Evidence against preserved syntactic comprehension in healthy aging

Charlotte Poulisse[1], Linda Wheeldon[2], Katrien Segaert[1,3]

[1]School of Psychology, University of Birmingham, Birmingham, United Kingdom

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

[3]Centre for Human Brain Health, University of Birmingham, Birmingham, United Kingdom

Contact: **K.Segaert@bham.ac.uk**

Abstract

We investigated age-related differences in syntactic comprehension in young and older adults. Previous research found no evidence of age-related decline in syntactic processing. We investigated elementary syntactic comprehension of minimal sentences (e.g I cook), minimizing the influence of working memory. We also investigated the contribution of semantic processing by comparing sentences containing real verbs (e.g I cook) versus pseudo-verbs (e.g I spuff). We measured speed and accuracy of detecting syntactic agreement errors (e.g., I cooks, I spuffs). We found that older adults were slower and less accurate than younger adults in detecting syntactic agreement errors for both real and pseudo-verb sentences. We thus found evidence for age-related decline in syntactic comprehension. The age-related decline in accuracy was smaller for the pseudo-verb sentences, and the decline in speed was larger for the pseudo-verb sentences, compared to real verb sentences. We suggest that syntactic comprehension decline is stronger in the absence of semantic information, which causes older adults to produce slower responses in order to make more accurate decisions. In line with these findings, accuracy for older adults was positively related to a measure of processing speed capacity. Taken together, we found evidence that elementary syntactic processing abilities decline in healthy ageing.

Keywords: healthy ageing, syntactic comprehension, processing speed, working memory

Introduction

Syntactic processing is often discussed in the literature as a key example of a cognitive function that is relatively resilient to age-related decline (Campbell et al., 2016; Samu et al., 2017; Shafto & Tyler, 2014). Studies investigating the effect of age on syntactic comprehension typically use sentences with a complex syntactic structure, such as garden path sentences with a temporary syntactic ambiguity (Samu et al., 2017), or relative clause manipulations that require disambiguation of referential choices (Payne et al., 2014). The interpretation of such complex syntactic structures does not exclusively rely on syntax, but also requires additional comprehension mechanisms including semantic and pragmatic processing. Consequently, such measures of complex sentence processing may not be ideal for measuring syntactic comprehension as an isolated component. Furthermore, processing sentences with a complex syntactic structure imposes a larger burden on working memory and processing speed. Given that age is associated with declines in working memory and processing speed (G. Waters & Caplan, 2007) the use of such computationally expensive sentences is problematic. In the present work, we aim to address these issues by reducing the complexity of our stimuli to simple two word sentences, in order to investigate comprehension of elementary syntactic structures. Consequently, contextual cues and working memory load are kept to a minimum. Moreover, we compare these elementary syntactic operations in word and pseudo-word sentences in order to investigate the contribution of meaning to syntactic comprehension. Lastly, we investigate whether individual differences in working memory, processing speed and physical health impact on decline in syntactic comprehension in healthy ageing.

Syntactic comprehension

Syntax plays a fundamental role in understanding spoken language. Syntactic information, in addition to other types of information, enables the listener to extract meaning in the incoming speech input. Syntactic processes are used in sentence comprehension in a number of ways, including structure building (e.g. combining words into larger units based on grammar rules and

word category information) and checking agreement (e.g. in English, the verb needs to agree in number and person with the subject; (Kaan & Swaab, 2002). Furthermore, syntax plays an important role in mapping thematic roles (e.g. mapping the agent ('doer') and patient ('doe-ee') onto certain positions in the sentence). The order of noun phrases to thematic role mapping strongly influences the complexity of the sentence structure and the number of syntactic operations needed to determine the meaning of a sentence. In sum, the level of syntactic processing required to understand spoken language can range from rather simple to very complex.

A considerable amount of research has focused on whether there is age-related decline in sentence comprehension. The emphasis in this line of research tends to be on complex sentence structures. Using a paradigm that capitalizes on syntactic ambiguity, Tyler and colleagues investigated syntactic processing during sentence comprehension in younger and older adults, in sentences varying in the level of syntactic processing required (Campbell et al., 2016; Davis, Zhuang, Wright, & Tyler, 2014; Meunier, Stamatakis, & Tyler, 2014; Samu et al., 2017; M. A. Shafto & Tyler, 2014; M. a Shafto et al., 2014). Specifically, unambiguous sentences have only one possible syntactic interpretation (e.g. …"*sneering boys*"), whereas ambiguous sentences have two possible interpretations: an interpretation that, given its higher frequency in the language is dominant, or more expected (e.g. ….."*cooking apples are*"), or a subordinate, less expected interpretation (e.g. …"*cooking apples is*"). Participants are asked to indicate whether the disambiguating word (*are* or *is* in the examples) is an acceptable or an unacceptable continuation of the sentence. For individuals without any language disorders, a conventional pattern of responding is to reject more (and respond more slowly to) subordinate sentences compared to dominant and unambiguous sentences, with little difference between the latter two sentences (Campbell et al., 2016). Tyler and colleagues repeatedly found no age-related differences in acceptability ratings (Davis et al., 2014; Meunier et al., 2014), or response times, in which the mean response time difference between the sentences requiring the most and the least syntactic processing (subordinate and unambiguous sentences) is used (Campbell et al., 2016; Samu et al., 2017). Another line of research has measured online syntactic processing with a

word-monitoring task to investigate younger and older adults' ability to develop syntactically and semantically coherent representations (Tyler et al., 2010). Participants listened to sentences and were instructed to press a response key whenever they heard a pre-specified target word. Word position of the target word varied from early to late across the sentences. The sentences differentially load on syntactic and semantic processing: normal prose sentences had a normal syntactic, semantic and pragmatic structure; anomalous prose sentences had a correct grammatical structure but lacked sentential meaning, and randomly ordered word strings lacked grammatical and sentential meaning. Response times increased at later word positions in both normal and anomalous prose. Comparing a group of young and older adults, this pattern of word position effects showed no age-related performance differences. Taken together, these results suggest that syntactic comprehension is preserved in the late years of adult life. However, all these studies have placed complex syntactic structures at the forefront. Since the manipulations in these studies potentially do not exclusively investigate the contribution of syntactic processes, is unclear to what extent the performance for processing these sentences also reflects additional (linguistic- and pragmatic) comprehension mechanisms.

Moreover, even though in a large number of studies it is concluded that syntactic comprehension performance is preserved in healthy ageing, there are also several studies that have found age-related syntactic comprehension decline. Obler, Fein, Nicholas, & Albert (1991) investigated age-related differences in the effect of syntactic complexity and semantic plausibility on sentence comprehension. Participants listened to sentences that were divided into six different syntactic types (active, passive, single negative, double negative, double embedded or comparative). Accuracy showed a general age-related decline and older adults were disproportionally less accurate at the harder sentenced types. In a sentence picture matching paradigm with sentences of increasing syntactic complexity Antonenko et al. (2013) found superior syntactic performance in younger compared to older adults. The paradigm consisted of sentences with three different levels of syntactic complexity. The easiest level did not have hierarchical embeddings (e.g. "*The tiger is crying,*

pulling the frog, and he is gray."), while the other two levels included one or two subordinate clauses (e.g. "*The tiger that is crying and pulling the frog is gray.*" and "*The tiger that is pulling the frog that is crying is gray"*). A correct picture matching decision required full understanding of the sentence structure. Older adults were less accurate and slower than younger adults in the task, however, the effect of syntactic complexity was not different between age groups. The behavioural results were related to brain function and structure. Syntactic abilities of young adults were associated with functional coupling in a dedicated, mainly left hemispheric syntax network. In contrast, the syntax network of the older adults included additional (frontal and parietal) regions supporting working memory as well as semantic processing. Indeed, numerous functional imaging studies have shown that older adults recruit different, or additional brain regions compared to younger adults to perform certain tasks, with some research suggesting these additional activity patterns are compensatory in nature (Cabeza, Anderson, Locantore, & McIntosh, 2002; Grossman et al., 2002). Crucially, the finding by Antonenko et. al (2003) that syntactic ability in older adults was related to the recruitment of regions supporting working memory as well as semantic processes, emphasizes the relevance of a behavioural measure that isolates the syntactic component in sentence comprehension.

The influence of semantic processing on syntactic comprehension

Another factor that may contribute to syntactic comprehension is the influence of semantics. Even though lexical-semantic processes and syntactic structure building are generally assumed to be separate processes (Fengler, Meyer, & Friederici, 2016), syntactic processing does interact with a word's semantic information (Deutsch, Bentin, & Katzttt, 1994). Like young adults, older adults are able to efficiently use the available information in language contexts to disambiguate word meanings (Hopkins, Kellas, & Paul, 1995; Hutchison et al., 2013) and aid word identification (Wingfield, Alexander, & Cavigelli, 1994). With regard to sentence comprehension, some findings indicate that older adults rely on morpho-syntactic information to a lesser degree than young adults

when other cues for sentence interpretation are available (Bates, Friederici, & Wulfeck, 1987). For example, Obler et al., (1991)(mentioned above) did not only investigate age-related decline in processing syntactic complexity, they also investigated whether semantic information can aid in processing syntactically complex sentences. Sentences were either semantically plausible or implausible. Older adults were disproportionally less accurate in acceptability judgements for more syntactic complex sentence types but also for implausible sentences. Obler et al., (1991) therefore suggested that older adults come to rely more on processing strategies that stress the plausibility of the semantics of the sentences in terms of their world knowledge rather than on a strict decoding of the syntactic structure. Several previous findings thus suggest that syntactic comprehension is aided by non-syntactic components such as semantics and pragmatics and that the role of these 'aiding processes' becomes more important with age.

