Healthy Ageing and Sentence Production: Impaired Lexical Access in the Context of Intact Syntactic Planning

Sophie M. Hardy[a]*

Katrien Segaert[a,b]

Linda Wheeldon[a,c]

[a] School of Psychology, University of Birmingham, Birmingham, UK

[b] Centre for Human Brain Health, University of Birmingham, Birmingham, UK

[c] Department of Foreign Languages and Translation, University of Agder, Kristiansand Norway

* Corresponding author: School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK. Email: smh649@bham.ac.uk

Contact information other authors: k.segaert@bham.ac.uk, linda.r.wheeldon@uia.no

ABSTRACT

Sentence production requires rapid syntax generation and word retrieval. We investigated how healthy ageing affects these processes. In Experiment 1, syntactic priming significantly facilitated sentence production speed for both young and older adults. In Experiment 2, participants produced sentences with initial coordinate or simple phrases (e.g., *the owl and the car move above the harp/the owl moves above the car and the harp*); on half the trials, the second picture (*car*) was previewed. Without preview, both age groups were slower to initiate sentences with larger coordinate phrases, suggesting a similar planning scope. Young adults displayed speed benefits of preview both within and outside the initial phrase; however, older adults only displayed speed preview benefits within the initial phrase, and preview outside the initial phrase significantly increased error rates. Thus, syntactic planning scope appears unaffected by age, but there is age-related decline in the integration of lexical items into syntactic structures.

Keywords: healthy ageing; sentence production; syntactic planning; syntactic priming; lexical retrieval.

Each day the average person will produce 16,000 words and spend 4.5 hours engaging in conversation (Mehl, Vazire, Ramirez-Esparza, Slatcher, & Pennebaker, 2007; Rozenthuler, 2013). The explicit purpose of human communication is to interact with others and to convey our thoughts in a meaningful manner; consequently, a breakdown in language abilities can lead to increased social withdrawal and loneliness (le Dorze & Brassard, 1995; Mick, Kawachi, & Lin, 2014; Ryan, Giles, Bartolucci, & Henwood, 1986). The profound importance of language skills throughout the lifespan has promoted a plethora of research into language production and comprehension in old age (for reviews, Abrams & Farrell, 2011; Burke & Mackay, 1997; Burke & Shafto, 2008; Shafto & Tyler, 2014). The elderly are the fastest growing age group worldwide and are predicted to account for 21% of the global population by 2050 (Harper, 2014). This highlights the importance of conducting novel studies into language and ageing in order to better understand the changes that occur to language with age, and crucially to identify the mechanisms that underlie these changes. In the present study, we investigated the mechanisms underlying lexical and syntactic production in old age.

Current research indicates that language production is particularly vulnerable to the ageing process. For example, older adults experience increased tip-of-the-tongue states in which they cannot access the phonology associated with a particular word (e.g., Burke, MacKay, Worthley, & Wade, 1991; Segaert et al., 2018; Shafto, Burke, Stamatakis, Tam, & Tyler, 2007) and are more error-prone in picture naming tasks (see Feyereisen, 1997, for a meta-analytical review). There is also an age-related decrease in the use of complex syntactic structures, such as embedded clauses, coupled with an increase in syntactic errors, such as the use of the incorrect tense, in spontaneous written and spoken sentence production (Kemper, 1987; Kemper, Greiner, Marquis, Prenovost, & Mitzner, 2001; Kemper & Sumner, 2001; Rabaglia & Salthouse, 2011, Studies 1 and 2). Moreover, older adults are even more error-prone when forced to use more complex syntactic structures in constrained production tasks, such as when incorporating more words into a sentence (Kemper, Herman, & Lian, 2003), imitating sentences with initial relative clauses (Kemper, 1986) or producing left-branching sentences in a stem completion task (Rabaglia & Salthouse, 2011, Study 3). It has been suggested that the language decline experienced by older adults is caused by other emerging cognitive deficits, such as a decline in working memory capacity (Kemper & Sumner, 2001; Waters & Caplan, 2001), inhibitory control (Hasher & Zacks, 1988; Sommers & Danielson, 1999) and processing speed (Pickora-Fuller, 2003; Salthouse, 1993). There is also emerging evidence that levels of cardiovascular fitness may contribute to the level of the language decline in old age (Segaert et al., 2018).

Most previous studies investigating language production and ageing have primarily used off-line measures, involving the assessment and coding of sentences after they have been produced (e.g., the number of syntactic errors). However, off-line measures involve an element of conscious decision-making, which places increased demand on other cognitive resources, and because assessment occurs after the sentence output, they do not provide information about the time-course of the underlying sentence generation processes (Marinis, 2010; Mertins, 2016). The present study therefore employs on-line sentence production paradigms, which measure speech onset latencies, as well as accuracy. Moreover, we measured individual differences in memory, vocabulary, processing speed, inhibition and physical health to investigate whether changes in these functions impact upon sentence generation processes in old age. In the following introduction, we will first review current models of sentence generation before discussing the two on-line paradigms employed in this study to investigate lexical and syntactic processing during sentence production in healthy ageing.

Modelling sentence generation processes

Producing a fluent and coherent sentence is a complex task involving multiple stages and processes that must be executed quickly and effectively for successful communication. A speaker must first form a conceptual representation of the information that they wish to convey; this is often referred to as the *message* (Bock & Levelt, 1994; Levelt, 1989). The message then triggers the formulation stage in which the message is turned into linguistic representations – this involves both the rapid retrieval of lexical items and the generation of an appropriate syntactic structure, which must then be integrated correctly to convey the appropriate message. According to the more traditional models of sentence production, 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 an appropriate syntactic structure that matches the lemma roles (Bock & Levelt, 1994; Garrett, 1980; Levelt, Roelofs, & Meyer, 1999; Pickering & Branigan, 1998). Alternatively, more computational models propose that there is a complete dissociation between syntax generation and lexical retrieval because, in order to able to use words within novel sentences, syntactic structure must be derived solely from conceptual structure (i.e., thematic roles) with lexical access occurring independently (Chang, 2002; 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 remaining sentence (Kempen & Hoenkamp, 1987; Levelt, 1989, 1992). An incremental sentence production system is beneficial as chunks can be rapidly released when planning is complete for that part of the sentence; this enables the processing load to be spread effectively across multiple components and reduces the demand on memory as production is spread across time (Levelt, 1989; Wheeldon, 2013). This is particularly advantageous when conversing with others as it enables a speaker to maintain a rapid speech production rate (150-190 words per minute; Tauroza & Allison, 1990) and the preferred minimal turn-taking gaps between speakers (about 200 ms; Stivers et al., 2009).

Nevertheless, there is considerably less agreement about the scope of planning that must occur before sentence production. Although Bock and Levelt (1994) proposed that incremental planning occurs in clausal units, evidence indicates that planning likely occurs at a more minimal phrasal level (e.g., Martin, Crowther, Knight, Tamborello, & Yang, 2010; Martin, Yan, & Schnur, 2014; Smith & Wheeldon, 1999; Wheeldon, Ohlson, Ashby, & Gator, 2013) or even at a highly incremental word-by-word level (e.g., Griffin, 2001; Zhao & Yang, 2016). It is possible however that planning scope is not rigidly fixed as it can vary due to multiple factors including: linguistics factors, such as the ease of syntactic processing (Konopka, 2012; Konopka & Meyer, 2014); situational factors, such as time pressure and task complexity (Ferreira & Swets, 2002; Wagner, Jescheniak, & Schriefers, 2010); and cognitive factors, such as working memory and production speed (Martin, Miller, & Vu, 2004; Slevc, 2011; Swets, Jacovina, & Gerrig, 2014; Wagner et al., 2010). In this present study, we investigated the effect of old age on on-line sentence generation at the syntactic and lexical level of processing.

Investigating on-line sentence generation

Although a speaker does not plan all of a sentence before beginning articulation, some critical amount of planning is required to enable fluent sentence production (Levelt, 1989). The amount of time that a speaker takes to begin a sentence is therefore indicative of 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. Consequently, on-line speech onset latency measures are often used to examine specific sentence generation mechanisms and aspects of incremental planning. In order to ensure that participants generate sentences entirely independently, studies have typically used picture description paradigms. In this section, we will review two such paradigms in detail, and explain how these can be used to test different hypotheses about age-related changes in sentence planning processes.

Using syntactic priming to investigate on-line sentence generation. Syntactic priming (also known as structural priming or persistence) refers to the facilitation of syntactic processing that occurs when a syntactic structure is repeated across otherwise unrelated prime and target trials (Bock, 1986; Pickering & Ferreira, 2008). Such facilitation can surface in multiple ways at both the selection and planning level of sentence production. *Choice syntactic priming* is the phenomena whereby speakers are more likely to produce a dispreferred syntactic alternative, such as a passive sentence, if they have just processed a passive prime compared to an active prime (see Mahowald, James, Futrell, & Gibson, 2016, for a meta-analytical review). *Onset latency syntactic priming* is 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, Smith and Wheeldon (2001) demonstrated that when a speaker must produce a given syntactic structure on a target trial (as in 1a), recent production of the structure on a previous trial (1b) increased the speed with which it was subsequently reproduced, compared to if a different structure (1c) had just been produced. The facilitatory effect of structural priming on onset latencies is particularly evident when the primed syntax is the generally preferred alternative (e.g., actives; Segaert et al., 2011, 2016).

