Poor coherence in older people's speech is explained by impaired semantic and executive processes

Paul HOFFMAN*

Ekaterina LOGINOVA

Asatta RUSSELL

Centre for Cognitive Ageing and Cognitive Epidemiology (CCACE), Department of Psychology, University of Edinburgh

* Correspondence to:
Dr. Paul Hoffman
Centre for Cognitive Ageing and Cognitive Epidemiology,
Department of Psychology, University of Edinburgh,
7 George Square, Edinburgh, EH8 9JZ, UK
Tel: +44 (0) 131 650 4654
Email: p.hoffman@ed.ac.uk

Acknowledgements

PH is supported by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and Medical Research Council (MRC) is gratefully acknowledged. We are grateful to Wing Yee Ho, Eszter Kalapos, Jasmine Kulay, Khushboo Mehra and Adam Ryde for assistance with data collection.

Abstract

The ability to speak coherently is essential for effective communication but declines with age: older people more frequently produce tangential, off-topic speech. The cognitive factors underpinning this decline are poorly understood. We predicted that maintaining coherence relies on effective regulation of activated semantic knowledge about the world, and particularly on the selection of currently relevant semantic representations to drive speech production. To test this, we collected speech samples and measures of executive and semantic ability from 60 young and older adults, using a novel computational method to quantify coherence. Semantic selection ability predicted coherence, as did level of semantic knowledge and domain-general executive ability. These factors fully accounted for the age-related coherence deficit. Our results indicate that maintaining coherence in speech becomes more challenging as people age because they accumulate more knowledge but are less able to effectively regulate how it is activated and used.

Introduction

Engaging in a conversation is a complex cognitive act, in which a speaker must settle on a topic for discussion, generate a series of appropriate, relevant and hopefully interesting statements and monitor their speech as the discourse unfolds to ensure that they remain on-topic. Discourse that successfully navigates these challenges is said to be coherent: it consists of a series of well-connected statements all related to a shared topic, making it easy to comprehend [1,2]. The ability to produce coherent speech tends to decline as people grow older. Older adults are more likely to produce tangential, off-topic utterances in conversation [1,3] and to provide irrelevant information when telling a story [4,5] or describing an object [6]. Such changes reduce the effectiveness of communication and the quality of older people's verbal interactions [7,8]. Indeed, less coherent speech is associated with higher levels of stress and less satisfaction in social interactions [3,8,9]. Researchers have often made a distinction between and local coherence (LC), the degree to which adjoining utterances related meaningfully to one another, and global coherence (GC), the degree to each utterance relates to the topic under discussion [1,10]. Most studies have reported larger declines in GC in later life, though reductions in LC have also been observed [1,5,11,12].

Explanations for age-related decline in coherence have typically focused on deterioration in domain-general cognitive control processes, which are assumed to be involved in the monitoring and selection of topics during speech [13]. This view is supported by a handful of studies reporting that coherence in speech is predicted by performance on non-verbal tasks requiring cognitive control [9,13,14]. One particular theory holds declines in coherence result from a reduced ability to inhibit irrelevant information, which means that older people are less able to prevent irrelevant or off-topic ideas from intruding into their discourse [3,15]. This view is supported by evidence that performance on cognitive tests requiring inhibition, such as the Stroop test and Trails test, predicts the level of coherence in older adults' speech [3,12]. However, this factor does not appear to provide a complete explanation for age-related coherence declines. This is likely because previous studies have overlooked the critical role that conceptual knowledge about the world (i.e., semantic knowledge) plays in the generation of meaningful, coherent speech. Here, we propose and test the hypothesis that coherence depends not only on domain-general executive resources, but specifically on the ability to regulate the activation of semantic knowledge, ensuring that only the most relevant concepts are selected for inclusion in speech.

All propositional speech relies on the retrieval and use of semantic knowledge. This is true at the lexical level, since the selection of words for production is guided by their meaning. But it is also true at the broader conceptual level, as the content of our speech is informed by our general semantic knowledge about the topic under discussion. For example, describing one's favourite season requires access to stored semantic knowledge about the typical characteristics of each time of year, the events associated with them and so on. Thus, the coherence of an individual's discourse is likely to be critically determined by (a) the quality of the semantic knowledge they have on the topic under discussion and (b) by their ability to retrieve and select the most appropriate information to talk about. It is important to note that these elements of semantic processing are served by distinct neural systems. Current theories hold that representations of semantic knowledge are centred on the anterior temporal cortices and that a separate "semantic control" system provides top-down regulation of the activation and selection of concepts from this store, based on current situational demands [16-21]. These principal components of semantic cognition – the store of representations and the control system – can be independently impaired following brain damage [17,18].

Importantly for the present study, recent evidence suggests some aspects of semantic control are impaired in later life. Hoffman [22] recently tested the verbal semantic abilities of 100 young and older adults. In common with many previous studies, older people were found to have a broader vocabulary, indicating a richer repository of semantic knowledge [23-28]. Unlike previous studies, we also probed semantic control abilities, using two experimental paradigms commonly used in cognitive neuroscience studies [29,30]. The first task probed the ability to engage in controlled search of the semantic knowledge store in order to detect weak associations between concepts. No age differences were found for this ability. The second tested the ability to select among competing semantic associations (hereafter termed *semantic selection*). The semantic selection task required participants to inhibit prepotent but irrelevant semantic information in favour of task-relevant aspects of knowledge. Older adults were less successful than young people, indicating age-related decline in this aspect of semantic control. Evidence for age-related decline in controlled selection of semantic information is consistent with a meta-analysis of 47 functional neuroimaging studies indicating that older adults show reduced activity in the left prefrontal region most strongly linked with this ability [31]. The ability to select task-relevant semantic representations may be crucial in speech production because

it may allow people to select the most relevant aspects of knowledge for use in speech and thus to avoid irrelevant shifts in topic.

Recent studies therefore suggest that ageing is associated with both positive and negative changes in the function of the semantic system. Here, we tested whether these changes could account for age-related declines in the coherence of speech. Young and older adults were asked to produce samples of speech in response to a series of prompts and the coherence of these samples was estimated using a novel computational approach. We hypothesised in particular that individuals with reduced semantic selection abilities would produce less coherent speech, since they would be less able to prevent irrelevant semantic information from influencing their responses. Importantly, we tested whether semantic abilities had unique effects on coherence, after accounting for the effects of domain-general executive function. We used the Trails test as a measure of domain-general executive function because it is a well-established task which draws on various aspects of executive control including task-switching and inhibition [32-34] and also because it has previously been linked to poor coherence in speech [3,12]. Finally, we also investigated the production of speech under conditions of divided attention, by including a dual-task condition in which participants completed a secondary manual task while speaking. We included this condition because a previous study has shown that people produce less coherent speech when their attention is divided and that this effect interacts with age [11].

Method

Participants: Thirty young adults, aged between 18 and 30 (mean = 19.3), were recruited from the undergraduate Psychology course at the University of Edinburgh and participated in the study in exchange for course credit. Thirty older adults, aged between 61 and 91 (mean = 76.0), were recruited from the Psychology department's volunteer panel. These participants were a subset of a group taking part in a larger study of semantic processing, some data from which have been reported elsewhere [22]. All participants reported to be in good health with no history of neurological or psychiatric illness. Demographic information for each group is shown in Supplementary Table 1. Young and older adults did not differ significantly in years of education completed ($t(58) = 0.93$, $p = 0.36$).