The moderating effect of individual differences

Although there a general picture of the cognitive decline in healthy aging, there is also a large amount of individual variability. In fact, the heterogeneity in performance tends to increase with age (Stones, Kozma & Hanna, 1990). Understanding what accounts for this inter-individual variability is an important issue in aging research.

It is well known that aging is associated with decline in working memory and processing speed (G. Waters & Caplan, 2007). Both are known also to contribute to language comprehension (Just & Carpenter, 1992; Salthouse, 1996). A study by Wingfield, Peelle & Grossman (2003) on the effects of speech rate and syntactic complexity in young and older adults established the moderating influence of processing speed on age differences in sentence comprehension. In this experiment, a group of younger and older adults heard short sentences that differed in syntactic complexity by using subject relative clauses ("*Men that assist women are helpful*") and object relative centre embedded clauses ("*Women that men assist are helpful*"). Furthermore, speech rate was time compressed to 80%, 65%, 50% or 35% of the original speaking time, varying the processing

challenge. Participants were asked to indicate whether the action was performed by either a male or female character. Accuracy was lower for the more complex object-relative clause sentences than the easier subject-relative sentences for both age groups, with older adults showing disproportionally poorer comprehension accuracy only at accelerated speech rates. While older adults were slower than younger adults at all speech rates, older adults had disproportionally longer response times for accelerated speech rates and more complex syntactic structures. In a similar vein, a number of studies have demonstrated that the influence of working memory on sentence processing is larger among older compared to younger adults. Payne et al. (2014) found that age differences in relative clause comprehension were largely modulated by individual differences in working memory and that this influence was exaggerated among older adults. Specifically, during comprehension of sentences introducing a temporary syntactic attachment ambiguity (e.g "*The son of the princess who scratched himself/herself in public was humiliated"*), poorer working memory in older adults was associated with increased processing time in sentences in which the reflexive pronoun referred to the object of the modifying prepositional phrase (*herself, the princess*). Payne et al. (2014) suggest that with increasing age, attentional control resources in working memory are recruited at progressively lower levels of difficulty in order to maintain comprehension. These findings illustrate the importance of investigating how individual differences in working memory and processing speed contribute to age-related differences in syntactic comprehension.

Another factor that has gained increasing attention is the person's physical health. Taking into account variability in health characteristics could explain a considerable proportion of variance that would otherwise be ascribed to age (Raz, 2009). In this context, Lara et al. (2015) have proposed a set of biomarkers of healthy aging, in which healthy aging was operationalised as preserved physical, cognitive, physiological, endocrine, immune and metabolic functions. Lifestyle variables such as regular physical activity and aerobic fitness have gained much attention in research focused on differential cognitive aging (Colcombe et al., 2004) and aerobic fitness levels have been shown to be associated with word production in healthy older adults (Segaert et al., 2018). We will measure

grip strength, because it is an established marker of a person's physical health (Lara et al., 2015) and it has previously been related to cognitive decline (Auyeung, Lee, Kwok & Woo, 2011). We will also administer a physical activity questionnaire.

Addressing the moderating influence of working memory, processing speed and physical health can leverage the predictive power of research on age differences in syntactic comprehension.

Current study

In the current study we investigate whether there is age-related decline in syntactic comprehension. Specifically, our aims are threefold. Firstly, we aim to test whether the comprehension of elementary syntactic structure is preserved in older age. Secondly, we aim to test whether lexical-semantic content aids syntactic comprehension and whether this changes with age. Thirdly, we aim to investigate whether individual differences in working memory, processing speed and physical health modulate syntactic comprehension and moreover, whether the impact of these increase with age.

We investigate syntactic comprehension in an auditory syntactic judgement task, in a group of younger and older participants. The complexity of our stimuli is reduced to simple two word phrases consisting of a pronoun and a verb (e.g. *"I walk").* Consequently, working memory load for processing these phrases is minimal. Lexical-semantic content is varied by using existing verbs versus pseudoverbs. A pseudoword follows the orthographic and phonological rules of a language, but has no meaning in the mental lexicon of that language. The pseudoverbs were used to create phrases of minimal semantic content (e.g. *"she ploffs"*), whereas the existing verbs were used to create semantically meaningful phrases (e.g. *"she cooks"*). The pseudoverbs and existing verbs formed two separate experiments, identical in all aspects but the use of the pseudoverb versus the real verb. We will refer to this the "Pseudoverb" and "Real Verb" experiment respectively. The task was to listen to the phrases and indicate whether it was morpho-syntactically correct (yes/no). In addition to

accuracy, we measured response time (RT) from the start of the response screen to the button press.

To investigate the impact of individual differences on syntactic comprehension, we measured important biomarkers of healthy aging (Lara et al., 2005): physical health was assessed as strength grip and a physical activity questionnaire; cognitive functioning was assessed through a working memory, processing speed and verbal IQ measure.

We predict the following. First, in line with most previous findings of preserved syntactic comprehension in aging, we predict that performance on the Real Verb experiment is equivalent for young and older adults. Second, we expect reduced performance on the Pseudoverb experiment for older adults, compared to young adults, in line with previous findings suggesting that older adults come to rely more on strategies involving semantic processing. We also expect a stronger influence of working memory and processing speed for older compared to young adults. Lastly, if a relationship exists between physical health and syntactic comprehension in older adults, we expect to find that age-related decline in syntactic comprehension is modulated by physical health, with higher levels of physical health associated with better performance in older adults.

Methods

Participants

50 young university undergraduates (45 women, mean age: 19, SD: 0.92, 5 men, mean age: 20 y, SD: 0.89) and 50 older adults (28 women, mean age: 71, SD: 5.79, 22 men, mean age: 72, SD: 5.68) participated in the study. Participants were recruited via the database of the School of Psychology of Birmingham University. All participants were native British English speakers with normal or corrected to normal hearing. Exclusion criteria included bilingualism, neurological disorders, speech or language disorders and dyslexia. To assess general cognitive function, the Montreal Cognitive Assessment test (MoCa; version 7.1) was administered to the elderly participants, resulting in 5 participants being excluded, as their scores were equal to or below the cut-off value of 26.

Consequently, 45 older participants (23 women, mean age: 71, SD: 5.66 and 22 men, mean age: 73, SD: 5.61) were included in the analyses. The older participants' education level ranged from Primary School (1 participant); O-levels/GCS2 (11); A levels/Vocational Course (6); Bachelors/Undergraduate level (21) and Masters degree or higher (10). All participants gave informed consent. Students were given university credits as compensation; older adults received monetary compensation. The research was conducted at the University of Birmingham and had full ethical approval.

Materials

A set of 20 English pseudoverbs created by Ullman et al. (1997) served as stimulus material for the Pseudoverb experiment: brop, crog, cug, dotch, grush, plag, plam, pob, prap, prass, satch, scash, scur, slub, spuff, stoff, trab, traff, tunch, vask. These pseudoverbs were all monosyllabic with an average word length of four letters. All pseudoverbs could be inflected according to regular grammar rules for verbs in English. They could be combined with six pronouns (I, you, he, she, we, they) or with 6 adverbs (daily, quickly, safely, early, promptly, rarely). This would yield minimal phrases, such as "I dotch", "he dotches", "they dotched", or "dotch quickly". In addition, a set of twenty common English verbs were selected to serve as stimulus material for the Real Verb experiment: chop, cook, cram, bake, drop, flap, skip, brew, rob, rush, scour, move, jog, slam, stir, tug, walk, pull, stack, reap. These were regular monosyllabic verbs, matched in length to the pseudoverbs. Like the pseudoverbs, these real verbs could be combined with a pronoun, or an adverb to form minimal phrases, such as "I chop", "she chops", "they chop", or "chop quickly". The same adverbs were used with both the pseudoverbs and real verbs. The adverbs were all disyllabic and care was taken to ensure that combining them with any of the real verbs would form a semantically meaningful combination.

Digital recordings of all stimuli were made using a male native speaker of English. All verbs were recorded in first, second and third singular and plural present tense. Each stimulus was pronounced three times, after which the clearest recording was selected. In order to equalize the

volume of the individual recordings, all audio files in wav format were normalized to 1db using the software program Adobe Audition.

Design

The order of the Real Verb and Pseudoverb experiments was counterbalanced across participants. Both experiments consist of the same four conditions (see Table 1). In the *Correct Integration condition* a (pseudo-)verb was paired with a pronoun, resulting in a morpho-syntactically correct combination (e.g. “she cugs”, “she walks”). In the *Incorrect Integration condition*, integration could be attempted, but the inflection of the verb/pseudoverb did not match the pronoun (“she cug”, “she walk”). In addition, two filler conditions were included. For the *No Integration filler condition*, the verb/pseudoverb was paired with another verb/pseudo (e.g. “dotch cugs”, “skip walks”). This combination of stimuli should not trigger integration processes at a morpho-syncatic level. The No Integration filler condition was included in the current experiment in order to verify that participants indeed read these as a pairing of two verbs/pseudoverbs and did not attempt to integrate them. The purpose of this condition (merely a filler condition in the present experiment) was to include it as a condition of interest (a baseline condition) in a follow-up EEG experiment. Finally, the *Adverb filler condition* consisted of a verb/pseudoverb paired with an adverb (e.g “cugs quickly”, “walks quickly”). The purpose of the Adverb fillers was to avoid any predictability in the engagement of integration processes for pairs beginning with verb/pseudoverb. Specifically, a word pair starting with a verb/pseudoverb had an equal chance of forming a syntactically correct or incorrect word pair. An overview of the stimulus sets for both experiments and examples of all conditions is provided in Table 1.

