Reductions in prefrontal activation predict off-topic utterances during speech production

Paul HOFFMAN

School of Philosophy, Psychology and Language Sciences, University of Edinburgh

Correspondence to:

Dr. Paul Hoffman

School of Philosophy, Psychology and Language Sciences,

University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK

Tel: +44 (0) 131 650 4654

Email: p.hoffman@ed.ac.uk

Acknowledgements

The study was funded 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. Imaging was carried out at the Edinburgh Imaging Facility (www.ed.ac.uk/edinburgh-imaging), University of Edinburgh, which is part of the SINAPSE collaboration (www.sinapse.ac.uk). I am grateful to Cyril Pernet and Sergio Della Sala for advice on study design and participant screening.

Abstract

The ability to speak coherently is essential for effective communication but declines with age: older people more frequently produce tangential, off-topic speech. Little is known, however, about the neural systems that support coherence in speech production. Few functional neuroimaging studies have investigated extended speech production and none have attempted to relate neural activity recorded during spontaneous speech to its coherence. To address these shortcomings, the present study used fMRI to investigate extended speech production in healthy older adults. Novel computational linguistic analyses were used to quantify the coherence of utterances produced in the scanner, allowing identification of the neural correlates of coherence for the first time. When people produced highly coherent speech, they showed increased activity in inferior prefrontal cortex (BA45), an area implicated in selection of task-relevant knowledge from semantic memory, and in rostrolateral prefrontal cortex (BA10), implicated more generally in the planning of complex goal-directed behaviours. Both effects were observed bilaterally, with BA10 effects preceding those in left BA45. These findings indicate that multiple prefrontal systems make distinct contributions to coherence in speech. More generally, they demonstrate that neural activity during spontaneous speech production can be predicted from formal analysis of speech content, paving the way for this approach to be adopted more widely.

Introduction

Although it often appears effortless, production of conversational speech involves complex cognitive processes. In order to speak coherently on a topic, individuals must constantly regulate their speech output to ensure that they produce statements that are informative and relevant to the topic under discussion. Numerous studies have found that this ability declines in later life (Kintz, Fergadiotis, & Wright, 2016). Older adults are more likely to produce tangential, off-topic utterances during conversation (Arbuckle & Gold, 1993; Glosser & Deser, 1992) and to provide irrelevant information when telling a story (Juncos-Rabadan, Pereiro, & Rodriguez, 2005; Marini, Boewe, Caltagirone, & Carlomagno, 2005) or describing an object (Long, Horton, Rohde, & Sorace, 2018). These changes reduce the effectiveness of communication and the quality of older people's verbal interactions (Arbuckle, Nohara-LeClair, & Pushkar, 2000; Pushkar et al., 2000). Indeed, less coherent speech is associated with higher levels of stress and less satisfaction in social interactions (Arbuckle & Gold, 1993; Gold, Andres, Arbuckle, & Schwartzman, 1988; Pushkar et al., 2000). Researchers have often made a distinction between local coherence, the degree to which adjoining utterances related meaningfully to one another, and global coherence, the degree to each utterance relates to the topic under discussion (Glosser & Deser, 1992; Kintsch & Vandijk, 1978). Most studies have reported larger declines in global coherence in later life (Glosser & Deser, 1992; Kemper, Schmalzried, Hoffman, & Herman, 2010; Marini et al., 2005; Wright, Koutsoftas, Capilouto, & Fergadiotis, 2014) and this aspect of coherence is the focus of the present work.

Despite its importance for effective communication, there is limited understanding of the cognitive and neural mechanisms involved in maintaining coherence during speech. Explanations for the age-related decline in coherence have focused on changes to domain-general cognitive control processes, rather than aspects of the language system per se (Kintz et al., 2016). A number of studies have reported that more coherent individuals perform better on tests of cognitive or executive control, such as the Stroop task (Arbuckle & Gold, 1993; Gold et al., 1988; Kintz et al., 2016; North, Ulatowska, Macaluso-Haynes, & Bell, 1986; Wright et al., 2014). One view is that declines in coherence result from a reduced ability to inhibit irrelevant information, which means that older people are less able to prevent tangential or off-topic ideas from intruding into their discourse (Arbuckle & Gold, 1993; Marini

& Andreetta, 2016). Supporting this idea, we recently demonstrated that the ability to select task-relevant aspects of semantic knowledge is a strong determinant of coherence (Hoffman, Loginova, & Russell, 2018a).

