Healthy Aging and Sentence Production:

Disrupted Lexical Access in the Context of Intact Syntactic Planning

Sophie M. Hardy

Katrien Segaert

Linda Wheeldon

University of Birmingham, UK

Univeristy of Agder, Norway

Author Note

Sophie M. Hardy, Katrien Segaert, School of Psychology, University of Birmingham; Katrien Segaert, Centre for Human Brain Heath, University of Birmingham; Linda Wheeldon, Department of Foreign Languages and Translation, University of Agder.

This research was partly supported by an Economic and Social Research Council (ESRC) studentship awarded to Sophie M. Hardy from the University of Birmingham Doctoral Training Centre (grant number: ES/J50001X/1). Sections of this study were presented at the Architectures and Mechanisms of Language Processing (AMLaP; Lancaster, September 2017) and the Experimental Psychology Society (EPS) meeting (Leicester, April 2018). A preprint of the manuscript was made available on the *bioRxiv* server from May 2018. We gratefully acknowledge the help of Denise Clissett, the coordinator of the Patient and Lifespan Cognition Database.

Correspondence concerning this article should be addressed to Sophie M. Hardy, School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK. Email: smh649@bham.ac.uk.

Abstract

Healthy aging does not affect all features of language processing equally. In this study, we investigated the effects of aging on different processes involved in fluent sentence production, a complex task that requires the successful execution and coordination of multiple processes. In Experiment 1, we investigated age-related effects on the speed of syntax selection using a syntactic priming paradigm. Both young and older adults produced target sentences quicker following syntactically related primes compared to unrelated primes, indicating that syntactic facilitation effects are preserved with age. In Experiment 2, we investigated age-related effects in syntactic planning and lexical retrieval using a planning scope paradigm: participants described moving picture displays designed to elicit sentences with either initial coordinate or simple noun phrases and, on half of the trials, the second picture was previewed. Without preview, both age groups were slower to initiate sentences with larger coordinate phrases, suggesting a similar phrasal planning scope. However, age-related differences did emerge relating to the preview manipulation: while young adults displayed speed benefits of preview in both phrase conditions, older adults only displayed speed preview benefits within the initial phrase (coordinate condition). Moreover, preview outside the initial phrase (simple condition) caused older adults to become significantly more error-prone. Thus, while syntactic planning scope appears unaffected by aging, older adults do appear to encounter problems with managing the activation and integration of lexical items into syntactic structures. Taken together, our findings indicate that healthy aging disrupts the lexical, but not the syntactic, processes involved in sentence production.

Keywords: healthy aging, sentence production, priming, syntactic planning, lexical retrieval.

Word count:9,995

Healthy Aging and Sentence Production:

Disrupted Lexical Access in the Context of Intact Syntactic Planning

Producing a fluent and coherent sentence is a complex task involving the coordination of multiple cognitive and neural mechanisms (Levelt, 1989; Mody, 2017). As we age, changes occur that can create challenges for language processing, such as a widespread reduction in grey matter volume (Good et al., 2001) and a decline in working memory capacity (Waters & Caplan, 2003). Nevertheless, older adults have a wealth of experience with language and are often able to adopt effective processing strategies, such as the recruitment of addition brain areas, to compensate for lost efficiency elsewhere (Reuter-Lorenz & Park, 2014; Wingfield & Grossman, 2006). This paints a multifactorial picture of language processing in old age in which some language skills decline because of age-related cognitive changes, but in which others are preserved because of the successful adoption of compensation strategies (for reviews, Burke & Shafto, 2008; Peelle, 2019). Investigating how different aspects of language processing are affected by old age is critical for understanding this complex balance between decline and preservation. In this study, we conducted two novel experiments investigating age-related changes in sentence production; specifically, we investigated the processes involved in syntax generation (Experiment 1), as well as sentence planning and lexical retrieval (Experiment 2). Our findings reveal a contrast between the preservation of syntactic skills, but the disruption of lexical access, in old age; this adds to the growing evidence that healthy aging does not affect all features of language processing to the same extent.

A number of previous studies have demonstrated age-related decline in language production. To first consider age-related changes at the word level, several studies have found older adults to be slower and more error-prone in picture naming tasks, particularly for low-frequency words (see Feyereisen, 1997, for a review), and to experience increased tip-of-the-tongue states, in which a speaker is certain that they know a word but is unable to produce it (Burke, MacKay, Worthley, & Wade, 1991; Segaert et al., 2018; Shafto, Burke, Stamatakis, Tam, & Tyler, 2007). This suggests an increased difficulty in retrieving the name of a lexical object and its corresponding phonological form, something which may be attributable to age-related atrophy in the left insula (Shafto et al., 2007). Age-related deficits are also found at the sentence level of production: with age, there is a decline in the production of complex syntactic structures, such as embedded clauses, coupled with an increase in syntactic errors, such as the use of the incorrect tense (Kemper, 1987; Kemper, Greiner, Marquis, Prenovost, & Mitzner, 2001; Kemper, Herman, & Liu, 2004; Kemper & Sumner, 2001; Rabaglia & Salthouse, 2011). This apparent decline in syntax production is often considered to arise from age-related decreases in the capacity or efficiency of working memory, a cognitive resource that is critical when producing complex sentences that contain multiple clauses and that require greater syntactic operations of movement (Abrams & Farrell, 2011; Kemper & Sumner, 2001; MacDonald & Christiansen, 2002).

In contrast, other aspects of language production are characterized by stability and even improvement with age. Most notably, vocabulary size and knowledge consistently increase with age (Verhaeghen, 2003). Older adults also appear to perform similarly to young adults in tasks where they must switch between formulating alternative syntactic structures, such as dative verb and transitive verb alternatives (Altmann & Kemper, 2006; Davidson, Zacks, & Ferreira, 2003). Moreover, in situations in which the task demands are reduced, minimal age differences are found; for example, Kemper, Herman, and Lian (2003) found that young and older adults produced similar responses when asked to incorporate intransitive (‘smiled’) or transitive (‘replaced’) verbs into their sentences, and age differences in fluency only emerged when participants were asked to incorporate more complex complement-taking verbs (‘expected’). This effect of task complexity on language production skills in old age can be best explained by Peelle's (2019) ‘supply and demand’ framework, which suggests that behavioral success reflects a complex balance between specific task requirements and the level of cognitive resources available to the speaker; specifically, if task requirements outweigh cognitive resources, processing efficiency will decline leading to poor performance. Due to overall neuroanatomical and cognitive changes that occur during healthy aging, it is no surprise that older adults’ neurocognitive capacity for any given language task is likely to be less than young adults. However, this does not necessarily mean that age differences will always emerge: older adults may still perform similarly to young adults when task requirements are sufficiently low (e.g., when producing simpler syntactic constructions) or they may adopt compensatory processing strategies (e.g., the recruitment of other brain areas). In this way, identical behavioral performance in young and older adults may not always reflect identical neural or cognitive processes.

The idea of neural compensation in aging has been most studied in terms of language comprehension, in which brain imaging studies have demonstrated that older adults engage additional brain areas in order to maintain high levels of accuracy (see Wingfield & Grossman, 2006, for a review). Likewise, older adults may employ different strategic approaches in order to compensate for processing deficits elsewhere, such as a greater reliance on discourse during reading (Stine-Morrow, Miller, Gagne, & Hertzog, 2008). These same principles of compensation can also be applied to production; for example, Altmann and Kemper (2006) suggested that the minimal age group differences they observed in their sentence generation task were the result of older adults adopting a different strategy to young adults (always assigned the initially presented item to the subject role). Overall, this highlights the importance of continuing to investigate the effect of aging on different aspects of language processing. Moreover, even when there appear to be no group differences, this does not necessarily mean that young and older adults are engaging the exact same processing networks.

The aim of our study was to investigate how the syntactic and lexical processes involved in sentence generation are affected by healthy aging, using paradigms that have not previously been used with older adults. In both experiments we employed on-line onset latency measures of sentence production in order to gain information about the incremental fashion in which sentences are planned and produced (i.e., speakers typically only plan a portion of a sentence before beginning speaking and plan the rest during articulation; Wheeldon, 2013). Most previous studies investigating sentence production and aging have predominantly used off-line measures, involving the assessment and coding of sentences after they have been produced (e.g., Kemper et al., 2001, 2003, 2004; Rabaglia & Salthouse, 2011), which while informative about syntactic choices and errors cannot provide insight into the time-course of the underlying sentence generation process (Marinis, 2010; Mertins, 2016). To our knowledge, only a handful of studies to date have investigated older adults' sentence production using on-line measures (Griffin & Spieler, 2006; Spieler & Griffin, 2006); hence, there remains a considerable gap in the aging literature regarding the timing of speech preparation and how different syntactic and lexical processes unfold during the course of sentence production.[1]

In Experiment 1, we used a syntactic priming paradigm (as in Smith & Wheeldon, 2001; Wheeldon & Smith, 2003) to investigate age-related differences in the speed of syntax generation. In Experiment 2, we used a planning scope paradigm with an embedded picture preview element (as in Smith & Wheeldon, 1999; Wheeldon, Ohlson, Ashby, & Gator, 2013) to investigate age-related differences in syntactic planning scope and lexical retrieval. All participants also completed a battery of eight additional measures designed to provide an indicator of their current cognitive and physical ability (see Supplementary Materials). We included these measures because there is widespread variation in the type and extent of cognitive and neuroanatomical changes that occur with healthy aging (Salthouse, 2012, 2016; Ziegler, Dahnke, & Gaser, 2012). This increase in heterogeneity with age means that all older adults are unlikely to adopt the same compensation strategies or perform exactly same way in any given language production task (Peelle, 2019). Investigating what kind of inter-individual variability accounts for individual differences in language performance is therefore an important consideration when aiming to fully understand the complex balance between decline and preservation in old age.

Experiment 1:

Examining the Effect of Aging on Latency Measures of Syntax Facilitation

The process of producing a sentence begins with the preparation of a non-linear preverbal message – this is a conceptual representation of the information that the speaker wishes to convey, which includes information about the selected concepts and thematic structure within the discourse context (Levelt, 1989). The preverbal message triggers the formulation stage in which the message is turned into linguistic representations, involving both the rapid retrieval of lexical items and the generation of an appropriate syntactic structure, which must be integrated correctly to convey the intended message. More traditional models of sentence production propose that grammatical encoding is lexically driven such that lemmas (representations of the syntactic and semantic properties of a word) are first selected and assigned grammatical roles (e.g., subject or object), which then drives the generation of a syntactic structure (Bock & Levelt, 1994; Levelt, Roelofs, & Meyer, 1999; Pickering & Branigan, 1998). Alternatively, computational models postulate that there is a complete dissociation between syntax generation and lexical retrieval, such that syntactic structure is derived solely from conceptual structure (i.e., thematic roles) with lexical access occurring independently (Chang, Dell, & Bock, 2006; Chang, Dell, Bock, & Griffin, 2000).

While there remains debate about the exact relationship between syntax generation and lexical retrieval (see Wheeldon, 2011, for a review of the evidence for both lexically mediated and lexically independent models), it is widely agreed that sentence production occurs incrementally, such that only a small amount of planning occurs prior to articulation and that planning continues to unfold after speech onset for the remainder of the sentence (Levelt, 1989, 1992). Consequently, the amount of time that a speaker takes to begin a sentence is informative about the amount of planning that has occurred prior to speech onset in terms of both the retrieval of lexical items and the generation of syntax (Levelt, 1989; Wheeldon, 2013). On-line onset latency measures can therefore be used to explore age-related differences in the type and amount of advanced planning, or scope, of the sentence generation process.

One paradigm that has been used to explore the processes involved in syntax generation is *syntactic priming*. Broadly speaking, syntactic priming refers to the facilitation of syntactic processing that occurs when a syntactic structure is repeated across an otherwise unrelated prime and target (Bock, 1986; Pickering & Ferreira, 2008). *Choice syntactic priming* is the phenomenon whereby speakers are more likely to repeat a syntactic structure that they have recently processed (see Mahowald, James, Futrell, & Gibson, 2016, for a meta-analytical review). In our study investigating the speed of syntax generation, we were interested in *onset latency syntactic priming:* the facilitated speed of syntactic processing that occurs when a syntactic structure is repeated across a prime and target (Corley & Scheepers, 2002; Segaert, Menenti, Weber, & Hagoort, 2011; Segaert, Weber, Cladder-Micus, & Hagoort, 2014; Segaert, Wheeldon, & Hagoort, 2016; Wheeldon & Smith, 2003). For example, using a picture description task, Smith and Wheeldon (2001) demonstrated that when a speaker must produce a given syntactic structure on a target trial (1a), this was initiated quicker (i.e., decreased speech onset latencies) following recent production of the same structure (1b), compared to when a different structure had just been produced (1c).