Table 1. Example stimuli in each condition for the Pseudoverb and Real Verb Experiment with trial number per condition.

Condition [number of trials]	**Example Pseudoverb experiment**	**Example Real Verb experiment**	**Correct sentence?**
Correct Integration Condition [36]	I ploff she ploffs we ploffed	I pull she pulls we pulled	Yes
Incorrect Integration Condition [36]	I ploffs he ploff	I pulls he pull	No
No Integration filler [36]	ploffs dotch ploff dotches ploff dotched	walks pull walk pulls walk pulled	No
Adverb filler [36]	ploff quickly ploffs quickly ploffed quickly	walk quickly walks quickly walked quickly	Yes

Task

Participants were tasked with detecting grammatical mistakes. The auditory presentation of the first word and the second word was followed by a response screen; the timing of each event in the trial is illustrated in Figure 1. Participants were instructed to indicate whether the word pair they just heard was grammatically correct by clicking the left and right mouse button to respond with 'yes' or 'no' respectively. The response screen showed the following text: *"Was this a grammatically correct sentence?".* The correct response for each condition is listed in Table 1. The experiments were run using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

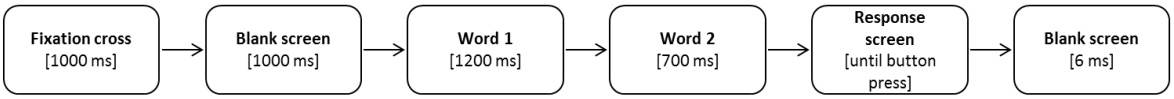

Figure 1. Timing of the components of one trial for both the Pseudoverb and the Real Verb experiment.

Experimental lists

As can be seen in Table 1, the Correct Integration condition can be formed with three possible pronoun – verb/pseudoverb combinations. That is, the verb/pseudoverb *stem* combined with either 'I', 'you', 'we' or 'they'; the verb/pseudoverb stem plus *–s* combined with 'he' or 'she', or the verb/pseudoverb stem plus *–ed* combined with each of the six pronouns. Each form occurred 12 times and the possible pronouns within each form occurred an equal number of times. This means

that each possible pronoun occurred 3 times in the *stem form*, 6 times in the *–s form* and 2 times in the *–ed form*. The verb/pseudoverb were randomly assigned to the pronouns, with the constraint that each verb would occur only once in each form. The Incorrect Integration word pairs were formed according to the same criteria. However, as no incorrect combination can be composed with the *–ed form*, only two forms were possible. To ensure an equal number of trials across conditions, both the *stem form* and *-s form* consisted of 18 trials in this condition, again ensuring that the possible pronouns occurred an equal number of times. The No Integration filler condition consisted of three possible forms, such that the second verb could either be *stem-form*, *-s form*, or *–ed form*, with 12 trials per form. To avoid repetition effects, the first word of the pair in this condition could neither be the same verb nor have the same ending as the second word of the pair. Lastly, the Adverb filler condition also consisted of three possible forms, with the first word being either in *stem-form*, *-s form*, or *–ed form*, followed by randomly assigned adverbs as the second word. There were 36 trials per condition, resulting in 144 trials in total for both experiments.

A unique randomized stimulus list was created for each participant and divided into three separate blocks, separated with self-paced breaks. The order of the Pseudoverb and Real Verb experiment was counterbalanced between participants. Each experiment was preceded by a unique list of 33 practice trials.

Inter-individual variability markers

A number of individual differences measures were collected to assess physical health and cognitive functioning of our participants.

Markers of cognitive function:

The Backward Digit Span task (Waters & Caplan, 2003) was administered to measure *working memory capacity*. Using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA), participants were instructed to attend to a series of visually presented digits of increasing length. After the presentation of the last digit, participants were instructed to enter the digits in the reverse

order by using the numbers on the keyboard. The task began at a length of two digits and went up to seven digits. There were 5 trials at each digit length. No practice trials were included. Span size was defined as the longest digit length at which a participant correctly recalled three out of five trials. If a participant recalled two out of five trials correctly at the longest digit length, half a point was added to the total score. The raw span size score were used in the analyses.

Using the WISC-IV Coding subtask (WAIS-IV; Wechsler, 2008), *processing speed* was assessed. In this task, the participant is asked to copy symbols that are paired with numbers within 120 seconds. A point is assigned for each correctly drawn symbol completed within the time limit. The total raw score is the number of correctly drawn symbols, with a maximum of 135. The raw scores were converted into age-scaled scores using the WAIS-IV manual.

Verbal IQ was assessed by means of the National Adult Reading Test (NART), based on Nelson and Willison (1991). The NART consists of 50 words with atypical phonemic pronunciation. Participants were instructed to slowly read aloud the list of words. Auditory recordings were made of the answers, which were scored on correct pronunciation according to Google translate pronunciation. Number of errors and number of correct items were summed. The Verbal IQ score that was used for analyses was calculated according to standard procedures: Estimated Verbal IQ = 129.0 – 0.919 X NART error score.

Markers of physical health:

We assessed grip strength using a standard adjustable hand dynamometer (Takei Scientific Instruments). Standing in upright position, the participant was instructed to hold the dynamometer towards the ceiling with a completely outstretched arm, so that the shoulder and elbow were fully flexed at 180 degrees, hand palm facing the gaze direction. From this starting position, the participant was instructed to move their arm downwards in three seconds while squeezing the dynamometer with maximum force. A total of three measurements was recorded for the dominant and non-dominant hand, which was preceded by three practice trials for each hand. The highest

value of the dominant hand was used for analyses. These raw scores were converted into standardised z-scores within age- and gender groups.

A physical activity questionnaire (New Zealand Physical Activities Survey Short Form; Sport and Recreation New Zealand, 2001) was included as a self-report measure of the participants' habitual practice of physical activity. A composite score, calculated by adding the duration (in minutes) of moderate activity and two times the duration of vigorous activity, was used for analyses.

Procedure

As mild hearing loss is a common condition in elderly people and the ability to clearly hear the stimuli is crucial for the aim of our study, the procedure started with a volume check. Participants listened to 20 randomly selected stimuli (10 real verbs and 10 pseudoverbs) through headphones and were asked to repeat what they heard. The experimenter paid special attention to correct pronunciation of the words' suffices. Volume settings were adjusted if necessary.

Half of the participants started with the Pseudoverb experiment and the other half started with the Real Verb experiment. Instructions were identical in both experiments. After the participant read the instructions, the experimenter briefly summarized the procedure. Participants wore headphones and used the computer mouse to give their responses. Both experiments started with a practice block, consisting of 33 trials, such that each possible word pair combination occurred three times. Participants received verbal feedback on their performance on the practice trials and only proceeded to the real experiment when they had a clear understanding of the task. The same procedure was repeated for the other language experiment. Participants were instructed that the second task was exactly the same as the previous one, only this time with real/pseudo- verbs.

Each language experiment took on average 30 minutes to complete, including the practice session and two self-paced breaks. Participants were then tested on a number of additional measurements which were conducted in the following order: the Backward Digit Span Task; the Hand Grip Strength; the Physical Activity questionnaire; the Coding task and lastly the NART.

Data analyses

The dependent variables are the response time (RT) and accuracy on the Correct-Integration and Incorrect-Integration trials. Trials with RTs below 300 milliseconds were counted as outliers and excluded from the analyses (12% of the response of the elderly group and 19% of young group). Following this, the RT data for each participant in each condition was subjected to a ± 2 standard deviation trim, resulting in an exclusion of 5% of the data points in both groups. Lastly, one elderly participant was removed from further analyses due to excessively long RT's (mean 2522, sd 1827, compared to the group mean 1164, sd 949). Only correct responses were included in the RT analyses. We analysed accuracy using a mixed-logit model in R (R Core Team, 2015), using the *lme4* package (Bates, Mächler, Bolker & Walker, 2015). This method is most suitable for analysing categorical responses while excluding the necessity to conduct separate participant and item analyses (Jaeger, 2008). RT was analysed with a linear mixed model. The use of mixed effects models offers the opportunity to estimate effects and interactions of the experimental manipulations, or fixed effects, while simultaneously estimating parameters of the variance and covariance components of individual subjects and items as random effects (Kliegl, Wei, Dambacher, Yan, & Zhou, 2011).

To avoid multicollinearity in the regression models, we computed the Pearson's correlation coefficients and p-values for our predictors using the corrplot package in R (Wei & Simko, 2016). Given that all correlations had a Spearman's rank correlation coefficient <0.3, all predictors were included in the models.

The regression models for predicting both RT and Accuracy were based on the following predictors: Verb type (Pseudoverb and Real Verb); Integration type (Correct and Incorrect); Age group (younger and older); Working Memory; Processing speed; Hand Grip; Physical Activity and Verbal IQ. Our categorical predictors Verb type, Integration and Age group were all sum coded, such

that the intercept of the model represents the grand mean (across all conditions) and the coefficients can directly be interpreted as main effects. Continuous variables were centred.