Assessments of cognitive and semantic ability: Participants completed the following tests of general cognitive function and executive ability: Mini-Mental State Examination, Trail-making task, verbal fluency (see Supplementary Information for details). As a measure of domain-general executive function, we computed the ratio of completion times for Trails part B to Trails part A. High ratios indicated disproportionately slow performance on part B, indicative of poor executive function. A ratio rather than a difference score was used as this measure minimises the influence of differences in general processing speed [33,34]. Participants also completed two tasks probed breadth of semantic knowledge: lexical decision [35] and synonym matching adapted Mill Hill vocabulary scale; [36] (for examples, see Supplementary Figure 1). As scores on these tasks were strongly correlated, they were averaged to give a single measure of breadth of semantic knowledge.

Semantic control was assessed using a 2 x 2 within-subjects experimental design that manipulated the need for semantic control in two different tasks following [29]. In the first task, participants made semantic decisions based on global semantic association. They were presented with a probe word and asked to select its semantic associate from either two or four alternatives (see Supplementary Figure 1 for examples). The strength of association between the probe and target was manipulated: the associate was either strongly associated with the probe (e.g., *town-city*) or more weakly associated (e.g., *iron-ring*). The Weak Association condition was assumed to place greater demands on controlled retrieval of semantic information, because automatic spreading of activation in the semantic network would not be sufficient to identify the correct response [37]. In the second task, participants were asked to select items that matched on particular features. At the beginning of each block, participants were given a feature to attend to (e.g., Colour). On each trial, they were provided with a probe and were asked to select the item that was most similar on the specified feature. Trials manipulated the semantic congruency of the probe and target. On Congruent trials, the probe and target shared a pre-existing semantic relationship, in addition to matching on the currently relevant feature (e.g., *cloud-snow*). In contrast, on Incongruent trials the probe and target shared no meaningful relationship, other than matching on the specified feature (e.g., *salt-dove*). Furthermore, on these trials one of the foils had a strong semantic relationship with the probe, although it did not match on the currently relevant feature (*salt-pepper*). Incongruent trials placed high demands on semantic selection processes for two reasons: first, because there was no pre-existing semantic relationship between probe and target to boost

activation of the target and second, because the strong but irrelevant relationship between the probe and foil had to be ignored.

Speech elicitation task: Samples of speech were elicited under conditions of undivided and divided attention. On *speech-only* trials, participants were asked to speak for 60 seconds at a time in response to a written prompt (for full list of prompts, see Supplementary Information). Prompts were designed to probe particular areas of semantic knowledge (e.g., *What sort of things do you have to do to look after a dog?*). Participants read each prompt on a computer monitor and pressed a key when ready to begin speaking. After 60s, a tone sounded to signal the end of the trial. Participants were instructed to continue speaking until they heard the tone. On *dual-task* trials, participants were asked to complete an attention-demanding secondary task while speaking [38]. On these trials, a horizontal array of four squares appeared on screen. Every 3s, a red circle appeared in one of the squares and participants pressed a key corresponding to its location. This task was performed continuously throughout the speech elicitation period. Seven speech samples were obtained in the speech-only condition and seven in the dual-task condition. Finally, to obtain a baseline measure of secondary task performance, there were five *secondary-only* trials where the secondary task was performed without speech for 60s. These trials were interspersed amongst the speech elicitation trials.

Spoken responses were digitally recorded for later transcription. The main dependent variables analysed were computed measures of global and local coherence (GC and LC), as described below. A number of other speech markers were also computed and were included in supplementary analyses (see Supplementary Information for details).

Coherence computations: Measures of local and global coherence were generated using an automated computational linguistic approach. Analyses were implemented in R; the code is publicly available and can easily be applied to new samples (https://osf.io/8atfn/). Our approach used latent semantic analysis (LSA) [39], one of a number of computational techniques in which patterns of word co-occurrence are used to construct a high-dimensional semantic space. The LSA method utilises a large corpus of natural language divided into a number of discrete contexts. The corpus is used to generate a co-occurrence matrix registering how often each word appears in each context. Data reduction techniques are then applied to this matrix, with the result that each word is represented as a high-dimensional vector. Words that are used in similar contexts (and are thus assumed to have related meanings) are assigned similar

vectors. Word similarities derived in this way are strong predictors of human judgements of semantic relatedness and human performance on a range of tasks [40,41].

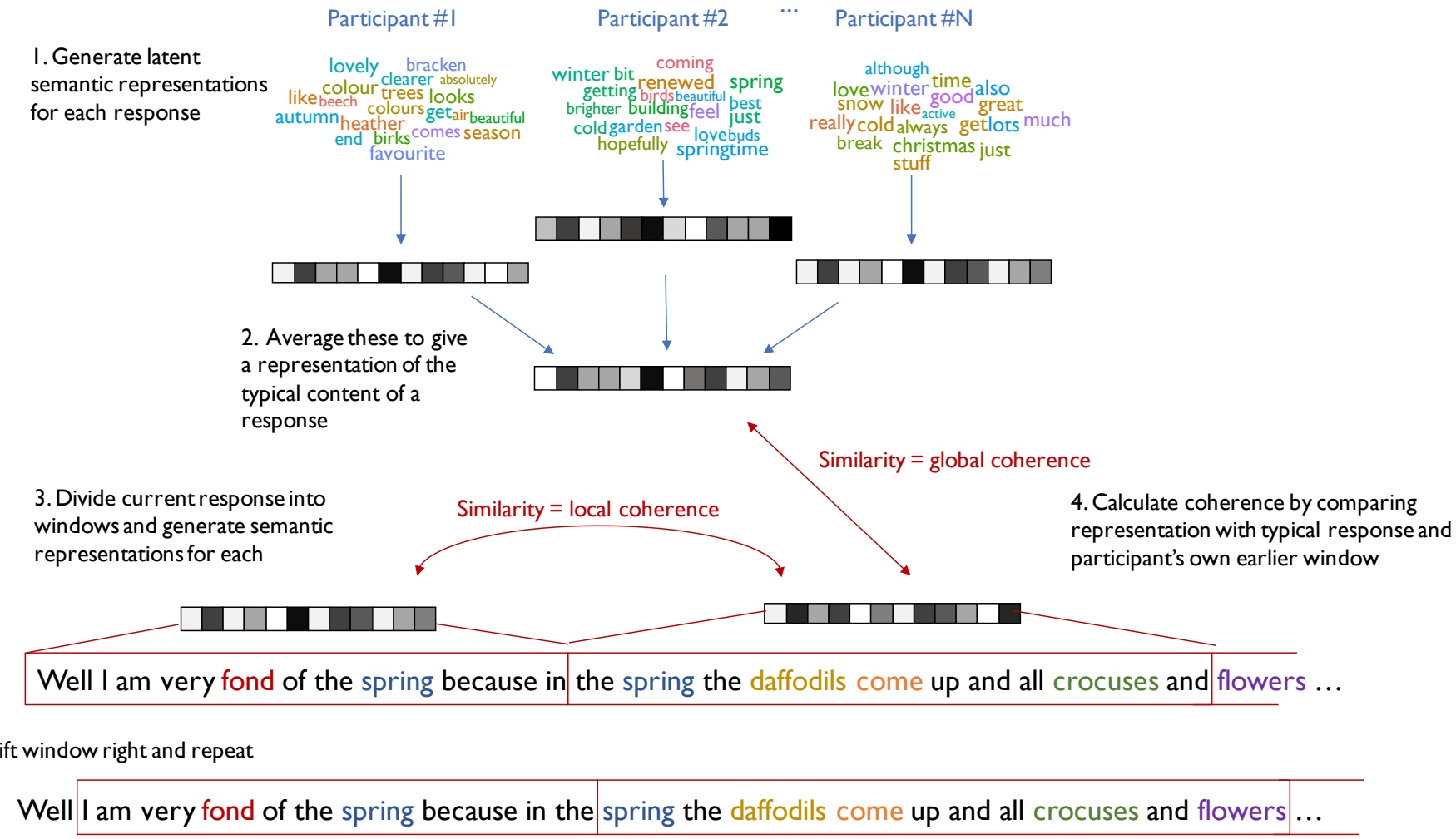
Importantly, vectors for individual words can be combined linearly to represent the meanings of whole sentences and passages of speech/text [42]. A number of researchers have used this property to generate estimates of coherence for texts or spoken samples based on LSA similarity measures [2,42-44]. The present work builds on this approach. The overall strategy we took was to divide each speech sample into smaller windows (of 20 words each) and to use LSA to generate vector representations of the semantic content of each window. Coherence was assessed by measuring the similarity of the vector for each window with that of the previous window (LC) and with a vector representing the typical semantics of responses to the same prompt (GC). This process is illustrated in Figure 1.