Less is known about the brain regions involved in promoting on-topic, coherent utterances during speech. This is in part due to a dearth of functional neuroimaging studies probing speech production beyond the single-sentence level. Collection of fMRI data during extended speech production has sometimes been considered problematic due to signal contamination caused by excessive head and jaw movements (Gracco, Tremblay, & Pike, 2005). A small number of neuroimaging studies have overcome these technical challenges, however. These indicate that, in addition to engaging areas involved in motor planning and production, extended speech production activates an left-lateralised network including ventral temporal and inferior parietal regions involved in the representation of semantic knowledge and prefrontal regions associated with planning and cognitive control (AbdulSabur et al., 2014; Awad, Warren, Scott, Turkheimer, & Wise, 2007; Blank, Scott, Murphy, Warburton, & Wise, 2002; Braun, Guillemin, Hosey, & Varga, 2001; Stephens, Silbert, & Hasson, 2010). The majority of participants in these studies were healthy young adults, however; no studies to date have used fMRI to investigate extended speech production in older people specifically.

Perhaps more importantly, no studies have investigated how brain activity during coherent speech differs from that associated with off-topic, incoherent speech. Identifying the neural predictors of coherence is critical if we are to understand the root causes of age-related declines in the ability to produce coherent speech. To address this issue, in the present study a group of older people were scanned with fMRI while they spoke about a series of topics for 50 seconds at a time. To ameliorate the effect of speech-related head movements, principal components analysis was used to partition genuine haemodynamic signal changes from movement-related artefact (Tierney et al., 2016) (for similar approaches, see Behzadi, Restom, Liau, & Liu, 2007; Silbert, Honey, Simony, Poeppel, & Hasson, 2014). Speech produced in the scanner was recorded, transcribed and subjected to novel computational linguistic analyses which quantified its coherence at each time-point during the scanning. This allowed investigation of (a) regions activated by older people during extended speech production, relative to a rote speech baseline and (b) regions whose activity varied as a

function of the coherence of the speech produced. Thus the analysis identified brain regions that support coherence for the first time.

The main prediction was that the left inferior frontal gyrus (IFG) would play an important role in the production of highly coherent speech. Left IFG has been identified as the key node in a network of regions that control the retrieval and selection of semantic information (Noonan, Jefferies, Visser, & Lambon Ralph, 2013; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). This area is thought to exert top-down regulation over the activation of semantic knowledge, based on current contextual and task demands (Hoffman, McClelland, & Lambon Ralph, 2018b; Jefferies, 2013). Different functions have been proposed for discrete regions within IFG (Badre, Poldrack, Pare-Blagoev, Insler, & Wagner, 2005; Badre & Wagner, 2007). Pars orbitalis (BA47) is thought to support controlled retrieval of weaker or less salient semantic associations from memory, while pars triangularis (BA45) is involved in selecting the most currently relevant aspects of retrieved knowledge and inhibiting activation of irrelevant knowledge ("semantic selection"). A recent behavioural study indicates that semantic selection ability is a strong predictor of coherence (Hoffman et al., 2018a). People who are skilled at selecting the most relevant aspects of their knowledge tend to be highly coherent when speaking because they are able to avoid producing tangential, off-topic information. For this reason, we predicted that activation of left BA45 in particular would be associated with production of highly coherent speech.

In addition, the study tested whether left and right IFG showed similar relationships with coherence. In young adults, IFG activation during verbal semantic processing is strongly left-lateralised, while bilateral IFG activation is observed more frequently in older people (Hoffman & Morcom, 2018). There is ongoing debate as to whether bilateral recruitment of IFG in later life is an adaptive strategy that boosts cognitive performance or whether it represents unhelpful dedifferentiation in neural activity (Cabeza, 2002; Logan, Sanders, Snyder, Morris, & Buckner, 2002; Morcom & Johnson, 2015; Reuter-Lorenz & Cappell, 2008). In light of this debate, the present study tested whether IFG involvement in coherent speech was bilateral in older people and whether more coherent older speakers displayed greater recruitment of right IFG.

Method

Participants: Fifteen healthy older adults were recruited from the psychology department's volunteer panel at the University of Edinburgh. All had previously taken part in a larger behavioural study investigating cognitive predictors of coherent speech (Hoffman, 2018; Hoffman et al., 2018a). Their mean age was 77.5 (s.d. = 8.4; range = 67-92) and they had completed an average of 15.0 years in education (s.d. = 2.9; range = 10-20). All reported to be in good health with no history of neurological or psychiatric illness. The Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975) was used to screen for possible cognitive impairments; all participants scored at least 28/30. Their scores on a range of neuropsychological tests are provided in Supplementary Table 1. Informed consent was obtained from all participants and the study was approved by the University of Edinburgh Psychology Research Ethics Committee (1-1516/1).

