Bilingualism: Language and Cognition

cambridge.org/bil

Bilingualism and ageing independently impact on language processing: evidence from comprehension and production

Eunice G. Fernandes[1], Katrien Segaert[2], Foyzul Rahman[3], Allison Wetterlin[4] and Linda Wheeldon[5]

[1]University of Agder, Kristiansand, Norway; [2]University of Birmingham, Birmingham, UK; [3]University of Birmingham, Birmingham, UK; [4]University of Agder, Kristiansand, Norway and [5]University of Agder, Kristiansand, Norway

Research Article

Cite this article: Fernandes, E.G., Segaert, K., Rahman, F., Wetterlin, A., & Wheeldon, L. (2024). Bilingualism and ageing independently impact on language processing: evidence from comprehension and production. *Bilingualism: Language and Cognition*, 1–15. https://doi.org/10.1017/S1366728924000245

Received: 12 May 2023
Revised: 29 February 2024
Accepted: 29 February 2024

Keywords:
Bilingualism; Ageing; Language comprehension; Language production

Corresponding author:
Eunice G. Fernandes;
Email: e.fernandes1977@gmail.com

Abstract

To examine the combined effects of ageing and bilingualism in language processing, we tested young and older mono- and bilingual speakers in L1 comprehension and production. In Experiment 1, bilinguals were slower to detect words than monolinguals in sentences with a low-constraint context, but not when a high-constraint context was provided. Older adults tended to outperform younger adults in high-constraint sentences. In Experiment 2, older speakers were slower than younger speakers to produce small-scope prepositional phrases (e.g., 'the cone **above** the grape), suggesting more extensive planning. Bilingual disadvantages were observed in larger-scope complex phrases (e.g., 'the cone **and** the pink grape'). Individual differences in language proficiency did not modulate the effects. The results support bilingual disadvantages in syntactic processing and age-preserved syntax, alongside semantic processing unaffected by either bilingualism or age. We found no interactions between age and bilingualism, suggesting that these two factors independently impact language processing.

Introduction

Both ageing and bilingualism can lead to more costly language processing. Despite the suggestions of potential similar disadvantages, most previous research was conducted separately for bilinguals and for older adults, and focused on either comprehending or producing language. Here, we investigate the impact of both age and bilingualism on production and comprehension of first language (L1), to examine their separate and combined impact on modality-specific and modality-general language processing.

Effects of bilingualism and of ageing on language processing

There is substantial evidence for at least some disadvantages in bilingual language processing. Such disadvantages have been documented mostly for second language (L2). For example, bilinguals are slower to name pictures and respond to written words in their L2 than in their first (L1) language (Gollan et al., 2011), and they have more difficulties than monolinguals when parsing complex structures in L2, such as in garden path sentences (Pozzan & Trueswell, 2016). Explanations for these L1-L2 differences appeal to different factors, including cognitive demands in L2 (e.g., Herbay et al., 2018); a stronger reliance on declarative memory in L2 (and the consequent larger difficulty with grammar than lexicon, e.g., Ullman, 2001; see also Clahsen & Felser, 2006; Weber-Fox & Neville, 1996); or the lower frequency of use of L2 relative to L1 (Gollan et al., 2005, 2008).

In contrast, little research has examined differences between mono- and bilinguals in L1, limiting our understanding of the general consequences of bilingualism for language processing – that is, of disadvantages that are not confounded with language dominance (i.e., L1-L2) effects. An exception is Gollan et al. (2011), who tested English monolinguals and English-dominant Spanish–English bilinguals in a production task (picture naming; see also Ivanova & Costa, 2008) and in a comprehension task (eye-tracking in reading) in English. Bilinguals were slower to name pictures, and gazed longer at target words in sentences, compared to monolinguals, indicating that the bilingual disadvantage is also observed in the dominant language. Two main explanations were put forward to account for these effects. The 'frequency lag' hypothesis (a.k.a., weaker links; Gollan et al., 2005, 2008, 2011) assumes that lexical representations are less accessible in the bilingual lexicon, because they are used less often (both for L2 and L1; Kroll & Gollan, 2014), impairing lexical access. The 'competition' hypothesis (which is not incompatible with the frequency lag hypothesis; e.g., Kroll et al., 2008), on the other hand, proposes that the disadvantage stems from the control processes that bilinguals use, at any time, to suppress the non-used (but active) language (Dijkstra & Van Heuven, 2002; Green, 1998). The activation of the two languages creates competition, which needs to be resolved, slowing down lexical access (e.g., Bialystok et al., 2008). Studies designed to test these hypotheses have used tasks that tap onto lexical access (e.g., Gollan et al., 2014), but have not assessed effects at higher-levels of linguistic processing such as parsing or semantic integration (though see Runnqvist et al., 2013; Sadat et al., 2012). In L2 processing, bilinguals seem to have more difficulties than monolinguals with syntax but not with semantics (e.g., Pozzan & Trueswell, 2016; Shook et al., 2015; Schwartz & Kroll, 2006; though see Martin et al., 2013), but how L1 processing by mono- and bilinguals might differ in those aspects remains to be determined.

Healthy ageing can also incur language processing costs and, here too, the extent and nature of such costs remains under investigation (see Abrams & Farrell, 2011; Peelle, 2018, for reviews). Like bilinguals, older adults are slower to produce and comprehend language (e.g., Caplan et al., 2011; Gollan et al., 2008). Some have attributed the costlier language processing in older age to the decline in functions such as working memory and processing speed (e.g., Caplan et al., 2011; Craik & Byrd, 1982; Salthouse, 1996). Yet, there is some evidence to suggest that different domains of language such as lexical-semantics or syntax are affected differently by age. For example, Tyler et al. (2010) investigated how ageing affects syntactic processing by asking young and older adults to detect words in normal sentences, meaningless but syntactically correct sentences, or random word order strings that violated syntactic rules. Older adults performed better in meaningless sentences compared to random word strings, suggesting that they relied on syntactic analysis to predict upcoming words (see also Payne et al. 2014; Waters & Caplan, 2001; for contrasting evidence see, e.g., Caplan et al., 2011; Poulisse et al., 2019).

Others have investigated the impact of age on the use of semantic or contextual information (for a review, see Burke et al., 2000). Like Tyler et al. (2010), Cohen and Faulkner (1983) showed that providing context prior to word presentation improved word recognition by older participants (but see Federmeier & Kutas, 2005; Federmeier et al., 2002, 2003). In addition, several studies report semantic priming effects in older adults (for meta-analysis see Laver & Burke, 1993). Such evidence is compatible with ageing models that highlight the knowledge that accumulates along the lifespan, also called crystallized intelligence, which can compensate for decline in cognitive resources (Burke et al., 2000; Opdebeeck et al., 2016).

The combined impact of bilingualism and ageing on language processing

Although both the bilingualism and ageing literature highlight the disadvantages in language processing, very little research has examined their combined impact (see Reifegerste, 2021, for a review), and only one previous study has examined L1 processing. Gollan et al. (2008) asked younger and older mono- and bilinguals to name pictures with low- or high-frequency names. Bilinguals showed slower naming than monolinguals, and this disadvantage was bigger for low- than high-frequency words. Larger frequency effects were attributed to reduced language use in bilinguals compared monolinguals. Gollan et al. also predicted that this effect should reduce with age, as frequency of use accrues. However, the bilingual disadvantage in L1 did not differ between young and older adults. To explain why frequency effects and the bilingual disadvantage were constant across age groups in the dominant language, Gollan et al. (2008) hypothesized that the increased frequency of use with age was not sufficiently powerful to counteract other ageing detriments (such as general slowing).

These results suggest that bilingual disadvantages remain unchanged with age. This might indicate that bilingualism- and age-related disadvantages in language processing, being of a similar nature, would not accumulate, possibly due to a ceiling effect in terms of a combined disadvantage of age and bilingualism. It is also possible, however, that the bilingualism- and age-related disadvantages in language processing are fundamentally of different natures. Importantly, Gollan et al. (2008) tested single word stimuli. In the present study, we tested young and older monolingual and bilingual speakers in their L1, targeting sentence-level language processes. The increased level of complexity may enable us to distinguish more clearly the effects of bilingualism from the effects of ageing. Moreover, Gollan et al. tested language production only. Disadvantages of bilingualism and ageing may be more apparent in production than comprehension (e.g., Burke et al., 2000; Gollan & Goldrick, 2019; Gollan et al., 2011), making comparisons across language modalities important for distinguishing modality-specific effects from general language processes. Therefore, we tested comprehension and production, tapping into two key features of language processing: predictive comprehension and speech planning scope. In addition, in order to characterise the linguistic aspects of processing that can be equally and/or differently affected by age and language status, we manipulated properties of the stimuli that are hypothesized to underlie effects of bilingualism and ageing in language processing i.e., syntactic complexity and semantic constraint. Below, we briefly review prior evidence from similar tasks before presenting the current study.

Investigating language comprehension in the context of bilingualism and ageing

Comprehension occurs incrementally, with speakers drawing on different information sources (syntactic, lexical, semantic) to build a representation and create expectations about upcoming linguistic material, which in turn facilitate lexical access (MacDonald, 1993; Marslen-Wilson, 1975; Tanenhaus et al., 1995). Expectation-based comprehension reflects the probabilistic nature of language processing (Jaeger, 2010; Levy, 2008) and is therefore closely linked to the effects of context in comprehension, which can make upcoming words more or less likely. Accordingly, words that follow semantically high-constraining contexts are recognized and produced faster than those following a low-constraining context (Rayner & Well, 1996).

Like monolinguals, bilinguals draw on semantic information for predictive comprehension (Gollan et al., 2011; see Shook et al., 2015, for related evidence in L2 listening comprehension). Likewise, Tyler et al. (2010) showed that older adults make as much use of contextual information as young adults, during comprehension, and that syntactic processing is also preserved in old age.

In Experiment 1, we considered the joint effects of ageing and bilingualism on the use of semantic and syntactic information to guide comprehension of spoken sentences in L1. We employed a listening comprehension task similar to Tyler et al. (2010) with random lists of words, low-constraint or high-constraint sentences. Differences between random lists and low-constraint, and between low- and high-constraint sentences, index the use of syntactic and semantic information, respectively. We may expect that both bilinguals and older adults make use of semantic information to facilitate comprehension. Note that Gollan et al. (2011) only contrasted low- and high-constraint sentences, which index the use of semantic context. There is, therefore, no evidence regarding a bilingual disadvantage in L1 syntactic processing, despite evidence for disadvantages in the use of structural information in L2 comprehension (Pozzan & Trueswell, 2016). In contrast, based on evidence for age-preserved syntax (Tyler et al., 2010), we may predict no age-related differences in syntactic processing. However, our main goal is to examine the combined impact of bilingualism and ageing, and how bilingual disadvantages are affected by ageing.