First, an LSA representation of each participant's response was computed by averaging the LSA vectors of all the words they produced in response to the same prompt (for details of the averaging method and vector space used, see Supplementary Information). These were averaged to give a composite vector that represented the typical semantic content produced in response to that prompt (this step excluded the target response). Next, the target response was analysed using a moving window approach. The target response was divided into windows of 20 words in length. An LSA vector was computed for each window. Local coherence was defined as the cosine similarity of the semantic vector for the current window with that of the previous window. Therefore, in common with other researchers [2,43], we define LC as the degree to which adjoining utterances convey semantically related content. A low LC value would be obtained if a participant switched abruptly between topics during their response.

Global coherence was defined as similarity of the vector for each window with the composite vector derived from the other participants' responses. Therefore, GC was a measure of how much the target response matched the typical semantic content of responses to that prompt. A low GC value would be obtained if a participant tended to talk about other topics that were semantically unrelated to the topic being probed. Thus, our measure of GC captured the degree to which participants maintained their focus on the topic under discussion, in line with the definition used by other researchers [1,12].

Once GC and LC had been calculated, the window moved one word to the right and the process was repeated, until all windows had been assessed. GC and LC values were averaged across windows to give overall values for each response, which were multiplied by 100 for ease of presentation.

Figure 1: Process for computing coherence in speech samples



5. Shift window right and repeat

Well | I am very fond of the spring because in the | spring the daffodils come up and all crocuses and flowers | …

Examples of responses with high vs. low coherence values are provided in Supplementary Table 2. The LSA-based coherence measures were validated by comparing them with judgements of coherence provided by human raters for a subset of speech samples (see Supplementary Information for details). There were strong correlations between rated coherence and LSA-based coherence for GC (r = 0.68) and LC (r = 0.37). Test-retest reliability was also high (see Supplementary Information).

Statistical analyses: A series of linear mixed effects models were used to investigate the effects of the experimental manipulations and individual differences in semantic and executive ability on characteristics of speech. The dependent variable in the first analysis was speech rate in words per minute (WPM). This was analysed in a linear mixed model with a 2 x 2 (age group x task) factorial design. We performed this analysis because previous studies have found that older people speak more slowly than young people and that speech rate is reduced under dual-task conditions [11,45]. It was important to investigate this possibility in our data as speech rate might have an impact on coherence. For example, participants who spoke very quickly could cover a wider range of topics in 60s, increasing the likelihood that their response would lose coherence. Since we found that speech rate was indeed influenced by both age and task, this variable was included as a covariate in later analyses.

Our main hypotheses were tested with a series of nested models which used GC and LC as dependent variables (in parallel). The first model included age group and task as predictors, as well as speech rate. Next, we added the Trails ratio score as an additional predictor, to test the hypothesis that general executive ability influences coherence. In the final model, we added three semantic task scores as predictors, to test the hypothesis that semantic abilities are an additional important determiner of coherence. The semantic task scores included were:

1. Semantic knowledge: the mean of accuracy on the lexical decision and synonym matching tasks.
2. Semantic selection: accuracy on the Incongruent condition of the feature association task, which required participants to select task-relevant aspects of semantic knowledge and inhibit irrelevant associations.
3. Weak association: accuracy on the Weak Association condition of the global association task, which required controlled retrieval of semantic knowledge to identify less salient semantic relationships.

Next, to establish whether the observed effects were specific to coherence, we investigated whether other characteristics of speech showed similar effects. We computed six measures of the lexical characteristics of the words produced in each speech sample (see Figure 4 and Supplementary Information). Principal components analysis was performed on these (along with the coherence measures) and used to extract four underlying factors, which were promax-rotated. Scores on each of these factors were then analysed using the same series of nested mixed models employed in the main analysis of GC and LC.

Finally, to analyse performance on the concurrent secondary task, we used a linear mixed model with group and task (secondary-only vs. dual-task) as predictors. The dependent variable was RT. We then added GC and LC values to the model as predictors to determine whether coherence was related to secondary task performance. All study data are available online (https://osf.io/8atfn/).

Mixed effects models were constructed and tested using the recommendations of Barr et al. [46]. We specified a maximal random effects structure, including random intercepts for participants and prompts as well as random slopes across participants for the effect of task and random slopes across prompts for task and age group. Continuous predictors were standardised prior to entry in the model. The significance of fixed effects was assessed by comparing the full model with a reduced model that was identical in every respect except for the exclusion of the effect of interest. Likelihood-ratio tests were used to determine whether the inclusion of the effect of interest significantly improved the fit of the model.