We began with a full model and then performed a step-wise "best-path" reduction procedure for the fixed effects to determine the simplest model that did not differ significantly from the full model in terms of variance explained (as described in Weatherholtz, Campbell-Kibler & Jaeger, 2014) using the drop1 function from the stats package (version 3.4.2). We used a maximum random effects structure, allowing us to include intercepts for participants and items ("random intercepts"), as well as well as by-participants and by-item random slopes for the fixed effects. When the model did not converge with the fully expressed random effects structure, we simplified the random effects structure removing first the interactions, followed by the slopes which contributed least to the variance explained (Barr, Levy, Scheepers, & Tily, 2013).

Results

A. Group differences on individual differences measures

Table 2 provides an overview of the additional measurements for the Younger and Older Age group. In accordance with typical findings, the young participants outperform the older participants in Working Memory capacity and Processing Speed. To disentangle the effect of age from Processing Speed, the scaled scores were used in the analyses. However, for the sake of completeness, the raw scores are reported as well. The older participants performed significantly better in terms of Verbal IQ. There was no difference in Physical Activity or Hand Grip strength between both groups.

Table 2. Means and Standard Deviations of Additional measurements for Young and Older Age group and the results of Comparisons between the Age groups (Independent Samples t-Test)

	Younger Age group (N=50)		**Older age group (N= 50)**		**Comparison**	
	mean	sd	mean	sd	t	p value
Working Memory	5.57	1.57	4.57	1.59	3.1417	0.002
Processing Speed	80.34	14.91	62.92	16.48	5.5424	< 0.001
Processing Speed scaled	11.52	2.91	12.52	16.48	-1.7233	0.088
Verbal IQ	28.08	4.70	39.60	5.32	-11.314	< 0.001
Physical Activity	122.28	121.22	131.20	102.49	-0.39736	0.692
Hand grip	25.92	7.54	26.99	8.85	-0.65062	0.517

Table 3. Coefficient estimates, standard errors (SE), associated t values and significance levels for all predictors in the generalized mixed model predicting accuracy

Coefficient	Estimate	Std. Error	z value	p	
(Intercept)	2.943468	0.184510	15.953	**< 0.001**	***
Working Memory	**0.158063**	**0.064261**	**2.460**	**0.014**	*
Age group	**1.459913**	**0.319516**	**4.569**	**< 0.001**	***
Verb type	**0.659749**	**0.204202**	**3.231**	**0.001**	**
Integration	0.261988	0.142930	1.833	0.067	.
Processing Speed	0.026363	0.034538	0.763	0.445	
Handgrip	0.045121	0.104507	0.432	0.666	
Verbal IQ	**0.106007**	**0.020429**	**5.189**	**< 0.001**	***
Working Memory*Age group	-0.094567	0.128716	-0.735	0.462	
Working Memory*Verb type	0.018615	0.056433	0.330	0.742	
Age group*Verb type	**-0.682971**	**0.267488**	**-2.553**	**0.011**	*
Age group*Integration	0.141092	0.291674	0.484	0.629	
Verb type*Integration	0.266403	0.168694	1.579	0.114	
Working Memory*Integration	0.052697	0.054719	0.963	0.336	
Age group*Processing Speed	0.018724	0.069091	0.271	0.786	
Verb type*Processing Speed	**0.067107**	**0.028438**	**2.360**	**0.018**	*
Integration*Processing Speed	0.048964	0.028984	1.689	0.091	.
Age group*Handgrip	0.209546	0.209083	1.002	0.316	
Verb type*Handgrip	0.019091	0.083233	0.229	0.819	
Integration*Handgrip	-0.012378	0.082838	-0.149	0.881	
Age group*Verbal IQ	0.016819	0.040700	0.413	0.679	
Verb type* Verbal IQ	**-0.100190**	**0.017126**	**-5.850**	**< 0.001**	***
Integration*Verbal IQ	-0.008386	0.017620	-0.476	0.634	
Working Memory*Age group*Verb type	**0.312953**	**0.112308**	**2.787**	**0.005**	**
Age group*Verb type*Integration	-0.085844	0.511238	-0.168	0.867	
Working Memory*Age group*Integration	**0.418367**	**0.109638**	**3.816**	**< 0.001**	***
Age group*Verb type*Processing Speed	**0.208321**	**0.058000**	**-3.592**	**< 0.001**	***
Age group*Integration*Processing Speed	**0.141760**	**0.057900**	**-2.448**	**0.014**	*
Verb type*Integration*Processing Speed	**0.120611**	**0.055971**	**2.155**	**0.031**	*
Age group*Verb type*Handgrip	-0.288745	0.166723	-1.732	0.083	.
Age group*Integration*Handgrip	-0.279916	0.165526	-1.691	0.090	.
Verb type*Integration*Handgrip	**0.454076**	**0.158186**	**2.871**	**0.004**	**
Age group*Integration*Verbal IQ	0.053463	0.034522	1.549	0.121	
Verb type*Integration*Verbal IQ	**-0.099581**	**0.033426**	**-2.979**	**0.003**	**
Age group*Verb type*Integration*Processing Speed	**-0.270340**	**0.112644**	**-2.400**	**0.016**	*
Age group*Verb type*Integration*Handgrip	**0.649676**	**0.317488**	**2.046**	**0.041**	*

*Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 ‘ ’ 1*
Model includes a random intercept for items and subjects.

B. Age differences in response accuracy for syntactic comprehension

Table 3 presents the results from the final mixed model predicting accuracy. This model was not significantly different from the full model (Full model = AIC: 5949.6, BIC 6256.2; Best model= AIC: 5943.0, BIC: 6220.3, p = 0.8534). Figure 2 (panel a) shows the shows the group average of the proportion of correct responses given by the younger and the older age group for each of the two experiments. The younger age group reached higher accuracy levels compared to the older age group in both experiments (p < 0.001), suggesting that indeed there is age-related decline in

syntactic comprehension accuracy. Generally, participants were less accurate in the Pseudo-Verb experiment compared to the Real Verb experiment ($p < 0.004$). The age-related decline in syntactic comprehension was stronger in the Real Verb experiment than the Pseudoverb experiment, as revealed by the significant Age group * Verb type interaction ($p = 0.021$).

In addition to these group effects, there was individual variation in performance accuracy for both groups (shown in panel 2b). Of particular interest are interactions between individual difference measures and Age group, which were found processing speed and for working memory. We turn to these next.

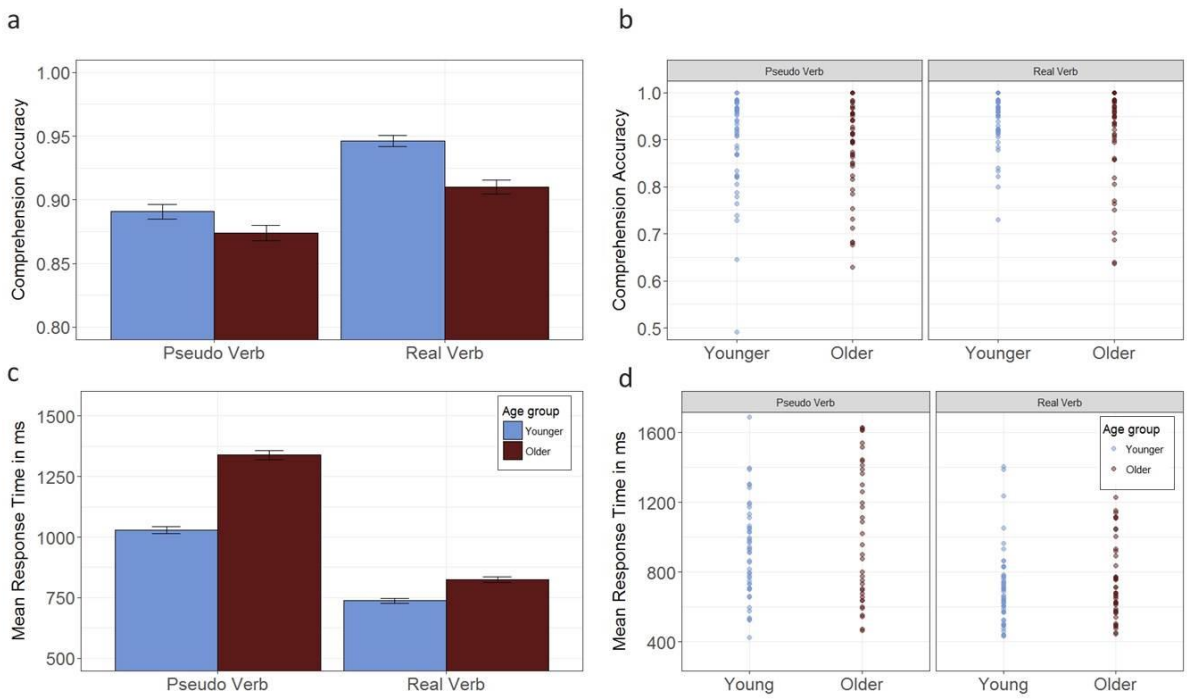

Figure 2. Age-related performance differences in accuracy (top row) and speed (bottom row) for syntactic comprehension. *Group average proportion of correct comprehension per age group (a) and individual means (b). Mean response times (RTs) to correct responses for the two age groups (c) and individual subjects (d). We have collapsed across the correct integration and incorrect integration condition in these graphs. Error bars are standard errors of the mean.*

