for more than half of the pictures in the Basque blocked condition of either the cued or voluntary language switching task (an additional four children). Spanish proficiency was high for all participants. The final sample consisted of 20 children (14 female; *M* age = 8.2, *SD* age = 0.7, range 7-9) and 30 teenagers (21 female; *M* age = 11.8, *SD* age = 0.4, range 11-12).

Since the age of two, the children and teenagers had been attending a trilingual school system in which approximately 60% of classes are taught in Basque, 20% in Spanish, and 20% in English. However, almost all children and teenagers were Spanish dominant and only or predominantly spoke Spanish at home with their parents. To match this language profile, we selected 25 young adults (*M* age = 25.6, *SD* = 4.3, range 19-35) from de Bruin et al. (2018; ten of these participants were also included in the analyses for Experiment 1). Initially, we selected the young adults with the lowest Basque objective proficiency and self-reported use scores. However, the vast majority of participants in this selection completed the cued task first. For the final analyses, we therefore selected 25 young adults with lower Basque proficiency/use scores while maintaining a more balanced count of cued versus voluntary task first. Importantly, the main results reported below were very similar regardless of the group of young adults selected.

All participants completed the 65-item picture-naming task to assess their proficiency in Basque and Spanish. In addition, the participants or parents (in the case of children and teenagers) completed a questionnaire including questions about age of acquisition, proficiency, exposure to, and use of Basque and Spanish (these data are missing for three children and for eight teenagers, and partly for an additional two teenagers). Repeated-measures ANOVAs with age group as the between-subject and language as within-subject variable showed significant differences between Basque and Spanish on all measurements reported in Table 8 (see Supplementary Table 2 for full statistics).

Matching children and adults on self-/parent-rated proficiency or complex proficiency measurements may not be overly informative because children are still developing their language skills. Indeed, children and teenagers were given significantly lower parent-rated proficiency scores than the adults' self-ratings, but this did not interact with language on any self-ratings other than understanding. We did, however, match the three age groups on their scores on the picture-naming proficiency task. This is a relatively simple vocabulary task that shows ceiling performance in the dominant language (Spanish), even in children, but is sensitive to proficiency differences in the non-dominant language (cf. de Bruin et al., 2017). Furthermore, this task is comparable in difficulty to the vocabulary required to complete the experimental task. By matching age groups on these proficiency scores, we aimed to minimise the possible influence of vocabulary gaps on language switching performance. In addition, the three age groups were comparable in both languages in terms of age of acquisition and their percentage of Spanish and Basque use (see Table 8). In terms of exposure to Basque and Spanish, there was a significant interaction between language and age group, showing that children received more exposure to Basque than teenagers and adults, with no differences for Spanish.

Parents were also asked to indicate how often their child switches languages, on a scale from 1 ('never') to 4 ('all the time'). No significant age effects were observed for switching on a daily basis (children M = 2.1, SD = 0.9; teenagers M = 2.2, SD = 0.8; adults M = 2.7, SD = 1.0; F(2,

60) = 2.73, p = 0.073); for switching within a conversation (children M = 2.0, SD = 0.9; teenagers M = 2.2, SD = 0.8; adults M = 2.2, SD = 0.9; $F(2, 59)$ = 0.19, p = 0.83); or within a sentence (children M = 1.8, SD = 0.8; teenagers M = 2.1, SD = 0.7; adults M = 2.0, SD = 0.9; $F(2, 59)$ = 0.73, p = 0.49).

All participants had normal or corrected-to-normal vision and hearing and no known neurological or language impairments. The study was approved by the BCBL ethics committee. The participants either gave informed consent themselves or, in the case of the children and teenagers, consent was given by the parents.

Table 8. *Summary of the language profiles of the 20 children, 30 teenagers, and 25 young adults. Means and (standard deviations) are provided for the different measurements of language proficiency and use for Spanish and Basque per age group. A single asterisk indicates a significant difference between age groups across languages; a double asterisk indicates a significant interaction between language and age group.*

	Spanish			**Basque**		
	Children	Teen-agers	Adults	Children	Teen-agers	Adults
AoA	0.0 (0.0)	0.0 (0.0)	0.4 (1.0)	1.9 (1.1)	2.0 (0.6)	1.7 (1.6)
Picture naming (0-65)	64.7 (0.7)	64.7 (0.5)	64.7 (0.7)	52.7 (8.1)	52.5 (5.6)	56.2 (6.9)
Self/parent-rated proficiency (0-10)						
*Speaking **	8.8 (0.8)	8.8 (0.8)	9.4 (0.9)	7.1 (1.0)	6.9 (0.9)	8.3 (1.7)
*Understanding *, ***	8.6 (1.1)	8.7 (0.9)	9.6 (0.7)	6.7 (1.0)	7.0 (1.1)	9.1 (1.1)
*Writing **	6.3 (1.6)	7.5 (1.8)	9.1 (1.1)	6.5 (1.1)	6.7 (0.9)	8.5 (1.3)
*Reading **	6.8 (1.8)	7.7 (1.2)	9.6 (0.6)	6.4 (1.1)	6.4 (1.0)	9.1 (1.1)
*General **	8.3 (1.2)	8.3 (1.1)	9.6 (0.8)	6.9 (0.9)	6.7 (1.1)	8.5 (1.4)
%exposure (0-100) **	51.0 (21.3)	60.0 (17.7)	60.4 (15.7)	38.4 (18.9)	22.4 (11.0)	23.6 (11.1)
%speaking (0-100)	68.2 (19.5)	79.3 (15.2)	64.4 (21.4)	22.6 (13.5)	17.4 (13.1)	25.2 (18.7)

Tasks

The participants completed the same experimental tasks as in Experiment 1. The children and teenagers completed the tasks in a lab space (BCBL JuniorLab) within their school building; the equipment used in this lab space is identical to the equipment used in the labs where the younger adults were tested. Data from one teenager were not recorded for the voluntary switching task. Data from another teenager had to be excluded from the cued analysis as data from the predictable and second blocked condition were not recorded. All participants completed the voluntary and cued switching task in two different sessions separated by a few days. Of the twenty children, 11 completed the cued task in the first session and 9 the voluntary task. Of the 30 teenagers, 15 completed the cued task first. Of the 25 adults, 15 completed the cued task first. The adults completed the background tasks at the end of the voluntary task while the children and teenagers completed these in an additional third session. The first blocked part was completed in Spanish by 12 children, 15 teenagers, and 14 young adults. Within the cued task, 8 children completed the predictable task first and 12 the unpredictable task; 15 teenagers completed the predictable task first and 15 the unpredictable task; in the group of young adults, 11 completed the predictable task first and 14 the unpredictable task. Similar to Experiment 1, we ran additional analyses including task order, language order, and predictability order. None of these interacted with trial type or trial type x age (the predictors of interest; see Supplementary Materials).

Data analysis

The data analysis approach was similar to the approach in Experiment 1. Age was coded orthogonally so that we could compare youths (children and teenagers together) to adults (children: -1/3; teenagers: -1/3; adults: 2/3) as well as children to teenagers (children: -0.5; teenagers: 0.5; adults: 0) to evaluate changes during development. The voluntary switching analysis converged with all participant slopes and the item slope for language x age; the voluntary mixing analysis converged with all participant slopes and the item slope for language x age. The cued switching analysis converged with all participant slopes apart from the three-way interaction and without item slopes; the cued mixing analysis converged with all participant slopes and the item slope for language x age.

In addition, we also analysed accuracy as performance from the children and teenagers was not at ceiling. Accuracy was analysed through generalized linear mixed-effect models. Considering that accuracy is a 0 or 1 score, a binomial distribution would have been most appropriate. However, due to the accuracy scores being at ceiling for the young adults but not for the other age groups, models with a binomial distribution yielded unreliable results (e.g., flipped signs). To at least get an estimate of age effects on accuracy, a Poisson distribution was used in the generalized linear mixed-effect analyses. In the voluntary mixed condition, the majority of incorrect responses were missing or late responses or words combining both languages (82%) and could thus not be classified as a Basque or a Spanish response. We therefore first compared the blocked condition as a whole to the mixed condition to assess general age effects on the mixing effect. This model converged without any slopes (subject and item intercepts only). Next, we compared the Spanish and Basque blocked conditions separately to the mixed condition to assess whether age effects were comparable for both languages. These models converged with the participant intercepts only. Models for the cued accuracy data were constructed in the same way as for the response times. The cued accuracy switching model converged with the participant and item intercepts and participant slope for language. The cued accuracy mixing model converged with participant and item intercepts only.

Results[2]

Voluntary task

Switching frequency

Two teenagers did not produce any non-switch trials in Spanish and were excluded from all analyses. In terms of switching frequency, children switched on 41.2 % of the trials (*SD* = 9.5),

[2] Similar to Experiment 1, ANOVAs on z-scored RT costs (to account for differences in overall response speed) showed similar result patterns as the reported mixed-effects models.

teenagers on 44.6% of the trials (*SD* = 5.3), and young adults on 42.9% of the trials (*SD* = 10.6, see Figure 1). The three age groups showed comparable switching frequencies (see Appendix A for the statistics). Of the trials classifiable as switch or non-switch trials, children named 47.6% (*SD* = 16.5) in Basque, teenagers 58.0% (*SD* = 10.3), and young adults 57.8% (*SD* = 9.4). Overall, a larger percentage of trials named in Spanish were switches compared to trials named in Basque. This was the case for both teenagers (% switch Basque: 40.1%, *SD* = 10.3%; % switch Spanish: 55.5%, SD = 11.2) and young adults (% switch Basque: 38.7%, *SD* = 12.3%; % switch Spanish: 52.2%, SD = 13.1). Children, in contrast, produced a higher number of switches in their Basque responses (48.0%, *SD* = 15.4) than in their Spanish responses (43.5%, *SD* = 15.7).