Results

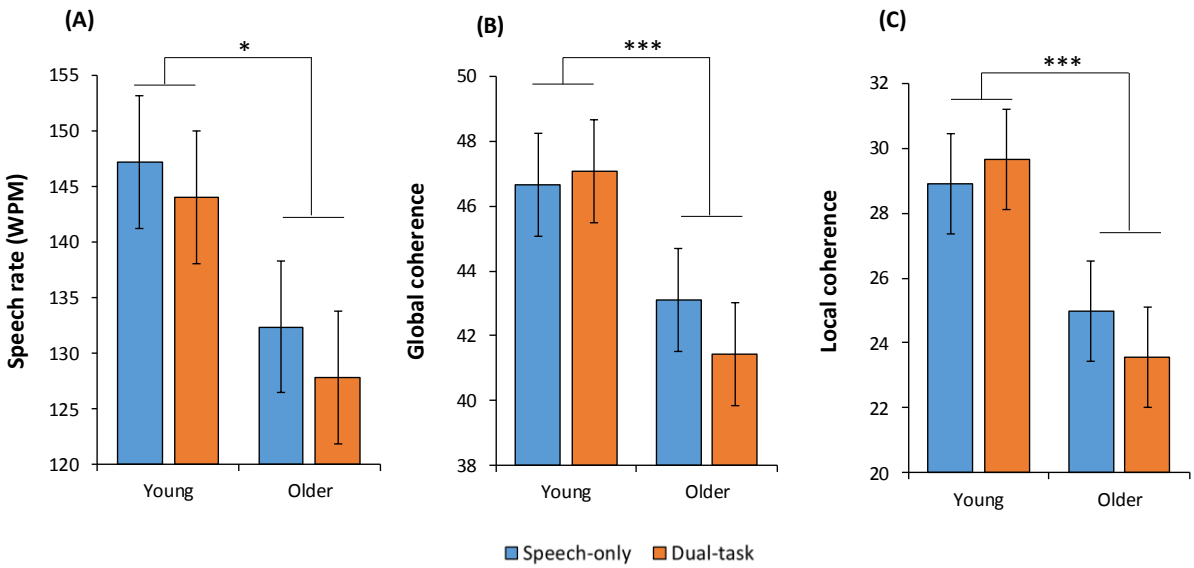
Assessments of cognitive and semantic ability: Mean scores on the background cognitive tests are reported in Supplementary Table 1. Young people were faster to respond in the Trails test and produced more items in category fluency. Older people produced slightly more words in letter fluency, however. There were no group differences in MMSE scores, with all participants scoring at least 26/30.

The full analysis of the semantic tasks is reported in Supplemental Information, with mean scores in each condition presented in Supplementary Figure 2. The older group scored significantly higher on the vocabulary tests of semantic knowledge, indicating that they had a broader set of verbal semantic information available to them. The manipulation of association strength had similar effects in young and older people, suggesting that the ability to retrieve

less salient semantic knowledge was equivalent in the two groups. However, older people showed a larger effect of the congruency manipulation during feature judgements; performing more poorly in the incongruent condition. This indicates that the older group had difficulty in selecting task-relevant semantic knowledge and inhibiting irrelevant associations.

Speech rate: We first considered the effect of our experimental manipulations on rate of speech production (number of words produced per minute). The results are shown in Figure 2A. Mixed effects modelling indicated that speech rate was influenced both by age group (B = -7.74, se = 3.81, p = 0.046), with older participants tending to speak more slowly, and by task (B = -1.93, se = 0.79, p = 0.016), with fewer words produced under dual-task conditions. We therefore included speech rate as a covariate in subsequent analyses, to ensure that effects on coherence were not attributable to this variable.

Figure 2: Effects of age and task on speech rate and coherence * = $p < 0.05$; *** = $p < 0.001$



Predictors of coherence: We began by assessing the internal reliability of GC and LC measures over the fourteen prompts used to elicit speech samples. Cronbach's alpha was high for both measures (GC = 0.83; LC = 0.79), indicating that stable individual differences in coherence were present over the various topics about which participants were asked to speak. GC and LC values were also strongly correlated with one another (r = 0.79), suggesting that both are

closely linked, as found in previous studies. In the older group, age was negatively correlated with GC (r = -0.64) and LC (r = -0.56).

Next we investigated the effects of our experimental manipulations on the GC of participants' speech. The first model included age group and task as predictors, with speech rate as an additional covariate (see Table 1 for results). Age group was a strong predictor of GC: as predicted, older participants produced markedly less coherent speech than young people (see Figure 2B). As shown in Table 1, the dual-task manipulation had no effect on GC. The interaction between age and task fell just short of statistical significance ($p = 0.052$). This suggests that there may be a weak tendency for the effect of the task manipulation to be larger in older people. Speech rate was a negative predictor of GC, indicating that participants who spoke more quickly showed a greater tendency to deviate from the topic being probed.

The addition of Trails ratio scores (Model 2) significantly improved the fit of the model ($\chi^2(1) = 11.9$, $p < 0.001$). As expected, participants with a smaller ratio of Trails B to A (indicating better executive ability) had higher GC values. However, this was not sufficient to explain the lower GC values of older people: a significant difference between the young and older groups remained. The inclusion of semantic test scores (Model 3) yielded a further improvement in model fit ($\chi^2(3) = 10.1$, $p = 0.018$). The estimated effects of the test scores on GC are plotted in Figure 3. Participants with higher scores on the semantic selection test produced more coherent speech. The effect of Trails ratio was also significant and there was also a tendency for individuals with higher semantic knowledge scores to produce *less* coherent speech. Weak association task scores were not a significant predictor of GC. Importantly, there was no remaining effect of age group in this model (see Table 1), suggesting that lower levels of GC in older adults can be explained in terms of their lower semantic selection and higher semantic knowledge scores.

LC measurements were subjected to the same sequence of analyses, with broadly similar results (see Table 1). The first model revealed an effect of age group with no effect of task and a non-significant interaction (see Figure 2C). In Model 2, Trails ratio was again a significant predictor of coherence, but a significant age effect remained. In contrast, the age group effect was not significant once semantic scores were included (Model 3). Scores on the semantic selection task were a significant predictor of LC, with participants who performed poorly on this task tending to be less coherent (see Figure 3).

Figure 3: Effects of semantic and executive abilities on coherence * = p < 0.05; ** = p < 0.01; *** = p < 0.001