(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”

Only a few studies have examined the effect of old age on syntactic priming and these studies have used off-line choice methodologies. While two studies found preserved priming of passives in older adults (Hardy, Messenger, & Maylor, 2017; Heyselaar, Wheeldon, & Segaert, 2017), others have not (Heyselaar, Segaert, Walvoort, Kessels, & Hagoort, 2017, footnote 2; Sung, 2015).[1] To date, no study has investigated age effects on onset latency syntactic priming; this is a noticeable gap as a complete model of sentence generation must account for age effects on both how a syntactic structure is chosen, as well the processes by which it is planned and produced.

Most models have primarily focused on just explaining choice syntactic priming by use of residual activation (Pickering & Branigan, 1998) or implicit learning mechanisms (Chang et al., 2006). However, 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 nodes with competing activation levels, and there are two sequential stages of production: first selection of one of the syntactic alternatives; and second planning, during which the selected syntax is generated and produced incrementally. While syntactic choice is determined solely at the selection stage, production speed is determined by the total 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 investigate the effect of age on onset latency syntactic priming effects. We removed the choice element of production as this allows us to tap more directly in 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.

Using planning scope to investigate on-line sentence generation. As discussed above, incrementality is an important feature of sentence production that enables fluent communication while minimalizing cognitive load (Kempen & Hoenkamp, 1987; Levelt, 1989). We therefore employ a second paradigm that specifically targets aspects of incremental sentence production and can therefore provide additional insight into age-related changes in the integration of syntactic and lexical information during sentence generation. In the *planning scope* paradigm, picture displays are typically used to elicit sentences with

[1] 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.

different syntactic structures, particularly in terms of the size of the initial phrase, and speech onset latencies are used as a measure of the amount of advanced planning that has occurred prior to articulation (see Wheeldon, 2013, for a review). For example, Smith and Wheeldon (1999) found that participants took longer to produce sentences with larger initial coordinate phrases (as in 2a) compared to smaller initial simple phrases (as in 2b), indicating 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; Wheeldon et al., 2013). 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 effect cannot be attributed to the fact that the second content word in the simple initial phrase (always the verb 'moves') may be easier to retrieve than in the coordinate initial phrase (always the second lexical item) as phrasal planning scope has been demonstrated in Japanese and Mandarin, both head-final languages in which the verb is always last regardless of the initial phrase type (Allum & Wheeldon, 2007, 2009; Zhao, Alariom, & Yang, 2015).

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

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

The paradigm is therefore a useful tool for investigating incremental planning scope in older adults. Given the proposed importance of working memory to sentence planning (Slevc, 2011; Swets et al., 2014), it may be that older adults perform differently to young adults. Indeed, 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., showed a markedly greater difference in the speed of production of larger initial phrases compared to smaller phrases); the authors attributed this effect 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. To our knowledge, no study to date has directly examined the effect of old age on incremental planning scope.[2] It is possible that due to emerging deficits in working memory,

[2] We note that some previous studies of incremental sentence production in aphasic patients have used non-young adults as controls; however, once again the samples are often small and the age ranges are large (e.g., Lee, Yoshida, & Thompson, 2015; Martin et al., 2004; Scott & Wilshire, 2010; Speer & Wilshire, 2013). It is therefore difficult to draw any firm conclusions about the effect of old age on incremental planning from these studies.

older adults will display a larger phrasal complexity effect than young adults (although this effect is unlikely be as exaggerated as that seen in aphasic patients). Alternatively, older adults may unconsciously choose to engage in a more extreme word-by-word incremental strategy. Ferreira and Swets (2002) found that when a time pressure element was applied, speakers engaged in considerably less advanced planning compared to when there was no time pressure, suggesting that incremental planning can be strategically controlled by the speaker (although time pressure does not affect planning at the phonological level; Damian & Dumay, 2007). Therefore, older adults may adopt a more minimal sentence planning strategy than young adults in order to maintain an acceptable speed of speech output.

In order to directly investigate lexical processing during sentence production, some planning scope studies have also included a picture preview element. 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 prioritised 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 (e.g., *"[the drum, the star and the* ***hat*** *move] above the crab"*). Thus, it may be that advanced lexical planning only encompasses a subset of the required nouns. Wheeldon et al. (2013) speculated that this is because attempting to retrieve and hold all three nouns within an initial phrase prior to articulation may lead to problems with buffering and maintaining a linearized output. For young adults, their preferred scope of lexical encoding appeared to be two lexical items within the initial phrase; however, given that older adults are known to have a reduced memory buffer for holding linguistic information (e.g., word and sentence span; Bopp & Verhaeghen, 2005; Waters & Caplan, 2003), their preferred limit may be even less. In particular, if older speakers typically only encode the first lexical item within a phrase prior to articulation, then, unlike young adults, they may not display the preview benefit of the second lexical item within a larger initial phrase.

Another interesting question relates to the effect of the preview on sentence production when it falls outside of the first phrase. Wheeldon et al. (2013) found that while the preview was not beneficial outside of the first phrase for young adults, neither was it problematic, as onset latencies were similar to when there was no preview. Nevertheless, there is reason to believe that this may not necessarily be the case for older adults. When the preview picture falls outside of the first phrase, it essentially serves as an unhelpful distractor as the participant will access some lexical information about the picture, which they must then inhibit while planning the initial phrase. Young adults appear to be fairly good at this; however, older adults have been shown to have deficits in inhibitory control (e.g., West & Alain, 2000; Williams, Ponesse, Schachar, Logan, & Tannock, 1999). This would suggest that previewing a lexical item outside of an older adult's planning scope (whether that be restricted to the first noun or the first phrase) may interfere with their sentence planning, potentially resulting in a slowing of onset latencies and an increase in errors.

The present study

We report two experiments designed to measure age-related changes in sentence planning using on-line measures of sentence production. We aimed to investigate changes in both syntactic and lexical processing and to test the relationship between any changes observed with individual differences in cognitive and physical functioning. In Experiment 1, young and older adults completed an onset latency syntactic priming task (similar to Smith & Wheeldon, 2001) to test changes in syntactic facilitation effects with age. In Experiment 2, participants completed a planning scope task with a picture preview element (similar to Wheeldon et al., 2013) to test for age-related changes in syntactic and lexical planning scope. In addition, participants completed seven additional cognitive and physical measures that are used as markers of the 'healthy ageing phenotype' (Lara et al., 2013), and have been related to individual differences in the magnitude of age-related language decline. We aimed to use participants' performances on the ageing markers to understand whether any age-related changes found in the sentence production tasks are related to specific cognitive deficits or whether language changes are the result of a general ageing decline (cognitive slowing; Cerella, 1985; Salthouse, 1996).

METHOD

Participants

Fifty young adults (36 female) aged 18-25 (*M* = 19.8; *SD* = 1.1) were recruited from the University of Birmingham student population and were compensated with course credits. Fifty-six community-dwelling older adults (37 female) aged 64-80 (*M* = 71.8; *SD* = 4.5) were recruited from the departmental Patient and Lifespan Cognition Database and were compensated monetarily. All older adults scored at least 26 out of 30 (M = 27.4*;* SD = 1.3) on the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), indicating that they were currently experiencing healthy ageing. 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[3] between young (M = 6.0, SD = 0.14) and older adults (M = 5.8, SD = 1.3), $t(104) = 1.36$, $p = .178$. Due to the length of the study, testing was split into two 90 minute test sessions, separated by 1-2 weeks. The study was approved by the University of Birmingham Ethical Review Committee and informed written consent was obtained prior to the first test session.

Experiment 1: Syntactic priming task

Design and materials. 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 (see Figure 1).

To create the stimuli, we used 134 simple photographic pictures of everyday concrete objects. All picture names were mono- or disyllabic, and particular care was taken to ensure that the objects could be identified and named quickly and easily. We used 40 of the pictures to generate 40 picture pairs for the target trials. Each one of the 40 pictures appeared in the leftmost position of one pair and in the rightmost position of a different pair; this ensured that each picture made a balanced contribution to production latencies for words at both positions. 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. Care was taken to ensure that there was no phonological or conceptual overlap between any of the four pictures within each experimental item.

The movement of each picture pair was controlled using E-Prime (Schneider, Eschman, & Zuccolotto, 2002). In all of the 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"*). However, in the prime trials the

[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).

picture movements was varied in the different conditions. 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"*). By contrast, 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"*).