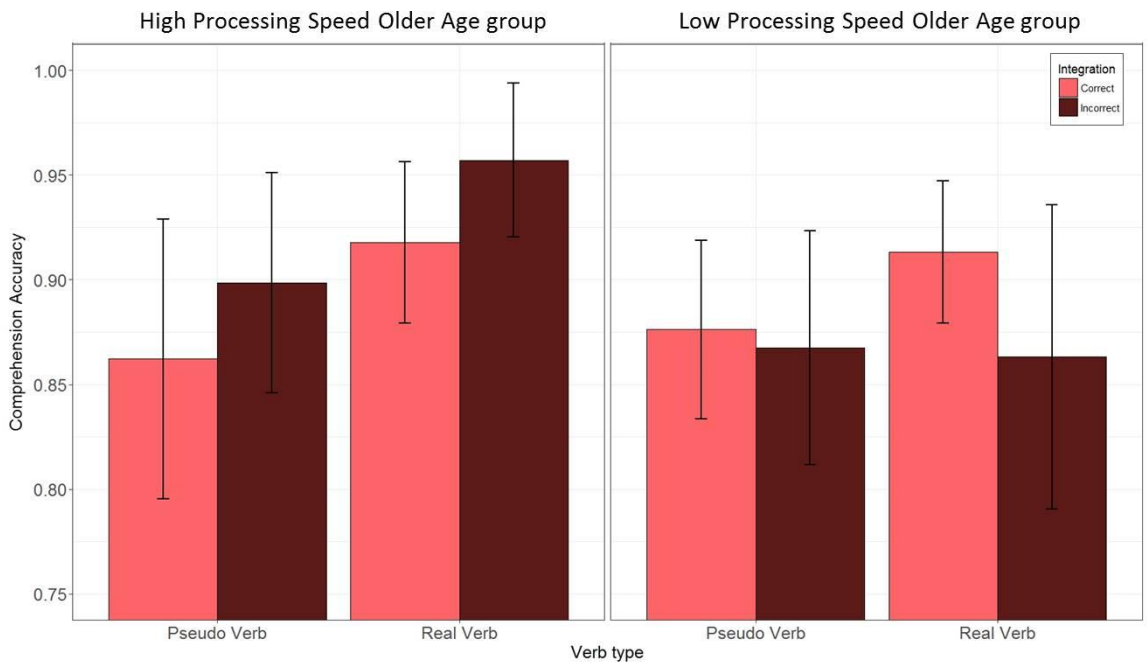

Figure 3 Processing speed modulates syntactic comprehension accuracy for the older Age group. Three-way interaction between Verb type, Integration and Processing Speed is depicted through a median split on Processing Speed. Older adults with high processing speed scores had higher accuracy scores than low processing speed older adults, particularly in real verb phrases in the incorrect integration condition.

Modulating effect of Processing Speed

There was a significant four-way interaction between Age group, Verb type, Integration and Processing Speed (p = 0.016). To understand the four-way interaction we analysed the two age-groups separately next. In this post hoc analysis the same model was applied to both age groups. This revealed that there was a non-significant Verb type * Integration * Processing Speed interaction in the younger age group (p = 0.845), but a significant Verb type * Integration * Processing Speed interaction in the older age group (p = 0.005). The results of this post hoc analysis are presented in Table 4. In order to visualize the three-way interaction with Processing Speed as a continuous predictor in the older adults, we performed a median split on Processing Speed and plotted the average accuracy for each verb type and for each condition for high and low performers separately (see Figure 3). The figure illustrates that older adults with low processing speed scores were less accurate than those with high processing speed scores. This performance difference was particularly

prevalent in the Real verb experiment in the incorrect integration condition. In older adults with high processing speed scores, accuracy depended on whether there was a pseudoverb or real verb phrase and on whether integration was correct or not. Note that we are using scaled scores so these effects cannot be attributed to effects of numerical age.

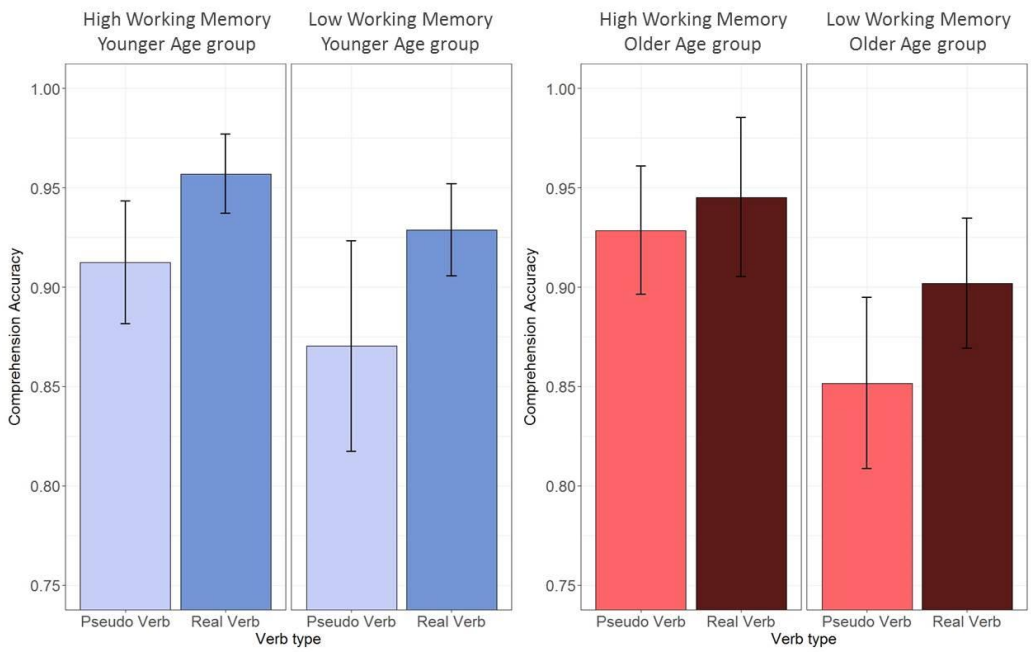

Figure 4 Working Memory differentially effects syntactic comprehension accuracy depending on age group. The three-way interaction between Age group, Verb type and Working Memory, depicted by a median split on Working Memory for both age groups separately. The left panel shows the younger age group, the right panel shows the older age group. Young adults with low working memory had lower accuracy scores than young adults with high working memory, particularly in pseudoverb phrases. There was no significant Verb type * Working Memory interaction in the older adults group.

Modulating effect of Working Memory

We found a significant three-way interaction between Age group, Working Memory and Verb type ($p = 0.005$). This interaction was due to the Working Memory * Verb type interaction reaching significance for the younger age group ($p = 0.022$), but not for the older age group ($p = 0.110$; see Table 4). Figure 4 shows the effects of Working Memory on accuracy for both Verb types at high and low Working Memory levels separately. In the younger age group (left panel of Figure 4), lower

Working Memory scores were related to lower accuracy, particularly in the Pseudoverb experiment. Notably, there was an additional significant three-way interaction between Age group, Working Memory and Integration ($p < 0.001$), which was due to a significant interaction between Working Memory and Integration for the younger age group ($p < 0.001$), but not for the older age group ($p = 0.068$). Young adults with low working memory scores were less accurate, particularly in the incorrect integration condition. Overall, this suggests that accuracy is moderated by working memory in young adults, such that a lower working memory span is associated with a relative disadvantage in performance, particularly in comprehending pseudoverb phrases and in identifying morpho-syntactically incorrect phrases. In contrast, neither the effect of verb type nor integration was significantly moderated by working memory in the older age group.

Modulating effect of Handgrip strength

We found a significant four way interaction between Age group, Verb type, Integration and Handgrip ($p = 0.041$). Post hoc analyses revealed this effect was driven by a significant interaction between Verb type, Integration and Handgrip in the young age group ($p = 0.001$) and a non-significant interaction between Verb type, Integration and Handgrip in the older age group ($p= 0.510$). In the younger age group, accuracy in the incorrect integration condition of the pseudo verb experiment was particularly low and modulated by variability in handgrip scores.

Table 4A. Coefficient estimates, standard errors, z values and p values of post hoc generalized mixed model predicting accuracy for the young age group. Model includes a random intercept for items and subjects.

Coefficient	Estimate	Std. Error	Z value	p	
(Intercept)	**3.13941**	**0.17116**	**18.342**	**< 0.001**	***
Working Memory	0.11820	0.07950	1.487	0.137	
Verb type	**1.12180**	**0.27104**	**4.139**	**< 0.001**	***
Integration	**0.37397**	**0.13038**	**2.868**	**0.004**	**
Processing Speed	0.03763	0.04182	0.900	0.368	
Handgrip	0.14485	0.12026	1.205	0.228	
Verbal IQ	**0.11570**	**0.02780**	**4.162**	**< 0.001**	***
Working Memory*Verb type	**0.17860**	**0.07827**	**2.282**	**0.022**	*
Verb type*Integration	**0.75944**	**0.26276**	**2.890**	**0.004**	**
Working Memory*Integration	**0.26598**	**0.07291**	**3.648**	**0.000**	***
Verb type*Processing Speed	-0.03573	0.04280	-0.835	0.404	
Integration*Processing Speed	-0.02258	0.04282	-0.527	0.598	
Verb type*Handgrip	-0.12489	0.11551	-1.081	0.280	

Integration* Handgrip	-0.15275	0.11513	-1.327	0.185	
Verb type* Verbal IQ	**-0.09286**	**0.02908**	**-3.193**	**0.001**	**
Integration*Verbal IQ	0.01206	0.02889	0.418	0.676	
Verb type*Integration*Processing Speed	-0.01626	0.08340	-0.195	0.845	
Verb type*Integration*Handgrip	**0.78459**	**0.23217**	**3.379**	**0.001**	***
Verb type*Integration*Verbal IQ	**-0.14693**	**0.05520**	**-2.662**	**0.008**	**

4B. Coefficient estimates, standard errors, z values and p values of post hoc generalized mixed model predicting accuracy for the older age group. Model includes a random intercept for items and subjects.