Investigating language production in the context of bilingualism and ageing

Similar to comprehension, language production occurs incrementally (Kempen & Hoenkamp, 1987; Levelt, 1989). Before uttering a sentence, speakers assemble the conceptual content of what they want to say (message generation). Yet, not all words are fully prepared before speech onset. Articulation of the first word of a sentence can begin before the rest of the grammatical and phonological structure of a sentence has been encoded (see Wheeldon & Konopka, 2023, for review).

The degree of preparation that occurs in advance of articulation, or planning scope, varies depending on the type of utterances to be produced. Smith and Wheeldon (1999; see also Levelt & Maassen, 1981) showed that, when producing sentences such as ‘The cup and the hat move above the chair’ and ‘The cup moves above the hat and the chair’, speakers planned only the first (verb argument) phrase prior to speech onset (i.e., ‘The cup and the hat’; ‘The cup’). Allum and Wheeldon (2007, 2009) further contrasted sentences containing an initial verb argument phrase of the same length but different syntactic complexity, such as ‘The flower and the dog are red’ and ‘The flower above the dog is red’. In both cases, the initial verb argument phrase is of the same length, but of different syntactic complexity (coordinate (CNP) vs. prepositional (PP) NPs: ‘The flower and the dog’; ‘The flower above the dog’). Speakers took longer to produce the simpler CNP compared to the complex PP, consistent with a greater planning scope for easier-to-construct syntactic representations, much as easy message-level representations encourage speakers to plan more (Konopka & Meyer, 2014; van de Velde & Meyer, 2014).

Little research has examined planning scope in bilingualism and old age. Li et al. (2022; see also Gilbert et al., 2020; Hardy et al., 2020) showed that bilinguals and older adults planned the initial (verb argument) phrases similarly to monolinguals and young adults. However, these studies compared initial phrases where syntactic complexity was confounded with length, as in the above examples from Smith and Wheeldon (1999), and assessed effects of bilingualism and of ageing separately.

In Experiment 2, we examined the effects of bilingualism and age on production planning scope. We used pairs of pictures to be described with either coordinate NPs (CNP, e.g., ‘The cone and the grape’) or NPs modified by prepositional phrases (PP, e.g., ‘The cone above the grape’) – that is, of the same length but different syntactic complexity – to assure that longer planning latencies for coordinate phrases would indicate planning of the whole utterance, and shorter latencies for prepositional phrases would indicate planning of the first NP only. Speakers might exhibit the same difference between planning coordinate and prepositional phrases as monolingual younger adults (Allum & Wheeldon, 2007), supporting the idea of a general mechanism whereby representations that are easier to construct result in longer planning scopes. Thus, we additionally manipulated the complexity of the second NP, which could be either simple or adjective modified (e.g., ‘The cone and the grape’ vs. ‘The cone and the pink grape’). This manipulation allows testing for effects of syntactic complexity to a further extent, and evaluation of the extent of second phrase processing in the PP scope condition.

Advanced planning sets a balance between fluency and memory costs. More extensive planning allows for more fluency once speech starts, as more of the sentence is prepared, reducing processing load and the likelihood of difficulties with unprepared words once speech has begun. However, advanced planning also increases the time taken to start speaking, and the demands on working memory needed to maintain the prepared words until they are articulated. Therefore, for both bilinguals and older adults, planning more in advance might be a strategy to reduce the processing load during speech. The processing load might naturally be higher for bilinguals and older adults as a result of activation of two languages or slowed processing, respectively. In both cases, however, more extensive planning would bring additional working memory costs to maintain representations until they can be articulated.

More importantly, we are interested in how syntactic complexity of the to-be-produced utterances will affect how long participants take to start speaking. Similarly to Experiment 1, we predict that effects of syntactic complexity will not be affected by age, but hypothesize that they may be affected by bilingualism. As outlined for Experiment 1, we are moreover interested in whether bilingual disadvantages in young age decrease, persist or increase as a function of ageing.

Current study

To summarize, there is evidence from separate studies on bilingualism and ageing for comparable disadvantages in language processing, but research on the combined effects of these two factors found no interaction between them. Further investigation of the combined impact of bilingualism and age in language processing will advance our understanding of how normal ageing affects bilingual language processing (e.g., Reifegerste, 2021). We therefore tested young and older monolingual and bilingual speakers in both comprehension (Experiment 1) and production (Experiment 2) of L1 sentences where we manipulated linguistic aspects that are thought to underlie disadvantages in language processing in bilinguals and older adults. Furthermore, we tested participants’ L1, in order to assess bilingualism effects that are not confounded with language dominance. If bilingualism and ageing have comparable effects, we might expect them to interact, and to observe similar interactions between these two factors and the linguistic variables manipulated. Failing to find interactions between bilingualism and age would, on the other hand, suggest that these factors impact language processing in different ways, which should be further illustrated by different interactions between each of the factors and the linguistic variables.

Experiment 1: Listening comprehension

Experiment 1 investigated bilingualism- and age-related differences in the use of syntactic and semantic information during sentence comprehension. We employed a speech monitoring task where participants listened to a spoken sentence and made a speeded button press when they heard a pre-specified target word (*spatula* below). Spoken sentences could be lists of words in random order (RWO) (example a), low-constraint (example b), or high-constraint sentences (example c).

(a) Tried I find to quickly the spatula without pancake it to flip the breaking.
(b) I tried to quickly find the spatula to flip the pancake without breaking it.
(c) I flipped the pancake with the spatula without breaking it.

Shorter word monitoring response times (RTs) reflect easier lexical access due to expectations created based on the different types of linguistic representations that comprehenders develop, on a word-by-word basis. In (b), but not in (a), comprehenders can build coherent syntactic representations. And in (c), the semantic information constraints interpretation more than it does in (b). Therefore, differences between RWO and low- constraint conditions index the use of syntactic information, whereas differences between low- and high- constraint sentences index the use of semantic information.

Method

Participants

The data for this study were collected as part of a larger project that was publicly registered on OSF (see https://osf.io/d7aw2/). The current study reports analyses on samples of each group: 40 young English-speaking monolinguals, 40 young Norwegian–English bilinguals, 40 older English-speaking monolinguals, and 40 older Norwegian–English bilinguals (see Table 1 for the demographics). Young monolinguals (aged 18-35) and bilinguals (aged 19-30) did not differ significantly in age (Welch Two Sample t-test, $t=-0.13$, $df=58.02$, $p=0.89$), nor did older monolinguals (aged 65-81) and bilinguals (aged 66-80) ($t=0.72$, $df=77.94$, $p=0.48$).

The monolinguals reported themselves to be native speakers of British English, this being the only language they spoke at home. It was also an eligibility criterion that monolinguals should not be able to hold a simple conversation in any other language. Bilinguals completed an adaptation of the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007), and participant selection was based on the following criteria: (i) Norwegian was the first acquired language; (ii) Norwegian was the dominant language; (iii) participants self-rated both their speaking and reading proficiency in English with at least 3 on a 0-10 scale with 0 being “none” and 10 being “perfect” (see Table 1 and Appendix 1).

Monolinguals were tested at the University of Birmingham, UK, and bilinguals at the University of Agder, Norway. The younger groups were students from these universities, and the older groups were recruited from the respective communities. Older participants were screened for mild cognitive impairment through the Montreal Cognitive Assessment questionnaire (MoCA, Nasreddine et al., 2005), and all achieved a score >= 23 (Table 1), the cut-off point recommended by Carson et al. (2018; Völter et al., 2023; see Engedal et al., 2022, for normative data for Norwegian older adults). All participants had normal or corrected-to-normal vision and gave written informed consent. Our study was approved by the University of Birmingham's Institutional Ethics Board (ERN_20-1107), the Norwegian Center for research data (Ref. 239577), and the

Table 1. Demographic characteristics of participants in Experiments 1 and 2.

Characteristic	Young Norwegian-English bilinguals (N = 40)	Older Norwegian-English bilinguals (N = 40)	Young English monolinguals (N = 40)	Older English monolinguals (N = 40)
Mean age (*SD*)	22.95 (2.69)	70.18 (3.85)	23.08 (5.27)	69.55 (3.95)
Education				
Compulsory, n (%)	0 (0)	3 (7)	0 (0)	11 (28)
Upper secondary, n (%)	21 (52)	7 (17)	0 (0)	14 (35)
Undergraduate degree, n (%)	18 (45)	19 (48)	33 (82)	7 (17)
Postgraduate degree, n (%)	1 (3)	11 (28)	7 (18)	8 (20)
Mean MoCA score (SD)	—	27.36 (1.51)	—	27.50 (1.60)
LEAP-Q[a]				
Mean self-rated proficiency in speaking L1 (SD)	9.60 (0.81)	9.28 (0.78)	—	—
Mean self-rated proficiency in reading L1 (SD)	9.38 (0.87)	9.38 (1.25)	—	—
Mean self-rated proficiency in speaking L2 (SD)	7.68 (1.40)	6.75 (1.54)	—	—
Mean self-rated proficiency in reading L2 (SD)	8.28 (1.20)	7.25 (1.63)	—	—
Age of Exposure to English (started hearing)	4.85 (1.83)	9.20 (2.91)	—	—
Age Acquisition of English (started speaking)	6.60 (1.68)	12.15 (3.51)	—	—

[a] Proficiency level based on self-ratings using a scale of 0-10 with 0 being “none” and 10 being “perfect”.

Committee for Medical and Healthcare Research Ethics in Norway (REK sør-østC, ref. 163931).

Language proficiency

To further control for the potential confound of language proficiency, we administered a reading comprehension task to all participants (mono- and bilinguals) in their L1, where they (self-paced) read sentences of different syntactic complexity followed by a comprehension question (See Appendix 2 for details). We restricted to trials with correct answers to the comprehension question (87.7%) and computed, for each participant, the average reading time across all sentence types, as a measure of proficiency capturing sentence parsing processes, which was entered as a co-variate in the analysis.