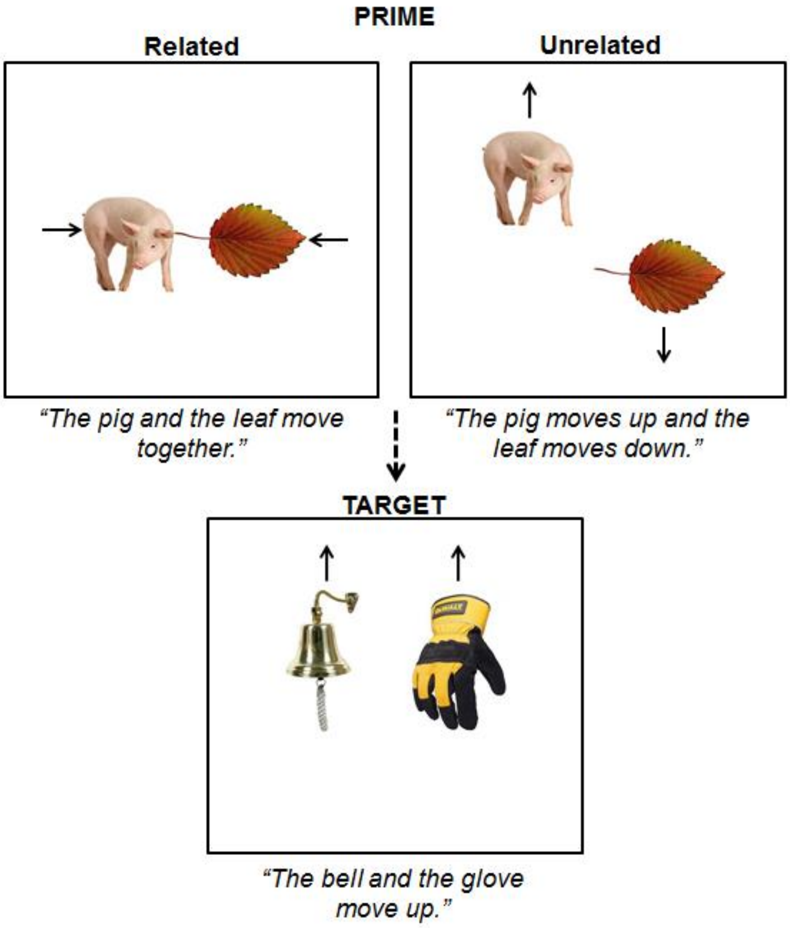

Figure 1. Syntactic priming task design.

We used a further 54 pictures to construct the filler trials. The main purpose of the fillers was to increase the variety of syntactic structures and therefore minimise the risk of the participant noticing the priming manipulation. We created 96 filler trials that elicited phrases such as: *"the X moves above the Y"* (one picture moves); *"there is an X and a Y"* (no picture movement); *"the Xs move up"* (two pictures of the same object move simultaneously) and *"there are no pictures"* (no pictures appear on screen). In addition, we created 24 filler trials that elicited phrases that were syntactically similar to the experimental trials. We reasoned that 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.

Lastly, we created two lists each containing 40 experimental items and 120 filler items. Both lists contained the same target and filler items; however, the prime item matched to each target item was rotated between the two lists, such that there were 20 related and 20 unrelated primes per list. We then divided each list into four blocks that each contained 10 *related* experimental items, 10 *unrelated* experimental items and 30 filler items. The distribution of items within each block was pseudorandomised with the constraint that two experimental items never occurred consecutively. The ordering of the blocks was rotated across participants.

Procedure. Each participant was randomly assigned to one of the two lists and was tested individually in a sound-attenuating booth facing the screen of a 17 inch *Dell* monitor. In front of the participant, there was a *Sony* microphone connected to an amplitude voice key that recorded his/her responses and onset latencies. The screen display consisted of a white background onto which the pictures were presented. Each trial began with a central fixation cross that remained on screen for 1000 ms; at its offset, the two pictures appeared on screen aligned centrally in the horizontal plane. The movement of the appropriate pictures began immediately in a smooth motion and was completed in 400 ms. The participant was instructed to describe the picture movements using specific sentence types and to begin doing so as soon as possible for each trial. The pictures were removed from the screen 1000 ms after response completion, or 4000 ms after the onset of the pictures if the participant did not produce a response. After an interval of 1500 ms, the next trail began.

To begin, there were 50 practice trials; each of the 80 experimental pictures appeared once in the practices and the sentences elicited resembled those in the experimental and filler trials. 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; use of a difference sentence structure; and disfluencies, such as stuttering and pausing.

Experiment 2: Planning scope task

Design and materials. 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 (see Figure 2). Critically, the previewed picture (always of the second upcoming lexical item) fell within the initial phrase in the coordinate condition, but fell outside of the initial phrase in the simple condition.

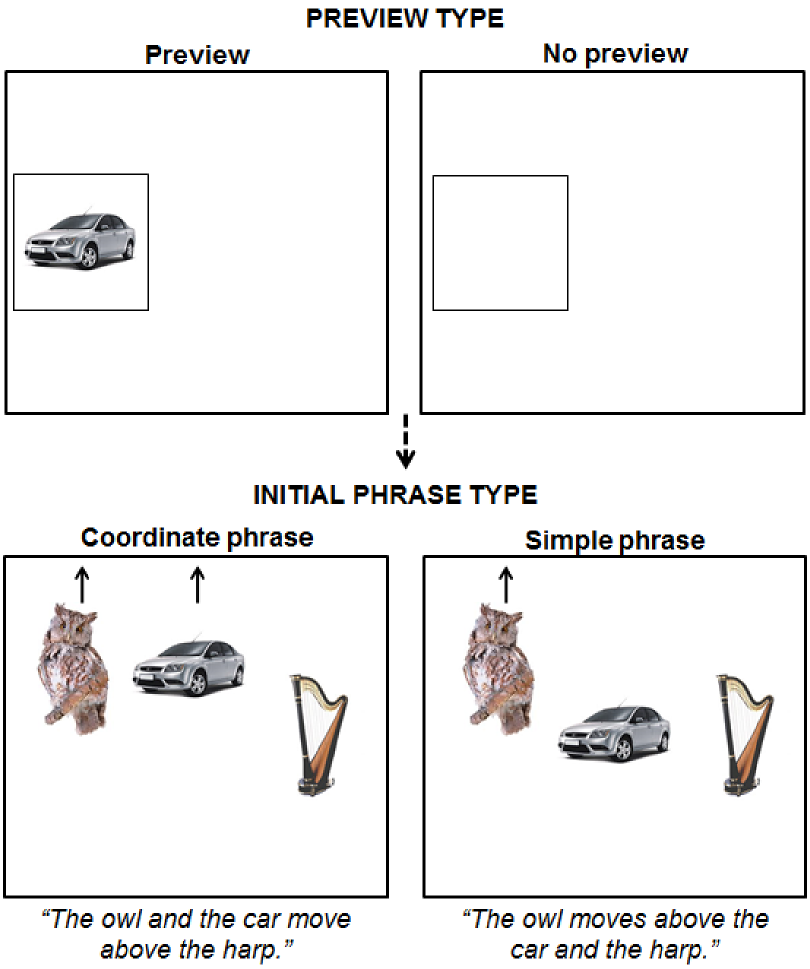

Figure 2. Planning scope task design.

For the stimuli we used 186 photographic pictures of everyday concrete objects; these pictures were different to those used in Experiment 1 but met the same criteria. We used 80 of the pictures to create the 80 experimental items, each consisting of three different pictures. Each of the 80 pictures appeared in the leftmost position in one item, in the central position in a second item, and in the rightmost position in a third item. We ensured that there was no phonological or conceptual overlap between the three pictures within each experimental item, and that each picture was combined with two different pictures in the three items in which it occurred. Similar to Experiment 1, the sentence descriptions of the items were elicited by controlling the movement of the three pictures (using E-prime) and participants were instructed to describe the picture movements from left to right using specific sentences. In the *simple* conditions, only the leftmost picture moved (either up or down) and the other two pictures remained stationary (*"the A moves above/below the B and the C"*). By contrast in the *coordinate* conditions, both the leftmost and the central picture moved simultaneously (either up or down) and only the rightmost 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 used the remaining 106 pictures to create 220 filler items. The purpose of the fillers was 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: *"the X, the Y and the Z move up"* (all pictures move simultaneously); *"there is an X, a Y and a Z"* (no picture movement); *"the Xs move up"* (two or three pictures of the same object move simultaneously); and *"there are no pictures"* (no pictures appear on screen). 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.

Lastly we constructed four lists, each containing 80 experimental items and 220 filler items. The same fillers were used in each list; however, the condition assigned to each experimental item was rotated across the four lists, such that there were 20 experimental items per condition within each list. We divided each list into five blocks that each contained 44 filler items and 16 experimental items (4 per condition). The distribution of the items within each block was pseudorandomised with the constraint that two experimental items never occurred consecutively. The ordering of the blocks was rotated across participants.