Coefficient	Estimate	Std. Error	Z value	p	
(Intercept)	**2.61915**	**0.16099**	**16.269**	**< 0.001**	***
Working Memory	**0.21467**	**0.10236**	**2.097**	**0.036**	*
Verb type	**0.48441**	**0.18714**	**2.588**	**0.010**	**
Integration	0.12582	0.10188	1.235	0.217	
Processing Speed	0.01684	0.05602	0.301	0.764	
Handgrip	-0.06671	0.17570	-0.380	0.704	
Verbal IQ	**0.09767**	**0.02988**	**3.269**	**0.001**	**
Working Memory*Verb type	0.13448	0.08414	-1.598	0.110	
Verb type*Integration	-0.10860	0.20172	-0.538	0.590	
Working Memory*Integration	-0.14941	0.08185	-1.825	0.068	.
Verb type*Processing Speed	**0.16979**	**0.03941**	**4.309**	**< 0.001**	***
Integration*Processing Speed	**0.11417**	**0.03895**	**2.932**	**0.003**	**
Verb type*Handgrip	0.15940	0.11976	1.331	0.183	
Integration*Handgrip	0.12711	0.11886	1.069	0.285	
Verb type*Verbal IQ	**-0.10477**	**0.02116**	**-4.951**	**< 0.001**	***
Integration*Verbal IQ	-0.03071	0.02095	-1.466	0.143	
Verb type*Integration*Processing Speed	**0.22555**	**0.07752**	**2.910**	**0.004**	**
Verb type*Integration*Handgrip	0.14232	0.21586	0.659	0.510	
Verb type*Integration*Verbal IQ	-0.07172	0.04196	-1.709	0.087	

*Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 ‘ ’ 1*

C. Age differences in response time for syntactic comprehension

Table 5 presents the results of the best linear mixed model predicting response times. This model was not significantly different from the full model (Full model = AIC: 154279 BIC 154697; Best model= AIC: 154697 BIC: 154584, p = 0.8948). Figure 2 (panel c) shows the mean response times in ms on the Pseudoverb and Real Verb experiment for both age groups.

The older age group took longer to respond than the younger age group ($p < 0.001$), suggesting that there is age-related decline in syntactic comprehension. In addition, response times were significantly longer in the Pseudoverb experiment compared to the Real Verb experiment ($p < 0.001$). Age-related decline in response times was larger for the Pseudoverb experiment compared to the Real Verb experiment, as revealed by the Age group * Verb type interaction ($p = 0.003$). Post hoc analyses within each age group revealed that the effect of Verb type exists in both age groups (see Table 6), however visual inspection of Figure 2c shows the effect to be larger in the older age group.

In addition to these group effects, there was individual variation in reaction times for both groups (shown in panel 2d). Of particular interests are interactions between individual difference measures and Age group, which were found for processing speed and for working memory. We turn to these next.

Table 5. Coefficient estimates, standard errors (SE), associated t values and p values for all predictors of linear mixed model predicting response time

Coefficient	Estimate	Std. error	t value	p value	
(Intercept)	**1277.8172**	**75.3091**	**16.968**	**< 0.001**	*******
Working Memory	-15.8049	33.4599	-0.472	0.637	
Age group	**-465.2187**	**115.0198**	**-4.045**	**< 0.001**	*******
Verb type	**-574.974**	**72.9218**	**-7.885**	**< 0.001**	*******
Integration	**126.8884**	**48.1706**	**2.634**	**0.008**	******
Processing Speed	-31.3743	19.4421	-1.614	0.107	
Handgrip	36.463	55.1778	0.661	0.509	
Physical Activity	0.2495	0.3099	0.805	0.421	
Verbal IQ	-37.9393	9.4003	-4.036	**< 0.001**	***
Working Memory* Age group	10.4149	45.9392	0.227	0.821	
Working Memory*Verb type	12.7928	34.7024	0.369	0.712	
Age group*Verb type	**367.3914**	**122.8109**	**2.992**	**0.003**	******
Working Memory*Integration	**58.7601**	**22.3174**	**2.633**	**0.008**	******
Age group*Integration	-137.9675	80.0278	-1.724	0.085	.
Verb type*Integration	**-141.531**	**42.9923**	**-3.292**	**0.001**	*******
Age group*Processing Speed	17.3717	25.3613	0.685	0.493	
Verb type*Processing Speed	**43.5975**	**19.9812**	**2.182**	**0.029**	*****
Integration*Processing Speed	-6.4694	8.5603	-0.756	0.450	
Age group*Handgrip	-63.137	63.8982	-0.988	0.323	
Verb type*Handgrip	-32.6503	40.4622	-0.807	0.420	
Integration*Handgrip	-0.3233	26.2521	-0.012	0.990	
Verb type*Physical Activity	0.2851	0.3296	0.865	0.387	
Integration*Physical Activity	-0.1497	0.2139	-0.7	0.484	
Age group*Verbal IQ	**31.6301**	**12.1551**	**2.602**	**0.009**	******
Verb type*Verbal IQ	**13.4997**	**7.8784**	**1.714**	**0.087**	**.**
Integration*Verbal IQ	-5.6674	5.0732	-1.117	0.264	
Working Memory* Age group*Verb type	-7.3029	48.2459	-0.151	0.880	
Working Memory*Age group*Integration	**-62.7746**	**31.1653**	**-2.014**	**0.044**	*****
Working Memory*Verb type*Integration	1.5218	23.0824	0.066	0.947	
Age group*Verb type*Integration	-20.5841	53.3838	-0.386	0.700	
Age group*Verb type*Processing Speed	**-53.3111**	**26.5436**	**-2.008**	**0.045**	*****
Verb type*Integration*Processing Speed	-13.8409	9.109	-1.519	0.129	
Verb type*Integration*Handgrip	-53.2151	27.8905	-1.908	0.056	.
Verb type*Integration*Physical Activity	**-0.4784**	**0.2265**	**-2.112**	**0.035**	*****
Working Memory* Age group*Verb type*Integration	59.0634	32.8179	1.8	0.072	.

Model includes a random intercept for items and subjects, a random slope for Integration for both items and subjects and a random slope for Verb type for subjects.
*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Modulating effect of Processing Speed

Similar to our accuracy analyses, we found an interaction between Age group, Verb type and Processing Speed (p = 0.045). However, post hoc analyses reveal that the Verb type * Processing Speed interaction is not significant, neither in the young age group (p = 0.439), nor in the old age group (p = 0.127, see table 6).

Moderating effect of Working Memory

There was a significant interaction between Age group, Working Memory and Integration. To understand what was driving this interaction, we examined the Working Memory * Integration interactions for each age group separately. There was a significant interaction between Working Memory and Integration for the older age group (p = 0.019), such that higher Working Memory was associated with shorter response times in the Correct Integration Condition. There was no significant interaction between Working Memory and Integration in the younger age group (p = 0.766).

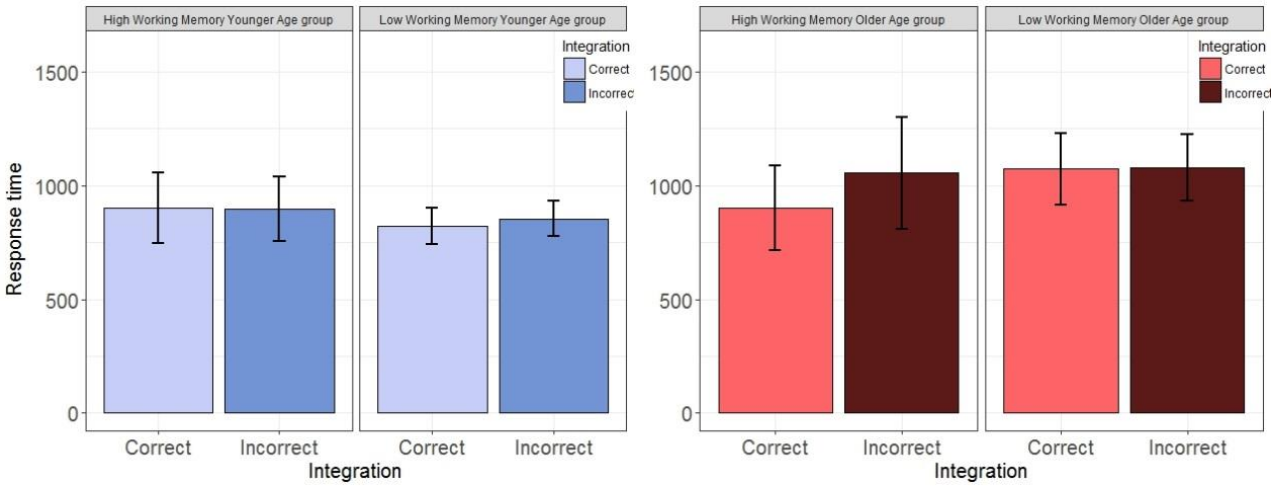

Figure 5 Working Memory differentially effects response time depending on age group. The three-way interaction between Age group, Integration and Working Memory, depicted by a median split on Working Memory for both age groups separately. The left panel shows the younger age group, the right panel shows the older age group. There was a significant decrease in response time in older adults with high working memory in the correct integration condition. There was no significant Integration * Working Memory interaction among the younger adults.