Design and materials

All participants completed the tasks in their L1. We created four different sets of stimuli[1], two sets in British English and two sets in Norwegian Bokmål[2]. Each set comprised 60 target word items embedded in sentences with low-constraining context, high-constraining context, or in the RWO condition (created by randomizing the order of words in the low-constraint sentence). To determine the extent of constraint of each sentence context, we conducted a pretest with a different set of participants completing a cloze task on 132 sentences, for each language, based on which we selected 120 items in each language, which were matched for length and mean cloze scores (see Appendix 4 for details).

For each set, each of the 60 items appeared in each of the three context conditions across three different lists (Latin Square Design). In addition to the experimental items, 12 fillers were created using different target words/ sentences. Therefore, each list contained 72 trials. These were divided into four blocks of 18 items (containing 15 experimental and 3 filler items). The order of trials within a block was pseudo-randomized so that no more than three trials of the same condition occurred after each other. Two versions of each list were created with a different order of presentation of the blocks. The final experiment therefore consisted of two unique sets with 6 lists each, per language. Participants were pseudo-randomly assigned to lists. The same four practice sentences were presented in the beginning of the experiment, that preceded the presentation of each individual list.

The sentences were recorded by female native speakers of Standard British English and Norwegian Bokmål in a sound-proof booth using a high-quality USB microphone (Røde NTG) at a sampling rate of 44.1 kHz. The sound files were equalized for intensity, cut to the exact length of the sentences at zero crossings using Praat (Boersma & Weenink, 2011), and were not further manipulated.

The design crossed two between-subjects variables (age group: old vs. young, and bilingualism group: bilingual vs. monolingual), and one within-subjects variable (context: RWO, low-constraint, high-constraint).

Procedure

The stimuli were presented using Presentation® (Version 20.1, Neurobehavioral Systems, Inc.). Spoken sentences were presented through speakers connected to the computer. Each trial began with a central fixation cross ('+') presented for 500ms, followed by a 1000ms blank screen and then the presentation of the target word, in the centre of the screen, for 1000ms (font 'Consolas', size 30). 500ms after target word offset, the spoken sentence started playing. Participants were instructed to pay attention to the written target word and to press the space bar on the keyboard as soon as they heard it in the spoken stimuli. Each trial ended 2000ms after the end of the audio file. Monitoring RTs were measured from the onset of each target word in the spoken stimulus. The task took around 15 minutes to complete.

Analysis

We first discarded RTs below 150ms or above 1500ms (3.18% of data; refer to OSF for an explanation of the trimming procedure and its impact on the analyses). To normalise the time responses and avoid the confound of potential differences in cognitive speed that do not reflect specific group differences in fundamental mental operations (such as, e.g., general slowing in older adults, but also potential cognitive speed differences between mono and bilinguals), we z-scored RTs, for each participant (Faust et al., 1999; each observation is subtracted from the mean and divided by the standard deviation of RT). We then excluded z-scores <-2.5 or >2.5 – that is, observations distancing more than 2.5*SD* from the mean – for each participant (2% of data).

For the analysis of (z-scored)RT, we ran linear mixed-effects models (LMMs; Baayen et al., 2008), which allow for simultaneous estimation of between-participants and between-items variance. We used LMMs as implemented by *blmer* in R. *blmer* closely follows the function *lmer*, the primary distinction being that *blmer* allows the user to do Bayesian inference or penalized maximum likelihood, with priors imposed on the different model components[3], avoiding 'singular fits' in the models when using *lmer* (Chung et al., 2013). We fitted full models (all main effects and interactions) with a maximal-random structure when justified by the design (Barr et al., 2013) and if convergence occurred. In particular, the random structure included intercepts by participant and item, but random slopes for each fixed-effect only by-participant (excluding Age Group, Bilingualism Group and Proficiency scores, as these are not justified) and not by item, as all of these resulted in non-convergence, or a worse fit, as assessed through Likelihood ratio tests[4] (refer to the summaries of models in the tables for its syntax). We report the predictors' coefficients (β values), *SE*, *t* (or *z*) values, and the derived *p* significance values (by treating the t-statistic using the standard normal distribution as a reference; e.g., Baayen et al., 2008, footnote 1).

We had participants (160) and items (240) as random effects in the analysis. For the fixed effects, we coded the context condition using forward difference coding, where each contrast compares adjacent levels (each level minus the next level): the first contrast is 'RWO' minus 'low-constraint' and the second one is 'low-constraint' minus 'high-constraint'. Other fixed effects were Bilingualism group (bilingual vs. monolingual) and Age group (old vs. young), both contrast coded by centering (refer to summary of the model for coefficients). In addition, reading scores were added as a predictor to account for individual differences in Proficiency (Prof). Finally, Education level (see Table 1) was added as a confounding variable (with four categorical levels: 'CompulsoryEducation', 'PostgraduateDegree', 'UndergraduateDegree', and 'UpperSecondary'; the first level was the reference level).

Results and discussion

Figure 1 presents the mean word monitoring RT (z-scored, refer to Appendix 6 for raw RTs), and Table 2 presents the fitted LMM model. Participants were faster to detect words in sentences with a low-constraint compared to random word lists (RWO). This effect was qualified by an interaction with bilingualism, indicating

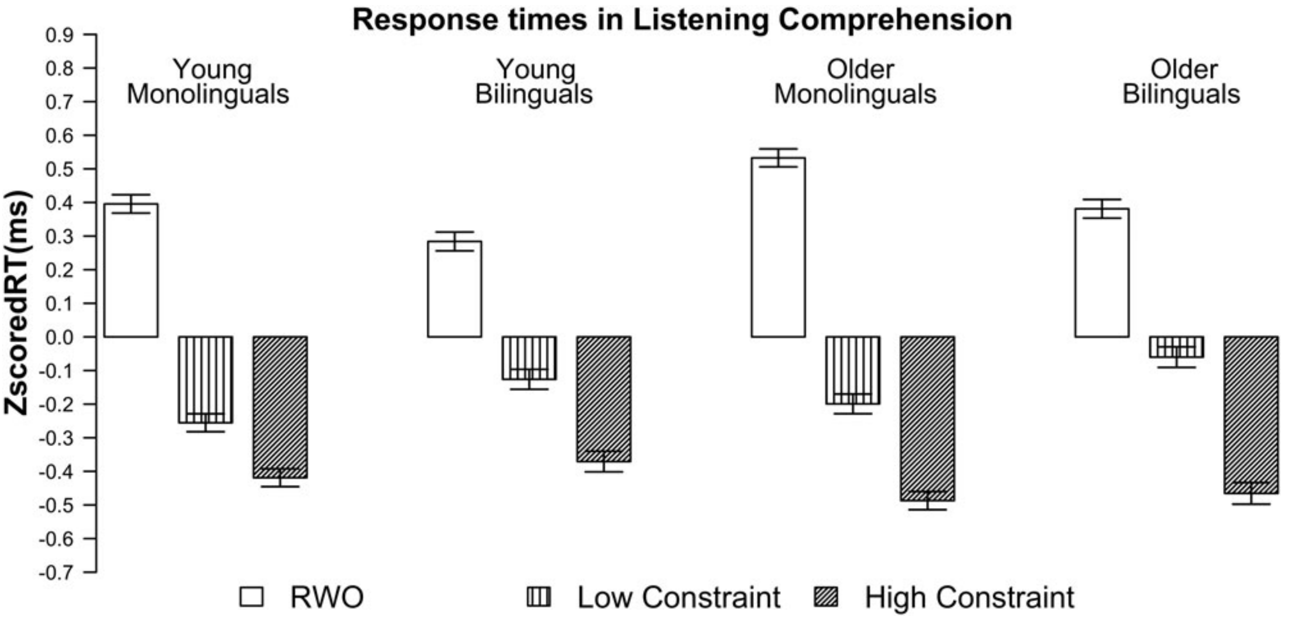

Figure 1. Mean word monitoring RT (z-scored, ms) for the Younger and Older Monolingual and Bilingual groups, in the three experimental conditions. Error bars represent standard errors on means.

that the difference was larger in monolinguals than bilinguals. To explore this interaction, we conducted a follow-up LMM on the data from these context conditions where we created a four-level variable combining the two context levels (RWO vs. low-constraint) with the two bilingualism levels (bilingual vs. monolingual). We found that monolinguals were faster than bilinguals in low-constraint sentences (Bilingualxlow_Monolingualxlow; $\beta = -0.13$, $SD = 0.06$, $p = 0.02$), suggesting they were more able to use basic syntactic information to predict upcoming words. This result shows that the bilingual disadvantages in L2 syntactic processing (e.g., Pozzan & Trueswell, 2016) are also observed for L1.

Interestingly, bilinguals were faster than monolinguals in RWO (BilingualxRWO_MonolingualxRWO; $\beta = 0.11$, $SD = 0.06$, $p = 0.05$). This result adds to previous research that only contrasted low- and high-constraint sentences (e.g., Gollan et al., 2011[5]). Our data suggest that bilinguals outperform monolinguals in detecting words that are embedded in linguistic material that does not allow listeners to construct a coherent syntactic or semantic analysis. In this case, comprehenders should benefit from ignoring the irrelevant linguistic stimuli of the random word lists, focussing attention only on the (pre-specified) target word. Although our study was not set to test this hypothesis, we suggest that bilinguals have an advantage in such circumstance, due to their improved language control (e.g., Bialystok & Craik, 2022).

Comprehenders were also faster in high- compared to low-constraint sentences, and this effect was qualified by two-way interactions with bilingualism and with ageing. To follow up the first interaction, we had a four-level variable combining the context (high- and low-constraint) with the two bilingualism levels. Monolinguals were faster than bilinguals in low-constraint (as reported already above), but there were no group differences when high-constraint context was provided (Bilingualxhigh_Monolingualxhigh; $p = 0.49$). Thus, bilinguals are not at a disadvantage when they have contextual information to guide comprehension. This result parallels findings for L2 comprehension, where differences between monolinguals and bilinguals were apparent in low-constraint contexts, but not when semantic information could be used for prediction (e.g., Schwartz & Kroll, 2006; Shook et al., 2015). This is also consistent with the findings for L1 processing in Gollan et al.'s (2011) Experiment 2 (but see their discussion on the possible trade-off between reading times and skipping rates).