Procedure. Each participant was randomly assigned to one of the four lists and tested in the same sound-attenuated booth used in Experiment 1 using the same equipment. Similarly, the screen display consisted of a white background onto which the pictures were presented. Each trial began with a fixation cross on the far left of the screen that remained for 1000 ms. This was replaced by a black square frame that either did or did not contain a preview picture; this remained on screen for 1000 ms. 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 on screen aligned centrally in the horizontal plane; the movement of the appropriate pictures began immediately in a smooth motion and was completed in 400 ms. The participant was instructed to describe the picture movements using specific sentence types and to begin doing so as soon as possible for each trial. The pictures were removed from the screen 1000 ms after the response completion, or 4000 ms after the onset of the pictures if the participant did not produce a response. After an interval of 1500 ms, the next trial began.

To begin, there were 40 practice trials; each of the 80 experimental pictures appeared once in the practices and the sentences elicited resembled those in the experimental and filler trials. The task then continued until all five blocks had been completed. As in Experiment 1, the experimenter listened from outside of the booth and noted down any errors.

Ageing markers

Each participant completed seven additional measures, designed to provide an indicator of how healthily he/she was ageing across cognitive, physical and physiological domains (Lara et al., 2013).

There were five cognitive tasks. As a measure of processing speed, the participant completed the Coding task from the Wechsler Adult Intelligence Scale (WAIS-IV; Wechsler, 2008). As a measure of short- and long-term explicit memory, the participant completed an immediate and delayed recall of 12 unrelated words taken from the Wechsler Memory Scale (WMS-III; Wechsler, 1997). As a measure of vocabulary, the participant completed the multiple choice part of the Mill Hill vocabulary test (Raven, Raven, & Court, 1988). As a measure of working memory, the participant completed a backward digit span task (Waters & Caplan, 2003); the participant's score was defined as the span length at which he/she could correctly recall the digits in reverse order on three out of five trials (an additional 0.5 was added if the participant was correct on two out of the five trials). As a measure of inhibition, the participant completed a stop-signal task in which they were instructed to respond to a 'go' stimulus as quickly as possible, but to withhold their response if the stop-signal (a red cross) appeared immediately after the 'go' stimulus (Logan & Cowan, 1984; Verbruggen & Logan, 2008). The participant's stop signal reaction time (SSRT), the measure of inhibitory control, was calculated by subtracting their final stop-signal delay (the interval between the presentation of the 'go' stimulus and the stop-signal that was varied dynamically using an online tracking system) from their average response time to the 'go' stimulus; a smaller SSRT score indicated better inhibitory control (see Williams et al., 1999, for an extensive explanation).

In addition, there was one physical and one physiological task. As a measure of physical capability, the participant completed a handgrip strength task (Cooper et al., 2011;

Cooper, Kuh, & Hardy, 2010) using a standard adjustable *Jamar* hand dynamometer (Lafayette Instrument Company, USA). The participant was instructed to hold the dynamometer upwards with a completely outstretched arm and then to move their arm downwards while squeezing with maximum effort for three seconds; the highest value across six trials (three per hand) was used for analysis. As a measure of physiological health, the participant completed a lung capacity task, designed to measure forced expiratory volume in one second (FEV1; Pathan et al., 2011; Richards, Strachan, Hardy, Kuh, & Wadsworth, 2005), using a standard *In2itive* spirometer (Vitalograph, UK). The participant was asked to take a deep breath in to fill their lungs to full capacity, make an airtight seal around the mouthpiece, and then blow as hard as possible into the spirometer until their lungs were empty, aiming to blow for at least six seconds; the highest value across three trials was used for analysis.

Data preparation and analyses

For both experiments, we excluded the data of participants whose error rates were above 50% on the experimental trials. This resulted in exclusion of five older adults in Experiment 1, and one older adult in Experiment 2.

Of the 4040 target responses in Experiment 1, individual trials were excluded from the analysis if: the participant made an error in the corresponding prime trial; the onset latency was below 300 ms or above 3000 ms; or the onset latency was 2.5 SDs above or below each participant's mean per experimental condition. In total, 223 (11.1%) of the young adults' trials and 350 (17.2%) of the older adults' trials were discarded. Of the 8400 experimental trials in Experiment 2, individual trials were excluded according to the same criteria (apart from the prime error specification). This resulted in the discarding of 124 (3.1%) of the young adults' trials and 166 (3.8%) of the older adults' trials. For both experiments, all remaining trials were included in the error analyses, but only correct response trials (87.4% and 81.7% of all trials in Experiments 1 and 2 respectively) were included in the onset latency analyses.

The data from Experiments 1 and 2 were analysed separately, but following the same method. 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 (Baayen, Davidson, & Bates, 2008; 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 interprets (random intercepts) with additional random adjustments to the fixed effects (random slopes).

For the Experiment 1 models, we entered age group (young vs. old) and prime type (related vs. unrelated) as fixed effects. For Experiment 2, we entered age group (young vs. older), initial phrase type (coordinate vs. simple) and preview type (no preview vs. preview) as fixed effects. In all models, we included random intercepts for participants and items, as well as by-participant and by-item random slopes for the fixed effects. We entered the standardized scores of the following ageing markers into all models as additional predictors: processing speed; short-term memory; long-term memory; vocabulary; inhibition; and lung capacity. Raw processing speed scores were converted to age-scaled scores using the WAIS-IV manual. For all other predictors, the raw scores were converted into standardized z-scores centred around 0 within age groups (and gender groups for lung capacity). We did not enter handgrip and working memory scores as predictors into the model as we did not find the typical age-related declines in these measures (see Table 1); hence, including the measures would not be informative about the effect of age changes in these domains on the experimental task performance.[4] We also modelled the young and older adult data separately for both experiments as we had *a priori* hypotheses about age group differences. The Variance Inflation Factor (VIF) of all models was < 2.8, indicating that there was limited multicollinearity between predictors (Jaeger, 2011).

Prior to analysis of each model, the numerical predictors were centred and 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

[4] We speculate that the lack of age differences in the working memory may be because we only used one measure. Use of multiple measures can potentially provide a better indicator of overall working memory ability because different tasks target different facets of working memory (Oberauer, Süß, Wilhelm, & Wittman, 2003; Waters & Caplan, 2003). We speculate that the lack of age differences in the handgrip measure may be related to the surprisingly low amount of physical activity reported by our young adult sample. It is recommended that every adult (young and older) engages in at least 150 minutes of moderate and vigorous physical activity per week (NHS, 2011). Overall, the young adults reported below this amount (*M* = 115 mins; *SD* = 105 mins); whereas the older adults reported above this amount (*M* = 165 mins; *SD* = 115 mins).

model converged (Barr et al., 2013; Jaeger, 2008). To ascertain the contribution of the different ageing marker predictors to the model, we then performed a step-wise "best path" reduction procedure, removing predictor interactions and then predictor main effects to locate the simplest model that did not differ significantly from the full model in terms of variance explained (Barr et al., 2013). To do this we used the *drop1* function of the *lme4* package that compares the Akaike's Information Criterion (AIC) values of the full model to a model with one interaction or main effect removed. We did not remove interactions and main effects of the fixed-effects because these were part of the experimental design. Significance *p* values for the linear mixed-effects model were calculated using the *car* package (Fox & Weisberg, 2011). Finally, on occasions when we found null effects, particularly of interactions involved age group, we used Bayesian analysis to quantify the likelihood of our non-significant results (as recommended by Gallistel, 2009). Following Wagenmakers (2007), we constructed a null mixed-effects model that did not include the effect of interest and an alternative model that did include the effect. We then used Bayesian Information criterion (BIC) values of the models to estimate the Bayes factor (BF) as $e^{(\text{Alternative BIC} - \text{Null BIC})/2}$. Inverse *BF* values <1 favour the null hypothesis, whereas values >1 favour the alternative (Jarosz & Wiley, 2014).

Table 1: *Means and standard deviations of the ageing markers for young and older adults, and the results of comparisons between the age groups (independent samples t-tests)*

	Young		**Older**		**Comparison**	
Measure	*M*	*SD*	*M*	*SD*	*t*(104)	*p*
Processing speed	80.90	14.77	67.61	14.62	4.65	< .001
Short-term memory	7.62	1.66	5.66	1.27	6.86	< .001
Long-term memory	5.14	1.60	3.16	1.51	6.54	<.001
Working memory	5.43	1.45	5.18	1.19	0.98	.331
Vocabulary	16.82	2.72	23.79	3.03	-12.39	<.001
Inhibition	207.7	57.1	255.4	57.2	-4.29	< .001
Handgrip	28.54	8.10	27.82	8.92	0.43	.667
Lung capacity	3.46	0.75	2.26	0.66	8.82	< .001

RESULTS

Experiment 1: Examining the effect of ageing on on-line syntactic priming

Figure 3 summarises the target error rates and the onset latencies across the two prime conditions for young and older adults.