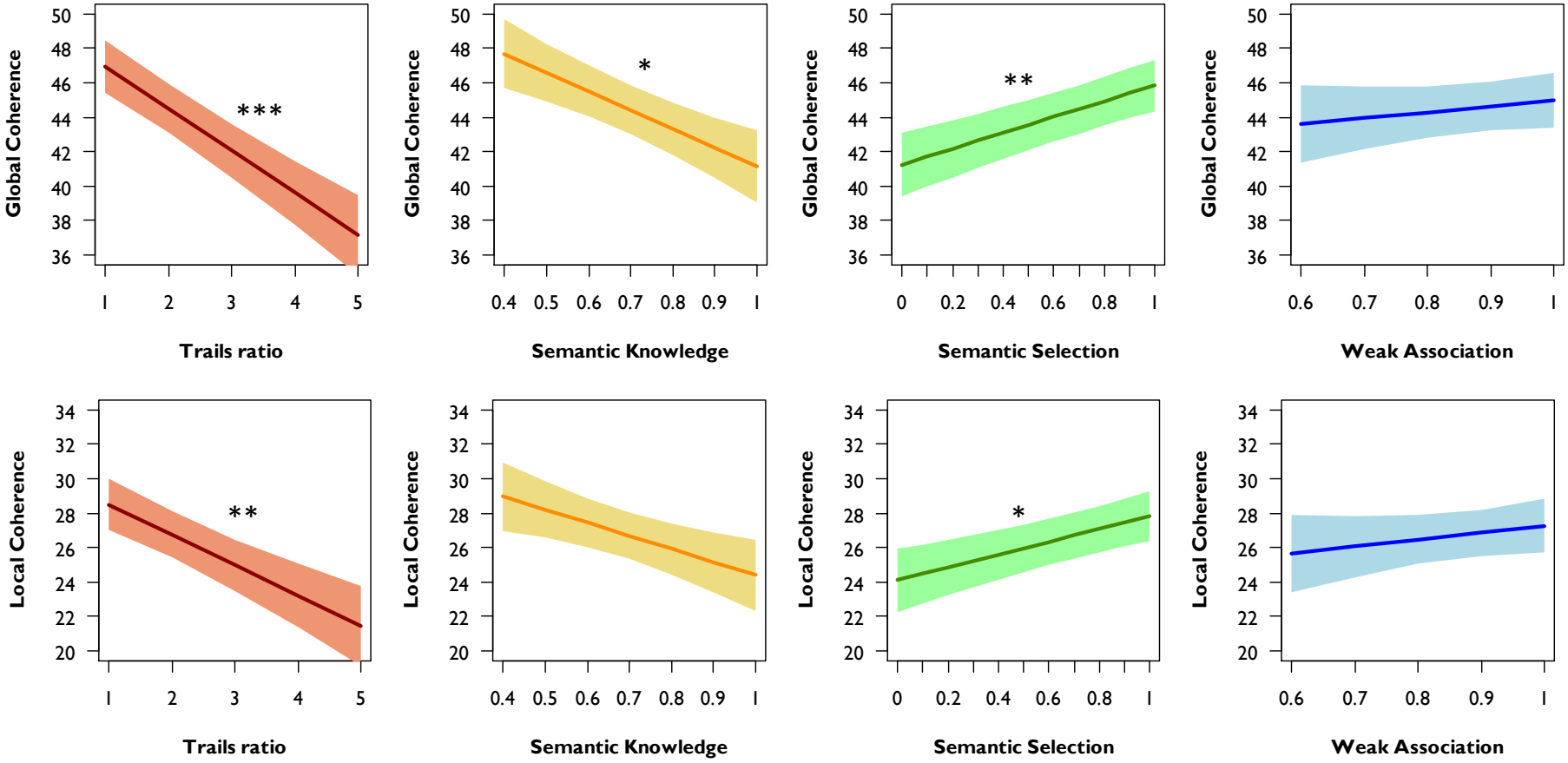


Table 1: Results of mixed effects models predicting global and local coherence

	Model 1			Model 2			Model 3		
	B	*se*	*p*	*B*	*se*	*p*	*B*	*se*	*p*
Global Coherence									
(Intercept)	44.6	1.45	<.001	44.6	1.43	<.001	44.6	1.42	<.001
Age	**-2.30**	**0.54**	**<.001**	**-1.97**	**0.50**	**<.001**	-0.68	0.73	.35
Task	-0.31	0.24	.20	-0.31	0.24	.20	-0.30	0.24	.21
Age*Task	-0.53	0.25	.056	-0.53	0.25	.052	-0.53	0.25	.052
Response length	**-0.97**	**0.38**	**.014**	**-0.90**	**0.36**	**.016**	**-0.86**	**0.35**	**.019**
Trails ratio				**-1.61**	**0.45**	**<.001**	**-1.71**	**0.42**	**<.001**
Semantic knowledge							**-1.63**	**0.73**	**.028**
Semantic selection							**1.16**	**0.42**	**.007**
Weak association							0.29	0.52	.58
Local Coherence									
(Intercept)	26.8	1.36	<.001	26.8	1.35	<.001	26.8	1.34	<.001
Age	**-2.51**	**0.58**	**<.001**	**-2.27**	**0.57**	**<.001**	-1.42	0.80	.081
Task	-0.16	0.34	.63	-0.16	0.33	.63	-0.16	0.34	.64
Age*Task	-0.55	0.32	.10	-0.55	0.32	.10	-0.55	0.32	.10
Response length	0.01	0.39	.98	0.09	0.38	.82	0.12	0.38	.74
Trails ratio				**-1.16**	**0.45**	**.012**	**-1.23**	**0.43**	**.006**
Semantic knowledge							-1.14	0.76	.14
Semantic selection							**0.94**	**0.43**	**.034**
Weak association							0.33	0.54	.54

Predictors of other speech measures: The purpose of this analysis was to establish whether the observed effects of executive and semantic abilities on coherence were specific to this aspect of speech, or whether they would be observed for other characteristics of speech. Principal components analysis was used to reduce the nine properties of speech into four latent factors, shown in Figure 4A. These were the only factors with eigenvalues greater than one and together they explained 81% of the variance. GC and LC loaded exclusively on Factor 2, confirming that coherence emerged as a discrete characteristic of speech. Factor 1 indexed the use of long, abstract, late-acquired words in speech, so appeared to reflect access to

complex vocabulary. High scores on Factor 3 were associated with use of low frequency, concrete nouns that were low in semantic diversity. This factor may reflect the degree to which speech referenced highly specific concepts, so we labelled it semantic specificity. Finally, high scores on Factor 4 were characterised by high type:token ratio and a low proportion of closed-class words, which are indicative of greater lexical diversity.

Scores on each factor were subjected to the same series of mixed effects analyses used for the analysis of GC and LC. The full results of these analyses are shown in Supplementary Table 3, while the effects of participants' semantic and executive scores on each factor are presented in Figure 4B. The results for Factor 2 (coherence) were the same as previously observed for GC and LC separately: lower coherence was associated with poorer Trails and semantic selection performance but with better semantic knowledge. Importantly, no other factor showed the same pattern. The only other significant effects were that semantic knowledge was positively correlated with scores on factors 1 (vocabulary) and 4 (lexical diversity), indicating that participants with broader semantic knowledge used a broader and more complex range of vocabulary when speaking. This analysis therefore confirms that the participants' executive and semantic selection abilities had a specific effect on their coherence but not on other aspects of their speech production.

Secondary task performance: Analysis of the secondary manual task is reported in Supplementary Information. In brief, older people had slower RTs and both groups were slower when the task was combined with speaking. Importantly, however, there was no interaction between these factors, indicating that the requirement to perform two tasks affected both groups equally (see Supplementary Figure 3). Participants' GC and LC scores did not predict performance on the secondary task, ruling out the possibility of a trade-off between secondary task performance and maintenance of coherence.