Accuracy

Accuracy (see Figure 5 and Table 9; the full statistics can be found in Appendix B) increased during childhood. There was an overall accuracy mixing benefit, with more correct voluntary than blocked responses, which decreased during development. Comparing the Spanish blocked condition to the mixed condition showed similar mixing effects for all age groups while comparing the Basque blocked condition to the mixed condition showed a larger Basque mixing benefit for youths than adults and for children than teenagers.

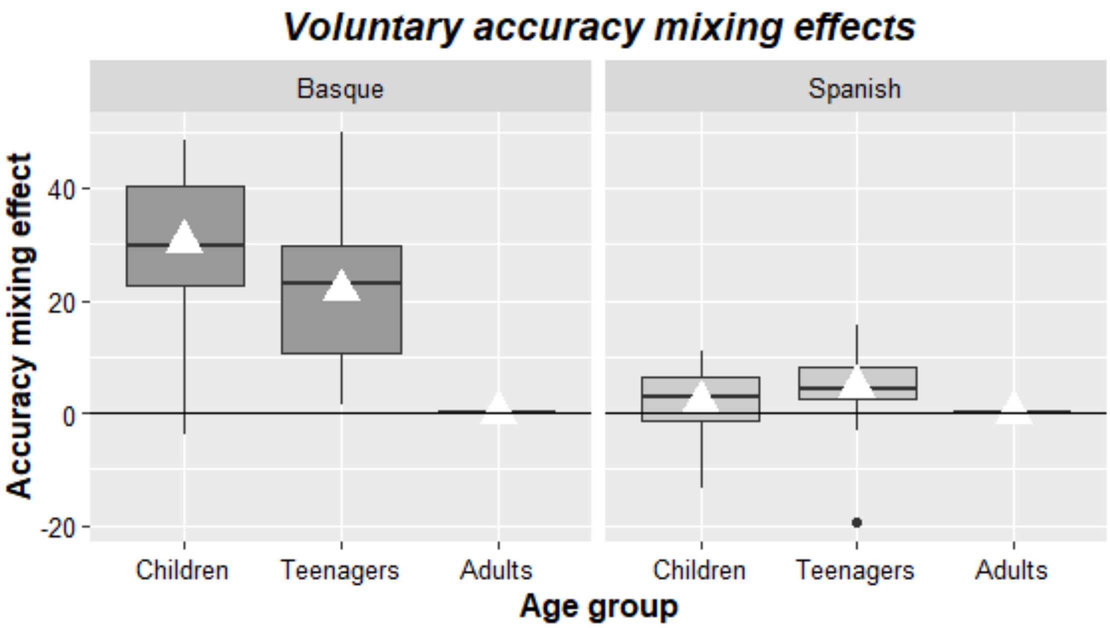

Figure 5. Boxplots showing the voluntary mixing effects (left: accuracy difference all voluntary trials – Basque blocked trials; right: accuracy difference all voluntary trials – Spanish blocked trials) for children, teenagers, and adults. Positive scores reflect a mixing benefit. The boxplot shows the interquartile range with the black dots representing the outliers falling outside 1.5*interquartile range. The median is indicated by the horizontal black line and the centres of the white triangles show the means.

Table 9. *Mean percentage correct (and standard deviations) for the children, teenagers, and adults in the voluntary task (Blocked per language and Voluntary mixed across languages) and the cued task (per language and trial type, but collapsed across predictability given that no differences were found between the predictable and unpredictable condition in accuracy). Note that in the cued task, trials preceded by a mistake were included and switch/non-switch refers to the pre-defined trial type.*

	Children	**Teenagers**	**Adults**
Voluntary			
Blocked			
Spanish	88.3 (6.9)	91.1 (5.9)	97.5 (3.7)
Basque	59.7 (13.9)	74.0 (14.0)	97.6 (3.0)
Mixed	90.1 (8.4)	95.7 (5.6)	99.4 (1.0)
Cued			
Blocked			
Spanish	87.8 (7.6)	93.1 (9.3)	97.8 (3.2)
Basque	62.8 (13.6)	75.5 (13.9)	95.2 (7.0)
Non-switch			
Spanish	79.9 (10.6)	90.7 (8.3)	98.4 (2.3)
Basque	62.3 (13.6)	74.3 (15.3)	97.9 (3.2)
Switch			
Spanish	78.3 (11.4)	87.4 (10.5)	96.6 (3.6)
Basque	60.6 (12.6)	73.4 (14.1)	96.7 (4.0)

Reaction times

RT analyses were conducted after removal of outliers following the same procedure as in Experiment 1 (1.6% of trials). Two models were constructed, one examining the switching effect (difference between switch and non-switch trials), and one examining the mixing effect (difference between blocked and mixed non-switch trials). Figure 6 shows the RT switching and mixing effects per language for children, teenagers, and young adults. Table 10 shows the RTs per trial type, language, and age group. Switching costs were similar for the three age groups. All age groups showed a mixing benefit and this benefit was larger in Basque for the youths.

Switching effects

The full results from the voluntary switching cost analysis are shown in Table 11. This analysis showed main effects of age reflecting the fact that children performed slowest (see Table 10). There was also a main effect of language, with faster responses in Basque than Spanish (see Table 10); the Basque RT advantage replicates the pattern found by de Bruin et al. (2018), despite Basque being the participants' non-dominant language. There was a main effect of trial type, reflecting faster responses to non-switch than switch trials. The switching cost did not interact with age, suggesting that switching costs were comparable for children, teenagers, and adults (Figure 6). Language did not interact with the switching cost or with age. There was also no three-way interaction between the switching cost, language, and age, confirming that switching costs were similar for the three age groups (see Figure 6).

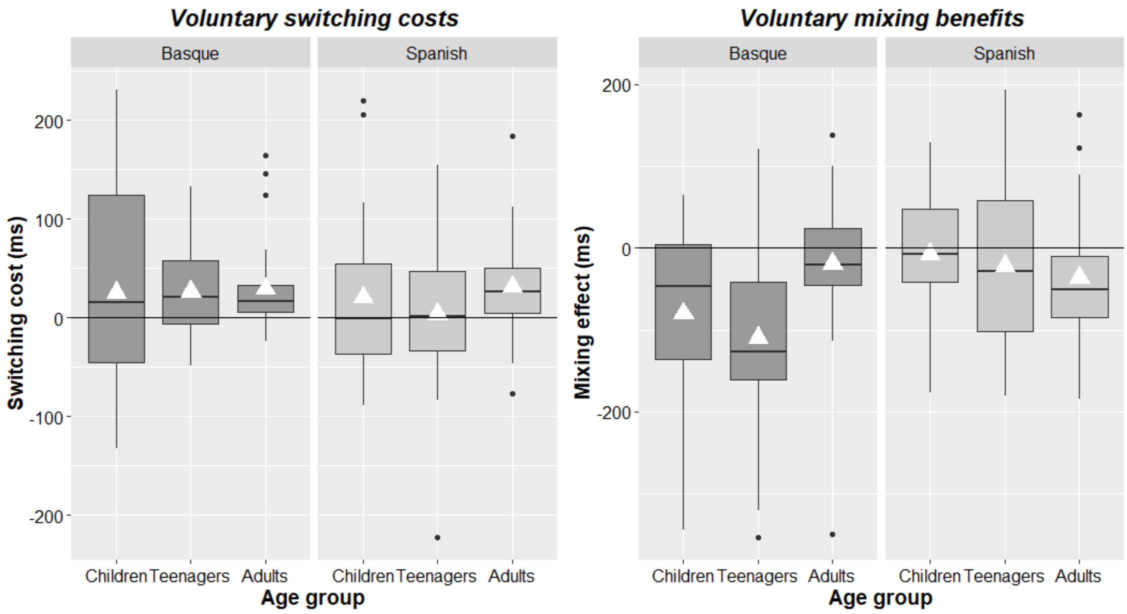

Figure 6. Boxplots showing the voluntary switching costs (left; RT difference between switch and non-switch trials) and mixing effects (right; RT difference between mixed non-switch and blocked trials) for children, teenagers, and adults in Basque (left panel) and Spanish (right panel). The boxplot shows the interquartile range with the black dots representing the outliers falling outside 1.5*interquartile range. The median is indicated by the horizontal black line and the centres of the white triangles show the means.