Error rates. The final best-fitting model of the error data is reported in Table 2A. Although older adults were significantly more error-prone than young adults (9.1% vs. 16.1%, $p < .001$), there was no main effect of prime type ($p = .432$) and no interaction between age group and prime type ($p = .487$); additional Bayesian analysis supported this lack of interaction (Inverse $BF = 0.02$). Further planned modelling of the young and older adult data separately also confirmed that prime type did not significantly affect error rates in either age group (Tables 2B and 2C). The final model also contained a significant interaction between prime type, age group and inhibition ($p = .029$); this indicates that participants' inhibitory control may have been affecting their error rates in the two prime conditions, but that this was different for young and older adults. The interaction between prime type and inhibition did not maintain significance when the young adults were considered separately ($p = .276$), but did for the older adults ($p = .033$). A median split of participants' scores revealed that older adults with poorer inhibition produced more errors in the unrelated than the related prime condition (17.9% vs. 13.1%), but those with better inhibitory control were more error-prone in the related prime condition (15.5% vs. 18.0%).

Onset latencies. The final best-fitting model of the onset latency data is reported in Table 3A. There was a main effect of prime type ($p < .001$), such that target responses were produced significantly quicker following related primes (953 ms) than following unrelated primes (994 ms), indicating an overall syntactic priming effect of 41 ms. Although older adults were significantly slower than young adults (898 ms vs. 1060 ms, $p < .001$), there was no interaction between age group and prime type ($p = .487$). This lack of interaction was supported by Bayesian analysis (Inverse $BF = 0.89$), suggesting that young and older adults were experiencing similar onset latency priming. Indeed, further planned modelling of the age groups separately confirmed that the priming effect was significant in both young (36 ms, 3.9% benefit, $p = .002$) and older (49 ms, 4.5% benefit, $p = .004$) adults (Tables 3B and 3C). The final model also included a main effect of long-term memory (as measured using delayed word recall, $p = .045$): a median split revealed that participants with better long-term memory skills were quicker overall than those with poorer skills (951 ms vs. 990 ms). However, this effect was no longer significant when the age groups were analysed separately.

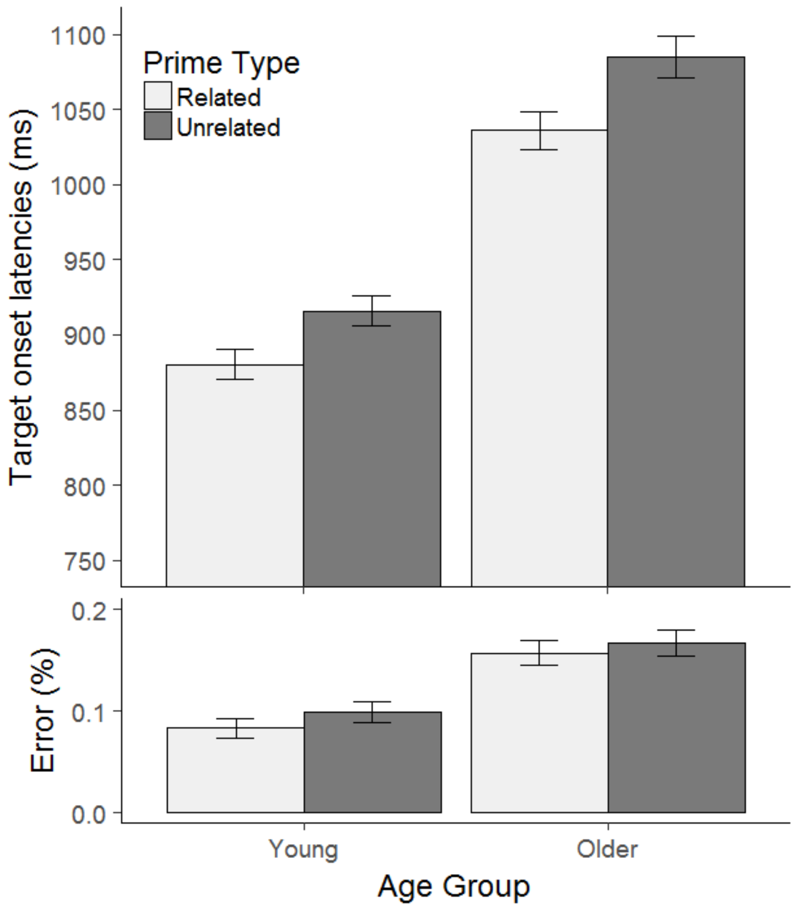

Figure 3. Mean target error rates and onset latencies for young and older adults following syntactically related and unrelated primes.

Table 2: *Summary of the best-fitted mixed-effects models for the Experiment 1 error data.*

Predictor	**Coefficient**	**SE**	**Wald Z**	***p***
A: all data†				
Intercept	2.33	0.16	14.71	< .001
Prime type	-0.11	0.14	-0.79	.432
Age group	0.76	0.20	3.77	< .001
Inhibition	0.06	0.08	0.77	.443
Prime type * Age group	-0.16	0.23	-0.70	.487
Prime type * Inhibition	-0.07	0.12	-0.63	.526
Age group * Inhibition	0.22	0.16	1.43	.153
Prime type * Age group * Inhibition	0.51	0.23	2.19	.029
B: young adults				
Intercept	2.70	0.20	13.83	< .001
Prime type	-0.09	0.23	-0.40	.691

Inhibition	0.17	0.12	1.44	.149
Prime type * Inhibition	0.19	0.18	1.09	.276
C: older adults				
Intercept	1.96	0.18	10.90	< .001
Prime type	-0.14	0.18	-0.79	.430
Inhibition	-0.05	0.10	-0.48	.632
Prime type * Inhibition	-0.33	0.15	-2.13	.033

† The final model did not differ significantly from the full model in terms of variance explained (full model AIC = 2476; final model AIC = 2450; *p* = .567). The model converged with random intercepts for participants and items with additional by-item and by-participant random slopes for the main effects of prime type and age group.

Table 3: *Summary of the best-fitted mixed-effects models for the Experiment 1 onset latency data.*

Predictor	**Coefficient**	**SE**	***t*-value**	***p***
A: all data†				
Intercept	982.97	23.47	41.88	< .001
Prime type	39.56	9.98	3.96	< .001
Age group	-173.23	42.63	-4.06	< .001
Long-term memory	-42.42	21.24	-2.00	.045
Prime type * Age group	-9.21	19.41	-0.47	.635
B: young adults				
Intercept	898.49	27.17	33.07	< .001
Prime type	35.05	11.55	3.03	.002
Long-term memory	-31.74	26.36	-1.20	.228
C: older adults				
Intercept	1074.01	35.81	29.99	< .001
Prime type	42.63	14.70	2.90	.004
Long-term memory	-55.57	33.53	-1.66	.097

† The final model did not differ significantly from the full model in terms of variance explained (full model AIC = 42178; final model AIC = 42143; *p* = .985). The model converged with a fully-expressed random intercepts and slopes structure.

Experiment 2: Examining the effects of ageing on on-line planning scope

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

Error rates. The final best-fitting model of the error data is reported in Table 4A – none of the ageing markers remained as predictors in the model following the "best path" reduction procedure. As in Experiment 1, older adults were significantly more error-prone than young adults (12.5% vs. 23.5%, $p < .001$). 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 sentence with initial simple phrases. Although the interaction between preview, initial phrase type and age group did not reach significance ($p = .285$), further planned analyses of the age groups separately indicated the experimental conditions may have had a greater effect on older adults' production of errors (Tables 4B and 4C). There were no main effects or interactions revealed in the young adult analysis, indicating that their error rates were fairly stable across all conditions. However, the older adult analysis revealed a significant interaction between preview and initial phrase type ($p = .016$): the presence of the preview resulted in an opposing patterns of errors when it fell within and outside of the initial phrase (as can be clearly seen in Figure 4). Further investigation revealed that the preview caused a significant 5.3% increase in older adults' errors in the simple condition ($\chi^2(1) = 7.08$, $p = .016$), but a 1.8% decrease in errors in the coordinate condition, although this difference was not significant ($\chi^2(1) = 0.53$, $p = .468$). [5]

Onset latencies. The final best-fitting model of the onset latency data is reported in Table 5A. As in Experiment 1, older adults were significantly slower than young adults (843 ms vs. 991 ms, $p < .001$).[6] There was also a main effect of initial phrase type, such that sentences with initial simple phrases were produced significantly quicker than sentences with

[5] The 'testInteractions' function in the *phia* package (de Rosario-Martinez, 2013) allows for the direct comparison of the contrasts specified within the model and adjusts *p* values for multiple comparisons using the Holm-Bonferroni correction.

[6] Due to the large speed differences between young and older adults, we also performed the modelling analysis with age-standardised 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.

larger initial coordinate phrases (895 ms vs. 935 ms, $p < .001$), indicating an overall phrasal planning effect of 40 ms. Moreover, the interaction between initial phrase type and age group was not significant (p =. 991; further supported by the Bayesian analysis, inverse $BF = 0.34$). This indicates that the incremental planning effect was similar for both young (40 ms, 4.6% benefit, $p < .001$) and older (41 ms, 4.0% benefit, $p < .001$) adults, as was further confirmed by the modelling of the age group separately (Tables 5B and 5C).