(1a) Target: "the spoon and the car move up"

(1b) Related prime: "the eye and the fish move apart"

(1c) Unrelated prime: "the eye moves up and the fish moves down"

This latency priming effect cannot have its source in conceptualization, lexical access or phonological planning as these factors were tightly controled within the experimental design (i.e., there was no prosodic, visual or lexical similarity between any of the corresponding primes and targets). Further experiments by Smith and Wheeldon (2001) also ruled out alternative explanations relating to overall sentence complexity (the effect persists when both the related and unrelated prime feature the same number of clauses as the target), as well as to visual perception and picture movement (the effect persists when the related and unrelated primes feature the exact same movement patterns, and when stationary written prime sentences are used). This indicates that the facilitation effect observed is specifically related to the repetition of syntactic structure between the prime and target.

The two most common theoretical accounts of structural priming relate to the residual activation of a prime syntactic structure (Pickering & Branigan, 1998) and implicit learning processes that occur when an unexpected prime is heard (Chang et al., 2006). However, these models only provide explanations of facilitation effects relating to syntactic choices and not to the speed of sentence production; thus, the models offer minimal insight into the mechanisms that underlie onset latency syntactic priming. By contrast, Segaert et al. (2016) proposed a two-stage competition model that explains the effect of syntactic priming on both choices and onset latencies (see also Segaert et al., 2011, 2014). According to the model, alternative syntactic structures (e.g., active vs. passive) are represented by competing nodes, with activation levels determined by the relative frequency of the structure. Sentence production begins with construction of the preverbal message and this is followed by two sequential stages. First is the selection stage during which a speaker selects one syntactic structure from competing alternatives. Next follows the planning stage during which the selected syntax is incrementally planned and produced. While syntactic choice is determined solely at the selection stage, production speed is determined by the additive time taken to complete both stages. Consequently, when the choice element is removed (as in Smith & Wheeldon, 2001), onset latencies are largely determined by processing at the planning stage with very minimal processing required at the selection stage as there are no competing syntactic alternatives. In this study, we therefore investigated age-related effects on onset latency syntactic priming without an additional choice element as this allowed us to tap more directly into the processes involved in sentence planning.

The magnitude of the onset latency syntactic priming effects observed in the older adults will be informative about age-related changes in syntactic planning and facilitation that occur during real-time sentence production. While no studies to date have examined age-related effects on onset latency priming, a few studies have investigated age effects on choice syntactic priming. However, this has produced mixed results with two studies finding preserved priming of passives in older adults (Hardy, Messenger, & Maylor, 2017; Heyselaar, Wheeldon, & Segaert, 2018), while others have not (Heyselaar, Segaert, Walvoort, Kessels, & Hagoort, 2017, footnote 2; Sung, 2015).[2] It is therefore difficult to make direct hypotheses about age-related effects on onset latency syntactic priming based on previous evidence. Nevertheless, hypotheses can be made by considering the two-stage competition model in combination with more general models of aging. The model of Segaert et al. (2016) includes an spreading activation architecture whereby recently processed syntactic structures are activated to an above-baseline level, which contributes to decreased selection and planning speed. However, according to Salthouse's (1996) general slowing model of aging, declines in overall processing speed with age can substantially decrease the speed of spreading activation throughout a cognitive or neural network. Similarly, the transmission deficit model postulates that aging weakens the strength of activation of different units and the connections amongst units, both critical for successful spreading activation (MacKay & Burke, 1990). Applied to syntactic priming, this may mean that when older adults' process a prime sentence, the syntactic information relating to the prime does not become available to a central processor quickly or strongly enough to sufficiently excite the representation of the syntactic structure to a level which may influence the speed of syntactic selection and planning. If this is the case, we might expect to observe greater onset latency priming effect in young adults (who possess a faster spreading activation network) compared to older adults (who generally display much slower processing speed; Salthouse, 2004).

Experiment 1: Method

Participants. We recruited 50 young adults (36 female) aged 18-25 (*M* = 19.8, *SD* = 1.1) from the University of Birmingham student population and 56 older adults (37 female) aged 64-80 (*M* = 71.8, *SD* = 4.5) from the Patient and Lifespan Cognition Database. All older adults scored above 26 out of 30 (*M* = 27.4; *SD* = 1.3) on the Montreal Cognitive Assessment (Nasreddine et al., 2005), indicating that they were currently experiencing healthy aging (scores < 26 indicate risk of mild cognitive impairment or dementia). All participants were native English speakers with normal or corrected-to-normal vision, and did not report any language disorders. There was no significant difference in education between age groups.[3] The study was approved by the University of Birmingham Ethical Review Committee and participants provided written informed consent. All participants completed Experiment 1 at the initial test session, followed by Experiment 2 3-7 days later.

Design. We used a 2 X 2 mixed design with one between-participant variable of age (young vs. old) and one within-participant variable of prime type (syntactically related vs. syntactically unrelated). Hence, there were two experimental conditions (Figure 1A).

Materials. To create the experimental items, we used 80 simple photographic pictures of everyday concrete objects. All picture names were mono- or disyllabic, and care was taken to ensure that the objects could be identified and named quickly and easily. Forty of the pictures were used to create the 40 picture pairs for the target trials; each picture appeared in two different pairs (once each in the left and right position). Using the same constraints, we constructed 40 picture pairs from another 40 pictures for the prime trials. We then paired each target pair with a prime pair to generate 40 experimental items. We ensured that there was no phonological or conceptual overlap between any of the four pictures within each experimental item; we did this to ensure that any effects we observed were related to syntactic processing, and not to semantic or pragmatic features.

The movement of each picture pair was controled using E-Prime (Schneider, Eschman, & Zuccolotto, 2002). In all target trials, both pictures moved in the same vertical direction (either up or down). Participants were instructed to describe the picture movements from left to right using specific sentences; hence, the target trials elicited a coordinate noun phrase (*"the A and the B move up/down"*). In the related prime condition, the pictures moved in opposing horizontal directions which elicited a sentence that was syntactically related to the target trials (*"the C and the D move together/apart"*). In the unrelated prime condition, the pictures moved in opposing vertical directions which elicited a sentence that was syntactically unrelated to the target trials (*"the C moves up/down and the D moves down/up"*). We then created two item lists that each contained the same 40 target sentences, but the prime condition matched to each target was rotated such that there were 20 related and 20 unrelated primes per list. Each participant was randomly assigned to one of the two lists and completed 20 experimental items (prime plus target pairs) from each condition (Table 1A).

Lastly, we used a further 54 pictures to construct 120 filler trials designed to increase the variety of syntactic structures produced by the participant and minimize the risk of them noticing the priming manipulation. We created 96 filler trials that elicited phrases such as: *"there is an X and a Y"* (no picture movement); *"the Xs move up"* (two repeat pictures move simultaneously) and *"there are no pictures"* (screen is blank). We also created 24 filler trials that elicited phrases that were syntactically similar to the experimental trials; without such 'decoy' fillers, experimental trials would always occur in pairs (i.e., prime and corresponding target) which may enable the participant to predict the upcoming movement of a target trial. All 120 fillers were added to each of the two items lists. We then divided each list into four blocks that each contained 5 related experimental items, 5 unrelated experimental items and 30 filler items. The distribution of items within each block was pseudorandomized with the constraint that two experimental items never occurred consecutively. The ordering of the blocks was rotated across participants.

Procedure. Each participant was tested individually in a sound-attenuating booth facing the screen of a 17 inch *Dell* monitor, in front of which was a *Sony* microphone connected to an amplitude voice key that recorded his/her responses and onset latencies. Figure 1B illustrates the sequence of stimuli presentation per trial. To begin, there were 50 practice trials; the sentences elicited resembled those in the experimental and filler trials and featured all 80 experimental pictures once. The task then continued until all four blocks had been completed. The experimenter listened from outside the booth via headphones and noted down any errors made by the participant. Errors included: incorrect picture naming (e.g., 'fish' instead of 'shark'); use of a difference sentence structure (e.g., *"the pig moves towards* *the leaf"* instead of *"the pig and leaf move together"*); and disfluencies, such as stuttering and pausing.

Data Preparation and Analyses. We excluded the data of participants whose error rates were above 50% on the experimental trials; this resulted in exclusion of five older adults. Of the 4040 target responses, we excluded trials in which the participant made an error on the corresponding prime (170 (8.5%) of young and 301 (14.7%) of older adult trials), and for which the target onset latency was below 300ms, above 3000ms or more than 2.5SD above/below the participants' mean per experimental condition (discarding 53 (2.9%) young and 49 (2.8%) older adult trials). All remaining trials were used in the error analyses, but only correct responses (87.4% of trials) were used in onset latency analyses.

All data were analysed in R (R Core Team, 2015) using mixed-effects models (*lme4* package; Bates, Mächler, Bolker, & Walker, 2014); this was the most suitable way to analyse the datasets as there were repeated observations for participants and items (Barr, Levy, Scheepers, & Tily, 2013; Jaeger, 2008). We fitted a logit mixed-effects model to the error data as the dependent variable was categorical (correct = 0, incorrect = 1), and a linear mixed-effects model to the onset latency data as the dependent variable was continuous. We used a maximal random effects structure as this allowed us to include per-participant and per-item adjustments to the fixed intercepts (random intercepts) with additional random adjustments to the fixed effects (random slopes). We entered age group (young vs. old) and prime type (related vs. unrelated) as fixed effects. We also included random intercepts for participants and items, as well as by-participant random slopes for within-participant fixed effects and by-item random slopes for within-item fixed effects. Prior to analysis, the fixed effects were sum-coded and transformed to have a mean of 0 and a range of 1. When a model did not converge with the maximal random effects structure, we simplified the random slopes, removing interactions before main effects in order of least variance explained until the model converged (Barr et al., 2013). Significance *p* values for the linear mixed-effects model were calculated using the *car* package (Fox & Weisberg, 2011).

Additional Measurements. See Supplementary Materials for extensive details about the measurements and the analysis procedure.

Table 1

Overview of the Different Items Used in the Experiments 1 and 2. Number of Items and Example Stimuli Completed by Each Participant are Provided.

Item Type	N	Example
A: Experiment 1		
Related	20	Prime: "the pencil and the orange move together" Target: "the clock and the drum move up"
Unrelated	20	Prime: "the cow moves up and the broom moves down" Target: "the apple and the goat move up"
Filler	120	"There are two houses"
B: Experiment 2		
Preview Initial Coordinate	20	Preview: spoon "The trumpet and the spoon move above the crab"
No Preview Initial Coordinate	20	Preview: NA "The skirt and the bell move above the carrot"
Preview Initial Simple	20	Preview: snail "The balloon moves above the snail and the pear"
No Preview Initial Simple	20	Preview: NA "The spanner moves above the monkey and the toaster"
Filler	220	"There are three stars"

Note. The condition to which each experimental item was assigned was rotated across lists (e.g., the picture trio of trumpet-spoon-crab would also have appeared in the three other conditions in Experiment 2 in alternative lists). This meant that, across all participants, each item appeared an equal number of times in each condition; therefore, lexical factors of individual words, such as age of acquisition, were not a concern.