Finally, we ran a similar four-level variable model contrasting age groups. No age differences were found in the low-constraint condition (youngerxlow_olderxlow; $p = 0.16$), supporting evidence for preserved syntactic processing with age (e.g., Tyler et al., 2010). Moreover, older adults tended to be faster than young to detect words in high-constraint contexts (youngerxhigh_olderxhigh; $\beta = -0.05$, $SD = 0.03$, $p = 0.09$). Although this effect did not reach significance, it would at least indicate the absence of a disadvantage in semantic processing in older adults, consistent with behavioural evidence for preserved semantic processing in old age. Tyler et al. (2010) found no age-differences in the processing of sentences like 'The church was broken into last night. Some thieves stole most of the LEAD off the roof' (their Normal prose condition). Note that this type of sentence is similar to the ones we used in our low-constraint condition, where we too found no age-differences. In contrast, we had an additional condition where coherent semantic information was not only present, but also strongly constrained the interpretation (high-constraint condition). Our results suggest that, in these circumstances, there is a trend for older adults to outperform young.

Language proficiency, as measured by reading scores, did not influence word monitoring responses.

Experiment 2: Phrase production

Experiment 2 tested bilingualism- and age-related differences in a production task assessing planning scope of utterances. Participants described two pictures using phrase types that are known to have different planning scopes in young monolinguals, coordinate NPs (CNP; e.g., 'The cone and the grape') and NPs modified by prepositional phrases (PP; 'The cone above the

Table 2. Summary of the linear mixed effects model fitted to the z-scored time to word monitoring (RT).

(Zscored-) RT to target, Listening Comprehension				
Predictors	*Est.*	*SE*	*t*	*p*
(Intercept)	−0.057	0.04	−1.49	0.14
RWO.lowConst	**0.578**	**0.03**	**18.19**	**<1.01**
lowConst.highConst	**0.237**	**0.03**	**8.75**	**<1.01**
Bilingualim [bil., -0.5; mon. 0.5]	−0.016	0.05	−0.33	0.74
Prof [scaled, continuous, -1.60 to 3.58]	0.000	0.01	−0.01	0.99
AgeGroup [old, -0.5; young, 0.5]	−0.031	0.02	−1.51	0.13
HighestEduPostgraduateDegree	−0.009	0.04	−0.24	0.81
HighestEduUndergraduateDegree	−0.011	0.03	−0.32	0.75
HighestEduUpperSecondary	−0.007	0.03	−0.20	0.84
RWO.lowConst:AgeGroup	−0.092	0.06	−1.46	0.15
lowConst.highConst:AgeGroup	**−0.105**	**0.05**	**−1.96**	**0.05**
Bilingualim:AgeGroup	−0.003	0.04	−0.07	0.94
RWO.lowConst:Bilingualim	**0.301**	**0.06**	**4.74**	**<1.01**
lowConst.highConst:Bilingualim	**−0.104**	**0.05**	**−1.93**	**0.05**
RWO.lowConst:Prof	0.007	0.03	0.23	0.82
lowConst.highConst:Prof	−0.020	0.03	−0.73	0.47
Bilingualim:Prof	0.013	0.02	0.64	0.52
Prof:AgeGroup	−0.002	0.02	−0.11	0.91
Bilingualim:Prof:AgeGroup	0.006	0.04	0.15	0.88
RWO.lowConst:Bilingualim:AgeGroup	0.004	0.13	0.03	0.98
lowConst.highConst:Bilingualim:AgeGroup	0.139	0.11	1.31	0.19
RWO.lowConst:Prof:AgeGroup	−0.019	0.07	−0.29	0.77
lowConst.highConst:Prof:AgeGroup	−0.088	0.06	−1.56	0.12
RWO.lowConst:Bilingualim:Prof	0.085	0.07	1.31	0.19
lowConst.highConst:Bilingualim:Prof	0.084	0.06	1.52	0.13
RWO.lowConst:Bilingualim:Prof:AgeGroup	0.180	0.13	1.36	0.17
lowConst.highConst:Bilingualim:Prof:AgeGroup	−0.019	0.11	−0.17	0.86

Note: The syntax of the model is: blmer(depM ∼ 1 + RWO.lowConst + lowConst.highConst + Bilingualim + Prof + AgeGroup + HighestEdu + RWO.lowConst:AgeGroup + lowConst.highConst:AgeGroup + Bilingualim:AgeGroup #+ compProf:AgeGroup + RWO.lowConst:Bilingualim + lowConst.highConst:Bilingualim + Bilingualim:AgeGroup:Prof + RWO.lowConst:Bilingualim:AgeGroup + lowConst.highConst:Bilingualim:AgeGroup + RWO.lowConst:Prof:AgeGroup + lowConst.highConst:Prof:AgeGroup + RWO.lowConst:Bilingualim:Prof + lowConst.highConst:Bilingualim:Prof + Prof:Bilingualim:AgeGroup + RWO.lowConst:Bilingualim:Prof:AgeGroup + lowConst.highConst:Bilingualim:Prof:AgeGroup + + (1 | subj) + (1 | item) + (0 + RWO.lowConst | subj) + (0 + lowConst.highConst | subj), data = dataset, control = lmerControl(optimizer =“Nelder_Mead”))

grape’), as indexed by longer production onsets on the first than on the latter (Allum & Wheeldon, 2007). We further manipulated complexity by having the second NP modified or not modified by an adjective (simple vs. complex; e.g., ‘The cone and the grape’ vs. ‘The cone and the pink grape’). Following previous studies, we take speech onset as a measure of the amount of preparation needed to start producing utterances, i.e., planning scope (e.g., Allum & Wheeldon, 2007; Konopka & Meyer, 2014).

Method

Participants

The participants were the same as in Experiment 1 and completed the phrase production task after finishing the listening comprehension experiment.

Language proficiency

To control for language proficiency effects in the production experiment, we used scores from a vocabulary task. Participants were presented with words and asked to choose between four options: one correct answer which was either a synonym or antonym of the word, and three foils. We computed, for each participant, the percentage of accurate responses (number of correct answers divided by number of trials; see Appendix 3 for further details). Lexical knowledge should modulate picture naming (e.g., Gollan et al., 2005), and we thus use the vocabulary scores as an additional predictor controlling for proficiency in the production experiment.

Design and materials

We adapted the multi-picture description task introduced by Smith and Wheeldon (1999) such that participants were presented, on each trial, with a four-picture display. Each experimental item was a 600 x 600-pixel image containing four individual pictures. On each display, two of the pictures (target word-pair) were surrounded by a rectangular red line frame, indicating the phrase type to be produced. Coordinate phrases were cued by a horizontal red frame (that could appear at the top or the bottom; Figure 2 (a) and (b)), and prepositional phrases were cued by a vertical red frame (that could appear at the right or left; Figure 2 (c) and (d)). The second NP on each phrase type could also be simple or modified by an adjective. In the latter case, the target picture corresponding to the second noun appeared twice, once as the target modified picture and once as the original non-modified picture (Figure 2 (b) and (d)). The design crossed phrase type (coordinate; prepositional) and complexity (simple: not modified; complex: adjective modified).

We created four different sets of stimuli[6], each using 20 different target pictures, from the MultiPic picture database (Duñabeitia et al., 2018), combined in 20 unique picture pairs (each picture occurred both as the first or second element of a different pair). In each set, each of the 20 word-pair items appeared in the four experimental conditions so that each participant experienced every item in every condition. Thus, each set comprised 80 experimental items distributed across four conditions, with each individual experimental picture appearing 8 times (rotated across the screen positions). The other two of the four pictures on each display were pictures that did not appear in another experimental item but only as part of the filler displays. In addition to the 80 experimental items, we had 48 filler items.

The length of the names of the target pictures did not differ significantly between English and Norwegian, as to the number of syllables and phonemes (see Appendix 5 for further details).

Two practice blocks were administered before the experiment started (containing the 20 experimental pictures re-arranged and 16 fillers). For each set of stimuli, we created two different orders of stimuli presentation. Therefore, we had 8 different lists, and participants were pseudo-randomly assigned to these.

Procedure

Pictures were named in speakers' L1. Following common practice in the picture naming literature, participants were first given a booklet to familiarise with the pictures and respective names, and they were encouraged to look carefully at each image-name pair, and to confirm that they recognized it. Participants were also shown previously examples of the grids of pictures that would appear on the screen, for which they would be asked to uniquely identify the red framed pictures. They should describe the pictures having a horizontal red frame around them from left to right using 'and' (e.g., 'the cone and the grape'; 'kjeglen og druen'), and the pictures having a vertical red frame around them from top to bottom using 'above' (e.g., 'the cone above the grape'; 'kjeglen over druen'). They were also told that one of the objects to be named could have a different colour or size but be of the same type as one object outside the red frame. In this case, they should use the adjectives to distinguish the picture within the red frame. Finally, they were told there were other cases where all the pictures would be identical within the display or where the display would be empty, in which case they should say 'all the images are the same/ alle bildene er like' or 'there are no pictures/ det er ingen bilder'. Participants were asked to start speaking as quickly and accurately as possible, after seeing each display.

The experiment was presented using Presentation® (Version 20.1, Neurobehavioral Systems, Inc.). Each trial began with a central fixation cross ('+') presented for 500ms, followed by a 500ms