Moderating effect of Verbal IQ

We found an interaction between Age Group and Verbal IQ, such that a higher Verbal IQ score was associated with faster response times for older adults, but not for young adults.

Table 6.A Coefficient estimates, standard errors (SE), associated t values and p values for all predictors of post hoc linear mixed model predicting response time for young adults. Model includes a random intercept for items and subjects, a random slope for Integration for both items and subjects and a random slope for Verb type for subjects.

Coefficient	Estimate	Std. error	t value	p value	
(Intercept)	**8.23E+02**	**7.30E+01**	**11.28**	**< 0.001**	*******
Working Memory	-6.54E+00	2.71E+01	-0.241	0.810	
Verb type	**-2.38E+02**	**6.92E+01**	**-3.439**	**< 0.001**	*******
Integration	1.04E+01	4.79E+01	0.218	0.828	
Processing Speed	-1.49E+01	1.39E+01	-1.073	0.283	
Handgrip	-2.18E+01	4.07E+01	-0.536	0.592	
Physical Activity	-4.34E-03	3.48E-01	-0.012	0.990	
Verbal IQ	-4.43E+00	9.21E+00	-0.481	0.630	
Working Memory*Verb type	9.74E+00	2.58E+01	0.378	0.705	
Working Memory*Integration	-5.22E+00	1.76E+01	-0.297	0.766	
Verb type*Integration	**-1.58E+02**	**3.88E+01**	**-4.071**	**< 0.001**	*******
Verb type*Processing speed	-1.02E+01	1.32E+01	-0.774	0.439	
Integration*Processing speed	-3.21E+00	8.95E+00	-0.359	0.719	
Verb type*Hand grip	-3.73E+01	3.86E+01	-0.966	0.334	
Integration*Handgrip	2.01E+01	2.62E+01	0.768	0.442	
Verb type*Physical Activity	4.94E-01	3.31E-01	1.491	0.136	
Integration*Physical Activity	-1.04E-01	2.25E-01	-0.463	0.643	
Verb type*Verbal IQ	8.18E+00	8.73E+00	0.937	0.349	
Integration*Verbal IQ	-2.94E+00	5.95E+00	-0.493	0.622	
Working Memory*Verb type*Integration	**6.16E+01**	**2.06E+01**	**2.991**	**< 0.001**	******
Verb type*Integration*Processing Speed	-1.17E+01	1.10E+01	-1.063	0.288	
Verb type*Integration*Hand grip	-2.56E+01	3.24E+01	-0.791	0.429	
Verb type*Integration*Physical Activity	**-7.10E-01**	**2.70E-01**	**-2.635**	**< 0.001**	******

6B. Coefficient estimates, standard errors (SE), associated t values and p values for all predictors of post hoc linear mixed model predicting response time for older adults. Model includes a random intercept for items and subjects, a random slope for Integration for both items and subjects and a random slope for Verb type for subjects.

Coefficient	Estimate	Std. error	t value	p value	
(Intercept)	**1055.411**	**55.3074**	**19.083**	**< 0.001**	*******
Working Memory	-14.1461	38.55634	-0.367	0.714	
Verb type	**-481.522**	**65.82336**	**-7.315**	**< 0.001**	*******
Integration	61.74264	42.69748	1.446	0.148	
Processing Speed	-31.0554	22.59838	-1.374	0.170	
Handgrip	25.70137	70.50421	0.365	0.715	
Physical Activity	0.68534	0.54891	1.249	0.212	
Verbal IQ	**-42.1103**	**11.68993**	**-3.602**	**< 0.001**	*******
Working Memory*Verb type	10.732	44.55644	0.241	0.810	
Working Memory*Integration	**67.78142**	**28.98878**	**2.338**	**0.019**	*****
Verb type*Integration	**-147.942**	**44.99419**	**-3.288**	**0.001**	******
Verb type*Processing Speed	39.89153	26.12684	1.527	0.127	
Integration*Processing Speed	-7.14032	17.00927	-0.42	0.675	
Verb type*Handgrip	-16.2572	81.54918	-0.199	0.842	
Integration*Handgrip	-30.9208	53.1573	-0.582	0.561	
Verb type*Physical Activity	-0.105	0.63407	-0.166	0.868	
Integration*Physical Activity	-0.12001	0.41203	-0.291	0.771	
Verb type*Verbal IQ	20.00802	13.49842	1.482	0.138	
Integration* Verbal IQ	-7.0546	8.7682	-0.805	0.421	
Working Memory*Verb type*Integration	12.01456	26.8349	0.448	0.654	
Verb type*Integration*Processing Speed	-16.1621	15.58484	-1.037	0.300	
Verb type*Integration*Handgrip	-93.1008	49.59537	-1.877	0.060	.
Verb type*Integration*Physical Activity	-0.08623	0.38033	-0.227	0.821	

*Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 ‘ ’ 1*

Discussion

Our study was designed to investigate whether there is decline in syntactic comprehension in healthy ageing. We investigate elementary syntactic comprehension of phrases such as “I cook” and “I spuff”. We demonstrated the following three key findings: 1) there is decline the syntactic comprehension of healthy older adults compared to young adults, in accuracy as well as response times; 2) the age-related decline in the accuracy of syntactic comprehension was stronger for phrases with real verbs, while the age-related decline in the response times of syntactic comprehension was stronger for phrases with pseudoverbs; 3) there was a high degree of individual variation in age-related decline, which was explained in part by differences in working memory and processing speed. Modulations of processing speed and working memory on syntactic comprehension present a complex picture, which can be summarized as follows. In older adults, processing speed determined comprehension accuracy, such that higher processing speed scores were associated with higher accuracy, particularly for phrases with real verbs. The response times were less sensitive to processing speed than were the accuracy scores. In addition, we found that the effect of working memory on comprehension accuracy depended on verb type and whether integration was correct or incorrect, but only in the younger age group. In addition, older adults with a higher working memory span took relatively less time to respond to syntactically correct phrases. We turn to a discussion of each of our three key findings next.

We have convincingly demonstrated that there is age-related decline in syntactic comprehension, when processing two-word phrases with real verbs as well as with pseudoverb. The effects were demonstrated in accuracy as well as response times: older adults were less accurate and slower than young adults. Previous literature on syntactic comprehension in older adults has predominantly used semantically meaningful sentences with complex syntactic structures. Most of these studies did not show age-related decline in processing these sentences (Campbell et al., 2016; Davis et al., 2014; Meunier et al., 2014; Samu et al., 2017; M. A. Shafto & Tyler, 2014; M. a Shafto et

al., 2014), although some studies did (Antonenko et al., 2013; Obler, Fein, Nicholas, & Albert, 1991b). Our results are in line with the latter set of studies.

We also found that there was a different pattern in age-related decline for phrases with real verbs versus for phrases with pseudoverbs. In the accuracy of syntactic comprehension the age-related decline in syntactic comprehension was stronger for phrases with real verbs than for phrases with pseudoverb. In the reaction times on the other hand, age-related decline in syntactic comprehension was stronger for phrases with pseudoverbs than for phrases with real verbs. One possible interpretation of this pattern of findings is that decline in syntactic comprehension is strongest in the absence of lexical-semantic information, which causes older adults to produce slower responses in order to make more accurate decisions. It has been suggested previously that older adults prioritize accurate responses over fast responses (Forstmann et al., 2011; Starns & Ratcliff, 2010). The stronger decline in response times when syntactically processing phrases in the absence of lexical-semantic information could shed some light on why some previous studies did not show any decline in syntactic processing. In these studies, syntactic comprehension is probed in the context of full sentence structures. Even when sentence length was deliberately kept short, these sentences were rich in semantic content and inevitably provide a more extensive context than would two word phrases. Our findings of reduced syntactic comprehension in a contextually deprived context suggest that the availability of additional lexical-semantic information reduces the decline in syntactic comprehension that comes with aging.

Our interpretation that syntactic decline is more pronounced in the absence of semantic information, is also in line with Peelle et al. (2003). The absence of semantic information can be considered a processing challenge. Peelle et al. (2003) investigated the influence of varying processing challenges on syntactic comprehension in older adults in a different way, by measuring syntactic comprehension of subject- and object relative clause sentences at varying speech rates. While older adults were slower than younger adults at all speech rates tested, this age difference became larger with increased speech rates. In other words, older adults took disproportionally

longer to give their comprehension responses at an increased level of processing challenge. Likewise, in the current study, the effect of processing challenge resulted in disproportionally increased response times in older adults when contextual constraints were reduced from a two word phrase with a meaningful content to a similar phrase structure without any representation in the mental lexicon. It should be noted that in the Peelle et. al study, comprehension accuracy only decreased in older adults at very fast speech rates, whereas in the current study, accuracy was already lower compared to young adults for the comprehension of real verb phrases, that is, when processing challenge was at relative minimum. However, as argued above, it could be that the minimal phrases used in the current study already provided a higher processing challenge than the semantically richer sentence structures used by Peelle et al. (2003).

A second explanation for the divergence in results of the current study compared to previous findings of preserved syntactic comprehension, is that the measure of syntactic comprehension used in the current study may draw on a different aspect of syntax. Studies that capitalize on syntactic ambiguity evaluate comprehension by asking questions about the thematic roles assigned to the agent or patient in the sentence (i.e 'who is doing what', e.g 'what is the gender of the agent in the sentence'). A correct response requires comprehension of the full sentence structure, which indirectly requires comprehension of the syntactic structure. In contrast, the measure of syntactic comprehension in the current study focuses on evaluating syntactic agreement. This study thus taps into a different aspect of syntactic processing. Grammatically judgements may provide a more direct measure of syntactic ability. Taken together, the current findings have important and novel implications for our understanding of age-related differences in syntactic comprehension.