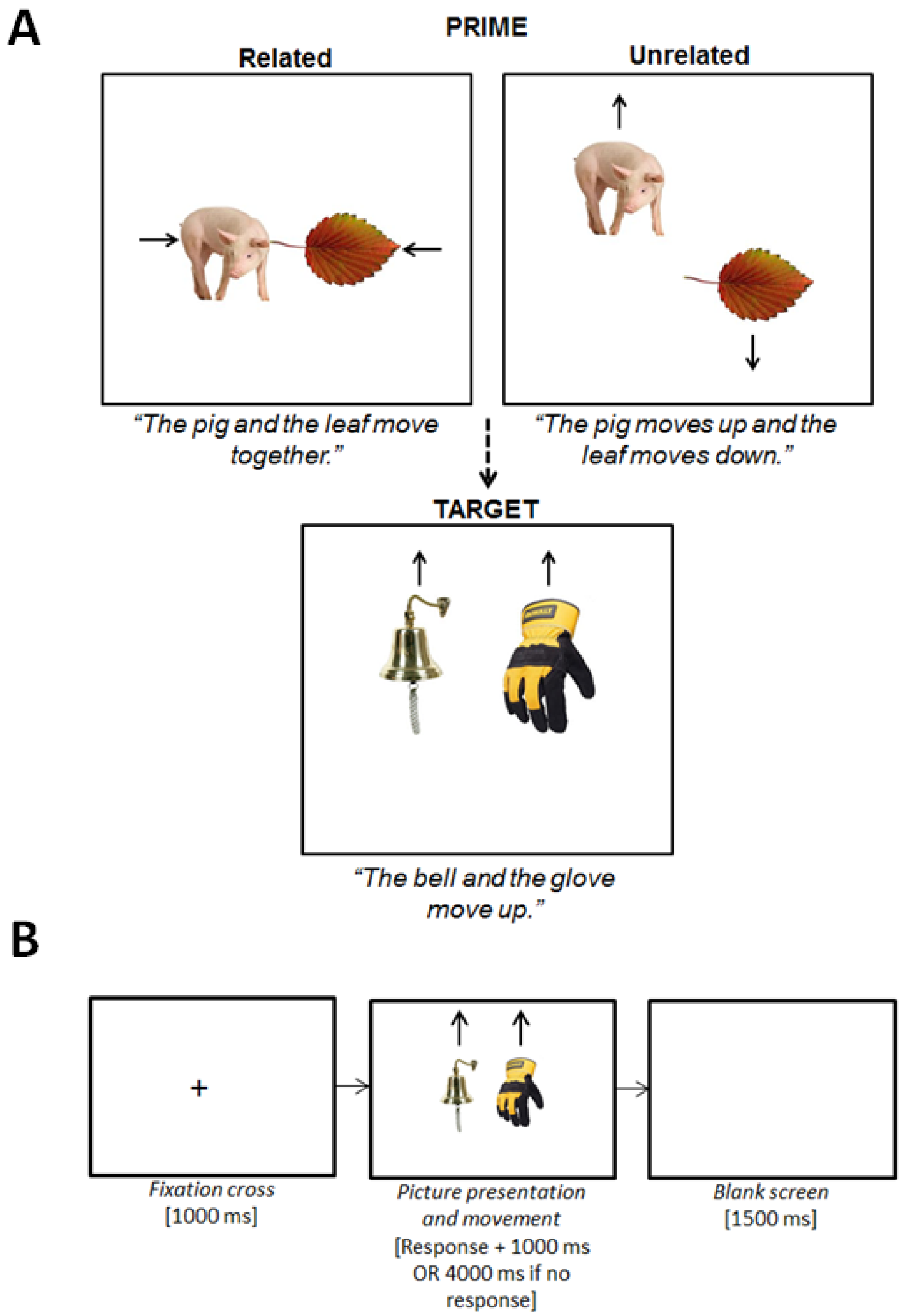

Figure 1. Experiment 1 syntactic priming task design (A) and stimuli presentation events per trial (B). The participant was instructed to begin describing the picture movement as soon as possible using specific sentence types. The stimuli presentation sequence was the same for prime and target trials, and primes were always immediately followed by the corresponding target (i.e., we used a 0-lag delay). Speech latencies on the target trials were recorded from the onset of the pictures to the participant beginning to speak.

Experiment 1: Results

Figure 2 summarizes the target error rates and onset latencies across the two prime conditions for young and older adults.

Error Rates. The best-fitting model of the error data is reported in Table 2A. Although older adults were significantly more error-prone than young adults (16.1% vs. 9.1%, $p < .001$), there was no main effect of prime type ($p = .369$) and no interaction between age group and prime type ($p = .868$), suggesting that neither young or older adults' production of errors on the target trials were affected by syntactic relatedness of the prime.

Onset Latencies. The best-fitting model of the onset latency data is reported in Table 2B. As expected, older adults were significantly slower than young adults (1060ms vs. 898ms, $p < .001$).[4] There was also a main effect of prime type ($p < .001$), such that target responses were produced significantly quicker following related primes (953ms) than following unrelated primes (994ms), indicating an overall syntactic priming effect of 41ms. Most interestingly, there was no interaction between age group and prime type ($p = .632$), indicating that the onset latency priming effect was similar for young (36ms, 3.9% benefit) and older (49ms, 4.5% benefit) adults.

Summary. The main findings of Experiment 1 are threefold: (1) older adults were slower and more error-prone when producing sentences compared to young adults; (2) our task produced a reliable latency priming effect on the production of target sentences; and (3) there was no age-related effect on the extent to which the speed of syntax generation benefited from repetition of syntactic structure. Together, this suggests that syntactic facilitation effects on onset latencies are preserved with age. In our additional analyses, we aimed to relate participants' scores on various measures of individual difference to their performance in Experiment 1 (see Supplementary Materials). However, we found the individual difference measures to have minimal influence on task performance; the one significant finding was that young adults with better inhibitory control were quicker to produce sentences overall.

Table 2

Summary of the Best-Fitted Mixed-Effects Models for the Experiment 1 Error and Onset Latency Data.

Predictor	Coefficient	SE	Wald Z	*p*
A: Error Data				
Intercept	2.34	0.16	14.69	< .001
Prime type	-0.14	0.15	-0.90	.369
Age group	0.76	0.20	3.74	< .001
Prime type * Age group	-0.05	0.28	-0.17	.868
B: Onset Latency Data				
Intercept	983.46	23.78	41.36	< .001
Prime type	39.65	9.98	3.97	< .001
Age group	-169.44	43.25	-3.92	< .001
Prime type * Age group	-9.28	19.41	-0.48	.632

Note. Both models converged with a fully-expressed random intercepts and slopes structure.

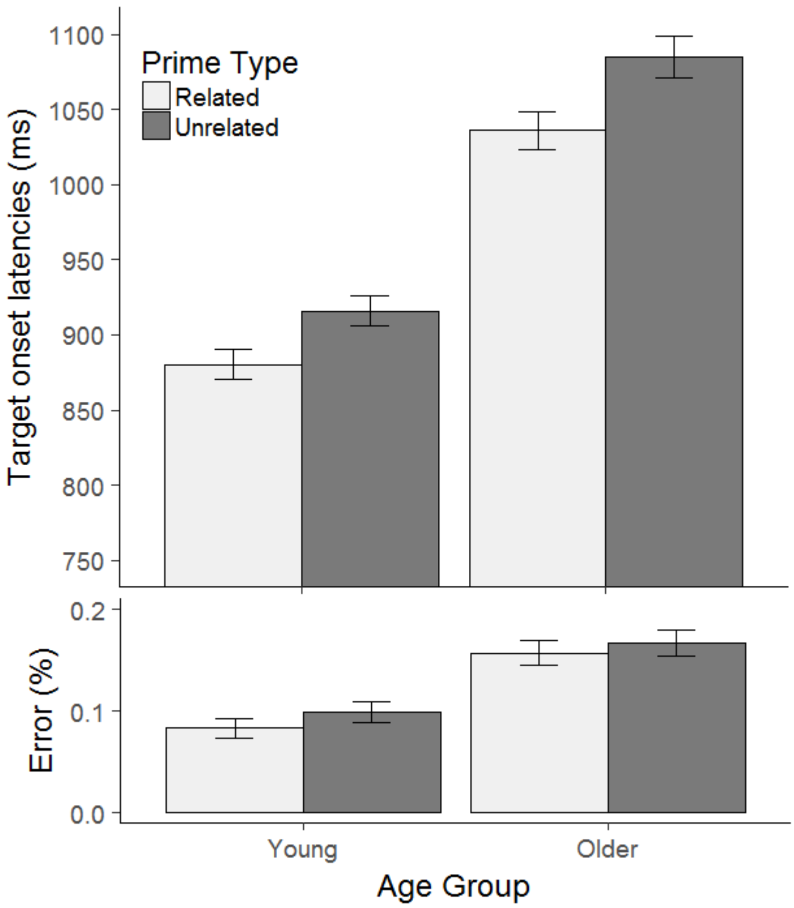

Figure 2. Experiment 1 mean target error rates and onset latencies for young and older adults following syntactically related and unrelated primes. Error bars denote ±1 the standard error of the mean.

Experiment 2:

Examining the Effect of Aging on On-line Planning Scope

In Experiment 1, we demonstrated that syntactic processing in both age groups was facilitated by the repetition of syntactic structure, which in turn benefited the speed of sentence production. This is specifically informative about age-related changes in the processes involved in syntactic facilitation at the planning level of sentence generation, as well as the mechanisms that underlie onset latency syntactic priming. In Experiment 2, we investigated older adults' sentence generation in unsupported situations in which sentence production is not primed and the speaker must generate a sentence entirely independently. Moreover, we employed a more complex sentence generation task in which participants produced sentences containing multiple phrases of varying length and complexity (this is in contrast to Experiment 1 where the target sentences all consisted of a single coordinate noun phrase). Within Experiment 2, we were therefore able to investigate age-related changes in incrementality in sentence production – the scope of sentence planning that occurs prior to articulation onset (Kempen & Hoenkamp, 1987; Levelt, 1989).

A number of studies have demonstrated that speakers do not plan all of what they wish to say before beginning speaking, but instead plan and produce a sentence incrementally in smaller word or phrasal units (see Wheeldon, 2013, for a review). An incremental system is beneficial as it allows for the rapid release of parts of the sentence as soon as planning is complete, reducing the demand for storage in working memory. Previous studies have shown that only a small amount of planning is required prior to speech onset, typically the first phrase (Martin, Crowther, Knight, Tamborello, & Yang, 2010; Martin, Yan, & Schnur, 2014; Smith & Wheeldon, 1999) or even as little as the first word (Griffin, 2001; Zhao & Yang, 2016). Moreover, incremental sentence production enables the processing load to be spread across multiple components and time, thereby further reducing demands on cognitive resources (Levelt, 1989; Wheeldon, 2013). One way to investigate the amount of planning that a speaker engages with prior to articulation is with the *planning scope* paradigm, in which picture displays are used to elicit sentences of different syntactic structures and speech onset latencies are used as an on-line measure of advanced planning. For example, Smith and Wheeldon (1999) found that participants took longer to initiate sentences with larger initial coordinate phrases (2a) compared to smaller initial simple phrases (2b), suggesting that planning scope occurs in phrasal units: when the first phrase is larger, speakers need longer to plan the syntax and retrieve the second lexical item before speech onset (see also Levelt & Maassen, 1981; Martin, Miller, & Vu, 2004; Wheeldon et al., 2013).

(2a) "[the dog and the hat move] above the fork"

(2b) "[the dog moves] above the hat and the fork"

Martin et al. (2010, 2014) ruled out an alternative explanation for this effect relating to the visual array (i.e., the grouping of objects moving together) as they found the same phrasal planning scope using stationary pictures arrays (e.g., *"the drum and the package are below the squirrel"*). Moreover, the phrasal planning effect cannot be attributed to the fact that, in English, the second content word in the simple initial phrase (always the verb 'moves'; 2b) may be easier to retrieve than in the coordinate initial phrase (the second lexical item; 2a) as the effect has been demonstrated when the verb changes from trial to trial (Martin et al., 2010), as well as in Japanese, a verb-final language in which the subject and the complement take the first two positions in the sentence regardless of initial phrase type (Allum & Wheeldon, 2007, 2009). Nevertheless, the size of speakers' planning scope is not rigidly fixed and can vary due to multiple factors including ease of syntactic processing (Konopka, 2012; Konopka & Meyer, 2014), task complexity (Ferreira & Swets, 2002; Wagner, Jescheniak, & Schriefers, 2010) and cognitive abilities, such as working memory and production speed (Martin et al., 2004; Slevc, 2011; Swets, Jacovina, & Gerrig, 2014;

Wagner et al., 2010). Our interest was in whether the scope of advanced sentence planning is also influenced by healthy aging.