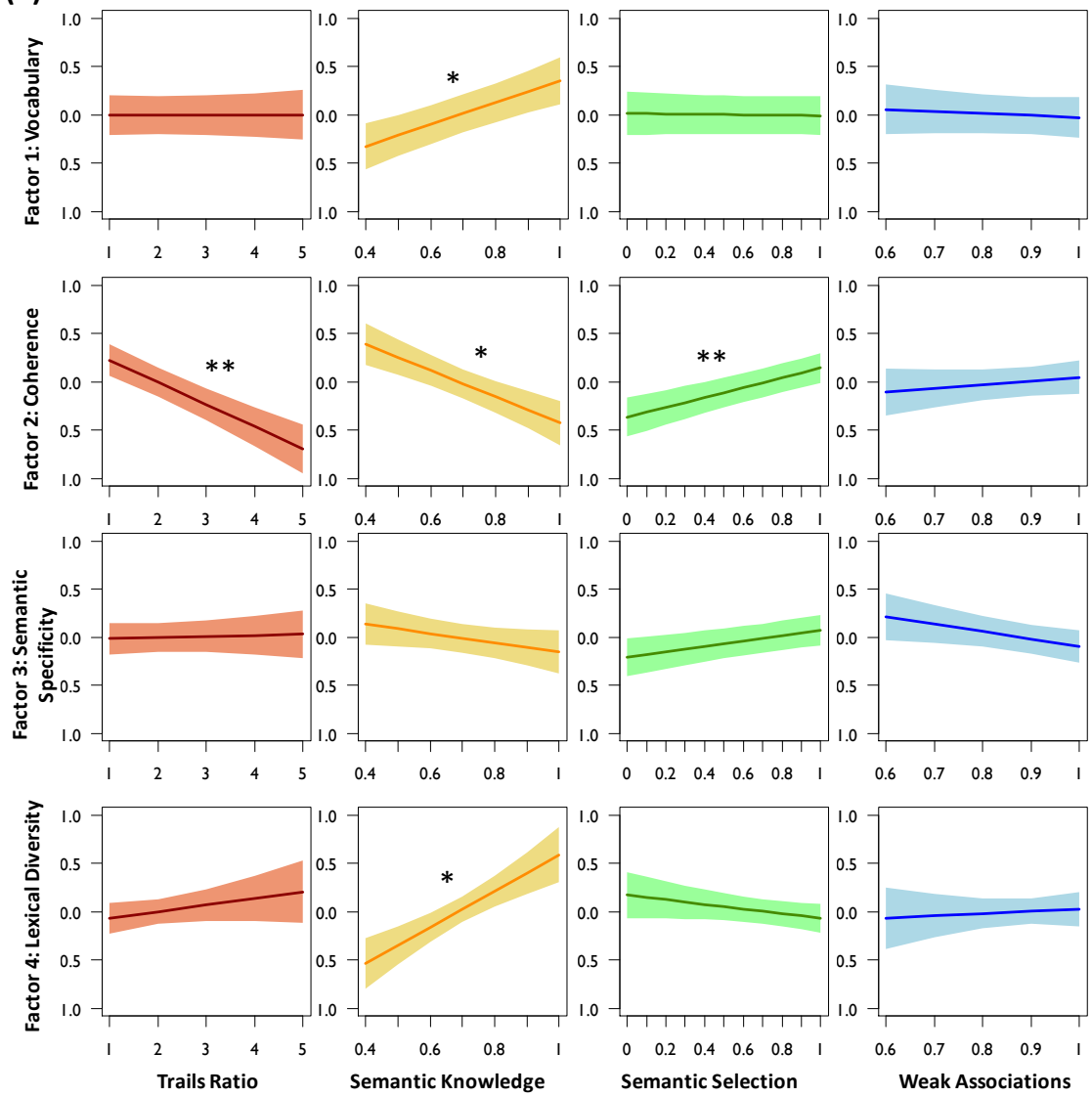
Figure 4: (A) Principal components analysis identifying four latent speech factors. (B) Effects of semantic and executive abilities on factor scores

** = p < 0.05; ** = p < 0.01; *** = p < 0.001*

(A)

	1: Vocabulary	2: Coherence	3: Semantic Specificity	4: Lexical Diversity
Nouns: Age of acquisition	.93	.00	.10	-.07
Nouns: Number of phonemes	.84	.05	.16	.07
Nouns: Concreteness	-.64	.00	.53	-.01
Local coherence	.04	.95	.03	.01
Global coherence	.01	.93	.01	.06
Nouns: Semantic diversity	-.35	.07	-.93	.00
Nouns: Frequency	.11	-.12	-.85	.01
% Closed class	.01	-.24	.06	-.91
Type: token ratio	.00	-.36	.08	.71

(B)



Discussion

The ability to produce coherent speech is critical for effective communication but tends to decline in later life. Here, we investigated cognitive factors that predict this decline, using computational linguistic techniques to quantify the coherence of speech produced by a large group of young and older adults. We replicated previous findings indicating that individuals with greater domain-general executive ability produce more coherent speech, but this effect did not fully account for age differences in coherence. However, when we included semantic abilities as additional predictors of coherence, the age group difference was eliminated. Semantic selection ability emerged as a positive predictor of coherence while breadth of semantic knowledge was a negative predictor. These effects were specific to coherence and not to other characteristics of speech, and were not attributable to differences in speech rate. Our results indicate that older people produce less coherent speech (a) because they are less skilled at selecting the most relevant aspects of semantic knowledge to include in their speech and (b) because they have a larger set of semantic knowledge to select from.

First and foremost, our results establish that the monitoring and control of discourse is influenced by the function of the semantic system, in addition to domain-general executive resources. In particular, we found that the ability to select task-relevant semantic information was a strong predictor of coherence in speech. The task we used to assess this ability is well-established as a measure of semantic control e.g., [29,30,47] and required participants to attend to specific semantic features of objects while inhibiting strong but irrelevant semantic associations. Our data indicate that similar selection and inhibitory demands are present during the production of discourse. A conversational cue, such as "what's your favourite season?", initially causes a wide range of knowledge to become activated in the semantic system. Some of this information will be useful in answering the question and some less so. Coherent communication requires the speaker to select the subset of that information which is directly relevant at the current time, while suppressing aspects of knowledge that are activated but less pertinent. These demands grow as the narrative develops and new associations are activated. Of course, the knowledge that drives speech production is not solely semantic in nature – specific episodic memories and more general autobiographical knowledge will often be triggered as well. The prompts used in the present study were designed to elicit general knowledge rather than specific personal experiences. However, episodic and semantic memories are mutually interdependent [48] and it was clear that

participants drew on both in generating their responses. It is likely that selection mechanisms for these distinct types of memory are shared to some degree. Indeed, the left mid ventrolateral prefrontal cortex, the brain region most closely associated with semantic selection, responds to selection demands in all three domains [37,49,50].

Another important question is the extent to which the selection processes involved in regulating semantic knowledge overlap or interact with other, domain-general executive functions. This is an area of active debate, with researchers proposing that some aspects of the regulation of semantic knowledge are performed by domain-general systems for competition resolution while others require more specialised neural resources [18,29,30,51]. In this study we employed a single measure of non-semantic executive ability, the Trails test, which is assumed to load on various aspects of executive control including task-switching and inhibition [32-34]. Although this task was a strong predictor of coherence, we found that semantic selection had an additional, independent effect. This indicates that the relationship between coherence and semantic selection cannot simply be attributed to poorer general executive ability. The present data do not permit a more fine-grained parcellation of non-semantic executive functions, however, and this is an important target for future work.

Our results indicate that that both domain-general cognitive control abilities and more specific selection processes make independent contributions to the planning and regulation of speech. However, declines in the ability to select among competing aspects of semantic knowledge appears to be responsible for age-related deficits in speech coherence. In fact, consistent with this conclusion, there is evidence for age-related structural and functional changes to brain regions involved in semantic selection. Selection among competing semantic representations has been linked most strongly with activity in the left mid ventrolateral prefrontal cortex [37,52]. The lateral prefrontal cortex exhibits the greatest reductions in cortical volume as a function of age [53,54]. In addition, a recent meta-analysis of 47 functional neuroimaging studies found that older adults activated this region less strongly than young people during semantic processing [31]. Therefore, the idea that older people's lack of coherence is due to reduced efficiency of semantic selection in consistent with age-related changes in the known neural correlates of this ability.