Table 10. *Mean RTs (and standard deviations) for the children, teenagers, and adults for blocked, voluntary non-switch, and voluntary switch trials per language.*

	Children	Teenagers	Adults
Blocked			
Spanish	1208.1 (173.3)	1081.7 (120.9)	894.1 (133.6)
Basque	1209.5 (217.1)	1061.2 (168.1)	812.9 (109.0)
Voluntary non-switch			
Spanish	1200.6 (192.9)	1060.3 (190.8)	858.0 (147.0)
Basque	1129.4 (186.2)	951.2 (158.1)	793.9 (118.6)
Voluntary switch			
Spanish	1221.1 (167.9)	1064.0 (166.6)	889.3 (157.1)
Basque	1155.0 (201.8)	977.8 (150.0)	823.3 (139.7)

Table 11. *Results of the voluntary switching model (Experiment 2). For each predictor, the estimate, standard error, and t value are given with an asterisk indicating a significant effect (in bold). Basque trials were coded as -0.5; Spanish trials as 0.5. Non-switch trials were coded as -0.5; switch trials as 0.5. For the comparison Youth vs Adult, the coding was (children: -1/3; teenagers: -1/3; adults: 2/3); for the comparison children vs teenagers the coding was (children: -0.5; teenagers: 0.5; adults: 0).*

Predictor	Estimate	SE	T value
Intercept	**6.862**	**0.019**	**363.29***
Trial type (switching)	**0.029**	**0.005**	**6.37***
Youth vs Adult	**-0.243**	**0.037**	**-6.60***
Child vs Teen	**-0.129**	**0.044**	**-2.96***
Language	**0.072**	**0.012**	**6.18***
Youth vs Adult x Trial type	0.007	0.009	0.79
Child vs Teen x Trial type	-0.013	0.011	-1.17
Trial type x Language	-0.005	0.011	-0.45
Youth vs Adult x Language	0.004	0.019	0.23
Child vs Teen x Language	0.034	0.021	1.60
Youth vs Adult x Trial type x Language	0.007	0.022	0.31
Child vs Teen x Trial type x Language	-0.015	0.027	-0.55

Table 12. *Results of the voluntary mixing model (Experiment 2). For each predictor, the estimate, standard error, and t value are given with an asterisk indicating a significant effect (in bold). Basque trials were coded as -0.5; Spanish trials as 0.5. Blocked trials were coded as -0.5; non-switch trials as 0.5. For the comparison Youth vs Adult, the coding was (children: -1/3; teenagers: -1/3; adults: 2/3); for the comparison children vs teenagers the coding was (children: -0.5; teenagers: 0.5; adults: 0).*

Predictor	Estimate	SE	T value
Intercept	**6.874**	**0.018**	**383.90***
Trial type (mixing)	**-0.045**	**0.009**	**-5.17***
Youth vs Adult	**-0.262**	**0.033**	**-7.92***
Child vs Teen	**-0.119**	**0.039**	**-3.02***
Language	**0.056**	**0.014**	**4.06***
Youth vs Adult x Trial type	0.025	0.018	1.35
Child vs Teen x Trial type	-0.013	0.022	-0.58
Trial type x Language	**0.034**	**0.012**	**2.86***
Youth vs Adult x Language	0.043	0.025	1.75
Child vs Teen x Language	0.035	0.029	1.20
Youth vs Adult x Trial type x Language	**-0.070**	**0.025**	**-2.83***
Child vs Teen x Trial type x Language	0.020	0.030	0.66

Mixing effects

The model examining age effects on the mixing effect included voluntary mixed non-switch and blocked trials. The full results are shown in Table 12. There was a main effect of trial type, showing an overall mixing benefit due to responses being faster on mixed non-switch than blocked trials. There was also an overall age effect as youths responded more slowly than adults. Similar to the model for the switching effect, responses were faster in Basque than in Spanish. The two-way interactions between age and trial type and between age and language were not significant. However, trial type interacted with language, such that the mixing benefit was larger for Basque. Furthermore, as shown in Figure 6, the larger mixing benefit for Basque than Spanish is mainly present in the groups of children and teenagers, which was also supported by the three-way interaction between age, trial type, and language. Further analyses per language showed a larger mixing benefit in Basque for youths than adults ($\beta = 0.061$, $SE = 0.022$, $t = 2.74$) but no difference between children and teenagers ($\beta = -0.026$, $SE = 0.027$, $t = -0.94$). In Spanish, the mixing benefit was comparable for youths and adults ($\beta = -0.007$, $SE = 0.022$, $t = -0.32$) and for children and teenagers ($\beta = -0.005$, $SE = 0.026$, $t = -0.20$).

To sum up the findings, children, teenagers, and young adults showed similar voluntary RT switching costs. Youths showed larger voluntary RT mixing benefits in Basque than adults but not in Spanish.

Cued switching task

Accuracy

Figure 7 shows the cued accuracy mixing and switching effects; these costs were small (see also Table 9). The full statistics can be found in Appendix B. First, we examined whether age affected accuracy in terms of the switching cost in the mixed condition[3]. Children and

[3] In this analysis, we excluded the first trial after a break, but included trials preceded by an error, because children and teenagers made more mistakes overall. Excluding correct responses preceded by a

teenagers made more errors than adults (see Table 9). Accuracy was furthermore higher in Spanish than Basque. The main effect of trial type did not reach significance, but accuracy was slightly higher in the non-switch than switch condition. These small switching costs were not affected by age. However, age interacted with language, such that the difference between Basque and Spanish accuracy decreased with age.

Similar to the model for switching costs, the model for mixing costs showed higher accuracy in Spanish than Basque and increased accuracy with development. Accuracy was somewhat higher in blocked than non-switch trials, but this mixing cost was not significant. The mixing effects were also similar for all three age groups (see Figure 7). Similar to the previous model, the accuracy difference between Basque and Spanish decreased with age.

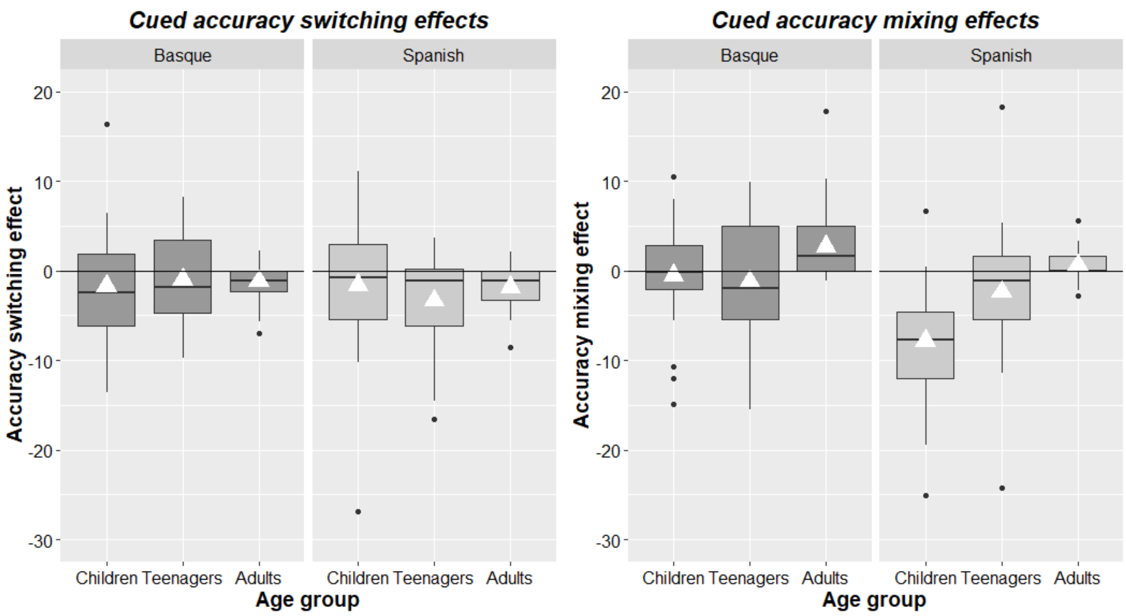

Figure 7. Boxplots showing the cued switching effect (left; accuracy switch trials – non-switch trials) and cued mixing effect (right; accuracy non-switch trials – blocked trials) for children, teenagers, and adults. Negative scores reflect a switching or mixing cost. Given that there were

mistake would inflate the percentage of errors in youths compared to young adults as more accurate trials would need to be removed for youths. As a consequence, trial type refers to the pre-specified trial type that may have been treated differently (i.e., even though a trial was supposed to be a switch trial, a mistake on the previous trial could have turned it into a non-switch trial). This approach was only used for the accuracy analysis, not for RT analyses in which trials preceded by a mistake were excluded.

no effects of predictability, the plots show means across the predictable and unpredictable conditions. The boxplot shows the interquartile range with the black dots representing the outliers falling outside 1.5*interquartile range. The median is indicated by the horizontal black line and the centres of the white triangles show the means.

Reaction times

Reaction time analyses were done after removal of outliers (1.2% of correct trials). The untransformed RT means per age group, trial type, and condition can be found in Table 13. Figure 8 shows the switching costs and Figure 9 the mixing costs per age group, language, and predictability condition. Foreshadowing the results, predictable switching costs decreased during childhood while unpredictable switching costs were not significantly affected by age. Spanish mixing costs were larger for youths than for adults.

Switching effects

First, we examined whether cued switching costs were modulated by age, language, or switching predictability. The full results from this analysis can be found in Table 14. The main effect of trial type confirmed that there was a switching cost, with longer RTs on switch than non-switch trials (see Table 13). The main effect of age showed that overall response times decreased from childhood to adulthood. Responses were also overall slower in Spanish than Basque but were comparable for the predictable and unpredictable condition.

The main interactions of interest were those indicating whether age affected the switching costs. The significant interaction between age, switching cost, and predictability suggested that switching costs decreased with development, but that these effects were modulated by predictability. Age effects were more pronounced in the predictable context (see Figure 8). Furthermore, switching costs were overall higher in the predictable than unpredictable condition. Lastly, switching costs were larger when switching to Basque than to

Spanish. This did not interact with age, suggesting that the larger switching cost to Basque was not affected by age.