Design and Procedure: The study contrasted extended speech production with a baseline consisting of automatic speech (reciting a nursery rhyme). The structure of trials is shown in Figure 1. During the extended speech task, participants were asked to speak for periods of 50s at a time in response to a series of prompts. There were 20 prompts, which were designed to probe particular a range of semantic knowledge (e.g., *What sort of things do you have to do to look after a dog?*; for a full list, see Supplementary Materials). Fourteen of the prompts were taken from a previous behavioural study (Hoffman et al., 2018a). Each trial began with a speech planning phase lasting 8s, in which the written prompt was presented on screen. Participants were instructed to prepare to speak about the subject given but not to begin speaking until the prompt was replaced by a green circle. They were asked to speak about the prompted subject for 50s, at which point the green circle would be replaced by a red cross. They were then instructed to stop speaking and wait for the next trial, which would begin after 7s. The structure of the baseline task was similar, except that the prompt was always "Humpty Dumpty". They were asked to begin reciting this nursery rhyme upon seeing the green circle. To reduce participant fatigue, baseline trials lasted for only 15s. If participants reached the end of the nursery rhyme before the 15s had elapsed, they were asked to start again from the beginning. The baseline condition therefore involved production of

grammatically well-formed continuous speech, but without the requirement to generate novel, meaningful utterances.

Functional imaging took place over two runs of approximately 16 minutes each, with each run including ten trials of each type in a pseudo-random order. The order in which prompts were presented was counterbalanced over participants. Prior to scanning, participants were presented with the words to Humpty Dumpty and given an opportunity to practice both tasks. The neuroimaging data for the study can be found at: https://osf.io/9ca5g/.