We also found that the breadth of participants' semantic knowledge influenced their coherence. Individuals with a wider range of lexical-semantic knowledge tended to be *less* coherent. This effect is consistent with the notion that selecting appropriately from activated knowledge is critical to maintaining coherence. This challenge becomes greater the more

information one has in one's semantic store, simply because more concepts are likely to be activated in response to a particular cue. Episodic memory studies have reported an analogous effect known as the "fan effect" [55,56]. These indicate that the more pieces of information one associates with a particular stimulus, the more difficult it is to retrieve a specific piece of information from the set. In a similar vein, our data indicate that being more knowledgeable in itself brings greater challenges in identifying the most relevant aspects of knowledge to use in speech. It is very well-established that older people have greater semantic and general world knowledge, as was the case in our study [26,27,57]. This factor, combined with older people's reduced ability to select among activated concepts, both explained their reduced coherence. It is also worth noting that breadth of semantic knowledge also predicted the use of more sophisticated vocabulary (i.e., more late-acquired, abstract nouns) and greater lexical diversity for a similar result, see [58]. Therefore it appears that quantity of semantic knowledge has diverse effects on the characteristics of speech, unlike semantic selection, which impacts specifically on coherence.

We found that conditions of divided attention had no overall effect on the coherence of speech, in contrast to the findings of Kemper et al. [11]. However, we note that our secondary task appeared less demanding that that used by Kemper and colleagues. Our secondary task produced a reduction in speech rate of around 5 WPM, compared with 20-40 WPM in Kemper et al. Despite this, we did find a non-significant trend (p = 0.056) towards an interaction of dual-task demands with age (for the GC measure). There was a suggestion that the secondary task may have had a small effect on the coherence of older people, while young people appeared unaffected (see Figure 2B). There remains a possibility therefore that divided attention has a particular detrimental effect on coherence in older adults. This could have important implications for conversations conducted in everyday situations in which speakers may simultaneously be engaged in other activities (e.g., talking while driving, shopping etc.). Future studies with more demanding concurrent tasks are needed to assess this possibility and its interaction with semantic abilities.

Previous studies have found varying effects of age on coherence, depending on the speech elicitation task used. Typically, narratives elicited from verbal prompts, as in the present study, reveal the greatest decrements in coherence while tasks that elicit speech using visual stimuli, such as picture descriptions or story-telling from comic strips, produce smaller effects [12,59]. These results fit well with our assertion that the ability to select relevant semantic content is a major determinant of coherence. When a pictured stimulus is used to cue speech,

it acts as a source of constraint over semantic activation. Upon analysing the image, knowledge related to the objects and events depicted automatically comes to mind and can be used to drive speech production. If any irrelevant concepts become activated during this process, they can easily be eliminated on the basis that they are not present in the image. In contrast, constructing a response to a brief verbal prompt is a trickier proposition, since a wide range of potentially relevant information may be activated and no external cues are available to guide selection. Of course, the monologues elicited in the present study are a rather extreme example of this phenomenon. In everyday conversational speech, environmental cues are often available to guide the selection of speech content. For instance, a look of confusion from the speaker's interlocutor can indicate when a loss of topic has occurred and a well-timed question could direct the speaker back to the topic in hand. Such cues can only be effective if speakers are sensitive to them, however, and evidence suggests that speakers with poor coherence are also less skilled at interpreting social cues [8].

Finally, it is important to consider an underlying assumption often made in the literature on coherence: that greater coherence is always a desirable characteristic for speech. Many situations do require specific information to be communicated quickly and efficiently and in these cases, it is beneficial to be able to provide the most germane information without digression to other topics. For example, there is evidence that individuals who are less coherent in conversation perform poorly at communicating task-related information to a partner in an experimental setting [7]. In other situations, however, a less focused approach to speech may have its advantages. When the goal of a conversation is to entertain, rather than to convey specific information, the ability to shift focus away from the original topic may be beneficial. Indeed, older people are generally considered to produce more enjoyable stories than young people [60]. One notable study collected responses of young and older people to questions about life events and asked judges to rate them on various dimensions [59]. Older people produced more off-topic speech than young people and their narratives were rated as less focused. However, while less coherent speakers were rated as less clear and focused, they were also considered to be more interesting and to have produced better stories. In summary, the ability to communicate coherently is critical in many but not all everyday conversations. It is possible that the most effective communicators are those who can tailor their selection of content to the current situational demands, focusing tightly on the subject at hand when required to but broadening their focus at other times. Little is known at present

about how coherence interacts with these situational demands and this is one area where more research is needed.

References

1 Glosser, G. & Deser, T. A comparison of changes in macrolinguistic and microlinguistic aspects of discourse production in normal aging. *J Gerontol* **47**, P266-272 (1992).

2 Foltz, P. W. Discourse coherence and LSA. *Handbook of latent semantic analysis*, 167-184 (2007).

3 Arbuckle, T. Y. & Gold, D. P. Aging, inhibition, and verbosity. *J Gerontol* **48**, P225-232 (1993).

4 Juncos-Rabadan, O., Pereiro, A. X. & Rodriguez, M. S. Narrative speech in aging: quantity, information content, and cohesion. *Brain Lang* **95**, 423-434, doi:10.1016/j.bandl.2005.04.001 (2005).

5 Marini, A., Boewe, A., Caltagirone, C. & Carlomagno, S. Age-related differences in the production of textual descriptions. *J Psycholinguist Res* **34**, 439-463 (2005).

6 Long, M. R., Horton, W. S., Rohde, H. & Sorace, A. Individual differences in switching and inhibition predict perspective-taking across the lifespan. *Cognition* **170**, 25-30 (2018).

7 Arbuckle, T. Y., Nohara-LeClair, M. & Pushkar, D. Effect of off-target verbosity on communication efficiency in a referential communication task. *Psychology and aging* **15**, 65 (2000).

8 Pushkar, D. *et al.* Social behavior and off-target verbosity in elderly people. *Psychology and aging* **15**, 361 (2000).

9 Gold, D., Andres, D., Arbuckle, T. & Schwartzman, A. Measurement and correlates of verbosity in elderly people. *Journal of gerontology* **43**, P27-P33 (1988).

10 Kintsch, W. & Vandijk, T. A. Toward a Model of Text Comprehension and Production. *Psychological Review* **85**, 363-394, doi:Doi 10.1037//0033-295x.85.5.363 (1978).

11 Kemper, S., Schmalzried, R., Hoffman, L. & Herman, R. Aging and the vulnerability of speech to dual task demands. *Psychology and aging* **25**, 949 (2010).

12 Wright, H. H., Koutsoftas, A. D., Capilouto, G. J. & Fergadiotis, G. Global coherence in younger and older adults: Influence of cognitive processes and discourse type. *Aging, Neuropsychology, and Cognition* **21**, 174-196 (2014).

13 Kintz, S., Fergadiotis, G. & Wright, H. H. in *Cognition, Language and Aging* (ed Heather Harris Wright) 81-106 (John Benjamins Publishing Company, 2016).

14 North, A. J., Ulatowska, H. K., Macaluso-Haynes, S. & Bell, H. Discourse performance in older adults. *The International Journal of Aging and Human Development* **23**, 267-283 (1986).

15 Marini, A. & Andreetta, S. in *Cognition, Language and Aging* (ed Heather Harris Wright) 55-79 (John Benjamins Publishing Company, 2016).

16 Badre, D. & Wagner, A. D. Semantic retrieval, mnemonic control, and prefrontal cortex. *Behavioral and Cognitive Neuroscience Reviews* **1**, 206-218 (2002).

17 Lambon Ralph, M. A., Jefferies, E., Patterson, K. & Rogers, T. T. The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience* **18**, 42-55, doi:10.1038/nrn.2016.150 (2017).

18 Jefferies, E. The neural basis of semantic cognition: Converging evidence from neuropsychology, neuroimaging and TMS. *Cortex* **49**, 611-625 (2013).