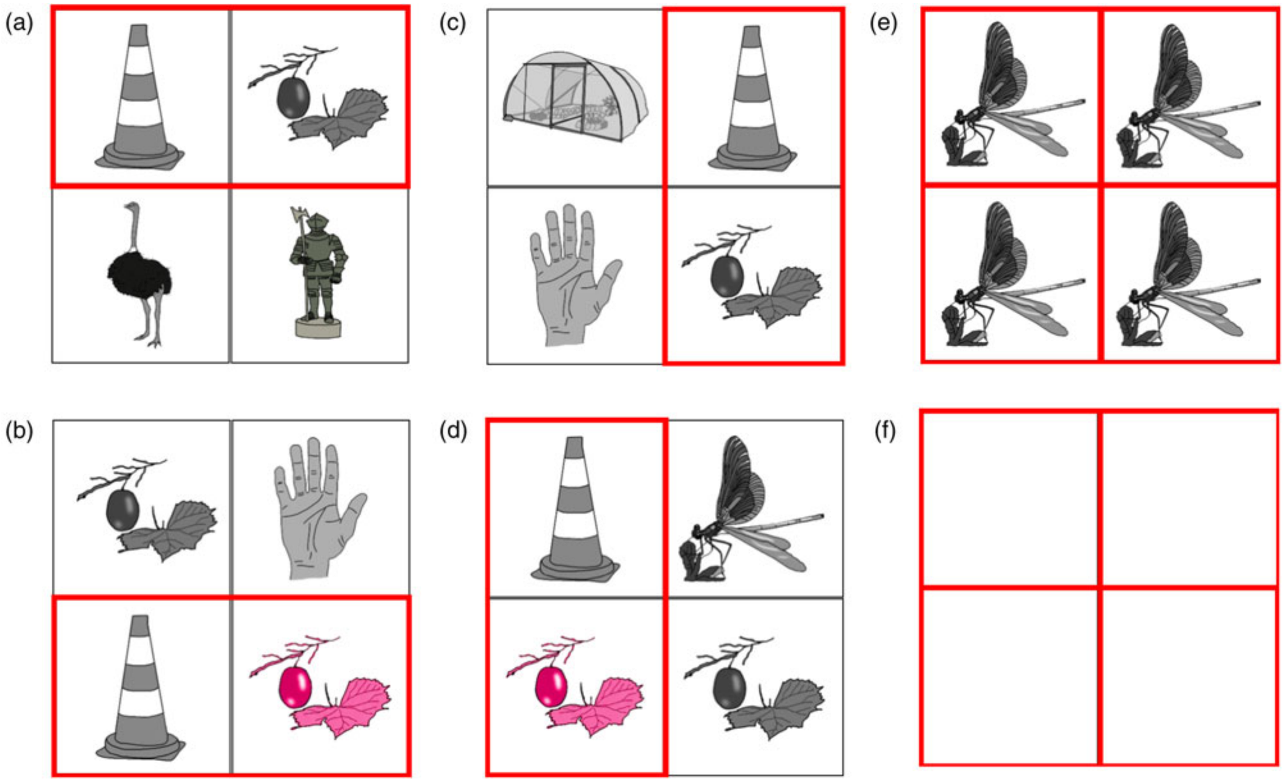

Figure 2. Example of the stimuli images. (a) to (d) illustrate an experimental item in the four conditions crossing phrase type and complexity, whereas (e) and (f) are examples of filler items. The utterances corresponding to (a) to (d) are, respectively (English version): 'The cone and the grape', 'The cone and the pink grape', 'The cone above the grape', and 'The cone above the pink grape'.

blank screen, after which the multi-picture display was presented and the recording of production started. The program registered speech onset, and an experimenter was present throughout the experiment to record the accuracy of responses (responses in which participants did not use the expected names or phrase type, where they did not mention the adjective, or disfluent responses were all categorized as errors). The picture-display disappeared after the participant finished speaking (or after 3000ms with no production), triggering the presentation of the next trial. The task took approximately 25 minutes.

Analysis

Our analyses of speech onset were restricted to trials with correct utterances (82.5% of data). Due to a technical problem, the voice onsets for one monolingual older adult were not registered and the data from this participant are missing from the analyses.

We first excluded extreme speech onset measures <250ms or >2500ms (2.1%). We then z-scored RTs to avoid spurious interactions between groups and conditions, due to general slowing (see *Analysis* for Experiment 1), and further discarded observations distancing more than 2.5*SD* from the mean, for each participant (i.e., zscores <-2.5 or >2.5, corresponding to 1.8% of data).

Speech onsets were analysed with LMMs, as in Experiment 1. Participants (159) and items (80) were entered as random factors. The fixed effects were bilingualism group (bilingual vs. monolingual), age group (old vs. young), phrase type (coordinate vs. prepositional), and complexity (complex vs. simple), all contrast coded by centering (refer to summaries of models for the coefficients). Moreover, vocabulary scores (voc; numeric, scaled) were added as a co-variate, to account for individual differences in language proficiency, and Education level was added as a confounding variable (as in Experiment 1). We fitted full models with a maximal-random structure when justified by the design and if convergence occurred (refer to the summaries of models for its syntax). We report the predictors' coefficients (β values), *SE*, *t* values, and the derived *p* significance values.

Results and discussion

Figure 3 presents the mean speech onset latencies (z-scored, refer to Appendix 7 for raw RTs), and Table 3 presents the fitted LMM model to (z-scored) RTs.

Participants were slower to produce coordinate than prepositional utterances, consistent with prior evidence for young monolinguals (Allum & Wheeldon, 2007). This confirms that, in the face of the syntactically simpler coordinates, speakers engage in longer planning, arguably preparing the whole coordinated phrase (CNP). Conversely, to produce more complex prepositional phrases (PP) speakers plan less, leaving preparation of (at least some aspects of) the second NP for later. Our data are the first to show that differences in the planning of coordinate and prepositional phrases are used in similar ways by monolinguals and bilinguals, and by young and older adults. Speakers also took longer to start producing utterances where the second NP was modified by an adjective, compared to utterances with non-modified NPs. This shows that the additional complexity of the adjective elicited longer preparation time, and that, despite a general strategy whereby, for more complex phrases (PP), the second NP is not fully prepared in advance, there is some level of processing of it before speech (see, Allum & Wheeldon, 2007, for discussion).

The effect of phrase type (CNP, PP) interacted with age group. In a follow-up analysis using a four-level variable combining the variables age and phrase type, we found that young and older participants did not differ in the onsets of coordinate utterances (olderxcoord_youngerxcoord; $p = 0.47$), but older adults were slower than young adults in prepositional utterances (olderxprep_youngerxprep; $\beta = -0.07$, $SD = 0.02$, $p < 0.01$). A possible interpretation is that older adults planned more in advance than young adults

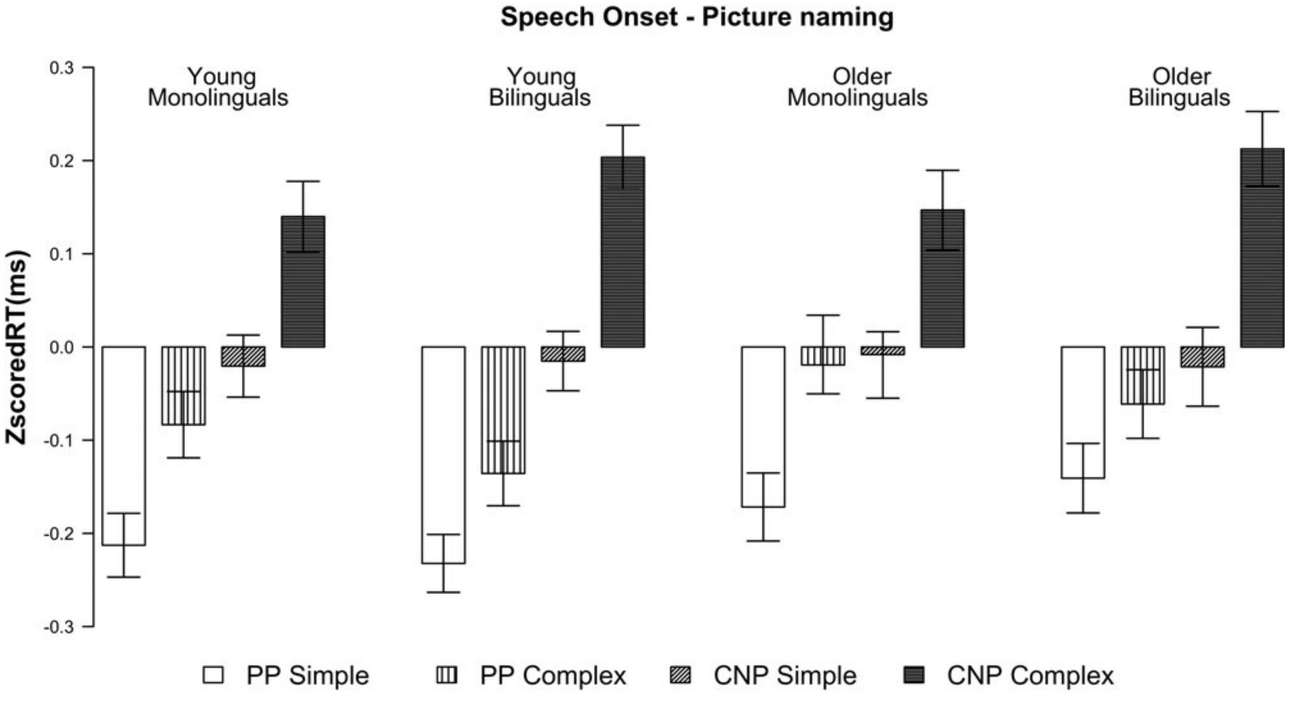

Figure 3. Mean onset latencies (z-scored, ms) for the Younger and Older Monolingual and Bilingual groups, in the four experimental conditions. Error bars represent standard errors on means.

Table 3. Summary of the linear mixed effects model fitted to the z-scored speech onset times (RT).

Speech onset time (Zscored-) RT				
Predictors	*Est.*	*SE*	*t*	*p*
(Intercept)	−0.005	0.05	−0.10	0.92
PhraseType [coord., -0.5; prep, 0.5]	**−0.212**	**0.03**	**−6.61**	**<1.01**
BilingualismGroup [bil., -0.5; monol., 0.5]	−0.009	0.03	−0.32	0.75
Complexity [complex, -0.5; simple, 0.5]	**−0.146**	**0.03**	**−5.27**	**<1.01**
voc [continuous, centred, -12 to 7]	−0.001	0.00	−0.29	0.77
AgeGroup [old, -0.5, young, 0.5]	−0.037	0.03	−1.18	0.24
HighestEduPostgraduateDegree	−0.026	0.05	−0.54	0.59
HighestEduUndergraduateDegree	−0.032	0.05	−0.70	0.49
HighestEduUpperSecondary	−0.033	0.04	−0.74	0.46
PhraseType:AgeGroup	**−0.129**	**0.07**	**−1.97**	**0.05**
BilingualismGroup:AgeGroup	0.042	0.06	0.68	0.50
Complexity:AgeGroup	0.025	0.06	0.44	0.66
voc:AgeGroup	0.002	0.01	0.43	0.67
PhraseType:voc	−0.007	0.01	−1.22	0.22
BilingualismGroup:voc	0.004	0.01	0.74	0.46
Complexity:voc	−0.001	0.01	−0.24	0.81
PhraseType:BilingualismGroup	0.062	0.05	1.33	0.18
PhraseType:Complexity	**0.076**	**0.04**	**2.14**	**0.03**
BilingualismGroup:Complexity	0.029	0.04	0.71	0.48
PhraseType:BilingualismGroup:AgeGroup	0.050	0.13	0.38	0.70
PhraseType:Complexity:AgeGroup	0.094	0.10	0.94	0.35
PhraseType:voc:AgeGroup	−0.011	0.01	−0.96	0.34
BilingualismGroup:Complexity:AgeGroup	−0.121	0.11	−1.08	0.28
BilingualismGroup:voc:AgeGroup	0.000	0.01	0.03	0.97
Complexity:voc:AgeGroup	0.002	0.01	0.24	0.81
PhraseType:BilingualismGroup:voc	−0.010	0.01	−0.90	0.37
PhraseType:Complexity:voc	0.015	0.01	1.60	0.11
BilingualismGroup:Complexity:voc	−0.013	0.01	−1.29	0.20
PhraseType:BilingualismGroup:Complexity	**−0.151**	**0.07**	**−2.02**	**0.04**
PhraseType:BilingualismGroup:Complexity:AgeGroup	0.018	0.15	0.12	0.90