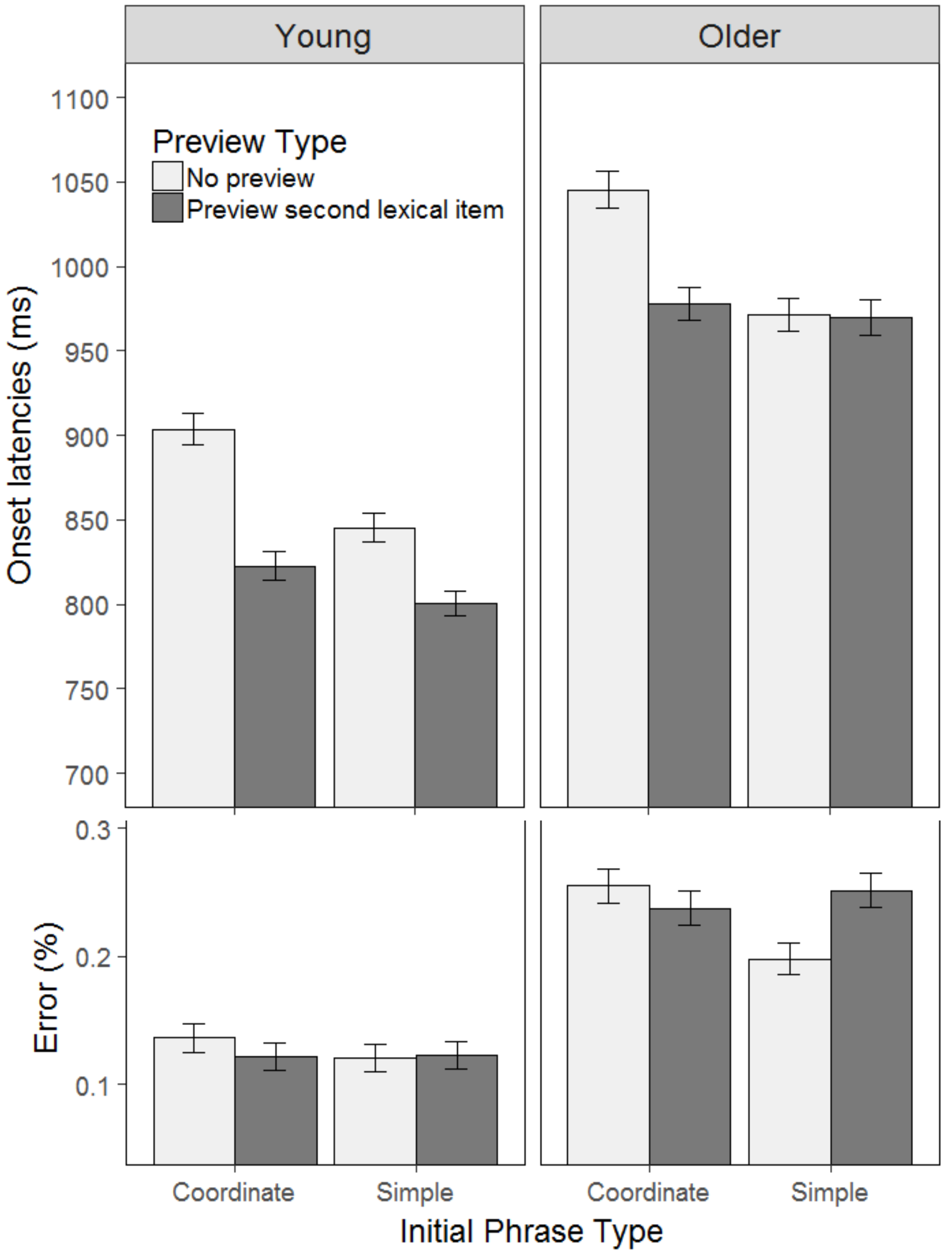

Figure 4. Mean error rates and onset latencies for young and older adults when producing sentences within initial coordinate and simple phrases following no preview or a preview of the second upcoming lexical item.

To now consider to the effect of preview on participants' onset latencies, there was a main effect of preview (see Table 5A), such that sentences were produced significantly quicker following preview of the second upcoming lexical item compared to no preview (940 ms vs. 980 ms, $p < .001$). Importantly, there was also a significant interaction between preview and initial phrase type ($p < .001$): while the overall preview benefit was 74 ms when it fell within the initial phrase (coordinate condition), this was reduced to 26 ms when it fell outside the initial phrase (simple condition). While the interaction between age group, preview and initial phrase type did not reach significance ($p = .250$), the additional Bayesian analysis indicates that there may be some age group differences as the inverses *BF* (1.28) was greater than 1. Indeed, although the interaction between preview and initial phrase type maintained significance when the young ($p = .011$) and older ($p < .001$) adult data was analysed separately (Tables 5B and 5C), inspection of Figure 4 would suggest that this interaction is representing something different for the two age groups. For young adults, while the preview benefit was greater when it fell within the initial phrase, the effect of preview was still significant in both the coordinate (81 ms (8.9%) preview benefit; $\chi^2(1) = 36.02$, $p < .001$) and simple (45 ms (5.3%) preview benefit; $\chi^2(1) = 12.90$, $p < .001$) conditions. By contrast, the difference in onset latencies between preview conditions was only significant for the older adults when it fell within the initial phrase (67 ms (6.4%) preview benefit; $\chi^2(1) = 20.57$, $p < .001$), but not outside of it (2 ms (0.2%) preview benefit; $\chi^2(1) = 0.43$, $p = .513$).

Interestingly, the final model (Table 5A) also contained the short-term memory predictor. Although there were no significant interactions involving short-term memory in the model, removing the predictor caused a significant change to the AIC values, suggesting it was accounting for some variance within the model (Barr et al., 2013). Significant effects involving short-term memory emerged when the age groups were analysed separately (Tables 5B and 5C). For the young adults, there was a significant interaction between preview and short-term memory ($p = .025$), such that those who scored above the group average on the short-term memory task displayed a greater preview effect overall than those who scored below average (93 ms and 36ms respectively). Whereas for the older adults, there was a significant 3-way interaction between preview, initial phrase type and short-term memory ($p = .045$). Closer inspection of the data would suggest that older adults who scored above average on the short-term memory task displayed a larger preview benefit within the initial phrase than those with below-average short-term memory (97ms and 25ms respectively).

Table 4: *Summary of the best-fitted mixed-effects model of the Experiment 2 error data.*

Predictor	**Coefficient**	**SE**	**Wald Z**	***p***
A: all 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	.098
Age group	0.89	0.16	5.70	< .001
Preview * Initial phrase type	-0.28	0.14	-2.05	.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: young adults				
Intercept	2.50	0.17	14.65	< .001
Preview type	-0.06	0.12	-0.51	.607
Initial phrase type	0.14	0.12	1.16	.245
Preview * Initial phrase type	-0.20	0.21	-0.93	.352
C: older adults				
Intercept	1.59	0.17	9.18	< .001
Preview type	-0.12	0.09	-1.43	.154
Initial phrase type	0.12	0.09	1.39	.163
Preview * Initial phrase type	-0.41	0.17	-2.51	.016

† The final model did not differ significantly from the full model in terms of variance explained (full model AIC = 6402; final model AIC = 6356; p = .756). The model 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 age group.

Table 5: *Summary of the best-fitted mixed-effects model for the Experiment 2 onset latency data.*

Predictor	**Coefficient**	**SE**	***t*-value**	***p***
A: all data†				
Intercept	924.27	17.03	54.30	< .001
Preview	-52.68	8.68	-6.07	< .001
Initial phrase type	-39.80	5.55	-7.18	< .001
Age group	-150.02	32.18	-4.66	< .001
Short-term memory	-5.15	16.06	-0.32	.662
Preview * Initial phrase type	49.43	10.80	4.57	< .001
Preview * Age group	-22.97	16.74	-1.37	.170
Initial phrase type * Age group	-0.28	11.05	-0.03	.991
Preview * Short-term memory	-12.28	8.57	-1.43	.162
Phrase type * Short-term memory	-0.55	5.55	-0.10	.947
Age group * Short-term memory	-19.45	32.11	-0.61	.611
Preview * Initial phrase type * Age group	-24.50	21.52	-1.14	.250
Preview * Initial phrase type * Short-term memory	13.97	10.81	1.29	.211
Preview * Age group * Short-term memory	-9.87	17.20	-0.57	.555
Initial phrase type * Age group * Short-term memory	17.39	11.10	1.57	.117
Preview * Initial phrase type * Age group * Short-term memory	-36.79	21.61	-1.70	.089
B: young adults				
Intercept	849.18	19.41	43.74	< .001
Preview	-63.06	10.62	-5.94	< .001
Initial phrase type	-41.19	7.77	-5.30	< .001
Short-term memory	-13.47	18.82	-0.72	.239
Preview * Initial phrase type	38.03	15.01	2.53	.011
Preview * Short-term memory	-23.30	10.46	-2.23	.025
Phrase * Short-term memory	7.97	7.70	1.04	.301

Preview * Initial phrase type * Short-term memory	-3.45	14.42	-0.24	.881
C: older adults				
Intercept	998.68	26.20	38.12	<.001
Preview	-41.30	13.53	-3.05	.003
Initial phrase type	-40.64	8.51	-4.78	< .001
Short-term memory	4.22	25.44	0.17	.809
Preview * Initial phrase type	61.87	16.59	3.73	< .001
Preview * Short-term memory	-7.86	16.59	3.73	.568
Phrase * Short-term memory	-9.32	8.14	-1.14	.253
Preview * Initial phrase type * Short-term memory	32.38	16.14	2.01	.045

† The final model did not differ significantly from the full model in terms of variance explained (full model AIC = 90826; final model AIC = 90784; p = .536). The model 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 and age group.