The aging process is typically associated with an increase in speech dysfluencies during sentence production, such as the use of non-lexical fillers ('uh' or 'um'), word repetitions and unnatural pauses (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001; Horton, Spieler, & Shriberg, 2010; Kemper, Rash, Kynette, & Norman, 1990). One significant factor that has been proposed to account for this age-related increase in speech dysfluencies is a reduction in the capacity and efficiency of working memory (Abrams & Farrell, 2011; Kemper & Sumner, 2001). This is because verbal working memory is essential for successfully being able to prepare more than one word before beginning articulation (Martin et al., 2004; Slevc, 2011) and for temporarily storing information that is needed for later syntactic processing, such as when producing an embedded clause sentence (Kemper, Kynette, Rash, O'Brien, & Sprott, 1989; Rabaglia & Salthouse, 2011). This suggests that incremental sentence planning processes may become less efficient with age (as the result of declining working memory) or that older adults may adopt different processing strategies when planning a sentence in order to compensate for age-related deficits in working memory. We therefore used the planning scope paradigm to investigate age-related changes in the amount of advanced planning that older speakers engage with prior to articulation.

Based on previous literature, we consider that there are two alternative hypotheses for age-related changes in planning scope. Firstly, a decline in working memory with age may disrupt older adults' ability to plan sentences with larger initial phrases. Martin et al. (2004) found that an aphasic patient with a semantic working memory deficit displayed a greater phrasal complexity effect than controls (i.e., a markedly greater difference in the speed of production of larger, compared to smaller, initial phrases), which they attributed to the patient attempting to plan both nouns in the initial phrase, but having difficulty doing so because of deficits at the lexical-semantic level (see also Lee & Thompson, 2011).[5] Although not as profound as aphasia patients, older adults also experience deficits in working memory (particularly at the verbal level; Bopp & Verhaeghen, 2005). Thus, one hypothesis is that older adults will display a larger phrasal complexity effect than young adults in the planning scope task. Alternatively, to compensate for decline in working memory, older adults may adopt a more extreme word-by-word incremental strategy (i.e., only plan the first word before speech onset regardless of the complexity of the initial phrase). Ferreira and Swets (2002) found that when time pressure was applied, speakers engaged in significantly less advanced planning, suggesting that incremental planning can be strategically controled by the speaker. This, combined with the evidence that older adults implement various strategies in other areas of language processing (Altmann & Kemper, 2006; Stine-Morrow et al., 2008), may mean that there is a strategic age-related decrease in the amount of advanced planning that occurs prior to articulation.

In Experiment 2 we further aimed to directly investigate age-related changes in the retrieval of lexical items and their integration into syntactic structures. Lexical retrieval and syntax generation do not rely on the exact same mechanisms, and may even be entirely dissociated (Chang et al., 2000, 2006). Thus, evidence of age effects in syntactic processing does not necessarily mean that age effects will also be observed in lexical processing (or vice-versa). One way to examine lexical processing during sentence production is to incorporate a picture preview element into the planning scope paradigm. Wheeldon et al. (2013) required participants to produce sentences similar to (2a) and (2b), but on some trials there was a preview of one of the upcoming pictures. They found that previewing the second to-be-produced lexical item (*hat* for the examples shown in 2) decreased onset latencies more when it fell within, rather than outside of, the initial phrase (see Allum & Wheeldon, 2009, for a similar study in Japanese using stationary visual displays). This suggests that the retrieval of lexical items within the first phrase is prioritized prior to speech onset. Nevertheless, the preview benefit was not reliably maintained when the phrase consisted of three nouns and participants previewed the third lexical item (*"[the drum, the star and the* ***hat*** *move] above the crab"*). Thus, it appears that advanced lexical planning only encompasses a subset of the required nouns and that this does not always align with scope of syntactic planning.

In Experiment 2 we therefore included a picture preview element within the planning scope task; the magnitude of the preview benefit displayed by older adults will be informative about age-related changes in lexical processing during sentence planning and production. Young adults' preferred scope of lexical encoding appeared to be two items (Wheeldon et al., 2013); however, we speculate that older adults' preferred limit may be less because they have a reduced memory buffer for holding linguistic information (Bopp & Verhaeghen, 2005; Waters & Caplan, 2003). Attempting to retrieve and hold an unmanageable number of lexical items prior to articulation can lead to problems with buffering and maintaining a linearized output (Slevc, 2011; Wheeldon et al., 2013). To overcome this and reduce demands on working memory, older adults may therefore only encode the first lexical item within a phrase prior to articulation; if this is the case, we may expect that, unlike young adults, older adults will not display the preview benefit of the second lexical item even when it falls within the initial phrase.

Experiment 2: Method

Participants. The same participants were used as described in Experiment 1.

Design. We used a 2 X 2 X 2 mixed design with one between-participant variable of age (young vs. older) and two within-participant variables of preview (no preview vs. preview) and initial phrase type (coordinate vs. simple). Hence, there were four experimental conditions (Figure 3A). Critically, the previewed picture (always of the second upcoming lexical item) fell within the initial phrase in the coordinate condition, but outside of the initial phrase in the simple condition.

Materials. To create the experimental items, we used 80 photographic pictures of everyday concrete objects (these were different to those used in Experiment 1, but meet the same criteria). We created 80 experimental items that each consisted of three different pictures that were conceptually and phonologically distinct: each of the 80 pictures appeared in three different experimental items (once in the left, central and right position). As in Experiment 1, the sentence descriptions of the items were elicited by controling the movement of the pictures (using E-prime) and participants were instructed to describe the picture movements from left to right using specific sentences. In the simple initial phrase conditions, only the left picture moved (either up or down) and the other two pictures remained stationary (*"the A moves above/below the B and the C"*). In the coordinate conditions, both the left and the central picture moved simultaneously (either up or down) and only the right picture remained stationary (*"the A and the B move above/below the C"*). In the preview trials, the preview was always of the central upcoming picture (i.e., object *B*). We created four item lists by evenly rotated the experimental condition assigned to each of the 80 experimental items. Each participant was randomly assigned to one of the four lists and completed 20 experimental items per condition (Table 1B).

Lastly, we used a further 106 pictures to create 220 filler items designed to prevent the participant from anticipating the location of the preview picture and building expectations to guide their response. The fillers elicited some experimental-type sentences and other sentences that differed from the experimental items in terms of the number of pictures and the type of movement, such as: *"there is an X, a Y and a Z"* (no picture movement); *"the Xs move up"* (three repeat pictures move simultaneously); and "*there are no pictures"*.

Importantly, we also varied the position of the preview pictures within the fillers, such that across all the experimental and filler items each screen position was previewed an equal number of times. All 220 filler items were added to each of the four item lists. We then divided each list into five blocks that each contained 44 fillers and 16 experimental items (4 per condition), and pseudorandomized the order of items using the same constraints as Experiment 1. The ordering of the blocks was rotated across participants.

Procedure. Each participant was tested using the same equipment set-up described in Experiment 1. Figure 3B illustrates the sequence of stimuli presentation per trial. To begin, there were 40 practice trials; the sentences elicited resembled those in the experimental and filler trials and featured all 80 experimental pictures once. The task then continued until all five blocks had been completed. Using the same criteria described in Experiment 1, the experimenter noted down any errors made by the participant.

Data Preparation and Analyses. One older adult was excluded from Experiment 2 because of error rates above 50% on the experimental trials. Of the 8400 experimental trials, we applied the same onset latency exclusion criteria described in Experiment 1, resulting in the discarding of 124 (3.1%) young and 166 (3.8%) older adult trials. All remaining trials were used in the error analyses, but only correct responses (81.7% of trials) were used in the onset latency analyses.

The data from Experiment 2 were analysed using the same mixed-effects modeling methods described in Experiment 1 (a logit model for error data and a linear model for the onset latency data). We entered age group (young vs. older), initial phrase type (coordinate vs. simple) and preview type (no preview vs. preview) into the models as fixed effects. We included random intercepts for participants and items, as well as by-participant and by-item random slopes appropriate for the design.

In Experiment 2, we also performed post-hoc pairwise comparisons to investigate the effect of preview on task performance in the different initial phrase conditions for young and older adults. We chose to conduct these post-hoc comparisons even when there were no higher-order interactions involving age group because visual inspection of the data suggested possible age group differences relating to the different experimental conditions (see Figure 4). These post-hoc comparisons should therefore be treated with caution; nevertheless, comparisons of factor level means may still be a useful (and acceptable) method for drawing meaningful conclusions in datasets in which main effects are significant but interactions are not (Wei, Carroll, Harden, & Wu, 2012). For our post-hoc analyses, we used the *'testInteractions'* function in the *phia* package (de Rosario-Martinez, 2015a, 2015b) which allows for the direct comparison of contrasts specified within an existing mixed-effect model; specifically, we compared the effect of preview vs. no preview within each phrase and age group condition. Importantly, the *'testInteractions'* function corrects *p* values for multiple comparisons using the Holm-Bonferroni method (adjusts the criteria of each individual hypothesis), thereby reducing the risk of discovering a false positive result (Aickin & Gensler, 1996; Holm, 1979).

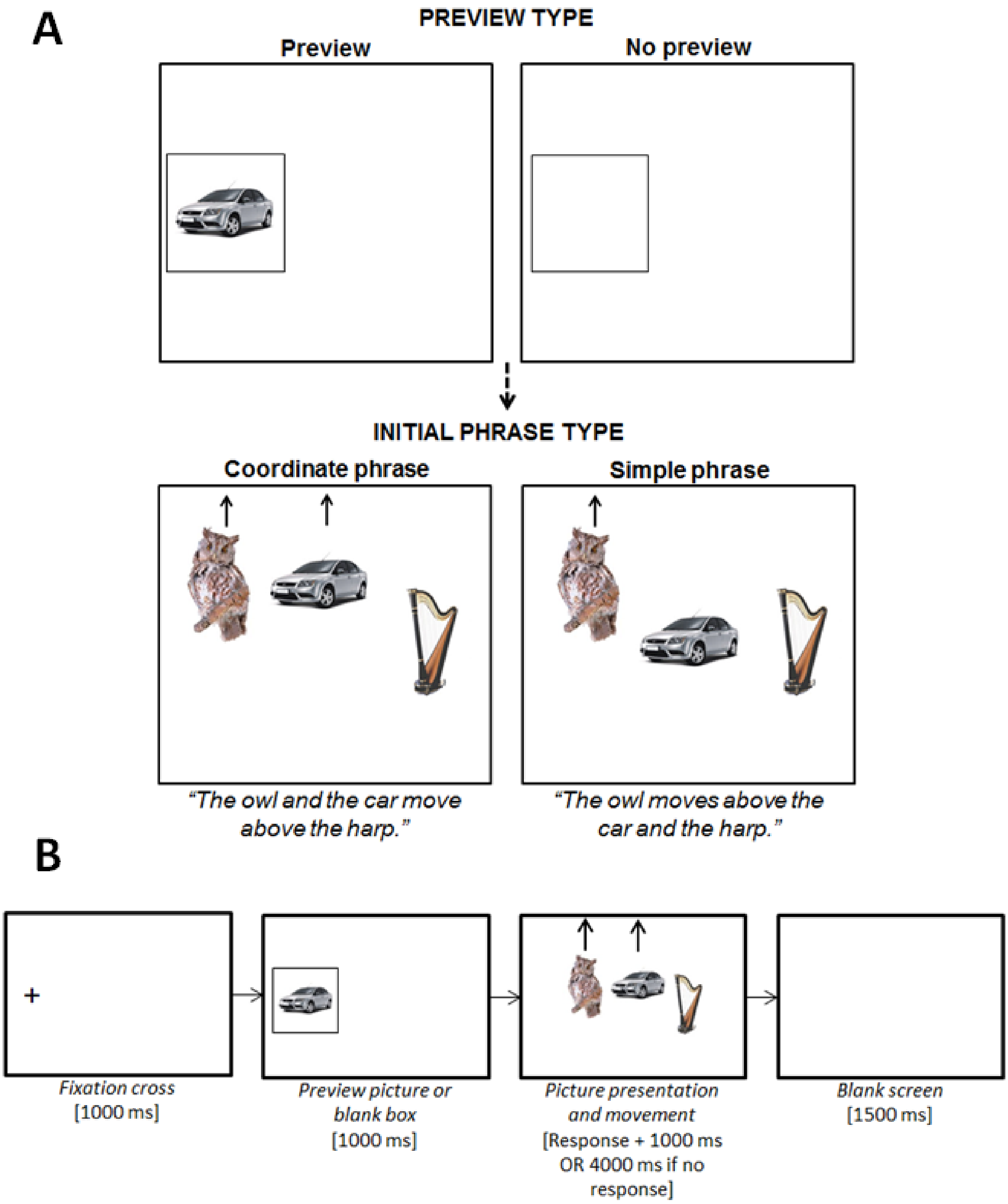

Figure 3. Experiment 2 planning scope design (A) and stimuli presentation events per trial (B). The participant was instructed to pay attention to the preview because it would appear in the upcoming trial, but not to name it aloud. The three pictures then appeared aligned centrally in the horizontal plane (importantly, the leftmost picture did not appear where the preview picture had just been, but in a more right-adjusted position). The participant was instructed to begin describing the picture movement as soon as possible using specific sentence types. Speech latencies were recorded from the onset of the pictures to the participant beginning to speak.