Note: The syntax of the model is: blmer(depM ∼ 1 + PhraseType + BilingualismGroup + Complexity + voc + AgeGroup + HighestEdu + PhraseType:AgeGroup + BilingualismGroup:AgeGroup + Complexity:AgeGroup + voc:AgeGroup + PhraseType:voc + BilingualismGroup:voc + Complexity:voc + PhraseType:BilingualismGroup + PhraseType:Complexity + BilingualismGroup:Complexity + PhraseType:AgeGroup:BilingualismGroup + PhraseType:AgeGroup:Complexity + PhraseType:AgeGroup:voc + BilingualismGroup:AgeGroup:Complexity + BilingualismGroup:AgeGroup:voc + Complexity:AgeGroup:voc + PhraseType:voc:BilingualismGroup + PhraseType:voc:Complexity + BilingualismGroup:voc:Complexity + PhraseType:BilingualismGroup:Complexity + PhraseType:BilingualismGroup:Complexity:AgeGroup + (1 | subj) + (1 | item) + (0 + Complexity | subj), data = dataset, control = lmerControl(optimizer =“Nelder_Mead”)

when the phrase to be uttered was syntactically more complex (PP). Yet, another possibility is that older adults engage in longer planning when possible (that is, in PP structures where some planning of the second NP could occur in parallel with planning of the first NP). In fact, the difference in syntactic complexity between coordinated and prepositional phrases is already reflected in the fact that, for prepositional phrases (the more complex structure), speakers prioritise preparation of the first noun phrase before speech. Therefore, we argue that this result indicates an overall larger planning scope in older age, whereby speakers prepared some aspects of the second NP (note that this would not affect CNP as in that case there is already full planning of the utterance). As we mentioned, advanced planning allows for more fluency once speech starts, and this should be particularly relevant for older adults, who have increased speech disfluencies (Cooper, 1990; Kemper, 1992; though see Spieler & Griffin, 2006, for contrasting evidence).

Phrase type also interacted with complexity, indicating that structures with adjective modified NPs evoked slower responses in coordinates than prepositional phrases, which confirms extensive pre-speech processing for coordinates but the prioritising of the first |NP in prepositional phrases. The interaction between phrase type and complexity was further qualified by a three-way interaction with Bilingualism. We conducted follow-up LMMs on the data from the complex and simple conditions, separately, and in each case, we created a four-level variable combining the two phrase type levels (coordinate vs. prepositional) with the two bilingualism levels (bilingual vs. monolingual). Bilinguals tended to be slower than monolinguals producing coordinate utterances that had a modified second NP (complex CNP; Bilingualxcoord_Monolingualxcoord; $\beta = -0.07$, $SD = 0.04$, $p = 0.069$; the effect of bilingualism was not observed in complex PPs, nor in any of the no adjective conditions: all $ps > 0.4$). This effect only approached significance, and therefore should be interpreted with caution. Yet, it is consistent with the idea that the effect of bilingualism is to slow speech onset, but only for the phrase productions that demand the longest planning – that is, coordinated noun phrases with the second NP modified by an adjective.

Finally, language proficiency, as measured by vocabulary scores, did not affect speech onset latencies.

General discussion

Many studies on bilingual language processing have focused on L2, and have examined comprehension and production separately. In the current study, we tested the same participants in comprehension and production of L1 (while controlling for language proficiency), and demonstrated comparable findings – namely, poorer syntactic processing in bilinguals compared to monolinguals. We thus provide evidence for bilingualism effects *per se*, that are modality-general. Moreover, we tested young and older adults, and observed age-preserved syntactic processing across modalities. In our comprehension experiment, we also tapped onto semantic processing, and we found no differences between mono and bilinguals, whereas older adults tended to outperform younger adults. Importantly, we found no interactions between ageing and bilingualism, suggesting that the two factors independently impact language processing. Any disadvantages in bilingual syntactic processing were not counteracted by the preservation of that ability in ageing, and thus the bilingual differences remained in old age. On the other hand, what seem to be advantages in semantic processing in old age are not larger for bilinguals than monolinguals, suggesting that a higher reliance on accrued knowledge by the older adults does not accumulate with, or is distinct in nature from, any potential increased reliance on lexical-semantics by bilinguals.

The results of Experiment 1 suggest that bilingualism adversely impacts syntactic processing during comprehension, a finding resembling evidence for L2 processing (e.g., Pozzan & Trueswell, 2016). On the contrary, syntax seems to be preserved in old age. Whereas some research reported age-related decline in processing of complex syntactic structures (e.g., Caplan et al., 2011) or morpho-syntactic agreement interpretation (e.g., Poulisse et al., 2019), we replicated Tyler et al. (2010) in finding no age-related differences in reaction times to detect words during online comprehension of sentences with 'normal' syntactic complexity.

Moreover, we used sentences where the target words were highly predictable given the sentence context to assess semantic processing. Consistent with some evidence (e.g., Gollan et al., 2011; Tyler et al., 2010), we found no processing differences between monolinguals and bilinguals, and older adults even tended to outperform young adults when sentences provided a high-constraint context. The reliance on semantic information might reflect strategies to compensate for detriments in other processes, such as general cognitive decline, in healthy ageing (Opdebeeck et al., 2016), and mirrors findings for L2 processing (e.g., Clahsen & Felser, 2006). We should note, however, that there is also physiological evidence against identical semantic processing mechanisms, both in bilingualism and in older age, from studies using sentences with low- and high-constraints (Federmeier & Kutas, 2005; Federmeier et al., 2002, 2003; Martin et al., 2013). This highlights that, even when performance between groups is comparable, the underlying brain function mechanisms may not be. For our present findings, it may also be the case that seemingly comparable behaviour across groups is supported by different neural mechanisms.

In Experiment 2, we examined how the grammatical complexity of to-be-produced phrases affected planning scope. We showed, for the first time, that similar planning scope patterns of larger planning for coordinated phrases than prepositional phrases (Allum & Wheeldon, 2007) are used across age and bilingualism groups. However, older adults were slower in producing prepositional phrases than younger adults, suggesting an overall larger planning scope in older age.

We further manipulated syntactic complexity by including phrases with an adjective modifying the second NP. We found that the adjective modification was particularly demanding when speakers planned coordinates, confirming the preparation of the whole utterance in advance. Thus, modified coordinated phrases required most preparation, as indicated by overall longer latencies. Complexity did not differentially affect old and young adults. This is in line with the absence of age-differences in syntactic comprehension (Experiment 1), and further supports our suggestion that older adults were slower in producing prepositional phrases because they engaged in more extensive planning overall, and not because of the additional syntactic complexity of these structures. However, bilinguals showed a disadvantage (longer speech onset latencies) relative to monolinguals in modified coordinated phrases. Therefore, syntactic complexity affected bilinguals more than monolinguals, which is in line with our findings from Experiment 1.

We were interested in the types of linguistic information that could underlie disadvantages in language processing in bilinguals and older adults. For that purpose, we used whole sentences and complex phrases as stimuli, and interpreted our results as reflecting syntactic and semantic processing effects on lexical access.

Lexical access has been at the heart of accounts for bilingual disadvantages. One of the main findings in bilingualism research refers to parallel activation of bilinguals' two languages in both comprehension (e.g., Schwartz & Kroll, 2006; Spivey & Marian, 1999) and production (e.g., Colomé, 2001; Costa et al., 2000). According to the competition hypothesis, the bilingual disadvantage results from the competition between activated lexical items and the need for conflict resolution during lexical access.

Semantically constraining contexts reduce the extent of dual language lexical activation during comprehension (e.g., Schwartz & Kroll, 2006) and, therefore, reduce language competition and the bilingual disadvantage (e.g., Gollan et al., 2011). We replicate this finding by showing a reduced bilingual disadvantage in comprehending words embedded in high-, compared to low-constraint sentences. The fact that syntactic and semantic information are used to guide comprehension is furthermore supported by the comparison between the low-constraint sentences and the random word order condition. When no coherent syntactic or semantic information is present, more lexical candidates are active, slowing processing. Yet, although both mono- and bilinguals were slower in this condition compared to normal sentences, bilinguals outperformed monolinguals in monitoring words embedded in random word lists. This effect is compatible with better attentional control in bilinguals, compared to monolinguals (Bialystok & Craik, 2022). Future research is required to determine the role of attention and executive function in more complex language processes.

On the other hand, the frequency lag hypothesis (Gollan et al., 2008, 2011) proposes that the bilingual disadvantage reflects frequency effects emerging from patterns of language use: bilinguals use word forms less frequently than monolinguals, and thus lexical items are less accessible, in both L1 and L2. Studies evaluating the frequency lag hypothesis have shown that the bilingual disadvantage is larger for low- than high-frequency words, and propose that the greater frequency of use in monolinguals leads to a ceiling in activation. Our study was not set to test this hypothesis, and does not speak directly to frequency effects. Also some aspects of our study would minimize the potential differences in frequency of use. First, our stimuli were matched for frequency in English and Norwegian. Second, we found no effects of language proficiency, which should correlate with language experience. Finally, and more importantly, our bilingual population had Norwegian as both their first acquired and dominant language (in contrast to switched-dominance bilinguals; see Hanulová et al., 2011, for discussion), they were early bilinguals (Table 1), and resident in Norway (the L1 speaking country). Although we had no direct measure of language exposure, these are conditions that promote an experience of L1 closer (though necessarily smaller) to that of monolinguals, and therefore our reported bilingual disadvantage likely reflects the (additional) influence of something other than a frequency lag.