DISCUSSION

In an on-line sentence production study, we investigated age-related changes in the syntactic and lexical processes involved in sentence generation. In Experiment 1, we found evidence of onset latency syntactic priming in young and older adults: participants produced target sentences quicker following recent production of syntactically related primes. This demonstrates that speed benefits of syntactic priming are preserved with age. In Experiment 2, we found that young and older adults initiated sentences quicker with smaller initial phrases compared to larger initial phrases. This demonstrates that planning scope for both age groups was influenced by phrasal structure, suggesting that the breath of incremental planning, at least at the syntactic level, is relatively unaffected by normal ageing. Age differences did emerge, however, in the preview conditions in Experiment 2. Whereas young adults displayed speed benefits of preview when it 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 significantly more dysfluent and error-prone. This suggests that age differences occur in the flexibility of lexical retrieval during sentence planning and the ability to integrate lexical information into syntactic structures. Taking Experiments 1 and 2 together, our study therefore demonstrates age effects of lexical, but not syntactic, processes on the speed and accuracy of sentence production.

Our robust finding of preserved onset latency syntactic priming in older adults is in line with previous research that has found no evidence for age effects on choice syntactic priming (Hardy et al., 2017; Heyselaar, Wheeldon, et al., 2017). Syntactic priming therefore facilitates both older adults' syntactic choices and speed of sentence planning. This is consistent with evidence that older adults maintain the benefit from priming in other areas of language processing. For example, both young and older adults experience benefits of semantic priming when performing a lexical decision task (see Laver & Burke, 1993, for a meta-analytical review), and prior processing of phonological-related primes decreases the number of tip-of-the-tongue states produced by older adults (James & Burke, 2000). Critically, older adults did not show a greater magnitude of facilitation than younger adults, despite being significantly slower and more error-prone in their sentence production. This stability in the syntactic priming effect with age could be attributable to either an age-related preservation of the residual activation and/or the implicit learning processes underlying syntactic priming (Chang et al., 2006; Pickering & Branigan, 1998). Whichever the underlying mechanisms, our experiment provides the first evidence that onset latency priming effects are preserved with age in a task specifically designed to tap into syntactic planning processes (Segaert et al., 2016).

In Experiment 2, the pattern of onset latencies and error rates observed in the older adults is similarly consistent with a preservation of syntactic planning skills, but there is also an age-related decline in lexical planning. These findings are particularly relevant to studies of sentence production in aphasia patients that often use non-young adults as controls without fully considering the effect of age on participants' linguistic processes (e.g., Cho-Reyes et al., 2016; Lee et al., 2015; Martin et al., 2004). At the syntactic level, we found robust evidence of a phrasal scope of planning in both age groups (i.e., speakers took longer to initiate sentences with larger initial phrases). This finding replicates previous research in young adults (e.g., Allum & Wheeldon, 2007; Martin et al., 2010, 2014; Smith & Wheeldon, 1999), and suggests that both age groups were prioritising 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) or engage in a more extreme word-by-word planning strategy (if this was the case, onset latencies would have been similar for sentence with simple and coordinate initial phrases). This indicates that older adults still maintain sufficient cognitive capacity to support the planning of an initial phrase containing at least two nouns. Engaging in such an incremental planning strategy is beneficial to older speakers as it allows for the fast release of formulated chunks of an utterance, enabling them to maintain a rapid and fluent communication rate (Levelt, 1989; Wheeldon, 2013). Nevertheless, it remains unclear whether older adults would continue to engage with a phrasal planning scope when producing a sentence containing a more complex initial phrase. Wheeldon et al. (2013) found that young adults continued to plan a phrasal level when the initial phrase consisted of three nouns. However, it is possible that three or more nouns may disrupt older adults' sentence planning as age-related effects did emerge due to the picture preview manipulation.

In Experiment 2, half the experimental trials were preceded by a picture of the upcoming second lexical item. When the previewed picture fell within the initial phrase (e.g., *"[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 (Allum & Wheeldon, 2009; Wheeldon et al., 2013). Moreover, we found that participants who scored above their age group average in the short-term memory task (immediate word recall) displayed a larger preview benefit than those who scored below average (evident across both initial phrase types for young adults, but only within the initial phrase for older adults). This is consistent with an influence of short-term memory on lexical planning scope, such that speakers with greater short-term memory capacity can retrieve and maintain more lexical items prior to articulation and are therefore better able to benefit more from the preview of the second to-be-produced word. Indeed, the ability to store and manage information in short-term memory has been regularly linked to language processing (e.g., Baddeley, 2003; MacDonald, 2016; Swets, 2015).

Nevertheless, age group differences did emerge in young and older adults' onset latencies and error rates when the previewed picture fell outside of the initial phrase (e.g., *"[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 phrase. This demonstrates that the young adults did prioritise the retrieval of lexical items within the first phrase prior to speech onset, but that they were also able to successfully manage the early activation of lexical items outside of their usual phrasal planning scope to benefit the overall speed of their sentence production. This indicates that there is flexibility within young adults' planning scope, such that they can adapt in situations in which additional lexical information is provided – this adds to the growing evidence that planning scope is flexible and can be influenced by the ease of syntactic and lexical processing (Ganushchak & Chen, 2016; 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 actually caused them to become significantly more error-prone. The onset latency and error data together therefore suggest that, unlike the young adults, the older adults did not benefit from premature access to lexical information and that this premature availability actually had a significant disruptive effect on their overall fluency. One explanation for this age 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. Evidence from comprehension indicates that older adults are more sensitive to phrasal boundaries than young adults, such that they show a stronger preference for segmenting sentences at phrasal pauses (Payne & Stine-Morrow, 2012, 2014; Stine-Morrow & Payne, 2016). This strong preference for phrasal segmentation may also apply to older speakers' sentence production, meaning that they are less able to successfully incorporate lexical information outside of the initial phrase into their sentence planning.

A second explanation for the age differences in lexical processing we found in Experiment 2 relates to the executive control required to successfully manage the premature access to lexical information. When there was a preview picture, the participant would have automatically accessed some lexical information about the item and this would then have been stored in their working memory. However, when the preview picture did not appear within the first phrase, the participant would have needed to momentarily inhibit this information to order to ensure that it did not interfere with their initial sentence planning (i.e., the retrieval of the first lexical item that was not previewed). Theoretical accounts of inhibition propose that ageing weakens the inhibitory processes that are responsible for regulating what information enters and leaves working memory, and a consequence of age-related declines 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 of sentence processing, such as older adults experiencing increased difficulty with ignoring visually-distracting information while reading (Connelly, Hasher, & Zacks, 1991), and naming pictures in the presence of semantically-related distractors (Taylor & Burke, 2002). Deficits in inhibitory control may also therefore provide a valid explanation of our findings as, if the older adults were less able to regulate the storage of information relating to the preview picture during sentence planning, this would have led to increased problems with buffering and maintaining a linearized output, resulting in increased errors and dysfluencies.

Against this explanation, however, is our lack of evidence of an effect on inhibitory control on participants' error rates and onset latencies in the planning scope task. Nevertheless, we did only use one measure of inhibition (a stop-signal task), and measuring individual differences within a factorial design is acknowledged to be inherently difficult (Hedge, Powell, & Sumner, 2017). Both of these factors may have limited the reliability of our measurement of inhibition and the sensitivity of our analyses to identify any effects of inhibition with the mixed-effects models. The administration of a more comprehensive battery of inhibition measures is required to gain a more accurate understanding of the effect of inhibition on age-related changes in syntactic and lexical planning scope. Future work is also needed to fully understand the disruption that occurs to older adults' lexical processes during sentence planning, particularly as to what extent the activation of difference lexical items causes interference during sentence production.