Experiment 2: Results

Figure 4 summarizes the error rates and the onset latencies across the four experimental conditions for young and older adults.

Error Rates. The best-fitting model of the error data is reported in Table 3A. As in Experiment 1, older adults were significantly more error-prone than young adults (23.5% vs. 12.5%, $p < .001$). While there were no main effects of preview ($p = .308$) or initial phrase type ($p = .097$), there was a significant interaction between the two variables ($p = .040$): the presence of the preview resulted in a 1.6% decrease in participants' errors when producing sentences with initial coordinate phrases, but a 2.9% increase in errors when producing sentences with initial simple phrases. The interaction between preview, initial phrase type and age group did not reach significance ($p = .285$); however, visual inspection of the data indicated some different patterns of errors in the young and older adults, especially relating to the effect of the preview picture (see Figure 4). Further post-hoc pairwise comparisons revealed no effect of preview on young adults' error rates in either the coordinate ($\chi^2(1) = 0.21$, $p = .649$) or simple ($\chi^2(1) = 0.20$, $p = .659$) initial phrase condition. For older adults, while there was no effect of preview on error rates when the it fell within the initial phrase (coordinate condition; $\chi^2(1) = 0.32$, $p = .570$), the presence of the preview caused a significant 5.3% increase in errors when it fell outside of the initial phrase (simple condition; $\chi^2(1) = 8.35$, $p = .003$).

Onset Latencies. The best-fitting model of the onset latency data is reported in Table 3B. As in Experiment 1, older adults were significantly slower than young adults (991ms vs. 843ms, $p < .001$). There was a main effect of initial phrase type, such that sentences with initial simple phrases were produced significantly quicker than sentences with initial coordinate phrases (895ms vs. 935ms, $p < .001$), indicating an overall phrasal planning effect of 40ms (4.5%). Furthermore, the interaction between initial phrase type and age group was not significant (*p* =. 994), indicating that the incremental planning effect was unaffected by healthy aging. Indeed, visual inspection of the data demonstrated that the phrasal planning benefit was similar for young (40ms, 4.6% benefit) and older (41ms, 4.0% benefit) adults.

The analyses further revealed a main effect of preview, such that sentences were produced significantly quicker following preview of the second upcoming lexical item compared to no preview (940ms vs. 980ms, $p < .001$). Moreover, there was a significant interaction between preview and initial phrase type ($p < .001$): the overall preview benefit was significantly greater when the preview picture fell within the initial phrase (coordinate condition; 74ms, 7.6%) compared to outside of it (simple condition; 26ms, 2.9%). Although, the interaction between age group, preview and initial phrase type did not reach significance ($p = .235$), visual inspection of the data again suggested possible age group differences relating to the effect of preview in the different initial phrase conditions (see Figure 4). For the young adults, pairwise comparisons revealed a significant benefit of preview in both the coordinate (81ms (8.9%), $\chi^2(1) = 33.48$, $p < .001$) and simple (45ms (5.3%), $\chi^2(1) = 10.30$, $p = .003$) phrase conditions, although the magnitude of the effect was distinctly larger when the preview fell within the initial phrase. By contrast, the difference in onset latencies between preview conditions was only significant for the older adults when it fell within the initial phrase (67ms (6.4%) preview benefit; $\chi^2(1) = 25.78$, $p < .001$), but not outside of it (2ms (0.2%) preview benefit; $\chi^2(1) = 0.51$, $p = .476$).

Summary. The main findings of Experiment 2 can be summarized as follows: (1) as in Experiment 1, older adults were slower and more error-prone than young adults; (2) our task elicited a reliable phrasal planning scope effect that was unaffected by healthy aging; and (3) while young adults' displayed speed benefit of preview in both phrase conditions, older adults only benefited when the preview fell within the initial phrase and produced significantly more errors when the previewed lexical item fell outside of the initial phrase.

Together, this suggests that there are age group differences in lexical processing during sentence planning, which only emerged when the preview fell outside of the initial phrase. In our additional analyses, we again found the individual difference measures to have minimal influence on task performance; the one significant result was an interaction between age group, phrase type and long-term memory score (see Supplementary Materials).

Table 3

Summary of the Best-fitted Mixed-effects Models for the Experiment 2 Error and Onset Latency Data.

Predictor	Coefficient	SE	Wald Z	*p*
A: Error Data				
Intercept	2.02	0.15	13.62	< .001
Preview	-0.07	0.07	-1.02	.308
Initial phrase type	0.12	0.07	1.66	.097
Age group	0.89	0.16	5.70	< .001
Preview * Initial phrase type	-0.28	0.14	-2.06	.040
Preview * Age group	0.14	0.14	1.05	.292
Initial phrase type * Age group	-0.04	0.14	-0.32	.747
Preview * Initial phrase type * Age group	0.29	0.27	1.07	.285
B: Onset Latency Data				
Intercept	924.22	16.89	54.71	< .001
Preview	-52.61	8.63	-6.09	< .001
Initial phrase type	-41.14	5.89	-6.98	< .001
Age group	-149.27	31.93	-4.68	< .001
Preview * Initial phrase type	49.62	10.80	4.60	< .001
Preview * Age group	-22.69	16.67	-1.36	.171
Initial phrase type * Age group	-0.02	11.03	0.00	.994
Preview * Initial phrase type * Age group	-25.53	21.51	-1.19	.235

Note. The model of the error data converged with random intercepts for participants and items with additional by-participant random slopes for the main effects of preview and initial phrase type, and a by-item random slope for the main effect of age group. The model of the onset latency data converged with random intercepts for participants and items with additional by-participant random slopes for the main effects of preview and initial phrase type, and by-item random slopes for the main effects of preview, initial phrase type and age group.

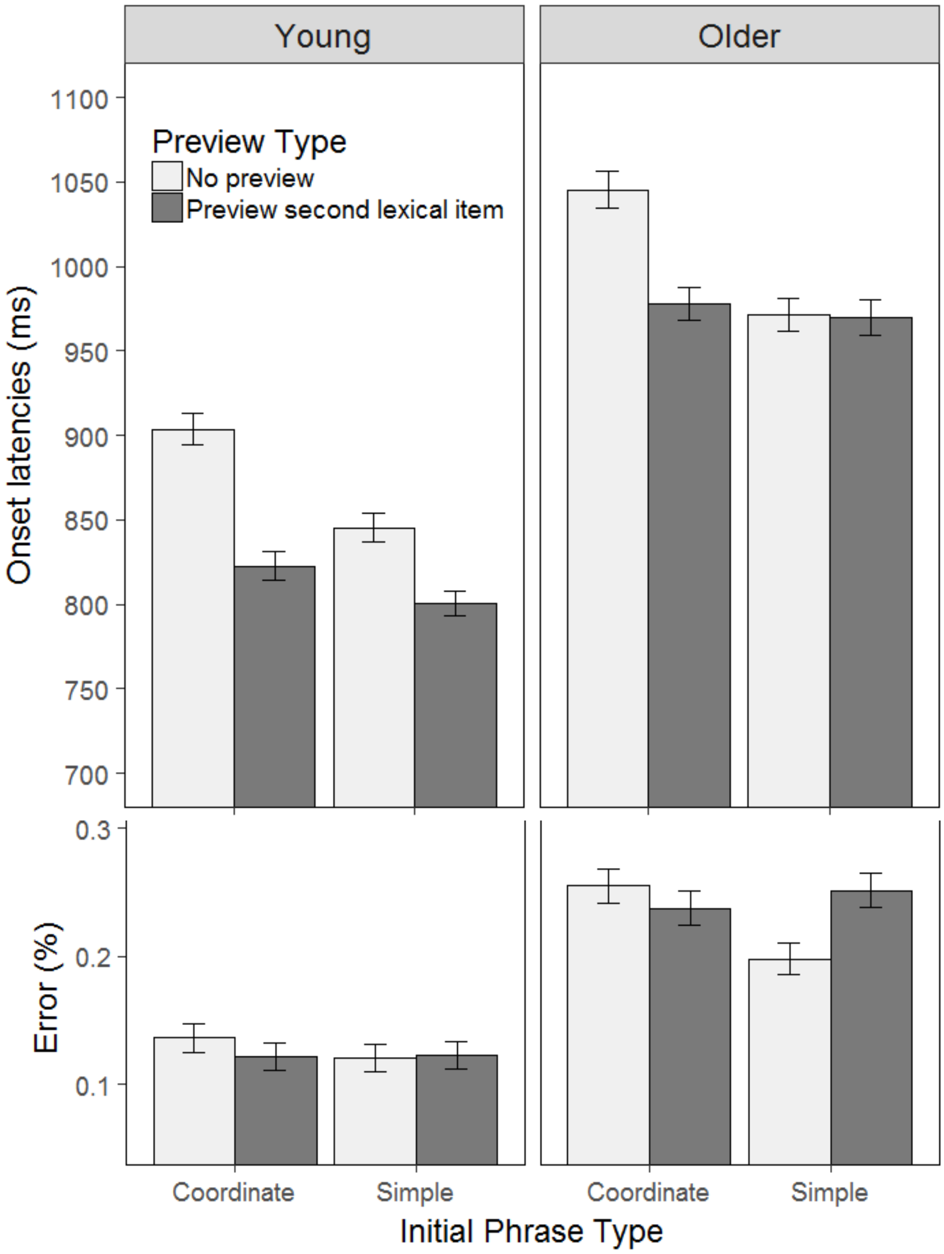

Figure 4. Experiment 2 mean error rates and onset latencies for young and older adults when producing sentences within initial coordinate and simple phrases following no preview or a preview of the second upcoming lexical item. Error bars denote ±1 the standard error of the mean.

General Discussion

Using two on-line experiments, we investigated age-related changes in the syntactic and lexical processes involved in sentence generation. In Experiment 1, both young and older adults produced target sentences quicker following syntactically related primes, demonstrating that speed benefits of syntactic priming are preserved with age, despite older adults' slower and more error-prone production. In Experiment 2, both young and older adults initiated sentences quicker with smaller, compared to larger, initial phrases, suggesting that planning scope, at least at the syntactic level, is unaffected by healthy aging. Age differences did emerge, however, in the preview conditions; whereas young adults displayed speed benefits of picture preview when the pictured word fell both within and outside the initial phrase, older adults only displayed speed benefits from the previewed picture when it fell within the initial phrase, and preview outside of the initial phrase caused them to become more error-prone. This suggests that there are age differences in the flexibility of lexical retrieval during sentence planning and in the ability to integrate lexical information into syntactic structures. Taking both experiments together, our study therefore suggests age-related effects of lexical, but not syntactic, processes on the speed and accuracy of sentence production.

Our robust finding of equal onset latency priming in both age groups in Experiment 1 provides the first evidence that syntactic facilitation effects are preserved with age in a task specifically designed to tap into the processes involved in the planning stage of sentence production. Applied to Segaert's et al. (2016) two-stage competition model, this suggests that older adults maintain the ability to quickly and efficiently generate previously activated syntactic structures. This is somewhat contrary to our initial hypothesis that overall decline in processing and transmission speed with age would result in decreases in the spreading activation architecture that supports syntactic priming (MacKay & Burke, 1990; Salthouse, 1996). Instead, the slowing associated with aging might not affect all cognitive networks equally (Fisher, Duffy, & Katsikopoulos, 2000; Fisk, Fisher, & Rogers, 1992). Thus, despite general slowing elsewhere, older adults appear to maintain sufficient cognitive resources to support successful syntactic priming. This is consistent with the evidence of priming benefits in older adults in other areas of language processing, such as morphological priming of both regularly-inflected verbs and transparent compounds (Clahsen & Reifegerste, 2017; Duñabeitia, Marín, Avilés, Perea, & Carreiras, 2009; Reifegerste, Elin, & Clahsen, 2018).