This leads us to our finding that there was individual variation in the age-related decline in syntactic comprehension. Processing speed explained the individual variation in the accuracy of syntactic comprehension of older adults. Specifically, individuals with a lower processing speed capacity had lower accuracy scores. The influence of processing speed on syntactic ability is

consistent with a large literature suggesting general processing speed impacts language processing (G. Waters & Caplan, 2007). Notably, this effect was present in the older age group only in our study. These findings are in line with the contention that the general slowing of processing speed that is associated with age impairs cognitive functioning (Salthouse, 1996). Critically, in the experiment that required the least processing load (the real verb phrases) and the condition that gave the highest accuracy (the incorrect integration condition) performance of the older adults with a higher processing speed capacity was equivalent to younger adults' performance.

We found that accuracy as well as response times for syntactic comprehension were modulated by working memory. Comprehension accuracy in young adults was modulated by working memory capacity, particularly when processing phrases with pseudoverbs. Remarkably, we did not find evidence in the older age group that higher working memory capacity was associated with higher accuracy. We saw, however, that working memory did modulate the response times in older adults, such that individuals with higher working memory capacity showed decreased response times in correctly identifying a syntactically correct phrase.

In summary, the results of the current study shed new light on the decline in syntactic comprehension in healthy ageing. Whereas previous studies have primarily investigated complex syntactic structures and focused on syntactic ambiguity, we investigated syntactic comprehension of the elementary building blocks of syntactic processing: syntactic agreement of pronoun and verb. Older adults were slower and less accurate compared to younger adults. This decline seems to increase in the absence of semantic contextual information, which causes older adults to produce slower responses in order to make more accurate decisions. In line with these findings, accuracy for older adults was positively related to processing speed capacity. Taken together our results provide very clear evidence that syntactic comprehension declines in healthy aging.

Acknowledgements

We warmly thank all our participants for their contributions to this research and Denise Clissett, the coordinator of Patient and Lifespan Cognition participant database at the University of Birmingham, for recruiting and scheduling participants. We thank Sian Beardsell, Jessica Low, Laura McCue, Zarin Shahabi and Emily Stone for their help with collecting the data.

Author contributions

C.P, L.W. and K.S. designed the study. C.P. analysed the data. C.P., L.W. and K.S. wrote the manuscript.

References

Antonenko, D., Brauer, J., Meinzer, M., Fengler, A., Kerti, L., Friederici, A. D., & Flöel, A. (2013). Functional and structural syntax networks in aging. *NeuroImage*, *83*, 513–523. https://doi.org/10.1016/j.neuroimage.2013.07.018

Auyeung, T. W., Lee, J. S. W., Kwok, T., & Woo, J. (2011). Physical frailty predicts future cognitive decline—a four-year prospective study in 2737 cognitively normal older adults. The journal of nutrition, health & aging, 15(8), 690-694.

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, E., Friederici, A., & Wulfeck, B. (1987). Comprehension in aphasia: a cross-linguistic study. *Brain and Language*, *32*(1), 19–67. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3651807

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging Gracefully: Compensatory Brain Activity in High-Performing Older Adults. *NeuroImage*, *17*(3), 1394–1402. https://doi.org/10.1006/nimg.2002.1280

Campbell, K. L., Samu, D., Davis, S. W., Geerligs, L., Mustafa, A., & Tyler, L. K. (2016). Robust Resilience of the Frontotemporal Syntax System to Aging. *The Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.4561-15.2016

Colcombe, S. J., Kramer, A. F., Erickson, K. I., Scalf, P., McAuley, E., Cohen, N. J., … Elavsky, S. (2004). Cardiovascular fitness, cortical plasticity, and aging. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(9), 3316–3321. https://doi.org/10.1073/pnas.0400266101

Davis, S. W., Zhuang, J., Wright, P., & Tyler, L. K. (2014). Age-related sensitivity to task-related modulation of language-processing networks. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2014.08.017

Deutsch, A., Bentin, S., & Katzttt, L. (1994). Lexical and Semantic Influences on Syntactic Processing. *Haskins Laboratories Status Report on Speech Research*, 119–120.

Fengler, A., Meyer, L., & Friederici, A. D. (2016). How the brain attunes to sentence processing: Relating behavior, structure, and function. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2016.01.012

Forstmann, B. U., Tittgemeyer, M., Wagenmakers, E.-J., Derrfuss, J., Imperati, D., & Brown, S. (2011). The speed-accuracy tradeoff in the elderly brain: a structural model-based approach. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(47), 17242–17249. https://doi.org/10.1523/JNEUROSCI.0309-11.2011

Grossman, M., Cooke, A., DeVita, C., Alsop, D., Detre, J., Chen, W., & Gee, J. (2002). Age-related changes in working memory during sentence comprehension: an fMRI study. *NeuroImage*, *15*(2), 302–317. https://doi.org/10.1006/nimg.2001.0971

Hopkins, K. A., Kellas, G., & Paul, S. T. (1995). Scope of Word Meaning Activation During Sentence Processing by Young and Older Adults. *Experimental Aging Research*, *21*(2), 123–142. https://doi.org/10.1080/03610739508254273

Hutchison, K. A., Balota, D. A., Neely, J. H., Cortese, M. J., Cohen-shikora, E. R., Tse, C., … Buchanan,

E. (2013). The semantic priming project, 1099–1114. https://doi.org/10.3758/s13428-012-0304-z

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

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

Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences*, *6*(8), 350–356. https://doi.org/10.1016/S1364-6613(02)01947-2

Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2011). Experimental effects and individual differences in linear mixed models: Estimating the relationship between spatial, object, and attraction effects in visual attention. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2010.00238

Lara, J., Cooper, R., Nissan, J., Ginty, A. T., Khaw, K.-T., Deary, I. J., … Mathers, J. C. (2015). A proposed panel of biomarkers of healthy ageing. *BMC Medicine*, *13*(1), 222. https://doi.org/10.1186/s12916-015-0470-9

Meunier, D., Stamatakis, E. A., & Tyler, L. K. (2014). Age-related functional reorganization, structural changes, and preserved cognition. *Neurobiology of Aging*, *35*(1), 42–54. https://doi.org/10.1016/j.neurobiolaging.2013.07.003

Obler, L. K., Fein, D., Nicholas, M., & Albert, M. L. (1991a). Auditory comprehension and aging: Decline in syntactic processing. *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716400005865

Payne, B. R., Grison, S., Gao, X., Christianson, K., Morrow, D. G., & Stine-Morrow, E. A. L. (2014). Aging and individual differences in binding during sentence understanding: Evidence from temporary and global syntactic attachment ambiguities. *Cognition*, *130*(2), 157–173. https://doi.org/10.1016/j.cognition.2013.10.005

Raz, N. (2009). Decline and compensation in aging brain and cognition: promises and constraints. Preface. *Neuropsychology Review*, *19*(4), 411–414. https://doi.org/10.1007/s11065-009-9122-1

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8759042

Samu, D., Campbell, K. L., Tsvetanov, K. A., Shafto, M. A., Brayne, C., Bullmore, E. T., … Tyler, L. K. (2017). Preserved cognitive functions with age are determined by domain-dependent shifts in network responsivity. *Nature Communications*. https://doi.org/10.1038/ncomms14743

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 language decline in healthy ageing. Retrieved from http://arxiv.org/abs/1801.01441

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

Shafto, M. a, Tyler, L. K., Dixon, M., Taylor, J. R., Rowe, J. B., Cusack, R., … Matthews, F. E. (2014). The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC Neurology*, *14*, 204. https://doi.org/10.1186/s12883-014-0204-1

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging*, *25*(2), 377–390. https://doi.org/10.1037/a0018022

Tyler, L. K., Shafto, M. A., Randall, B., Wright, P., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Preserving syntactic processing across the adult life span: The modulation of the frontotemporal language system in the context of age-related atrophy. *Cerebral Cortex*, *20*(2), 352–364. https://doi.org/10.1093/cercor/bhp105

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., … Pinker, S.

(1997). A Neural Dissocation within Language: Evidence That the Mental Dictionary Is Part of Declarative Memory, and That Grammatical Rules Are Processed by the Procedural System. *Journal of Cognitive Neuroscience*, *9*(2), 266–276. https://doi.org/10.1162/jocn.1997.9.2.266

Waters, G., & Caplan, D. (2007). The relationship between age, processing speed, working memory capacity, and language comprehension. https://doi.org/10.1080/09658210344000459

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

Weatherholtz K, Campbell-Kibler K, Jaeger TF (2014) Socially-mediated syntactic alignment. *Language Variation and Change, 26*:387–420

Wingfield, A., Alexander, A. H., & Cavigelli, S. (1994). Does Memory Constrain Utilization of Top-Down Information in Spoken word Recognition? Evidence from Normal Aging. *Language and Speech*, *37*(3), 221–235. https://doi.org/10.1177/002383099403700301

Wingfield, A., Peelle, J.E., & Grossman, M. (2003). Speech Rate and Syntactic Complexity as Multiplicative Factors in Speech Comprehension by Young and Older Adults. *Aging, Neuropsychology and Cognition*, *10*(4), 310-322. https://doi.org/10.1076/anec.10.4.310.28974