In summary, our study is the first to specifically examine on-line sentence production in older adults; our findings should therefore be considered in parallel with off-line studies of language and ageing in order to gain a more complete understanding of the effect of old age on language processing. Specifically, our study provides evidence for the age-related preservation of syntactic processing during on-line sentence production, but an increased difficulty with lexical retrieval and management with age. We interpret this age-related decline in lexical processing to be attributable 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 and/or to declines in executive control that mean that older speakers are less able to cope with premature lexical activation beyond the first phrase.

ACKNOWLEDGEMENTS

We gratefully acknowledge the help of Denise Clissett, the coordinator of the Patient and Lifespan Cognition Database. This research was partly supported by an Economic and Social Research Council (ESRC) studentship awarded to SMH from the University of Birmingham Doctoral Training Centre (grant number: ES/J50001X/1).

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

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

Baayen, H., Davidson, D. J., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

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

Burke, D. M., & Mackay, D. G. (1997). Memory, language, and ageing. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *352*(1363), 1845–1856. https://doi.org/10.1098/rstb.1997.0170

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

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin*, *98*(1), 67–83. https://doi.org/10.1037/0033-2909.98.1.67

Chang, F. (2002). Symbolically speaking: A connectionist model of sentence production. *Cognitive Science*, *26*(5), 609–651. https://doi.org/10.1016/S0364-0213(02)00079-4

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

Connelly, S. L., Hasher, L., & Zacks, R. T. (1991). Age and reading: The impact of distraction. *Psychology and Aging*, *6*(4), 533–541. https://doi.org/10.1037//0882-7974.6.4.533

Cooper, R., Hardy, R., Sayer, A. A., Ben-Shlomo, Y., Birnie, K., Cooper, C., … Kuh, D. (2011). Age and gender differences in physical capability levels from mid-life onwards: The harmonisation and meta-analysis of data from eight UK cohort studies. *PLos ONE*, *6*(11), e27899. https://doi.org/10.1371/journal.pone.0027899

Cooper, R., Kuh, D., & Hardy, R. (2010). Objectively measured physical capability levels and mortality: systematic review and meta-analysis. *British Medical Journal*, *341*, c4467. https://doi.org/10.1136/bmj.c4467

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

Damian, M. F., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language*, *57*(2), 195–209. https://doi.org/10.1016/j.jml.2006.11.001

de Rosario-Martinez, H. (2013). Analysing interactions of fitted models. *CiteSeer*, 1–30. https://doi.org/10.1007/s13398-014-0173-7.2

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

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

Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review*, *116*(2), 439–453. https://doi.org/10.1037/a0015251

Ganushchak, L. Y., & Chen, Y. (2016). Incrementality in planning of speech during speaking and reading aloud: Evidence from eye-tracking. *Frontiers in Psychology*, *7*, 33. https://doi.org/10.3389/fpsyg.2016.00033

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (pp. 117–220). London: Academic Press.

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

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

Harper, S. (2014). Economic and social implications of aging societies. *Science*, *346*(6209), 587–591. https://doi.org/10.1126/science.1254405

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. *Psychology of Learning and Motivation - Advances in Research and Theory*, *22*, 193–225. https://doi.org/10.1016/S0079-7421(08)60041-9

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. (2017). 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

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/

James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1378–1391. https://doi.org/10.1037//0278-7393.26.6.1378

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting bayes factors. *The Journal of Problem Solving*, *7*(1), 2–9. https://doi.org/10.7771/1932-6246.1167

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. (1986). Imitation of complex syntactic constructions by elderly adults. *Applied Psycholinguistics*, *7*(3), 277–288. https://doi.org/10.1017/S0142716400007578

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., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging*, *16*(2), 312–22. https://doi.org/http://dx.doi.org/10.1037/0882-7974.16.2.312

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

Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychology and Aging*, *8*(1), 34–43. https://doi.org/10.1037/0882-7974.8.1.34

le Dorze, G., & Brassard, C. (1995). A description of the consequences of aphasia on aphasic persons and their relatives and friends, based on the WHO model of chronic diseases. *Aphasiology*, *9*(3), 239–255. https://doi.org/10.1080/02687039508248198

Lee, J., Yoshida, M., & Thompson, C. K. (2015). Grammatical planning units during real-time sentence production in speakers with agrammatic aphasia and healthy speakers. *Journal of Speech Language and Hearing Research*, *58*(4), 1182. https://doi.org/10.1044/2015_JSLHR-L-14-0250

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. (2016). Speak, act, remember: The language-production basis of serial order and maintenance in verbal memory. *Current Directions in Psychological Science*, *25*(1), 47–53. https://doi.org/10.1177/0963721415620776

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

Mehl, M. R., Vazire, S., Ramirez-Esparza, N., Slatcher, R. B., & Pennebaker, J. W. (2007). Are women really more talkative than men? *Science*, *317*(5834), 82.

https://doi.org/10.1126/science.1139940

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.

Mick, P., Kawachi, I., & Lin, F. R. (2014). The association between hearing loss and social isolation in older adults. *Otolaryngology - Head and Neck Surgery*, *150*(3), 378–384. https://doi.org/10.1177/0194599813518021

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

NHS. (2011). *Start active, stay active: A report on physical activity for health from the four home countries' Chief Medical Officers*. Retrieved from https://www.sportengland.org/media/2928/dh_128210.pdf

Oberauer, K., Süß, H. M., Wilhelm, O., & Wittman, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence, 31*(2), 167–193. https://doi.org/10.1016/S0160-2896(02)00115-0

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

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

Pickora-Fuller, M. K. (2003). Processing speed and timing in aging adults: Psychoacoustics, speech perception, and comprehension. *International Journal of Audiology*, *42*(1), 59–67. https://doi.org/https://doi.org/10.3109/14992020309074625

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

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/

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

Richards, M., Strachan, D., Hardy, R., Kuh, D., & Wadsworth, M. (2005). Lung function and cognitive ability in a longitudinal birth cohort study. *Psychosomatic Medicine*, *67*, 602–608. https://doi.org/10.1097/01.psy.0000170337.51848.68

Rozenthuler, S. (2013). *Life-changing conversations: 7 strategies for talking about what matters most*. London: Watkins Publishing.

Ryan, E. B., Giles, H., Bartolucci, G., & Henwood, K. (1986). Psycholinguistic and social psychological components of communication by and with the elderly. *Language and Communication*, *6*(1), 1–24. https://doi.org/10.1016/0271-5309(86)90002-9

Salthouse, T. A. (1993). Speed and knowledge as determinants of adult age differences in verbal tasks. *Journal of Gerontology: Psychological Sciences*, *48*(1), 29–36. https://doi.org/10.1093/geronj/48.1.P29

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

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

Scott, R. M., & Wilshire, C. E. (2010). Lexical competition for production in a case of nonfluent aphasia: Converging evidence from four different tasks. *Cognitive Neuropsychology*, *27*(6), 505–538. https://doi.org/10.1080/02643294.2011.598853

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

Shafto, M. A., & Tyler, L. K. (2014). Language in the aging brain: The network dynamics of cognitive decline and preservation. *Science*, *346*(6209), 583–587. https://doi.org/10.1126/science.1254404

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

Speer, P., & Wilshire, C. E. (2013). What's in a sentence? The crucial role of lexical content in sentence production in nonfluent aphasia. *Cognitive Neuropsychology*, *30*(7–8), 507–543. https://doi.org/10.1080/02643294.2013.876398

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

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., … Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation.

Proceedings of the National Academy of Sciences of the United States of America, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

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. (2015). Psycholinguistics and planning: A focus on individual differences. In S. Fuchs, D. Pape, C. Petrone, & P. Perrier (Eds.), *Speech production and perception: Individual differences in speech production and perception* (pp. 89–122). Frankfurt, Germany: Peter Lang International Academic Publishers.

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*(2014), 12–44. https://doi.org/10.1017/langcog.2013.5

Tauroza, S., & Allison, D. (1990). Speech rates in British English. *Applied Linguistics*, *11*(1), 90–105. https://doi.org/10.1093/applin/11.1.90

Taylor, J. K., & Burke, D. M. (2002). Asymmetric aging effects on semantic and phonological processes: naming in the picture-word interference task. *Psychology and Aging*, *17*(4), 662–676. https://doi.org/10.1037/0882-7974.17.4.662

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

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, *12*(11), 418–424. https://doi.org/10.1016/j.tics.2008.07.005

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

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. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging*, *16*(1), 128–144.

https://doi.org/10.1037//0882-7974.16.1.128

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.

West, R., & Alain, C. (2000). Age-related decline in inhibitory control contributes to the increased Stroop effect observed in older adults. *Psychophysiology*, *37*(2), 179–189. https://doi.org/10.1017/S0048577200981460

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

Zhao, L.-M., Alariom, F.-X., & Yang, Y.-F. (2015). Grammatical planning scope in sentence production: Further evidence for the functional phrase hypothesis. *Applied Psycholinguistics*, *36*(5), 1059–1075. https://doi.org/10.1017/S0142716414000046

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