19 Hoffman, P., McClelland, J. L. & Lambon Ralph, M. A. Concepts, control and context: A connectionist account of normal and disordered semantic cognition. *Psychological Review* **125**, 293-328, doi:http://dx.doi.org/10.1037/rev0000094 (2018).

20 Yee, E. & Thompson-Schill, S. L. Putting concepts into context. *Psychonomic bulletin & review* **23**, 1015-1027 (2016).

21 Jefferies, E. & Lambon Ralph, M. A. Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. *Brain* **129**, 2132-2147, doi:10.1093/brain/awl153 (2006).

22 Hoffman, P. An individual differences approach to semantic cognition: Divergent effects of age on representation, retrieval and selection. *bioRxiv*, https://doi.org/10.1101/182170, doi:https://doi.org/10.1101/182170 (2017).

23 Park, D. C. *et al.* Models of visuospatial and verbal memory across the adult life span. *Psychology and aging* **17**, 299-320, doi:10.1037//0882-7974.17.2.299 (2002).

24 Nilsson, L. G. Memory function in normal aging. *Acta Neurologica Scandinavica* **107**, 7-13 (2003).

25 Nyberg, L., Bäckman, L., Erngrund, K., Olofsson, U. & Nilsson, L.-G. Age differences in episodic memory, semantic memory, and priming: Relationships to demographic, intellectual, and biological factors. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* **51**, P234-P240 (1996).

26 Salthouse, T. A. Localizing age-related individual differences in a hierarchical structure. *Intelligence* **32**, 541-561 (2004).

27 Rönnlund, M., Nyberg, L., Bäckman, L. & Nilsson, L.-G. Stability, growth, and decline in adult life span development of declarative memory: cross-sectional and longitudinal data from a population-based study. *Psychology and aging* **20**, 3 (2005).

28 Grady, C. The cognitive neuroscience of ageing. *Nature Reviews Neuroscience* **13**, 491-505 (2012).

29 Badre, D., Poldrack, R. A., Pare-Blagoev, E. J., Insler, R. Z. & Wagner, A. D. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* **47**, 907-918 (2005).

30 Whitney, C., Kirk, M., o'Sullivan, J., Lambon Ralph, M. A. & Jefferies, E. Executive-semantic processing is underpinned by a large-scale neural network: Revealing the contribution of left prefrontal, posterior temporal and parietal cortex to controlled retrieval and selection using TMS. *Journal of Cognitive Neuroscience* **24**, 133-147 (2011).

31 Hoffman, P. & Morcom, A. M. Age-related changes in the neural networks supporting semantic cognition: A meta-analysis of 47 functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews* **84**, 134-150 (2018).

32 Sanchez-Cubillo, I. *et al.* Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society* **15**, 438-450 (2009).

33 Arbuthnott, K. & Frank, J. Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *Journal of clinical and experimental neuropsychology* **22**, 518-528 (2000).

34 Salthouse, T. A. What cognitive abilities are involved in trail-making performance? *Intelligence* **39**, 222-232 (2011).

35 Baddeley, A., Emslie, H. & Smith, I. N. *The Speed and Capacity of Language Processing test.* (Thames Valley Test Company, 1992).

36 Raven, J. C., Raven, J. E. & Court, J. H. *Mill Hill vocabulary scale*. (Psychological Corporation, 1989).

37 Badre, D. & Wagner, A. D. Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia* **45**, 2883-2901 (2007).

38 Craik, F. I. M., Govoni, R., NavehBenjamin, M. & Anderson, N. D. The effects of divided attention on encoding and retrieval processes in human memory. *Journal of*

Experimental Psychology-General **125**, 159-180, doi:Doi 10.1037/0096-3445.125.2.159 (1996).

39 Landauer, T. K. & Dumais, S. T. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review* **104**, 211-240 (1997).

40 Recchia, G. & Jones, M. N. More data trumps smarter algorithms: Comparing pointwise mutual information with latent semantic analysis. *Behavior Research Methods* **41**, 647-656 (2009).

41 Bullinaria, J. A. & Levy, J. P. Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods* **39**, 510-526 (2007).

42 Foltz, P. W., Kintsch, W. & Landauer, T. K. The measurement of textual coherence with latent semantic analysis. *Discourse Processes* **25**, 285-307, doi:10.1080/01638539809545029 (1998).

43 Elvevag, B., Foltz, P. W., Weinberger, D. R. & Goldberg, T. E. Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia. *Schizophrenia research* **93**, 304-316, doi:DOI 10.1016/j.schres.2007.03.001 (2007).

44 Graesser, A. C., McNamara, D. S., Louwerse, M. M. & Cai, Z. Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods* **36**, 193-202 (2004).

45 Kemper, S., Herman, R. E. & Lian, C. H. The costs of doing two things at once for young and older adults: Talking while walking, finger tapping, and ignoring speech of noise. *Psychology and aging* **18**, 181 (2003).

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

47 Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K. & Farah, M. J. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *P Natl Acad Sci USA* **94**, 14792-14797 (1997).

48 Binder, J. R., Desai, R. H., Graves, W. W. & Conant, L. L. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex* **19**, 2767-2796 (2009).

49 Dobbins, I. G. & Wagner, A. D. Domain-general and domain-sensitive prefrontal mechanisms for recollecting events and detecting novelty. *Cerebral Cortex* **15**, 1768-1778 (2005).

50 King, J., Hartley, T., Spiers, H., Maguire, E. & Burgess, N. Anterior prefrontal involvement in episodic retrieval reflects contextual interference. *Neuroimage* **28**, 256-267 (2005).

51 Nagel, I. E., Schumacher, E. H., Goebel, R. & D'Esposito, M. Functional MRI investigation of verbal selection mechanisms in lateral prefrontal cortex. *Neuroimage* **43**, 801-807 (2008).

52 Badre, D. & Wagner, A. D. Frontal lobe mechanisms that resolve proactive interference. *Cerebral Cortex* **15**, 2003-2012 (2005).

53 Raz, N. *et al.* Aging, sexual dimorphism, and hemispheric asymmetry of the cerebral cortex: replicability of regional differences in volume. *Neurobiol Aging* **25**, 377-396 (2004).

54 Fjell, A. M. *et al.* High consistency of regional cortical thinning in aging across multiple samples. *Cerebral cortex*, bhn232 (2009).

55 Anderson, J. R. & Reder, L. M. The fan effect: New results and new theories. *Journal of Experimental Psychology-General* **128**, 186-197 (1999).

56 Anderson, J. R. Retrieval of propositional information from long-term memory. *Cognitive psychology* **6**, 451-474 (1974).

57 Verhaeghen, P. Aging and vocabulary score: A meta-analysis. *Psychology and aging* **18**, 332-339, doi:10.1037/0882-7974.18.2.332 (2003).

58 Kemper, S. & Sumner, A. The structure of verbal abilities in young and older adults. *Psychology and aging* **16**, 312 (2001).

59 James, L. E., Burke, D. M., Austin, A. & Hulme, E. Production and perception of" verbosity" in younger and older adults. *Psychology and aging* **13**, 355 (1998).

60 Ryan, E. B., See, S. K., Meneer, W. B. & Trovato, D. Age-based perceptions of language performance among younger and older adults. *Communication Research* **19**, 423-443 (1992).