Nevertheless, it is important to consider that we found evidence of preserved latency priming effects in older adults in a task in which the demands were relatively low: participants only needed to dedicate minimal cognitive resources to syntactic selection (because we removed the choice element) and we did not manipulate the ease of lexical encoding. According to Peelle (2019), the relationship between cognitive supply and task demands would still therefore have been balanced in favor of good behavioral performance in older adults, despite likely declines in overall cognitive capacity. It therefore remains unclear whether linguistic priming effects would continue to be observed in older adults in a task in which demands are increased (e.g., by manipulating the codability of the nouns). Moreover, the consideration of task demands vs. cognitive supply may also be necessary for clarifying the mixed findings within the existing choice syntactic priming and aging literature (Hardy et al., 2017; Heyselaar et al., 2017, 2018; Sung, 2015). There are minimal methodological differences between the various studies (e.g., all used a picture description production task); however, it remains possible that differences in the characteristics of the samples, such as education level and native language use, may have resulted in differences in processing efficiency of the older adult groups, leading to different behavioral findings between studies (Peelle, 2019). Unfortunately, this information is unavailable for previous studies, meaning such a comparison is not possible. This highlights why it is important for future research to collect individual differences data, as well as age group information, when investigating what determines latency and choice syntactic priming (see Supplementary Materials for a characterization of our sample and our investigation of individual difference effects).

Turning now to the findings of Experiment 2, the pattern we observed in the onset latencies is similarly consistent with an age-related preservation of syntactic processing skills as we found robust evidence of a phrasal scope of planning in both age groups: speakers took longer to initiate sentences with larger initial phrases. This replicates previous research in young adults (e.g., Martin et al., 2010, 2014; Smith & Wheeldon, 1999), and suggests that both age groups prioritized the generation of syntax within the first phrase prior to articulation. It is notable that older speakers did not experience disproportionate difficulty in planning the larger initial phrases (as has been observed in aphasia patients; Martin et al., 2004), indicating that although aging is associated with decline in general cognitive function, this is not substantial enough to cause age-related deficits in incremental sentence production. Moreover, our findings demonstrate that older adults do not actively engage in a more extreme word-by-word planning strategy (if this was the case, latencies would have been similar for simple and coordinate initial phrases), further suggesting that older adults maintain sufficient cognitive capacity to support the planning of an initial phrase containing at least two nouns. Spieler and Griffin (2006) also found no differences in the sentence planning strategy used by young and older adults; however, they found that both age groups planned in single word, not phrasal, units. This apparent contrast to our findings can likely be explained by the different measurements used; specifically, while our use of onset latency measures provided insight into the preparation time before sentence articulation, Spieler and Griffin's (2006) use of eye-tracking focused more on the gaze shifts that occur during articulation and which are tightly locked to individual word onset. Nevertheless, both findings indicate that there are minimal age group differences in on-line syntactic processing, as has been found in other studies in which participants are presented with different words on screen and asked to formulate a sentence (Altmann & Kemper, 2006; Davidson et al., 2003).

An important point to make, however, is that the minimal age group differences we observed in syntactic planning do not necessarily mean that young and older adults were engaging the exact same cognitive networks when performing the task. While young adults may be predominantly relying on activity in the left anterior temporal lobe to support incremental sentence planning (Brennan & Pylkkänen, 2017), older adults may be recruiting additional areas outside of the core language network to support performance (in the same way as has been observed for other aspects of language processing; Peelle, Troiani, Wingfield, & Grossman, 2010; Wingfield & Grossman, 2006). Further work is therefore needed to fully understand the age-related changes in the neural networks that underlie incremental sentence planning. Indeed, age group differences did emerge due to the picture preview manipulation, suggesting that young and older adults may be adopting different strategies relating to lexical processing, a finding we turn to next.

In Experiment 2, half of the experimental trials were preceded by a picture of the upcoming second lexical item. When the previewed picture fell within the initial phrase (*"[the owl and the **car** move] above the harp"*), both young and older adults were quicker to initiate the sentence compared to when there was no preview, suggesting that the prior retrieval of the lexical item was significantly benefiting their sentence planning at the lexical encoding level (Allum & Wheeldon, 2009; Wheeldon et al., 2013). This further adds to the evidence that older adults are engaging in a phrasal scope of advanced planning because if they had only planned the first word prior to articulation then previewing the second lexical item would not be beneficial. Nevertheless, some interesting group differences did emerge in participants' onset latencies and error rates when the previewed picture fell outside of the initial phrase (*"[the owl moves] above the **car** and the harp"*). Young adults continued to display speed benefits of preview outside the initial phrase, albeit to a lesser extent than when it fell within the initial phrase. This demonstrates that the young adults prioritized the retrieval of lexical items within the first phrase prior to speech onset, but they were also able to successfully manage the early activation of lexical items outside of their usual phrasal planning scope to benefit their overall speed of sentence production. This evidence of adaptability within young adults' planning scope adds to the growing evidence that planning scope is flexible and can be influenced by the ease of syntactic and lexical processing (Konopka, 2012; Konopka & Meyer, 2014; van de Velde & Meyer, 2014).

In contrast, older adults did not display any speed benefits of preview outside of the initial phrase, and the presence of the picture preview outside their preferred phrasal planning scope caused them to become significantly more error-prone. The onset latency and error data together suggest that, unlike young adults, older adults did not benefit from this early access to lexical information and that instead this premature availability had a disruptive effect on their overall fluency. One explanation for this age-related difference is that older adults' planning scope may be more rigidly fixed to phrasal boundaries and so they are less adaptable when it comes to integrating new lexical information into syntactic structures. Indeed, older adults show less parafoveal preview effects across syntactic pauses than young adults during sentence comprehension, suggesting an age-related segmentation strategy designed to aid syntactic processing (Payne & Stine-Morrow, 2012, 2014; Stine-Morrow & Payne, 2016). This segmentation strategy may also apply to older speakers' sentence production; specifically, in an attempt to decrease processing demands, older adults may strategically choose to only attend to lexical information when it is relevant (i.e., only when it is contained within the next to-be-produced phrase). Thus, older adults are less able to successfully incorporate lexical information outside of the initial phrase into their sentence planning. This contrast between the flexible sentence planning approach observed in the young adults and the rigid approach in the older adults further highlights how apparently similar behavior in both age groups (i.e., both displayed a phrasal scope of planning) may be supported by different cognitive networks and processing strategies.

A second explanation for the age-related difference in lexical processing that we observed involves the executive control required to successfully manage the premature access to lexical information. During picture preview, participants automatically access lexical information about the pictured item which would be stored in their working memory. Participants must then temporarily inhibit this information in order to retrieve the first lexical item required for sentence onset. Young adults appear to be good at this as they even benefit from the preview information when it is required in the planning of the second phrase. In contrast, older adults only benefit when the previewed picture is required in the construction of the first phrase. Moreover, when the preview is required for the second phrase, and more inhibition is required, older adults become more error-prone. Theoretical accounts propose that aging weakens the inhibitory processes that are responsible for regulating what information enters and leaves working memory, and a consequence of age-related decline in inhibition is increased interference effects (Hasher, Lustig, & Zacks, 2007; Hasher & Zacks, 1988). Indeed, deficits in inhibitory control have been used to explain other age effects on language processing, such as older adults having difficulty ignoring distracting or irrelevant information during comprehension and production (Britt, Ferrara, & Mirman, 2016; Sommers & Danielson, 1999; Tun, O'Kane, & Wingfield, 2002). Deficits in inhibitory control may also explain our findings as, if the older adults were less able to inhibit irrelevant lexical information during the planning of the first phrase, this would lead to increased problems with formulating a linearized output, resulting in increased errors.

Critically, the effect of poor inhibitory control on the fluency of older adults' sentence planning only emerged when we increased the task demands. Although previewing information typically decreases task demands, this does not seem to be the case for older adults when the preview related to lexical information outside of the initial phrase. While previewing lexical information may facilitate lexical retrieval, it also increases demand on linearization when the previewed lexical item falls outside of the initial phrase; this places increased demand on cognitive resources as inhibition of the premature lexical information is now required. Thus, if older adults are less able to engage the required level of inhibitory control, the balance between processing efficiency and task demands will move to favor the latter, resulting in poorer behavioral performance (Peelle, 2019).

We now turn to an additional aim of our study, which was to investigate what kind of inter-individual variability accounted for individual differences in performance in the language tasks (see Supplementary Materials for more detail). Integrating measures of individual differences is important for understanding the cognitive factors that underlie different aspects of language processing and how this affected by the broad variation in the aging process (Kidd, Donnelly, & Christiansen, 2018; Peelle, 2019). However, throughout our study, we found little evidence to suggest that participants' scores on the eight additional factors we measured were accounting for any individual variation in performance in Experiments 1 and 2. Significant effects were limited to an influence of inhibitory control on the young adults' onset latencies in Experiment 1 and an influence of long-term memory ability on the magnitude of the phrasal planning scope effect in Experiment 2. Most notably, however, we observed no evidence of a relationship between our inhibitory control measure and preview effects in the planning scope task – such a relationship would be predicted by our explanation for the age-related differences in the preview benefit. We speculate that our lack of significant findings may be due to the fact that we only used a single measurement per construct, which would have impacted on measurement reliability. This limitation was the result of the broad range of measures we employed, which led to time constraints in testing to avoid participant fatigue. Moreover, there are inherent difficulties involved in measuring individual difficulties within a factorial design, particularly one that involve measures of speed (Hedge, Powell, & Sumner, 2017; Miller & Ulrich, 2013). Further research using a stronger battery of measures is therefore required to test our prediction of a relationship between inhibitory control and lexical planning in healthy aging.

In summary, our study is the first to examine age-related changes in syntactic and lexical processing during sentence production using on-line onset latency measures. Specifically, our study provides evidence for the age-related preservation of syntactic processing (as evident in the syntactic priming and phrasal planning scope effects we observed in both age groups), but an increased difficulty with lexical retrieval and integration with age. We attribute this apparent age-related decline in lexical processing to a decline in the flexibility of sentence planning processes. This may be related to older speakers' stronger preference for segmentation at phrasal boundaries when planning a sentence (a strategic approach designed to minimize processing demands) and/or to a decline in executive control making older speakers less able to cope with premature lexical activation beyond the first phrase. Our findings should be considered in parallel with off-line studies of language and aging in order to gain a more complete picture of language processing in old age, in terms of which processes are preserved and which decline.

Footnotes

[1] We note that many other studies have employed on-line measures of speech production to investigate age-related differences at the single word level (see Mortensen, Meyer, & Humphreys, 2006, for a review). While production of single words requires the retrieval of lexical information, it does not require the incorporation of the lexical items into a syntactic structure, as is required during sentence production (Levelt, 1989). Thus, it is difficult to apply these single word findings to age-related effects on sentence production (Kavé & Goral, 2017).

[2] Note, some other studies have tested non-young adults as controls for clinical patients; however, the samples are small and the age ranges are large. While Ferreira, Bock, Wilson, and Cohen, (2008, n = 4 aged 50-58) and Cho-Reyes, Mack, and Thompson (2016; n = 13 aged 33-76) found evidence of choice syntactic priming in controls, Hartsuiker and Kolk (1998; n = 12 aged ~28-67) did not.

[3] Education was scored according to the International Standard Classification of Education (United Nations, 2011), which classifies education on a scale of 0 (pre-primary school) to 8 (university doctorate). There was no significant difference in scores between young (M = 6.0, SD = 0.1) and older (M = 5.8, SD = 1.3) adults, $t(104) = 1.36$, p = .178.

[4] Due to the large speed differences between young and older adults, we also performed the modeling analysis with age-standardized onset latencies (using z-score adjustments within age groups). This produced the same effects (expect for the main effect of age) seen in the non-adjusted onset latencies analyses for both Experiments 1 and 2.