The analysis showed that predictability modulated the RTs in two main ways: Switching costs were larger in the predictable than unpredictable condition and predictability modulated the age effects on switching costs. We therefore examined age effects separately for the predictable and unpredictable conditions. The model examining the predictable condition showed a significant interaction between age and trial type (Youth vs Adult: $\beta = -0.034$, $SE = 0.011$, $t = -3.02$), showing that switching costs decreased during childhood and adolescence in the predictable condition (see Figure 8). Predictable switching costs were also somewhat larger for children than teenagers ($\beta = -0.030$, $SE = 0.015$, $t = -1.94$). The unpredictable condition showed no interaction between youth vs adult and switching costs ($\beta = 0.003$, $SE = 0.012$, $t = 0.22$) or between children vs teenagers and switching costs ($\beta = 0.013$, $SE = 0.016$, $t = 0.81$). This confirmed that children, teenagers, and young adults had comparable switching costs in the unpredictable condition.

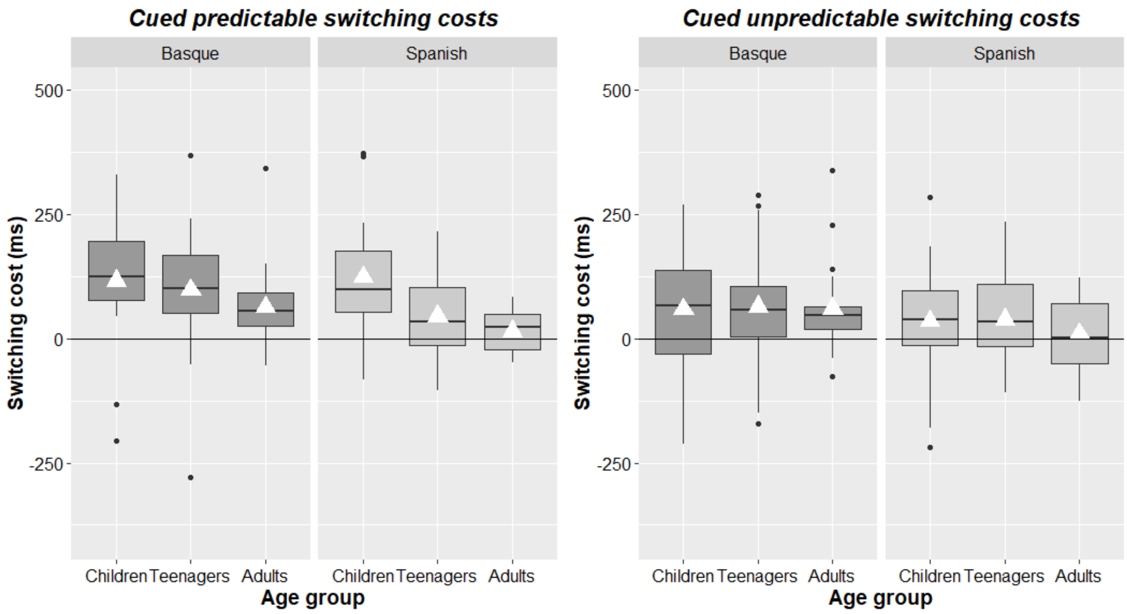

Figure 8. Boxplots showing the switching effects (RT difference between switch and non-switch trials) for children, teenagers, and adults in the cued predicable (left) and unpredictable condition (right). Per condition, the left panel shows the switching effects in Basque and the right panel in Spanish. The boxplot shows the interquartile range with the black dots representing the outliers falling outside 1.5 *interquartile range. The median is indicated by the horizontal black line and the centres of the white triangles show the means.

Table 13. *Mean RTs (and standard deviations) for the children, teenagers, and adults in the blocked condition per language and in the mixed condition per trial type, language, and predictability condition (P = Predictable; UP = Unpredictable).*

	Children		**Teenagers**		**Adults**	
Blocked						
Spanish	1178.0 (143.1)		1123.4 (151.5)		894.1 (97.5)	
Basque	1243.4 (176.9)		1075.2 (128.2)		824.1 (108.6)	
	P	**UP**	**P**	**UP**	**P**	**UP**
Mixed non-switch						
Spanish	1350.0	1373.0	1204.2	1216.0	936.4	926.1
	(194.6)	(172.1)	(161.8)	(235.8)	(126.6)	(141.3)
Basque	1277.7	1312.6	1092.0	1126.9	852.6	835.6
	(170.8)	(178.5)	(163.3)	(200.0)	(114.6)	(114.9)
Mixed switch						
Spanish	1473.0	1407.9	1248.4	1254.2	950.1	935.2
	(183.7)	(199.7)	(171.3)	(204.3)	(128.4)	(130.7)
Basque	1392.7	1371.4	1189.2	1191.4	915.6	895.2
	(196.9)	(205.8)	(175.6)	(211.3)	(147.2)	(154.4)

Table 14. *Results of the cued switching model (Experiment 2). For each predictor, the estimate, standard error, and t value are given with an asterisk indicating a significant effect (in bold). Basque trials were coded as -0.5; Spanish trials as 0.5. Non-switch trials were coded as -0.5; switch trials as 0.5. Predictable trials were coded as -0.5; unpredictable trials as 0.5. For the comparison Youth vs Adult, the coding was (children: -1/3; teenagers: -1/3; adults: 2/3); for the comparison children vs teenagers the coding was (children: -0.5; teenagers: 0.5; adults: 0).*

Predictor	Estimate	SE	T value
Intercept	**6.998**	**0.017**	**420.18***
Trial type (switching)	**0.050**	**0.004**	**11.30***
Youth vs Adult	**-0.331**	**0.032**	**-10.50***
Child vs Teen	**-0.143**	**0.037**	**-3.85***
Language	**0.059**	**0.009**	**6.28***
Predictability	-0.007	0.009	-0.69
Youth vs Adult x Trial type	-0.017	0.009	-1.92
Child vs Teen x Trial type	-0.008	0.011	-0.74
Trial type x Language	**-0.027**	**0.010**	**-2.72***
Youth vs Adult x Language	0.014	0.019	0.71
Child vs Teen x Language	0.027	0.024	1.17
Trial type x Predictability	**-0.028**	**0.008**	**-3.47***
Language x Predictability	-0.006	0.009	-0.69
Youth vs Adult x Predictability	-0.018	0.020	-0.93
Child vs Teen x Predictability	0.015	0.024	0.62
Youth vs Adult x Trial type x Language	-0.030	0.020	-1.49
Child vs Teen x Trial type x Language	-0.022	0.026	-0.83
Youth vs Adult x Trial type x Predictability	**0.032**	**0.016**	**2.03***
Child vs Teen x Trial type x Predictability	0.042	0.021	1.99
Youth vs Adult x Language x Predictability	0.015	0.018	0.82
Child vs Teen x Language x Predictability	-0.002	0.024	-0.09
Trial type x Language x Predictability	0.011	0.015	0.71
Youth vs Adult x Trial type x Language x Predictability	-0.009	0.029	-0.30

Child vs Teen x Trial type x Language x Predictability	0.062	0.040	1.55

Mixing effects

Next, we examined whether age affected the cued language mixing cost by only including non-switch and blocked trials. The full results can be found in Table 15. Similar to the model on switching costs, children responded slowest, followed by teenagers and young adults, and Basque responses were faster overall than Spanish responses. There was also a mixing cost, with responses to non-switch trials being slower than to blocked trials.

The mixing cost decreased from childhood to adolescence and from adolescence to adulthood (see Figure 9). This age effect on the mixing cost was furthermore modulated by language (see Table 13). As can be seen in Figure 9, age effects on the mixing cost were larger for Spanish than Basque, as was confirmed by follow-up analyses per language. On Spanish trials, there was a significant interaction between age and mixing costs (Youth vs Adult: β = -0.067, *SE* = 0.011, *t* = -6.21), such that Spanish mixing costs decreased with development; these Spanish mixing costs were also larger for children than for teenagers (β = -0.073, *SE* = 0.013, *t* = -5.50). In contrast, Basque mixing costs were not affected by age (Youth vs Adult: β = -0.011, *SE* = 0.012, *t* = -0.93; Children vs Teenagers: β = -0.015, *SE* = 0.016, *t* = -0.93).

Overall, mixing costs were larger for Spanish than Basque. In addition, across trial type, adults showed larger RT differences between Basque and Spanish than younger participants. As can be seen in Table 13, adults showed a Basque advantage on both blocked and non-switch trials, whereas this was only the case on non-switch trials for children, thus leading to an overall larger Basque advantage for adults.

No effects of predictability were observed, although age effects on mixing costs were somewhat larger in the unpredictable than predictable condition (see Figure 9).

To sum up, the cued task showed age effects on both switching and mixing costs. These age effects on switching costs were largest in the predictable condition; age effects on mixing costs were largest in Spanish.