We also argued that syntactic complexity affected bilinguals in Experiment 2, where modification of the second NP reflected higher syntactic complexity. In a previous study, Sadat et al. (2012) asked mono- and bilinguals to name pictures by producing either single words (e.g., 'airplane') or adjective modified NPs (e.g., 'the red airplane'). Bilinguals always took longer to start speaking than monolinguals, but the bilingual disadvantage in speech onset was of similar magnitude for single word and modified-NP utterances. Although production is arguably more complex for modified NPs than bare nouns, it is possible that Sadat et al.'s (2012) stimuli were not complex enough to elicit differences. In contrast, we elicited utterances comprising two NPs. When these were coordinated (CNP; 'The cone and the grape'), speakers across groups planned more extensively, but the additional complexity of having the second NP adjective-modified ('The cone and the pink grape') only affected the bilingualism group, where bilinguals showed a cost in speech onset relative to monolinguals. Of course, in the CNP condition, there are more lexical items to retrieve, which makes our result compatible with a bilingual disadvantage being observed due to lexical competition. However, this finding is also compatible with an explanation based on the increased complexity of the syntactic processes needed to produce these utterances.

It has been argued that the bilingual disadvantage due to a frequency lag may be restricted to low-frequency lexical items and syntactic structures. Again our study does not speak directly to this issue, as we tested structures that are frequent in the tested languages (see Runnqvist et al., 2013, for related evidence on a bilingual disadvantage in production of infrequent structures and the teasing apart of lexical and syntactic retrieval). Crucially, the fact that we did find reliable bilingual disadvantages in processing frequent stimuli, for early and non-switched dominant bilinguals, suggests that the effects of being bilingual on language processing go beyond effects of frequency of experience with a language.

More important for the contrast between the competition and the frequency lag theories is the lack of interactions between bilingualism and age that we found. According to the frequency lag hypothesis, increased language experience should result in reduced bilingual disadvantages in older compared to younger adults, as frequency of use of words and structures accrues with age, potentially reaching the ceiling effects of activation observed for monolinguals (Gollan et al., 2008, 2011; Runnqvist et al., 2013). However, we observed no interactions between ageing and bilingualism, against the predictions of the weaker links account. Consistent with prior evidence, in our additional proficiency tasks, older adults were slower in reading comprehension but had better vocabulary scores than younger adults, across bilingualism groups. In addition, bilinguals scored lower in the vocabulary task, but this difference was only significant in the young group (see Appendices 2 and 3). However, these individual measures of proficiency did not affect comprehension and production, further suggesting that the effects assessed by our tasks were not mediated nor modulated by language competence.

On the other hand, accounts of bilingualism appealing to control processes (Green, 1998) do not make clear predictions about how bilingualism would interact with age. We might assume also a smaller bilingual disadvantage in older age, if mechanisms like inhibition ameliorate with age (Markiewicz et al., 2024; Veríssimo et al., 2022), but it is also conceivable that other cognitive resources involved in language, which decrease with age (e.g., working memory), would counteract any potential advantages (as suggested by Gollan et al., 2008).

Whereas the frequency lag hypothesis is appealing – namely, in providing a single explanation for both language dominance and bilingualism effects – it seems to be insufficient for a comprehensive account of the bilingual disadvantage, particularly in the face of the evidence for a bilingual advantage in executive control tasks (e.g., Bialystok et al., 2012), which could result from bilinguals negotiating cross-language competition (Kroll et al., 2008). Future research should examine how age and bilingualism effects in language processing are affected by frequency of use, language competition, and other measures of broader cognitive function.

Our study joins research that failed to find interactions between bilingualism and age in language processing (Gollan et al., 2008), and our findings suggest that bilingualism and ageing are two factors that independently impact on language processing. Nevertheless, given the lack of studies examining combined effects of bilingualism and ageing, further research is important for elucidating the underlying mechanisms involved, and for evaluating the possibilities for 'connecting models' of language processing in bilingualism and healthy older age (Rossi & Diaz, 2016).

Moreover, future studies could address some limitations of the research conducted so far. In the current study, we made every attempt to make the materials comparable across languages (English and Norwegian). However, this is always difficult to achieve, and ultimately there is always a potential confound of language.

Different studies have used varying tasks and stimuli to tap into language processes, which may be differently affected by both age and bilingualism. For example, whereas our comprehension experiment measured comprehension of target words in spoken sentences, it would also be important to investigate the interaction between age and bilingualism in written comprehension of complex sentences. Likewise, we accounted for effects of proficiency, but neither reading times nor vocabulary are catch-all measures of all domains that affect language proficiency, and so it is possible that the effects would be modulated by aspects of proficiency or language use that we did not measure.

Finally, bilinguals are a diverse population, and researchers have been encouraged to consider more in-depth differences in bilingual profile (e.g., Rothman et al., 2022). In the current study, we tested Norwegian–English bilinguals, who can be thought to be a relatively homogeneous population concerning language use and exposure, but it would be important to test other bilingual populations to determine how variability in bilingual profile can affect processing.

Supplementary Material. For supplementary material accompanying this paper, visit https://doi.org/10.1017/S1366728924000245

Acknowledgements. We would like to thank the research assistants who helped to conduct the data collection: Jack Feron, Ali Kalib Ali, Renate Gjetnes, Consuelo Vidal Gran, Amelie Grandjean, Nicolas Hayston Lecaros, Karethe Nilsen, Helene Øye, Øystein Søren Østerhus, and Ellinor Skjerli. We also thank Jan Zandhuis for technical support. This work was funded by the Norwegian Research Council (FRIPRO 300030).

Competing interests. The authors declare none.

Notes

[1] The second set of sentences was created for the purpose of re-testing participants, in another study, after a physical exercise intervention.

[2] Norwegian has two official written languages - Bokmål and Nynorsk. Bokmål was used in this study since the vast majority of writing is done in bokmål (85-90%): Language Council of Norway (https://www.sprakradet.no/Spraka-vare/Norsk/fakta-om-norsk/

[3] We used the default priors of *blmer* function in R, which are 'wishart' for covariance, and 'NULL' for both fixed effects and residual variance.

[4] The only random slope by item that did not make the model to fail to converge was the one for the fixed effect of proficiency in Experiment 1. However, the model including this random slope was not significantly better than the one excluding it (χ^2 (1) = 1.37, p = 0.24).

[5] Gollan et al. (2011) had a 'no-context' condition but it consisted of the processing of words in isolation. They found (Experiment 2) that monolinguals were only marginally faster than bilinguals, in their L1, and both groups were equally affected by the manipulation of target word frequency, suggesting that there is small or no difference in how monolinguals and bilinguals comprehend words in isolation.

[6] The creation of four sets served for testing bilinguals also in their L2, and for re-testing older participants after a physical exercise intervention, for the purpose of two different studies.

References

Abrams, L., & Farrell, M. T. (2011). Language processing in normal aging. In J. Guendouzi, F. Loncke, & M. J. Williams (Eds.), *The handbook of psycholinguistic and cognitive processes: Perspectives in communication disorders* (pp 49–73). Psychology Press. https://doiorg/104324/9780203848005ch3

Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning Memory and Cognition*, *33*, 791–810. https://doi.org/10.1037/0278-7393.33.4.791

Allum, P. H., & Wheeldon, L. R. (2009). Scope of lexical access in spoken sentence production: Implications for the conceptual-syntactic interface. *Journal of Experimental Psychology: Learning Memory and Cognition*, *35*, 1240–1255. https://doi.org/10.1037/a0016367

Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. https://doiorg/101016/jjml200712005

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. https://doi.org/10.1016/j.jml.2012.11.001

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. Journal of experimental psychology. *Learning, memory, and cognition*, *34*(4):859–73. 10.1037/0278-7393.34.4.859 Erratum in: J Exp Psychol Learn Mem Cogn. 2009 May;35(3):828. PMID: 18605874.

Bialystok, E., & Craik, F. I. (2022). How does bilingualism modify cognitive function? Attention to the mechanism. *Psychonomic Bulletin & Review*, *29*, 1246–1269. https://doiorg/103758/s13423-022-02057-5

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Science*, *16*(4):240–50. https://doi.org/10.1016/j.tics.2012.03.001

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer Online: http://wwwpraatorg/

Burke, D. M., MacKay, D. G., & James, L. E. (2000). Theoretical approaches to language and aging. In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging* (pp 204–237). Oxford University Press.

Caplan, D., Dede, G., Waters, G., Michaud, J., & Tripodis, Y. (2011). Effects of age, speed of processing, and working memory on comprehension of sentences with relative clauses. *Psychology and Aging*, *26*(2), 439–50. https://doi.org/10.1037/a0021837

Carson, N., Leach, L., & Murphy, K. J. (2018). A re-examination of Montreal Cognitive Assessment (MoCA) cutoff scores. *International Journal of Geriatric Psychiatry* *33*(2), 379–388. https://doiorg/101002/gps4756

Chung, Y., Rabe-Hesketh, S., Dorie, V., Gelman, A., & Liu, J. (2013). A Nondegenerate penalized likelihood estimator for variance parameters in multilevel models. *Psychometrika*, *78*, 685–709. https://doi.org/10.1007/s11336-013-9328-2

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3–42. https://doi.org/10.1017/S0142716406060024

Cohen, G., & Faulkner, D. (1983). Word recognition: Age differences in contextual facilitation effects. *British Journal of Psychology*, *74*(2), 239–251. https://doi.org/10.1111/j.2044-8295.1983.tb01860.x

Colomé, A. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language*, *45*, 721–736. https://doi.org/10.1006/jmla.2001.2793

Cooper, P. V. (1990). Discourse production and normal aging: Performance on oral picture description tasks. *Journal of Gerontology*, *45*(5), 210–214. https://doi.org/10.1093/geronj/45.5.P210

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283–1296. https://doi.org/10.1037/0278-7393.26.5.1283

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik, & S. Trehub (Eds.), *Aging and Cognitive Processes* (pp 191–211). New York. NY: Plenum.