[5] We note that both of these studies did include non-young adults as controls for aphasia patients; however, the sample are small and the age ranges large (Lee & Thompson, 2011, n = 9 aged 48-73; Martin et al., 2004, n = 10 aged ~55-66), making it difficult to draw any firm conclusions about age-related effects on incremental sentence planning.

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. 49–74). New York: Psychology Press. https://doi.org/10.4324/9780203848005.ch3

Aickin, M., & Gensler, H. (1996). Adjusting for multiple testing when reporting research results: The Bonferroni vs Holm methods. *American Journal of Public Health*, *86*(5), 726–728. https://doi.org/10.2105/AJPH.86.5.726

Allum, P. H., & Wheeldon, L. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 791–810. https://doi.org/10.1037/0278-7393.33.4.791

Allum, P. H., & Wheeldon, L. (2009). Scope of lexical access in spoken sentence production: Implications for the conceptual–syntactic interface. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(5), 1240–1255. https://doi.org/10.1037/a0016367

Altmann, L. J. P., & Kemper, S. (2006). Effects of age, animacy and activation order on sentence production. *Language and Cognitive Processes*, *21*(1–3), 322–354. https://doi.org/10.1080/0169096054400006

Barnes, J. N. (2015). Exercise, cognitive function, and aging. *Advances in Physiology Education*, *39*, 55–62. https://doi.org/10.1152/advan.00101.2014

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*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Bock, K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *Journals of Gerontology: Psychological Sciences*, *60B*(5), 223–233. https://doi.org/10.1093/geronb/60.5.P223

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech*, *44*(2), 123–147. Retrieved from https://journals.sagepub.com/doi/pdf/10.1177/00238309010440020101

Brennan, J. R., & Pylkkänen, L. (2017). MEG evidence for incremental sentence composition in the anterior temporal lobe. *Cognitive Science*, *41*(6), 1515–1531. https://doi.org/10.1111/cogs.12445

Britt, A. E., Ferrara, C., & Mirman, D. (2016). Distinct effects of lexical and semantic competition during picture naming in younger adults, older adults, and people with aphasia. *Frontiers in Psychology*, *7*, 813. https://doi.org/10.3389/fpsyg.2016.00813

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word-finding failures in young and older adults. *Journal of Memory and Language*, *30*, 542–579. https://doi.org/https://doi.org/10.1016/0749-596X(91)90026-G

Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 373–443). New York: Psychology Press. https://doi.org/10.1080/03610738508259280

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, *29*(2), 217–229. https://doi.org/10.1023/A:1005101313330

Cho-Reyes, S., Mack, J. E., & Thompson, C. K. (2016). Grammatical encoding and learning in agrammatic aphasia: Evidence from structural priming. *Journal of Memory and Language*, *91*, 202–218. https://doi.org/10.1016/j.jml.2016.02.004

Clahsen, H., & Reifegerste, J. (2017). Morphological processing in old-age bilinguals. In M. Libben, M. Goral, & G. Libben (Eds.), *Bilingualism: A framework for understanding the mental lexicon* (pp. 217–248). Amsterdam: John Benjamins Publishing Company.

Corley, M., & Scheepers, C. (2002). Syntactic priming in English sentence production: Categorical and latency evidence from an internet-based study. *Psychonomic Bulletin & Review*, *9*(1), 126–131. https://doi.org/10.3758/BF03196267

Davidson, D. J., Zacks, R. T., & Ferreira, F. (2003). Age preservation of the syntactic processor in production. *Journal of Psycholinguistic Research*, *32*(5), 541–566. https://doi.org/10.1023/A:1025402517111

de Rosario-Martinez, H. (2015a). Analysing interactions of fitted models. *CiteSeer*, 1–30. https://doi.org/10.1007/s13398-014-0173-7.2

de Rosario-Martinez, H. (2015b). *phia: Post-hoc interaction analysis. R package version 0.2-1*. Retrieved from https://cran.r-project.org/package=phia

Duñabeitia, J. A., Marín, A., Avilés, A., Perea, M., & Carreiras, M. (2009). Constituent priming effects: Evidence for preserved morphological processing in healthy old readers. *European Journal of Cognitive Psychology*, *21*(2–3), 283–302. https://doi.org/10.1080/09541440802281142

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*(1), 57–84. https://doi.org/10.1006/jmla.2001.2797

Ferreira, V. S., Bock, K., Wilson, M. P., & Cohen, N. J. (2008). Memory for syntax despite memory amnesia. *Psychological Science*, *19*(9), 940–946. https://doi.org/10.1111/j.1467-9280.2008.02180.x.Memory

Feyereisen, P. (1997). A meta-analytic procedure shows an age-related decline in picture naming: Comments on Goulet, Ska, and Kahn (1994). *Journal of Speech, Language, and Hearing Research*, *40*(6), 1328–1333. https://doi.org/10.1044/jslhr.4006.1328

Fisher, D. L., Duffy, S. A., & Katsikopoulos, K. V. (2000). Cognitive slowing amoung older adults: What kind and how much? In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging* (pp. 87–124). Oxford, UK: Oxford University Press.

Fisk, A. D., Fisher, D. L., & Rogers, W. A. (1992). General slowing alone cannot explain age-related search effects: Reply to Cerella (1991). *Journal of Experimental Psychology: General*, *121*(1), 73–78. https://doi.org/10.1037/0096-3445.121.1.73

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression*. Thousand Oaks, CA: Sage.

Geda, Y. E., Roberts, R. O., Knopman, D. S., Christianson, T. J. H., Pankratz, V. S., Ivnik, R. J., … Rocca, W. A. (2010). Physical exercise, aging, and mild cognitive impairment a population-based study. *Archives of Neurology*, *67*(1), 80–86. https://doi.org/10.1001/archneurol.2009.297

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N. A., Friston, K. J., & Frackowiak, R. S. J. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*, *14*, 21–36. https://doi.org/10.1006/nimg.2001.0786

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, *82*(1), 1–14. https://doi.org/https://doi.org/10.1016/S0010-0277(01)00138-X

Griffin, Z. M., & Spieler, D. H. (2006). Observing the what and when of language production for different age groups by monitoring speakers' eye movements. *Brain and Language*, *99*(3), 272–288. https://doi.org/10.1016/j.bandl.2005.08.003

Hardy, S. M., Messenger, K., & Maylor, E. A. (2017). Aging and syntactic representations: Evidence of preserved syntactic priming and lexical boost. *Psychology and Aging*, *32*(6), 588–596. https://doi.org/10.1037/pag0000180

Hartsuiker, R. J., & Kolk, H. H. J. (1998). Syntactic facilitation in agrammatic sentence production. *Brain and Language*, *62*(2), 221–254. https://doi.org/10.1006/brln.1997.1905

Hasher, L., Lustig, C., & Zacks, R. T. (2007). Inhibitory mechanisms and the control of attention. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory* (pp. 228–249). New York: Oxford University Press.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 22, pp. 193–225). New York: Academic Press.

Hedge, C., Powell, G., & Sumner, P. (2017). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*. https://doi.org/10.3758/s13428-017-0935-1

Heyselaar, E., Segaert, K., Walvoort, S. J. W., Kessels, R. P. C., & Hagoort, P. (2017). The role of procedural memory in the skill for language: Evidence from syntactic priming in patients with amnesia. *Neuropsychologia*, *101*, 97–105. https://doi.org/https://doi.org/10.1016/j.neuropsychologia.2017.04.033

Heyselaar, E., Wheeldon, L., & Segaert, K. (2018). Structural priming is supported by different components of non-declarative memory: Evidence from priming across the lifespan. *BioRxiv*, *190355*. https://doi.org/10.1101/190355

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Horton, W. S., Spieler, D. H., & Shriberg, E. (2010). A corpus analysis of patterns of age-related change in conversational speech. *Psychology and Aging*, *25*(3), 708–713. https://doi.org/10.1037/a0019424

Hoskyn, M., & Swanson, H. L. (2003). The relationship between working memory and writing in younger and older adults. *Reading and Writing*, *16*(8), 759–784. https://doi.org/10.1023/A:1027320226283

Howell, D. C. (2010). *Statistical methods for psychology*. Belmont, CA: Wadsworth Cengage Learning.

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Jaeger, T. F. (2011). *Diagnosing collinearity in mixed models from lme4*. Retrieved from https://hlplab.wordpress.com/2011/02/24/diagnosing-collinearity-in-lme4/

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149. https://doi.org/10.1037/0033-295X.99.1.122

Kavé, G., & Goral, M. (2017). Do age-related word retrieval difficulties appear (or disappear) in connected speech? *Aging, Neuropsychology, and Cognition*, *25*(5), 508–527. https://doi.org/10.1080/13825585.2016.1226249

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, *11*(2), 201–258. https://doi.org/10.1016/S0364-0213(87)80006-X

Kemper, S. (1987). Life-span changes in syntactic complexity. *Journal of Gerontology*, *42*(3), 323–328. https://doi.org/10.1093/geronj/42.3.323

Kemper, S., Greiner, L. H., Marquis, J. G., Prenovost, K., & Mitzner, T. L. (2001). Language decline across the life span: Findings from the Nun Study. *Psychology and Aging*, *16*(2), 227–239. https://doi.org/10.1037/0882-7974.16.2.227

Kemper, S., Herman, R. E., & Lian, C. H. T. (2003). Age differences in sentence production. *Journal of Gernotology: Psychological Sciences*, *58*(5), 260–269. https://doi.org/10.1093/geronb/58.5.P260

Kemper, S., Herman, R. E., & Liu, C. J. (2004). Sentence production by young and older adults in controlled contexts. *Journals of Gerontology: Series B*, *58B*, 220–224. https://doi.org/10.1093/geronb/59.5.P220

Kemper, S., Kynette, D., Rash, S., O'Brien, K., & Sprott, R. (1989). Life-span changes to adults' language: Effects of memory and genre. *Applied Psycholinguistics*, *10*(1), 49–66. https://doi.org/10.1017/S0142716400008419

Kemper, S., Rash, S., Kynette, D., & Norman, S. (1990). Telling stories: The structure of adults' narratives. *European Journal of Cognitive Psychology*, *2*(3), 205–228. https://doi.org/10.1080/09541449008406205

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging*, *16*(2), 312–322. https://doi.org/http://dx.doi.org/10.1037/0882-7974.16.2.312

Kemper, S., Thompson, M., & Marquis, J. G. (2001). Longitudinal change in language production: Effects of aging and dementia on grammatical complexity and propositional content. *Psychology and Aging*, *16*(4), 600–614. https://doi.org/10.1037/0882-7974.16.4.600

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences*, *22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006

Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language*, *66*, 143–162. https://doi.org/10.1016/j.jml.2011.08.003

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

Lara, J., Godfrey, A., Evans, E., Heaven, B., Brown, L. J. E., Barron, E., … Mathers, J. C. (2013). Towards measurement of the Healthy Ageing Phenotype in lifestyle-based intervention studies. *Maturitas*, *76*, 189–199. https://doi.org/10.1016/j.maturitas.2013.07.007

Lee, J., & Thompson, C. K. (2011). Real-time production of arguments and adjuncts in normal and agrammatic speakers. *Language and Cognitive Processes*, *26*(8), 985–1021. https://doi.org/10.1080/01690965.2010.496237

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

Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, *42*(1–3), 1–22. https://doi.org/10.1016/0010-0277(92)90038

Levelt, W. J. M., & Maassen, B. (1981). Lexical search and order of mention in sentence production. In W. J. M. Levelt & W. Klein (Eds.), *Crossing the boundaries in linguistics* (pp. 221–252). Dordrecht, Netherlands: D. Reidel Publishing Company.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75. https://doi.org/10.1017/S0140525X99001776

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, *91*(3), 295–327. https://doi.org/10.1037/a0035230

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*(1), 35–54. https://doi.org/10.1037/0033-295X.109.1.35

MacKay, D. G., & Burke, D. M. (1990). Cognition and aging: A theory of new learning and the use of old connections. In T. M. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp. 213–263). New York: Elsevier Science. https://doi.org/10.1016/S0166-4115(08)60159-4

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, *91*, 5–27. https://doi.org/10.1016/j.jml.2016.03.009

Marinis, T. (2010). Using on-line processing methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 139–162). Amsterdam: John Benjamins Publishing Company.