Table 15. *Results of the cued mixing model (Experiment 2). For each predictor, the estimate, standard error, and t value are given with an asterisk indicating a significant effect (in bold). Basque trials were coded as -0.5; Spanish trials as 0.5. Trial type was coded as -2/3 (blocked), 1/3 (predictable non-switch), and 1/3 (unpredictable non-switch); for predictability we coded these levels respectively as 0, -0.5, and 0.5. For the comparison Youth vs Adult, the coding was (children: -1/3; teenagers: -1/3; adults: 2/3); for the comparison children vs teenagers the coding was (children: -0.5; teenagers: 0.5; adults: 0).*

Predictor	Estimate	SE	T value
Intercept	**6.961**	**0.017**	**412.91***
Trial type (mixing)	**0.062**	**0.007**	**8.92***
Youth vs Adult	**-0.320**	**0.029**	**-11.01***
Child vs Teen	**-0.130**	**0.034**	**-3.78***
Language	**0.043**	**0.018**	**2.48***
Predictability	0.009	0.010	0.92
Youth vs Adult x Trial type	**-0.042**	**0.014**	**-2.95***
Child vs Teen x Trial type	**-0.044**	**0.018**	**-2.51***
Trial type x Language	**0.045**	**0.012**	**3.87***
Youth vs Adult x Language	**0.067**	**0.028**	**2.36***
Child vs Teen x Language	**0.065**	**0.030**	**2.12***
Language x Predictability	-0.010	0.012	-0.84
Youth vs Adult x Predictability	-0.035	0.021	-1.67
Child vs Teen x Predictability	-0.009	0.026	-0.34
Youth vs Adult x Trial type x Language	**-0.059**	**0.024**	**-2.52***
Child vs Teen x Trial type x Language	**-0.059**	**0.029**	**-2.01***
Youth vs Adult x Language x Predictability	0.015	0.024	0.62
Child vs Teen x Language x Predictability	-0.032	0.032	-0.98

Figure 9. Boxplots showing the mixing effects (RT difference between mixed non-switch and blocked trials) for children, teenagers, and adults in the cued predicable (left) and unpredictable condition (right). Per condition, the left panel shows the mixing effects in Basque and the right panel in Spanish. The boxplot shows the interquartile range with the black dots representing the outliers falling outside 1.5*interquartile range. The median is indicated by the horizontal black line and the centres of the white triangles show the means.

Discussion

Apart from overall slower responses, no negative effects of age were observed on the voluntary switching task. On the cued task, both switching costs (especially in the predictable condition) and mixing costs (especially in Spanish) were larger for youths than adults. Together, these results suggest that children and teenagers experience increased difficulties when they need to control their languages and switch between them in response to external cues, but not when they can freely use their languages.

In terms of differences between children and teenagers, all analyses showed slower overall responses for children than teenagers. The voluntary task showed no differences between children and teenagers beyond overall RTs. That is, the switching costs and larger Basque mixing benefits were similar for children and teenagers. Some differences between children and teenagers were observed in the cued task, which showed a more gradual decrease in mixing and switching costs from children to teenagers and to young adults. In terms of switching costs, predictable costs were somewhat larger for children than for teenagers (with no difference in unpredictable costs). Furthermore, the Spanish mixing costs were not only larger for youths than for adults but also for children than for teenagers (with no difference in Basque mixing costs). These patterns suggest an ongoing development in cued (but not voluntary) language control, with twelve year olds showing improved control compared to 7-9 year olds.

No negative age effects were observed in the voluntary task. In terms of accuracy and RTs, mixing benefits were found to be larger for youths than adults, but only in Basque. This is possibly related to the children's and teenagers' ongoing development of Basque, their non-dominant language. For instance, while adults showed high accuracy in both Spanish and Basque in the blocked condition, children and teenagers performed much worse on the Basque blocked condition, suggesting that they encountered more difficulties naming the items in Basque. The larger mixing benefit in Basque likely reflects the increased flexibility that the voluntary mixed condition offers, namely the freedom to avoid naming certain words that you are not familiar with and the option to choose words that are faster to produce. While children and teenagers may have benefitted from this in their non-dominant language in particular, these advantages were comparable in both languages for adults who had already acquired both languages to a high level. However, this is not to say that children and teenagers avoided using their non-dominant language. All age groups switched frequently between their languages: Children used both languages approximately equally often, while teenagers and adults actually used Basque more frequently.

Apart from an overall age effect, accuracy in the cued task did not reveal further age effects. However, considering that the switching and mixing effects on accuracy were very small in the first place, the accuracy measure may not have been sensitive enough to detect further age effects. In terms of response times, mixing costs decreased between childhood and adulthood on the cued task, in particular for Spanish. Spanish mixing costs were found to be larger overall, which could reflect increased proactive inhibition of Spanish (the dominant language) and/or increased activation of Basque (the non-dominant language) in the mixed condition. Asymmetrical language effects, with larger costs for the dominant language, have been observed in previous language switching studies, both with respect to the switching costs (e.g., Meuter & Allport, 1999) as well as the mixing costs (e.g., de Bruin, Roelofs, Dijkstra, & FitzPatrick, 2014). This asymmetry in mixing costs was present for adults and was even more pronounced for children and teenagers, who slowed down relatively more in Spanish than Basque in the mixed condition.

One possible explanation is that the younger participants proactively needed to suppress Spanish more strongly in the mixed condition in order to name half of the pictures in Basque. Previous work (Davidson et al., 2006) has suggested that it is easier for children to continuously suppress the dominant response than to be in a situation in which the dominant response has to be suppressed only part of the time. In line with this reasoning, the youths in the current study may have suppressed Spanish proactively from the start of the task rather than reactively on individual trials. An alternative or additional explanation is that the youths proactively activated Basque more strongly in the mixed condition than the adults to be able to use their non-dominant language more easily in the cued dual-language condition. Cross-language intrusions in the accuracy data also suggest that Basque was more active in the cued task (either through suppression of Spanish or enhanced activation of Basque). For children, of the mixed items that had to be named in Basque, on average 5.1 items (*SD* = 6.5) were instead incorrectly named in Spanish. Of the items that were supposed to be named in Spanish, 9.0 (*SD*

= 9.8) were instead incorrectly named in Basque. Similar patterns were found for teenagers (Basque items incorrectly named in Spanish: M = 3.4 items; Spanish items incorrectly named in Basque: M = 5.6 items). Children and teenagers thus showed a relatively high intrusion of Basque responses when they were supposed to use Spanish (their dominant language).

The larger mixing costs for Spanish than Basque in this second experiment is not entirely limited to younger participants, as the older adults in Experiment 1 also showed numerically larger mixing costs in Spanish than Basque. Overall increased activation of Basque (in all age groups) could also explain the larger switching costs to Basque, if interpreted in light of the Inhibitory Control model (Green, 1998). Proactively increasing Basque activation would come with the cost of needing more reactive inhibition of Basque during Spanish trials, which could then have led to relatively higher switching costs to Basque.

With respect to age effects on switching costs, these costs decreased from childhood to adulthood on the cued task, but only in the predictable condition. This is consistent with the findings from Experiment 1, as the age effects of older adults were also somewhat larger in the predictable than unpredictable condition (although predictability was not a significant modulator of age effects in Experiment 1). Again, switching costs were found to be largest for the predictable condition. We will discuss the role of predictability in more detail in the General Discussion.

General Discussion

This study assessed language switching across the lifespan by comparing older and younger adults (Experiment 1) and children, teenagers, and young adults (Experiment 2) on voluntary and cued language switching tasks. Within the cued task, we included a predictable and unpredictable switching condition. We focused on age effects on mixing costs and switching costs. The voluntary task showed overall mixing benefits that were not negatively affected by age. Voluntary switching costs were observed that were larger for the older adults than the younger adults, but not affected by age in the experiment with children and teenagers. While detrimental age effects were relatively limited in the voluntary task, the cued task showed that switching and mixing costs decreased during development and increased again in later adulthood. For children and teenagers, the age effects on mixing costs were more pronounced in the dominant language while the effects on switching costs were only observed in the predictable switching condition. Overall, the results demonstrate different patterns for voluntary versus cued language switching: Age effects across the lifespan are limited to larger switching costs with cognitive ageing during voluntary language use but are more widespread when the two languages need to be used in response to external task demands. This suggests that it is not the proactive control related to using languages in a dual-language versus single-language context as such that is more effortful for youths and older adults. Rather, it is the need to use and maintain multiple languages in response to external cues in a dual-language context.

Bilingual language control across the lifespan

Cued language switching

The overall pattern of age effects was very similar for children and older adults, especially on the cued task. Both mixing costs and switching costs decreased during childhood and increased again during later adulthood. This suggests that both children and older adults experience greater difficulties with the proactive control needed to maintain and use two languages (as indicated by increased mixing costs). The absence of age effects on the mixing effect in the voluntary task suggests that age effects on mixing costs in the cued tasks are related to the need to use two languages in response to external cues and not purely to the use of two languages as compared to using one. While older adults showed larger voluntary and cued switching costs, suggesting increased difficulties with reactive control, proactive control as measured through mixing effects appeared to be only affected by age when two languages were used in response to external demands. These age effects on mixing costs are consistent with the main findings from the task-switching literature (Wasylyshyn et al., 2011). They also align with previous language-switching studies comparing older to younger adults (e.g., Hernandez & Kohnert, 1999; Weissberger et al., 2012) and those comparing children and adults (Kohnert et al., 1999).

Cued switching costs too were affected by age. This is in contrast to results from meta-analyses on task switching (Wasylyshyn et al., 2011), but in line with some previous work on language switching (e.g., Weissberger et al., 2012). Both children and older adults needed more time to respond to switch trials than young adults did. Crucially, the current study shows that, contrary to non-linguistic task-switching (cf. meta-analysis by Wasylyshyn et al., 2011), age can affect the ability to switch between two languages. This suggests that age effects on language switching are grounded in more reactive components of executive control, potentially related to inhibitory control. Moreover, very similar patterns on the cued task were observed for children and older adults despite differences in their language profiles, suggesting that bilingual language switching may operate similarly during cognitive development and decline.