Dijkstra, T., & Van Heuven, W. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition* *5*(3), 175–197. doi:101017/S1366728902003012

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816. doi:101080/1747021820171310261

Engedal, K., Gjøra, L., Benth, J. Š., Wagle, J., Rønqvist, T. K., & Selbæk, G. (2022). The Montreal Cognitive Assessment: Normative Data from a Large, Population-Based Sample of Cognitive Healthy Older Adults in Norway-The HUNT Study. *Journal of Alzheimer's disease*, *86*(2):589–599. https://doi.org/10.3233/jad-215442

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information processing rate and amount: Implications for group difference in response latency. *Psychological Bulletin*, *125*(6), 777–799. https://doi.org/10.1037/0033-2909.125.6.777

Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, *42*(2), 133–141. https://doiorg/101111/j1469-8986200500274x

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, *39*, 133–146. https://doi.org/10.1111/1469-8986.3920133

Federmeier, K. D., van Petten, C., Schwartz, T. J., & Kutas, M. (2003). Sounds, words, sentences: Age-related changes across levels of language processing. *Psychology and Aging*, *18*, 858–872. https://doi.org/10.1037/0882-7974.18.4.858

Gilbert, A. C., Cousineau-Perusse, M., & Titone, D. (2020). L2 exposure modulates the scope of planning during first and second language production. *Bilingualism: Language and Cognition*, *23*, 1093–1105. https://doi.org/10.1017/S1366728920000115

Gollan, T. H., Ferreira, V. S., Cera, C., & Flett, S. (2014). *Translation-priming effects on tip-of-the-tongue states, Language, Cognition and Neuroscience*, *29*:3, 274–288. https://psycnet.apa.org/doi/10.1080/01690965.2012.762457

Gollan, T. H., & Goldrick, M. (2019). Aging deficits in naturalistic speech production and monitoring revealed through reading aloud. *Psychology and Aging*, *34*(1), 25–42. https://doi.org/10.1037/pag0000296

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always a means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory and Cognition*, *33*, 1220–1234. https://doiorg/103758/BF03193224

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). *Frequency drives lexical access in reading but not in speaking: the frequency-lag hypothesis Journal of Experimental Psychology: General*, *140*(2):186–209. https://doi.org/10.1037/a0022256

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

Hanulová, J., Davidson, D. J., & Indefrey, P. (2011). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Language and Cognitive Processes*, *26*:7, 902–934. https://doi.org/10.1080/01690965.2010.509946

Hardy, S., Segaert, K., & Wheeldon, L. (2020). Healthy aging and sentence production: Disrupted lexical access in the context of intact syntactic planning. *Frontiers in Psychology*, *11*, 257. https://doi.org/10.3389/fpsyg.2020.00257

Herbay, A. C., Gonnerman, L. M., & Shari, R. B. (2018). How Do French–English Bilinguals Pull Verb Particle Constructions Off? Factors Influencing Second Language Processing of Unfamiliar Structures at the Syntax-Semantics Interface. *Frontiers in Psychology*, *9*, 1–16. https://doi.org/10.3389/fpsyg.2018.01885

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2):277–88. https://doi.org/10.1016/j.actpsy.2007.06.003

Jaeger, T. F. (2010). Redundancy and reduction: speakers manage syntactic information density. *Cognitive Psychology*, *61*(1), 23–62. https://doi.org/10.1016/j.cogpsych.2010.02.002

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, *11*(2), 201–258. https://doi.org/10.1207/s15516709cog1102_5

Kemper, S. (1992). Adults' sentence fragments: Who, what, when, where, and why. *Communication Research*, *19*(4), 444–458. https://doi.org/10.1177/009365092019004003

Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, *73*, 1–40. https://doi.org/10.1016/j.cogpsych.2014.04.001

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychologica*, *128*(3), 416–30. https://doi.org/10.1016/j.actpsy.2008.02.001

Kroll, J. F., & Gollan, T. H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In J. Goldrick, V. Ferreira, & M. Miozzo (Eds..), *The Oxford handbook of language production* (pp. 165–181). Oxford University Press.

Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychology and Aging*, *8*, 34–43.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.

Levelt, W. J. M., & Maassen, B. (1981). Lexical search and order of mention in sentence production. In W. Klein, & W. J. M. Levelt, (Eds.), *Crossing the Boundaries in Linguistics* (pp 221–252). Reidel, Dortrecht.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition 106*(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006

Li, C., Ferreira, V. S., & Gollan, T. H. (2022). Language control after phrasal planning: Playing Whack-a-mole with language switch costs. *Journal of Memory and Language*, *126*, 1–15. https://psycnet.apa.org/doi/10.1016/j.jml.2022.104338

MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language 32*(5), 692–715. https://doi.org/10.1006/jmla.1993.1035

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doiorg/101044/1092-4388(2007/067)

Markiewicz, R., Rahman, F., Fernandes, E. G., Limachya, R., Wetterlin, A., Wheeldon, L., & Segaert, K. (2024). Effects of healthy ageing and bilingualism on attention networks. *Bilingualism: Language and Cognition*, 1–14. https://doi.org/10.1017/S1366728924000154

Marslen-Wilson, W. (1975). Sentence Perception as an Interactive Parallel Process. *Science*, *189*, 226–228. https://doi.org/10.1126/science.189.4198.226

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588. https://doiorg/101016/jjml201308001

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699 https://doiorg/101111/j1532-5415200553221x

Opdebeeck, C., Martyr, A., & Clare, L. (2016). Cognitive reserve and cognitive function in healthy older people: a meta-analysis. *Aging, Neuropsychology, and Cognition*, *23*(1), 40–60. https://doi.org/10.1080/13825585.2015.1041450

Payne, B. R., Grison, S., Gao, X., Christianson, K., Morrow, D. G., & Stine-Morrow, E. A. L. (2014). Aging and individual differences in binding during sentence understanding: Evidence from temporary and global syntactic attachment ambiguities. *Cognition*, *130*, 157–173. http://dxdoiorg/101016/jcognition201310005

Peelle, J. E. (2018). Language and aging. In G. I. de Zubicaray, & N. O. Schiller (Eds.), *The Oxford Handbook of Neurolinguistics* (pp 295–316). Oxford University Press.

Poulisse, C., Wheeldon, L. R., & Segaert, K. (2019). Evidence against preserved syntactic comprehension in healthy aging. *Journal of Experimental Psychology Learning, Memory and Cognition*, *45*(12), 2290–2308. https://doi.org/10.1037/xlm0000707

Pozzan, L., & Trueswell, J. C. (2016). Second language processing and revision of garden-path sentences: a visual word study A visual word study. *Bilingualism: Language and Cognition 19*(3), 636–643. https://doi.org/10.1017/s1366728915000838

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin and Review*, *3*, 504–509. https://doiorg/103758/BF03214555

Reifegerste, J. (2021). The effects of aging on bilingual language: What changes, what doesn't, and why. *Bilingualism: Language and Cognition*, *24*(1), 1–17. https://doiorg/101017/S1366728920000413

Rossi, E., & Diaz, M. T. (2016). How aging and bilingualism influence language processing: theoretical and neural models. *Linguistic Approaches to Bilingualism*, *6*(1-2), 9–42. https://doi.org/10.1075/lab.14029.ros

Rothman, J., Bayram, F., DeLuca, V., Di Pisa, G., Duñabeitia, J. A., Gharibi, K., Hao, J., Kolb, N., Kubota, M., Kupisch, T., Laméris, T., Luque, A., van Osch, B., Pereira Soares, S. M., Prystauka, Y., Tat, D., Tomić, A., Voits, T., & Wulff, S. (2022). Monolingual comparative normativity in bilingualism research is out of "control": Arguments and alternatives. *Applied Psycholinguistics*, *44*(3):316–329. doi:10.1017/S0142716422000315

Runnqvist, E., Gollan, T. H., Costa, A., & Ferreira, V. S. (2013). A disadvantage in bilingual sentence production modulated by syntactic frequency and similarity across languages. *Cognition*, *129*(2):256–63. https://doi.org/10.1016/j.cognition.2013.07.008

Sadat, J., Martin, C. D., Alario, F. X., & Costa, A. (2012). Characterizing the bilingual disadvantage in noun phrase production. *Journal of Psycholinguist Research*, *41*(3), 159–179. https://doi.org/10.1007/s10936-011-9183-1

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, *55*(2), 197–212. https://doiorg/101016/jjml200603004

Shook, A., Goldrick, M., Engstler, C., & Marian, V. (2015). Bilinguals show weaker lexical access during spoken sentence comprehension. *Journal of Psycholinguistic Research*, *44*(6), 789–802. https://doiorg/101007/s10936-014-9322-6

Smith, M., & Wheeldon, L. R. (1999). High level processing scope in spoken sentence production. *Cognition*, *73*(3), 205–246. https://doi.org/10.1016/s0010-0277(99)00053-0

Spieler, D. H., & Griffin, Z. M. (2006). The influence of age on the time course of word preparation in multiword utterances. *Language and Cognitive Processes*, *21*, 291(1-3)-321. https://doi.org/10.1080/01690960400002133

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. https://doi.org/10.1111/1467-9280.00151

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634. https://doi.org/10.1126/science.7777863

Tyler, L. K., Shafto, M. A., Randall, B., Wright, P., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Preserving syntactic processing across the adult life span: the modulation of the frontotemporal language system in the context of age-related atrophy. *Cerebral Cortex*, *20*(2), 352–64. https://doi.org/10.1093/cercor/bhp105

Ullman, M. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, *4*(2), 105–122. https://doi.org/10.1017/S1366728901000220

van de Velde, M., & Meyer, A. S. (2014). Syntactic flexibility and planning scope: the effect of verb bias on advance planning during sentence recall. *Frontiers in Psychology*, *5*, 1174. https://doi.org/10.3389/fpsyg.2014.01174

Veríssimo, J., Verhaeghen, P., Goldman, N., Weinstein, M., & Ullman, M. T. (2022). Evidence that ageing yields improvements as well as declines across attention and executive functions. *Nature Human Behaviour*, *6*(1), 97–110. https://doi.org/10.1038/s41562-021-01169-7

Völter, C., Fricke, H., Faour, S., Lueg, G., Nasreddine, Z. S., Götze, L., & Dawes, P. (2023). Validation of the German Montreal-Cognitive-Assessment-H for hearing-impaired. *Frontiers in Aging Neuroscience*, *15*, Article 1209385. https://psycnet.apa.org/doi/10.3389/fnagi.2023.1209385

Waters, G. S., & Caplan, D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging*, *16*, 128–144. http://dxdoiorg/101037/0882-7974161128

Weber-Fox, C., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, *8*(3), 231–256. https://doiorg/101162/jocn199683231

Wheeldon, L., & Konopka, A. (2023). *Grammatical Encoding for Speech Production (Elements in Psycholinguistics)*. Cambridge University Press. https://doi.org/10.1017/9781009264518