Martin, R. C., Crowther, J. E., Knight, M., Tamborello, F. P., & Yang, C. L. (2010). Planning in sentence production: Evidence for the phrase as a default planning scope. *Cognition*, *116*(2), 177–192. https://doi.org/10.1016/j.cognition.2010.04.010

Martin, R. C., Miller, M., & Vu, H. (2004). Lexical-semantic retention and speech production: Further evidence from normal and brain-damaged participants for a phrasal scope of planning. *Cognitive Neuropsychology*, *21*(6), 625–644. https://doi.org/10.1080/02643290342000302

Martin, R. C., Yan, H., & Schnur, T. T. (2014). Working memory and planning during sentence production. *Acta Psychologica*, *152*, 120–132. https://doi.org/10.1016/j.actpsy.2014.08.006

Maylor, E. A. (2005). Age-related changes in memory. In M. L. Johnson, V. L. Bengtson, P. G. Coleman, & T. B. L. Kirkwood (Eds.), *The cambridge handbook of age and ageing* (pp. 200–208). Cambridge, UK: Cambridge University Press.

Mertins, B. (2016). The use of experimental methods in linguistic research: Advantages, problems and possible pitfalls. In T. Anstatt, A. Gattner, & C. Clasmeier (Eds.), *Slavic Languages in Psycholinguistics* (pp. 15–33). Tübingen, Germany: Narr Francke Attempto Verlag GmnH + Co.KG.

Miller, J., & Ulrich, R. (2013). Mental chronometry and individual differences: Modeling reliabilities and correlations of reaction time means and effect sizes. *Psychonomic Bulletin and Review*, *20*(5), 819–858. https://doi.org/10.3758/s13423-013-0404-5

Mody, M. (2017). *Neural mechanisms of language*. New York: Springer. https://doi.org/10.1007/978-1-4939-7325-5

Mortensen, L., Meyer, A. S., & Humphreys, G. W. (2006). Age-related effects on speech production: A review. *Language and Cognitive Processes*, *21*(1–3), 238–290. https://doi.org/10.1080/01690960444000278

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal cognitive assessment, MoCA: A brief screening for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*, 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Pathan, S. S., Gottesman, R. F., Mosley, T. H., Knopman, D. S., Sharrett, A. R., & Alonso,

A. (2011). Association of lung function with cognitive decline and dementia: The Atherosclerosis Risk in Communities (ARIC) Study. *European Journal of Neurology*, *18*(6), 888–898. https://doi.org/10.1111/j.1468-1331.2010.03340.x

Payne, B. R., & Stine-Morrow, E. A. L. (2012). Aging, parafoveal preview, and semantic integration in sentence processing: Testing the cognitive workload of wrap-up. *Psychology and Aging*, *27*(3), 638–649. https://doi.org/10.1037/a0026540

Payne, B. R., & Stine-Morrow, E. A. L. (2014). Adult age differences in wrap-up during sentence comprehension: Evidence from ex-gaussian distributional analyses of reading time. *Psychology and Aging*, *29*(2), 213–228. https://doi.org/10.1037/a0036282

Peelle, J. E. (2019). Language and aging. In G. I. de Zubicaray & N. O. Schiller (Eds.), *The Oxford handbook of neurolinguistics*. Oxford, UK: Oxford University Press. https://doi.org/10.1016/B978-0-12-809324-5.01889-7

Peelle, J. E., Troiani, V., Wingfield, A., & Grossman, M. (2010). Neural processing during older adults' comprehension of spoken sentences: Age differences in resource allocation and connectivity. *Cerebral Cortex*, *20*(4), 773–782. https://doi.org/10.1093/cercor/bhp142

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*(4), 633–651. https://doi.org/10.1006/jmla.1998.2592

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*(3), 427–459. https://doi.org/10.1037/0033-2909.134.3.427

Poulisse, C., Wheeldon, L., & Segaert, K. (2019). Evidence against preserved syntactic comprehension in healthy aging. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, in press. https://doi.org/10.1101/299883

R Core Team. (2015). *A language and enviroment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Rabaglia, C. D., & Salthouse, T. A. (2011). Natural and constrained language production as a function of age and cognitive abilities. *Language and Cognitive Processes*, *26*(10), 1505–1531. https://doi.org/10.1080/01690965.2010.507489

Raven, J., Raven, J. C., & Court, J. H. (1988). *Manual for Raven's progressive matrices and vocabulary scales.* London: H. K. Lewis.

Reifegerste, J., Elin, K., & Clahsen, H. (2018). Persistent differences between native speakers and late bilinguals: Evidence from inflectional and derivational processing in older speakers. *Bilingualism: Language and Cognition*, 1–16. https://doi.org/10.1017/S1366728918000615

Reuter-Lorenz, P. A., & Park, D. C. (2014). How Does it STAC Up? Revisiting the scaffolding theory of aging and cognition. *Neuropsychology Review*, *24*(3), 355–370. https://doi.org/10.1007/s11065-014-9270-9

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

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, *13*(4), 140–144. https://doi.org/https://doi.org/10.1111/j.0963-7214.2004.00293.x

Salthouse, T. A. (2012). Are individual differences in rates of aging greater at older ages? *Neurobiology of Aging*, *33*(10), 2373–2381. https://doi.org/10.1016/j.neurobiolaging.2011.10.018

Salthouse, T. A. (2016). Contributions of the individual differences approach to cognitive aging. *Journals of Gerontology: Psychological Sciences*, *72*(1), 7–15. https://doi.org/10.1093/geronb/gbw069

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh, PA:

Psychology Software Tools Inc.

Schoot, L., Heyselaar, E., Hagoort, P., & Segaert, K. (2016). Does syntactic alignment effectively influence how speakers are perceived by their conversation partner? *PloS One*, *11*(4), e0153521. https://doi.org/10.1371/journal.pone.0153521

Segaert, K., Lucas, S. J. E., Burley, C. V, Segaert, P., Milner, A. E., Ryan, M., & Wheeldon, L. (2018). Fit to speak - Physical fitness is associated with reduced langauge decline in healthy ageing. *Scientific Reports*, *8*, 6715. https://doi.org/10.1038/s41598-018-24972-1

Segaert, K., Menenti, L., Weber, K., & Hagoort, P. (2011). A paradox of syntactic priming: Why response tendencies show priming for passives, and response latencies show priming for actives. *PLoS ONE*, *6*(10), e24209. https://doi.org/10.1371/journal.pone.0024209

Segaert, K., Weber, K., Cladder-Micus, M., & Hagoort, P. (2014). The influence of verb-bound syntactic preferences on the processing of syntactic structures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1448–1460. https://doi.org/10.1037/a0036796

Segaert, K., Wheeldon, L., & Hagoort, P. (2016). Unifying structural priming effects on syntactic choices and timing of sentence generation. *Journal of Memory and Language*, *91*, 59–80. https://doi.org/10.1016/j.jml.2016.03.011

Shafto, M. A., Burke, D. M., Stamatakis, E. A., Tam, P. P., & Tyler, L. K. (2007). On the tip-of-the-tongue: Neural correlates of increased word-finding failures in normal aging. *Journal of Cognitive Neuroscience*, *19*(12), 2060–2070. https://doi.org/10.1162/jocn.2007.19.12.2060

Slevc, L. R. (2011). Saying what's on your mind: Working memory effects on sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1503–1514. https://doi.org/10.1037/a0024350

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition*, *73*(3), 205–246. https://doi.org/10.1016/S0010-0277(99)00053-0

Smith, M., & Wheeldon, L. (2001). Syntactic priming in spoken sentence production – an online study. *Cognition*, *78*(2), 123–164. https://doi.org/10.1016/S0010-0277(00)00110-4

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging*, *14*(3), 458–472. https://doi.org/10.1037/0882-7974.14.3.458

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*(1–3), 291–321. https://doi.org/10.1080/01690960400002133

Sternäng, O., Reynolds, C. A., Finkel, D., Ernsth-Bravell, M., Pedersen, N. L., & Dahl Aslan, A. K. (2016). Grip strength and cognitive abilities: Associations in old age. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *71*(5), 841–848. https://doi.org/10.1093/geronb/gbv017

Stine-Morrow, E. A. L., Miller, L. M. S., Gagne, D. D., & Hertzog, C. (2008). Self-regulated reading in adulthood. *Psychology and Aging*, *23*(1), 131–153. https://doi.org/10.1016/B978-0-12-811448-3.00002-4

Stine-Morrow, E. A. L., & Payne, B. R. (2016). Age differences in language segmentation. *Experimental Aging Research*, *42*(1), 83–96. https://doi.org/10.1080/0361073X.2016.1108751

Sung, J. E. (2015). Age-related changes in sentence production abilities and their relation to working-memory capacity: Evidence from a verb-final language. *PLoS ONE*, *10*(4). https://doi.org/10.1371/journal.pone.0119424

Swets, B., Jacovina, M. E., & Gerrig, R. J. (2014). Individual differences in the scope of speech planning: Evidence from eye-movements. *Language and Cognition*, *6*, 12–44. https://doi.org/10.1017/langcog.2013.5

Tun, P. A., O'Kane, G., & Wingfield, A. (2002). Distraction by competing speech in young and older adult listeners. *Psychology and Aging*, *17*(3), 453–467. https://doi.org/10.1037//0882-7974.17.3.453

United Nations. (2011). *International standard classification of education*. Montreal, Canada: UN Institute for Statistics. Retrieved from http://uis.unesco.org/en/topic/international-standard-classification-education-isced

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

van Dyke, J. A. (2012). The role of memory in language and communication. In R. Peach & L. Shapiro (Eds.), *Cognition and acquired language disorders: A process-oriented approach* (pp. 94–120). St Louis, MO: Mosby Elsevier.

Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging*, *18*(2), 332–339. https://doi.org/10.1037/0882-7974.18.2.332

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 423–440. https://doi.org/10.1037/a0018619

Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers*, *35*(4), 550–564. https://doi.org/10.3758/BF03195534

Wechsler, D. (1997). *Wechsler memory scale - third edition manual*. San Antonia, TX: Pearson.

Wechsler, D. (2008). *Wechsler adult intelligence scale - fourth edition*. San Antonio, TX: Pearson.

Wei, J., Carroll, R. J., Harden, K. K., & Wu, G. (2012). Comparisons of treatment means when factors do not interact in two-factorial studies. *Amino Acids*, *42*(5), 2031–2035. https://doi.org/10.1007/s00726-011-0924-0

Wheeldon, L. (2011). Generating spoken sentences: The relationship between words and syntax. *Linguistics and Language Compass*, *5*(6), 310–321. https://doi.org/10.1111/j.1749-818X.2011.00281.x

Wheeldon, L. (2013). Producing spoken sentences: The scope of incremental planning. In S. Fuchs, M. Weirich, D. Pape, & P. Perrier (Eds.), *Speech production and perception: Speech planning and dynamics* (pp. 97–118). Frankfurt, Germany: Peter Lang International Academic Publishers.

Wheeldon, L., Ohlson, N., Ashby, A., & Gator, S. (2013). Lexical availability and grammatical encoding scope during spoken sentence production. *The Quarterly Journal of Experimental Psychology*, *66*(8), 1653–1673. https://doi.org/10.1080/17470218.2012.754913

Wheeldon, L., & Smith, M. (2003). Phrase structure priming: A short-lived effect. *Language and Cognitive Processes*, *18*(4), 431–442. https://doi.org/10.1080/01690960244000063

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., & Tannock, R. (1999). Development of inhibitory control across the life span. *Developmental Psychology*, *35*(1), 205–213. https://doi.org/10.1037/0012-1649.35.1.205

Wingfield, A., & Grossman, M. (2006). Language and the aging brain: Patterns of neural compensation revealed by functional brain imaging. *Journal of Neurophysiology*, *96*(6), 2830–2839. https://doi.org/10.1152/jn.00628.2006

Zhao, L.-M., & Yang, Y.-F. (2016). Lexical planning in sentence production is highly incremental: Evidence from ERPs. *PLoS ONE*, *11*(1), e0146359. https://doi.org/10.1371/journal.pone.0146359

Ziegler, G., Dahnke, R., & Gaser, C. (2012). Models of the aging brain structure and individual decline. *Frontiers in Neuroinformatics*, *6*, 3. https://doi.org/10.3389/fninf.2012.00003