The greater switching costs in older adults and children may be the result of multiple underlying cognitive processes, including activating the new target language, overcoming previously applied inhibition on the new target language, applying inhibition on the previously used language, and preparing for the switch. Future studies will need to clarify which of these components are affected by age, for example by manipulating the interval between cue and target and between response and stimulus. Studies with bimodal bilinguals have suggested that switching away from a language might be more effortful than activating the new target language (Emmorey, Chuchu, Petrich, & Gollan, 2019). Furthermore, work with unimodal older bilinguals has argued that it is not the application of inhibition that declines with age but rather overcoming the previously applied inhibition (Ivanova, Murillo, Montoya, & Gollan, 2016). In line with this, our age effects on cued switching may reflect difficulties when it comes to releasing inhibition of the new target language.

Voluntary language switching

Both children and older adults showed voluntary mixing benefits that were not negatively affected by age. This, in combination with the age effects on cued mixing costs, suggests that A) voluntary language mixing may be less effortful than having to use one specific language and that B) voluntary language use is less effortful than cued language use. In contrast to these mixing benefits, voluntary switching costs were observed; these costs were larger for older than younger adults. This suggests that language switching, even when done voluntarily, comes with a temporal delay and that older adults may experience greater difficulties than younger adults. Various previous studies on voluntary language switching have found voluntary switching costs (e.g., de Bruin et al., 2018; Gollan & Ferreira, 2009; but see Blanco-Elorrieta & Pylkkänen, 2017, Kleinman & Gollan, 2016). It has previously been suggested that these switch costs might reflect reactive inhibition (e.g., de Bruin et al., 2018). That is, even though voluntarily using two languages was less costly than being in a single-language condition, switching away from a previously used language and to a new target language might still require reactive control. This interpretation is further supported by the larger voluntary switching costs in older adults, who might have deficits releasing inhibition when switching to another language in both cued and voluntary paradigms.

Interestingly, however, larger voluntary switching costs were found for older adults, but not for children and teenagers. It is possible that the ability to overcome previously applied inhibition was already developed to some extent in our youngest bilinguals, at least sufficiently to switch voluntarily. In addition, the older adults switched less often in the voluntary task than younger adults, whereas the three age groups in Experiment 2 were comparable in their switching frequency. Gollan and Ferreira (2009) noted that older bilinguals who switched less often in their task also showed higher switching costs. In the current study, there was a similar correlation between switching frequency in the task and switching costs ($r(45) = -0.50$, $p <$ .001) reflecting the fact that both younger ($r(23) = -0.47$, $p = 0.019$) and older ($r(20) = -0.46$, $p = 0.032$) adults with a higher switching frequency had smaller switching costs. Similar findings were found for Experiment 2 (children: $r(18) = -0.46$, $p = 0.043$; teenagers: $r(25) = -0.20$, $p = 0.32$; young adults: $r(23) = -0.41$, $p = 0.040$). It thus remains unclear whether the older bilinguals switched less often in the voluntary task because they experienced greater difficulty or whether they had larger costs because they switched less often for another reason[4].

Switching sequence predictability

Switching costs were larger in the predictable than unpredictable cued switching condition and age effects were largest when switches were predictable. This is contrary to some previous studies (e.g., Kray et al., 2002), but in accord with Hernandez and Kohnert (1999) who also observed larger age effects in predictable switching sequences. Kray (2006) furthermore observed no age effects on switching costs in a predictable task but *smaller* switching costs for older adults in an unpredictable switching task. Predictability was confounded with various other variables in previous studies (e.g., the presence of cues or percentage of switch trials), whereas the current study made the predictable and unpredictable tasks as similar as possible by including the same number of switch trials, always presenting cues, including the same number of practice trials, using the same stimuli, and comparing the two within the same participants.

Why did these age effects predominantly emerge in the predictable condition? Studies comparing younger to older adults have found that older adults are less capable of learning predictable sequences (e.g., Dennis, Howard, & Howard, 2003; Howard & Howard, 1997).

[4] The older adults also reported lower levels of within-sentence switching in daily life. While this could explain the lower percentage of switching in the task, there was no significant correlation between the two ($r(47) = -0.196$, $p = 0.196$).

Participants were told about the order of the language cues in the predictable condition. However, it is possible that older adults and youths were less able to use this information to their benefit. This may be related to the 'hyper-binding' phenomenon showing that older adults bind too much information together and have difficulties focusing on task-relevant information while ignoring distracting information (e.g., Campbell, Hasher, & Thomas, 2010; Campbell, Trelle, & Hasher, 2014). As a consequence of binding task-irrelevant information, older adults may experience greater difficulties extracting regularities. It is possible that children also experience hyper-binding of task-irrelevant information. For instance, studies testing 7 and 8 year old children (i.e., similar to the age group tested in the current study) suggest that they are more distracted by novel information than young adults (e.g., Wetzel, Widmann, & Schröger, 2009). While young adults may have benefited from the predictable switching pattern, older adults and children may not have been able to sufficiently suppress distracting information, to fully encode the predictable information, and to prepare for the upcoming predictable switches.

However, this interpretation does not explain why all age groups including the young adults showed smaller switching costs in the unpredictable than predictable condition. These findings also cannot be explained by differences in trial sequences between the predictable and unpredictable condition (see Supplementary Materials). These findings seem counter-intuitive and are not in line with some previous studies observing that switching costs increase with task uncertainty (e.g., Cooper, Garrett, Rennie, & Karayanidis, 2015). At the same time, other studies (e.g., Kray, 2006) have also observed *larger* switching costs in predictable than unpredictable conditions. Tornay and Milán (2001) also found that unpredictable switching costs can be smaller than predictable costs, at least when participants are given sufficient time to prepare for the switch. They suggested that the smaller costs found under unpredictable, more difficult conditions might arise if these conditions engaged attentional or control mechanisms more strongly, leading to more controlled task processing. While predictable sequences may recruit more automatic reconfiguration processes to switch between languages, unpredictable sequences may have led to greater recruitment of attention. As soon as the response was given, the bilinguals may have suppressed the language currently used or activated the current non-target language in order to prepare for a possible switch, leading to relatively smaller costs for unpredictable sequences. Thus, the larger predictable switch costs might not indicate that predictable switching is more difficult. Rather, it might occur if bilinguals recruit more resources to prepare and execute unpredictable switches.

Some EEG and fMRI studies have suggested that older adults recruit more proactive and reactive control than younger adults while they are switching between tasks (e.g., Jimura & Braver, 2009; Karayanidis, Whitson, Heathcote, & Michie, 2011; Whitson et al., 2014). It is possible that all age groups applied additional resources during the unpredictable task, but that this was done more strongly by the children and older adults. In addition, older adults and children may have used their resources in a different manner than the young adults. The young adults (see Table 5 for Experiment 1 and Table 13 for Experiment 2) were faster overall in the unpredictable condition, but especially so on the switch trials. This suggests that the young adults may have especially engaged more attentional resources to respond to switch trials. The children and older adults also show faster RTs to switch trials in the unpredictable than predictable condition. However, this is in combination with *slower* responses to unpredictable non-switch trials than predictable non-switch trials. While young adults may have the flexibility to update their decision criterion on a trial by trial basis, thus recruiting additional resources more strongly in response to switch trials, children and older adults may not have this ability (cf. Davidson et al., 2006). As a consequence, they may have treated all trials as switch trials and may have continuously prepared for a switch by suppressing the just used language and activating the other language (cf. Karayanidis et al., 2011). This would have led to faster responses on switch trials, but relatively slower responses on non-switch trials, producing the relatively small age effects in the unpredictable condition.

Cognitive development and ageing

In this study, cognitive development and ageing were studied in two separate experiments. While our older adults (and matching group of younger adults) were relatively balanced bilinguals with a high proficiency and use of both languages, the youths (and matching young adults) were more unbalanced bilinguals with a higher proficiency in and use of Spanish. Considering this difference in language profile, we are not able to directly compare cognitive development and ageing in one analysis. The comparison between youths and younger adults showed some effects of language that might be related to the language background of this sample. Youths showed a larger mixing benefit in Basque (their L2) than young adults while older and younger adults showed comparable mixing benefits in both languages. In the cued task, youths also showed a larger mixing cost in Spanish (their L1) than young adults. These language-related effects are likely to be related to the development of (control over) the stronger versus weaker language during ongoing language acquisition in childhood. The other striking difference between youths and older adults concerned the larger voluntary switching cost in the group of older adults. As discussed above, this might be related to the older adults' language-switching patterns in the task and daily life. Future studies with more comparable language profiles across youths and older adults will need to establish whether these differences are related to different mechanisms underlying development versus ageing or differences in the participants' language profile.

The overall picture, however, showed similar patterns during development and ageing. In the cued task, both switching and mixing effects decreased during childhood and increased again during adulthood. In the voluntary task, a mixing benefit was observed for all age groups. Overall, the study suggests that the proactive control associated with using two languages in a dual-language context in response to cues develops during childhood and declines during later adulthood.

Conclusion

In conclusion, this study shows that bilingual language control and language switching change across the lifespan, but that the effects of age depend on the way bilinguals have to use their languages. In daily life, bilinguals experience different interactional contexts that may require different forms and levels of language control. In some circumstances, strict language control is needed to avoid interference from the non-target language and to ensure that switching takes place at the right moment. In these contexts, children and older adults may have greater difficulty maintaining their two languages and switching between them in response to external cues. However, some bilinguals, especially those living in a bilingual society, do not always require strict language control and are sometimes able to freely switch between their languages. Our study shows that when language use is voluntary, language mixing does not need to be effortful in any of the age groups. Thus, under free switching circumstances, using two languages is not more demanding than using one language for children and older adults.

Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement number 743691. Support was provided by a Ministerio de Ciencia e Innovación Grant #PSI2017-82563, #PGC2018-097145-B-I00, RED2018-102615-T, Grant PIBA18-29 from the Basque Government, Grant H2019/HUM-5705 from the Comunidad de Madrid, and by Ayuda Centro de Excelencia Severo Ochoa SEV-2015-0490. The authors would like to thank the BCBL research assistants for their help with data collection.

References

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*(2), 71-82.

Baayen, R. H. (2008). Analyzing Linguistic Data: A practical introduction to statistics using R. Cambridge University Press, Cambridge, U. K.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255-278.

Basnight-Brown, D. M., & Altarriba, J. (2007). Code-switching and code-mixing in bilinguals: Cognitive, developmental, and empirical approaches. In A. Ardila & E. Ramos (Eds.), *Speech and Language Disorders in Bilinguals* (pp. 69-89). New York: Nova Science Publishers.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R package version*, *1*(7), 1-23.

Blanco-Elorrieta, E., & Pylkkänen, L. (2017). Bilingual Language Switching in the Laboratory versus in the Wild: The Spatiotemporal Dynamics of Adaptive Language Control. *Journal of Neuroscience*, *37*(37), 9022-9036.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology*, *7*, 1116.

Bugg, J. M., Zook, N. A., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2006). Age differences in fluid intelligence: contributions of general slowing and frontal decline. *Brain and Cognition*, *62*(1), 9-16.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron*, *33*(2), 301-311.

Butler, K. M., & Weywadt, C. (2013). Age differences in voluntary task switching. *Psychology and Aging*, *28*(4), 1024-1031.

Calabria, M., Branzi, F. M., Marne, P., Hernández, M., & Costa, A. (2015). Age-related effects over bilingual language control and executive control. *Bilingualism: Language and Cognition*, *18*(1), 65-78.

Campbell, K. L., Hasher, L., & Thomas, R. C. (2010). Hyper-binding: A unique age effect. *Psychological Science*, *21*(3), 399-405.

Campbell, K. L., Trelle, A., & Hasher, L. (2014). Hyper-binding across time: Age differences in the effect of temporal proximity on paired-associate learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 293-299.

Castel, A. D., Balota, D. A., Hutchison, K. A., Logan, J. M., & Yap, M. J. (2007). Spatial attention and response control in healthy younger and older adults and individuals with Alzheimer's disease: Evidence for disproportionate selection impairments in the simon task. *Neuropsychology*, *21*(2), 170-182.

Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. (2001). Changes in executive control across the life span: examination of task-switching performance. *Developmental Psychology*, *37*(5), 715-730.

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences*, *106*(14), 5529-5533.

Cooper, P. S., Garrett, P. M., Rennie, J. L., & Karayanidis, F. (2015). Task uncertainty can account for mixing and switch costs in task-switching. *PloS one*, *10*(6), e0131556.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491-511.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037-2078.

de Bruin, A., Carreiras, M., & Duñabeitia, J. A. (2017). The BEST dataset of language proficiency. *Frontiers in Psychology*, *8*.

de Bruin, A., & Della Sala, S. (2018). Effects of age on inhibitory control are affected by task-specific features. *The Quarterly Journal of Experimental Psychology, 71(5), 1219-1233.*

de Bruin, A., Roelofs, A., Dijkstra, T., & FitzPatrick, I. (2014). Domain-general inhibition areas of the brain are involved in language switching: FMRI evidence from trilingual speakers. *NeuroImage*, *90*, 348-359.

de Bruin, A., Samuel, A.G., & Duñabeitia, J.A. (2018). Voluntary language switching: When and why do bilinguals switch between their languages? *Journal of Memory and Language,* 103, 28-43.

Declerck, M., Koch, I., & Philipp, A. M. (2012). Digits vs. pictures: The influence of stimulus type on language switching. *Bilingualism: Language and Cognition*, *15*(4), 896-904.

Declerck, M., & Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review*, *22*(6), 1630-1645.

Dennis, N. A., Howard Jr, J. H., & Howard, D. V. (2003). Age deficits in learning sequences of spoken words. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *58*(4), 224-227.

Diamond, A. (2006). The early development of executive functions. *Lifespan cognition: Mechanisms of Change*, *210*, 70-95.

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2017). MultiPic: A standardized set of 750 drawings with norms for six European languages. *The Quarterly Journal of Experimental Psychology*, 1-24.

Emmorey, K., Li, C., Petrich, J., & Gollan, T. H. (in press). Turning languages on and off:

Switching into and out of code-blends reveals the nature of bilingual language control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

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*(3), 189-198.

Frank, A. F. (2011) R-hacks/mer-utils.R. https://github.com/aufrank/R-hacks/blob/master/mer-utils.R. Accessed October 2017.

Fricke, M., & Kootstra, G. J. (2016). Primed codeswitching in spontaneous bilingual dialogue. *Journal of Memory and Language*, *91*, 181-201.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel hierarchical models* (Vol. 1). New York, NY, USA: Cambridge University Press.

Gollan, T. H., & Ferreira, V. S. (2009). 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*, *35*(3), 640-665.

Gollan, T. H., Kleinman, D., & Wierenga, C. E. (2014). 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*, *143*(6), 2167.

Grange, J.A. (2015). trimr: An implementation of common response time trimming methods. R package version 1.0.1. https://cran.r-project.org/web/packages/trimr/index.html

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67-81.

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

Gross, M., & Kaushanskaya, M. (2015). Voluntary language switching in English–Spanish bilingual children. *Journal of Cognitive Psychology*, *27*(8), 992-1013.

Gross, M., & Kaushanskaya, M. (2018). Contributions of nonlinguistic task-shifting to language control in bilingual children. *Bilingualism: Language and Cognition*, *21*(1), 181-194.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), The psychology of learning and motivation, (Vol. 22, pp. 193-225). San Diego, CA: Academic Press.

Hernandez, A. E., & Kohnert, K. J. (1999). Aging and language switching in bilinguals. *Aging, Neuropsychology, and Cognition*, *6*(2), 69-83.

Hernandez, A. E., & Kohnert, K. J. (2015). Investigations into the locus of language-switching costs in older adult bilinguals. *Bilingualism: Language and Cognition*, *18*(1), 51-64.

Horoufchin, H., Philipp, A. M., & Koch, I. (2011). The dissipating task-repetition benefit in cued task switching: Task-set decay or temporal distinctiveness?. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 455-472

Howard Jr, J. H., & Howard, D. V. (1997). Age differences in implicit learning of higher order dependencies in serial patterns. *Psychology and Aging*, *12*(4), 634-656.

Jevtović, M., Duñabeitia, J. A., & de Bruin, A. (in press). How do bilinguals switch between languages in different interactional contexts? A comparison between voluntary and mandatory language switching. *Bilingualism: Language and Cognition*.

Jia, G., Kohnert, K., Collado, J., & Aquino-Garcia, F. (2006). Action naming in Spanish and English by sequential bilingual children and adolescents. *Journal of Speech, Language, and Hearing Research*, *49*(3), 588-602.

Jimura, K., & Braver, T. S. (2009). Age-related shifts in brain activity dynamics during task switching. *Cerebral Cortex*, *20*(6), 1420-1431.

Karayanidis, F., Whitson, L. R., Heathcote, A., & Michie, P. T. (2011). Variability in proactive and reactive cognitive control processes across the adult lifespan. *Frontiers in Psychology*, *2*, 318.

Kleinman, D., & Gollan, T. H. (2016). Speaking two languages for the price of one: Bypassing language control mechanisms via accessibility-driven switches. *Psychological Science*, *27*(5), 700-714.

Kohnert, K. J., Bates, E., & Hernandez, A. E. (1999). Balancing bilinguals: Lexical-semantic production and cognitive processing in children learning Spanish and English. *Journal of Speech, Language, and Hearing Research*, *42*(6), 1400-1413.

Kray, J. (2006). Task-set switching under cue-based versus memory-based switching conditions in younger and older adults. *Brain Research*, *1105*(1), 83-92.

Kray, J., Eber, J., & Lindenberger, U. (2004). Age differences in executive functioning across the lifespan: The role of verbalization in task preparation. *Acta Psychologica*, *115*(2-3), 143-165.

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*, *15*(1), 126-147

Kray, J., Li, K. Z., & Lindenberger, U. (2002). Age-related changes in task-switching components: The role of task uncertainty. *Brain and Cognition*, *49*(3), 363-381.

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure?. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(3), 575-599.

Mayr, U. (2001). Age differences in the selection of mental sets: the role of inhibition, stimulus ambiguity, and response-set overlap. *Psychology and Aging*, *16*(1), 96-109

Mayr, U., & Kliegl, R. (2003). Differential effects of cue changes and task changes on task-set selection costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(3), 362-372.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*(1), 25-40.

Mosca, M., & Clahsen, H. (2016). Examining language switching in bilinguals: The role of preparation time. *Bilingualism: Language and Cognition*, *19*(2), 415-424.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8-13.

Proctor, R. W., Pick, D. F., Vu, K. P. L., & Anderson, R. E. (2005). The enhanced Simon effect for older adults is reduced when the irrelevant location information is conveyed by an accessory stimulus. *Acta Psychologica*, *119*(1), 21-40.

Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859-862.

Reimers, S., & Maylor, E. A. (2005). Task switching across the life span: effects of age on general and specific switch costs. *Developmental Psychology*, *41*(4), 661-671.

Rey-Mermet, A., & Gade, M. (2018). Inhibition in aging: What is preserved? What declines? A meta-analysis. *Psychonomic Bulletin & Review*, *25*(5), 1695-1716.

Roux, F., Armstrong, B. C., & Carreiras, M. (2017). Chronset: An automated tool for detecting speech onset. *Behavior Research Methods*, *49*(5), 1864-1881.

Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1477-1491.

Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, *2*(3), 179-183.

Scarmeas, N., Zarahn, E., Anderson, K. E., Habeck, C. G., Hilton, J., Flynn, J., ... & Moeller, J. R. (2003). Association of life activities with cerebral blood flow in Alzheimer disease: implications for the cognitive reserve hypothesis. *Archives of Neurology*, *60*(3), 359-365.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*(3), 281-284.

Tornay, F. J., & Milán, E. G. (2001). A more complete task-set reconfiguration in random than in predictable task switch. *The Quarterly Journal of Experimental Psychology*, *54*(3), 785-803.

Van Asselen, M., & Ridderinkhof, K. R. (2000). Shift costs of predictable and unexpected set shifting in young and older adults. *Psychologica Belgica*, *40*(4), 259-273.

Verhaeghen, P. (2011). Aging and executive control: Reports of a demise greatly exaggerated. *Current Directions in Psychological Science*, *20*(3), 174-180.

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: a meta-analysis. *Psychology and Aging*, *18*(3), 443-460

Wasylyshyn, C., Verhaeghen, P., & Sliwinski, M. J. (2011). Aging and task switching: a meta-analysis. *Psychology and Aging*, *26*(1), 15-20.

Weissberger, G. H., Wierenga, C. E., Bondi, M. W., & Gollan, T. H. (2012). Partially overlapping mechanisms of language and task control in young and older bilinguals. *Psychology and Aging*, *27*(4), 959-974.

Wetzel, N., Widmann, A., & Schröger, E. (2009). The cognitive control of distraction by novelty in children aged 7–8 and adults. *Psychophysiology*, *46*(3), 607-616.

Whitson, L. R., Karayanidis, F., Fulham, R., Provost, A., Michie, P. T., Heathcote, A., & Hsieh, S. (2014). Reactive control processes contributing to residual switch cost and mixing cost across the adult lifespan. *Frontiers in Psychology*, *5*, 383.

Appendix A. Switching frequency analyses

Table A1. *Results of voluntary switching frequency analysis (Experiment 1). For each predictor, the estimate, standard error, and z value are given with an asterisk indicating a significant effect (in bold). The DV trial type was coded as 0 (non-switch trials) and 1 (switch trials). Basque trials were coded as -0.5; Spanish trials as 0.5. Younger adults were coded as -0.5; older adults as 0.5*

Predictor	Estimate	SE	Z value
Intercept	**-0.552**	**0.071**	**-7.83***
Language	**0.455**	**0.094**	**4.82***
Age	**-0.282**	**0.138**	**-2.04***
Age x language	-0.089	0.196	-0.45

Table A2. *Results of voluntary switching frequency analysis (Experiment 2). For each predictor, the estimate, standard error, and z value are given with an asterisk indicating a significant effect (in bold). The DV trial type was coded as 0 (non-switch trials) and 1 (switch trials). Basque trials were coded as -0.5; Spanish trials as 0.5. For the comparison Youth vs Adult, the coding was (children: -1/3; teenagers: -1/3; adults: 2/3); for the comparison children vs teenagers the coding was (children: -0.5; teenagers: 0.5; adults: 0).*

Predictor	Estimate	SE	Z value
Intercept	**-0.169**	**0.037**	**-4.52***
Language	**0.349**	**0.111**	**3.14***
Youth vs Adult	-0.047	0.076	-0.62
Child vs Teen	0.112	0.090	1.24
Youth vs Adult x language	0.353	0.232	1.52
Child vs Teen x language	**0.831**	**0.276**	**3.01***

Appendix B. Accuracy analyses

Table B1. *Results of voluntary accuracy analysis (Experiment 2). For each predictor, the estimate, standard error, and z value are given with an asterisk indicating a significant effect (in bold). The DV accuracy was coded as 0 (incorrect) and 1 (correct). Blocked trials were coded as -0.5; voluntary trials were coded as 0.5. For the comparison Youth vs Adult, the coding was (children: -1/3; teenagers: -1/3; adults: 2/3); for the comparison children vs teenagers the coding was (children: -0.5; teenagers: 0.5; adults: 0).*

Comparing voluntary versus blocked trials across languages

Predictor	Estimate	SE	Z value
Intercept	**-0.113**	**0.008**	**-13.90***
Trial type	**0.121**	**0.014**	**8.97***
Youth vs Adult	**0.145**	**0.016**	**9.15***
Child vs Teen	**0.085**	**0.020**	**4.32***
Youth vs Adult x Trial type	**-0.154**	**0.027**	**-5.63***
Child vs Teen x Trial type	-0.050	0.035	-1.44

Comparing voluntary versus blocked trials in Spanish

Predictor	Estimate	SE	Z value
Intercept	**-0.067**	**0.009**	**-7.38***
Trial type	0.030	0.017	1.73
Youth vs Adult	**0.075**	**0.019**	**4.05***
Child vs Teen	**0.046**	**0.023**	**2.03***
Youth vs Adult x Trial type	-0.016	0.036	-0.45
Child vs Teen x Trial type	-0.028	0.043	-0.65

Comparing voluntary versus blocked trials in Basque

Predictor	Estimate	SE	Z value
Intercept	**-0.167**	**0.010**	**-16.15***
Trial type	**0.229**	**0.019**	**12.01***
Youth vs Adult	**0.226**	**0.021**	**11.04***
Child vs Teen	**0.137**	**0.027**	**5.14***
Youth vs Adult x Trial type	**-0.317**	**0.038**	**-8.43***
Child vs Teen x Trial type	**-0.155**	**0.050**	**-3.10***

Table B2. *Results of cued accuracy analysis (Experiment 2). For each predictor, the estimate, standard error, and z value are given with an asterisk indicating a significant effect (in bold). The DV accuracy was coded as 0 (incorrect) and 1 (correct). For the comparison Youth vs Adult, the coding was (children: -1/3; teenagers: -1/3; adults: 2/3); for the comparison children vs teenagers the coding was (children: -0.5; teenagers: 0.5; adults: 0).*

Switching effect

Predictor	Estimate	SE	Z value
Intercept	**-0.206**	**0.019**	**-10.75***
Trial type (switching)	-0.023	0.014	-1.65
Youth vs Adult	**0.263**	**0.024**	**10.81***
Child vs Teen	**0.152**	**0.029**	**5.15***
Language	**0.151**	**0.018**	**8.55***
Predictability	-0.002	0.014	-0.18
Youth vs Adult x Trial type	0.007	0.028	0.25
Child vs Teen x Trial type	-0.002	0.036	-0.04
Trial type x Language	-0.012	0.028	-0.44
Youth vs Adult x Language	**-0.220**	**0.036**	**-6.10***
Child vs Teen x Language	-0.066	0.045	-1.48
Trial type x Predictability	0.002	0.028	0.06
Language x Predictability	0.014	0.028	0.49
Youth vs Adult x Predictability	0.014	0.028	0.51
Child vs Teen x Predictability	-0.015	0.036	-0.43
Youth vs Adult x Trial type x Language	-0.0003	0.056	-0.006

Child vs Teen x Trial type x Language	-0.028	0.072	-0.39
Youth vs Adult x Trial type x Predictability	0.004	0.056	0.07
Child vs Teen x Trial type x Predictability	-0.028	0.072	-0.40
Youth vs Adult x Language x Predictability	-0.025	0.056	-0.44
Child vs Teen x Language x Predictability	-0.050	0.072	-0.70
Trial type x Language x Predictability	0.027	0.056	0.48
Youth vs Adult x Trial type x Language x Predictability	-0.030	0.112	-0.27
Child vs Teen x Trial type x Language x Predictability	0.016	0.144	0.11

Mixing effect

Predictor	Estimate	SE	Z value
Intercept	**-0.188**	**0.020**	**-9.25***
Trial type (mixing)	-0.017	0.015	-1.12
Youth vs Adult	**0.239**	**0.022**	**10.91***
Child vs Teen	**0.142**	**0.027**	**5.31***
Language	**0.168**	**0.015**	**10.98**
Predictability	-0.003	0.020	-0.17
Youth vs Adult x Trial type	0.055	0.031	1.76
Child vs Teen x Trial type	0.031	0.039	0.78
Trial type x Language	-0.033	0.031	-1.09
Youth vs Adult x Language	**-0.226**	**0.031**	**-7.36***
Child vs Teen x Language	-0.077	0.039	-1.97
Language x Predictability	0.002	0.039	0.05
Youth vs Adult x Predictability	0.013	0.039	0.32
Child vs Teen x Predictability	-0.001	0.050	-0.02
Youth vs Adult x Trial type x Language	0.028	0.062	0.45
Child vs Teen x Trial type x Language	0.071	0.079	0.90
Youth vs Adult x Language x Predictability	-0.011	0.079	-0.14
Child vs Teen x Language x Predictability	-0.059	0.101	-0.58