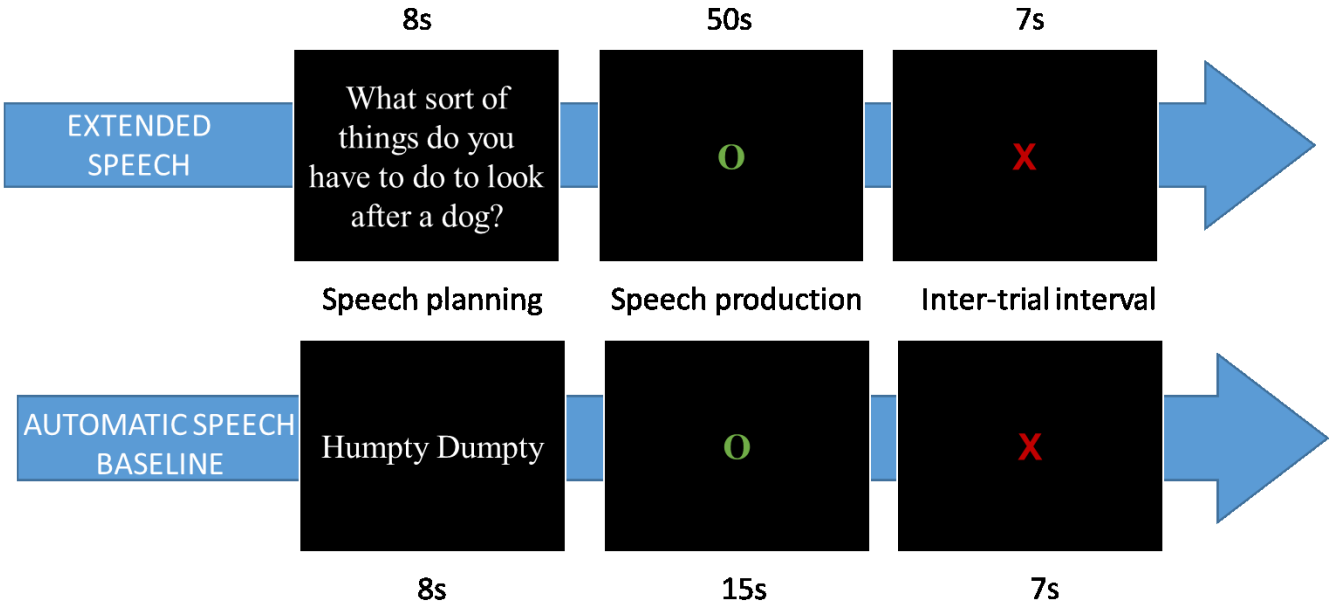


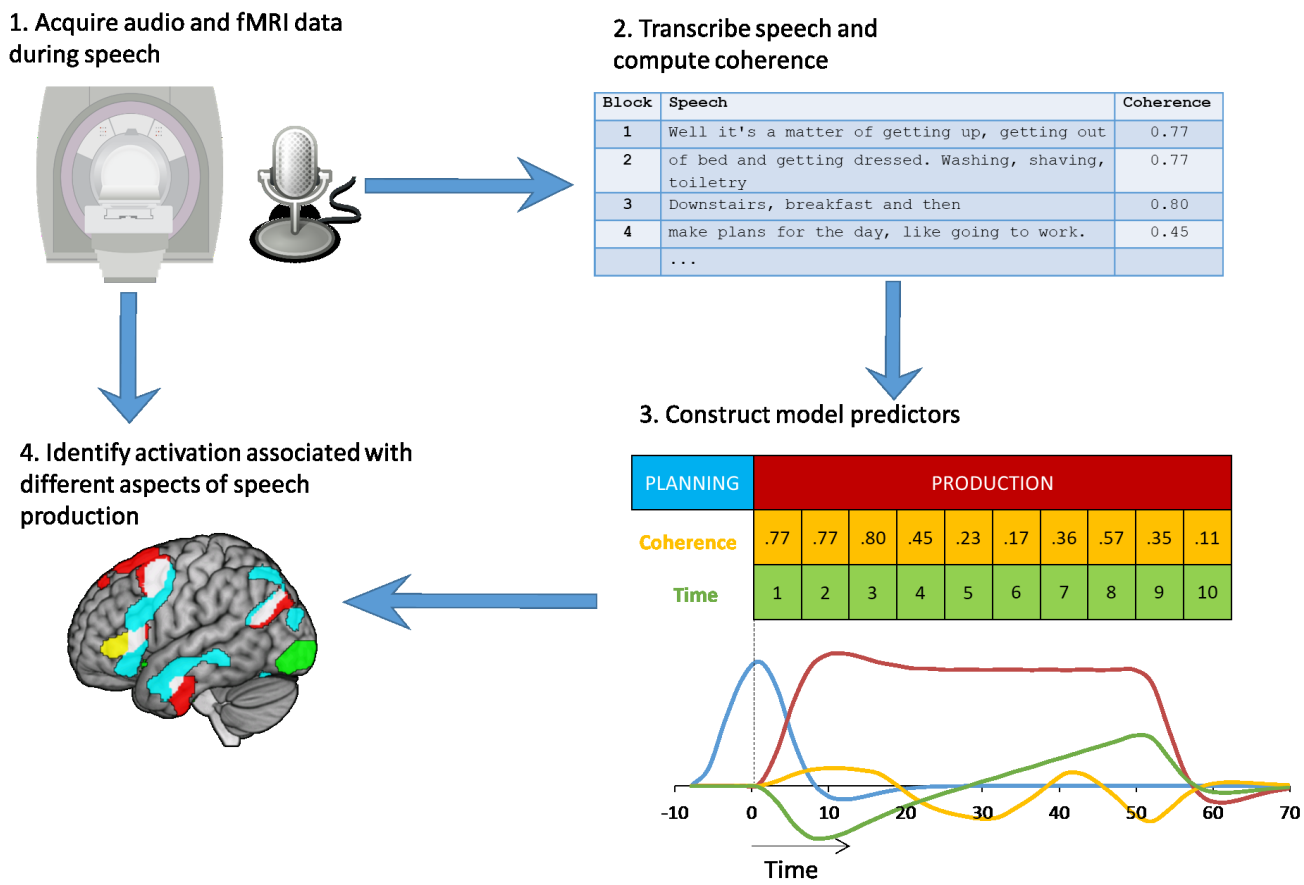
Figure 1: Illustration of single trials

Processing of speech samples: Data analysis steps are illustrated in Figure 2. Speech was recorded using an MR-compatible microphone and audio files processed with noise cancellation software to reduce scanner noise (Cusack, Cumming, Bor, Norris, & Lyzenga, 2005). Responses to each prompt were transcribed and measures of global coherence were obtained by subjecting speech samples to computational analyses based on latent semantic analysis (LSA) (Landauer & Dumais, 1997). LSA provides the user with vector-based representations of the meanings of words, which can be combined linearly to represent the meanings of whole passages of speech or text (Foltz, Kintsch, & Landauer, 1998). We used these representations to characterise the content of each speech sample, defining coherence

as the similarity of a sample's semantic content to the prototypical semantics typically produced in response to the same prompt. Coherence calculated in this way has high internal reliability and test-retest reliability and is highly correlated with human ratings of coherence (Hoffman et al., 2018a).

The methods used to compute coherence have been described in detail elsewhere (Hoffman et al., 2018a) and are summarised here. Analyses were implemented in R; the code is publicly available and can easily be applied to new samples (https://osf.io/8atfn/). For a response to a given prompt, coherence was calculated as follows. First, an LSA representation was computed for all other participants' responses to the same prompt, excluding the response currently under analysis. These were averaged to give a composite vector that represented the typical semantic content that people produced when responding to the prompt. For example, the composite vector for the prompt "Which is your favourite season?" would be close to the vectors for *summer, autumn, weather, sunshine* and so on, as these concepts were frequently used in responses to this prompt. Next, the target response was analysed using a moving window approach. For each word in the response, a 20-word window was created consisting of the current word and the 19 preceding words. LSA vectors were computed for each window and compared with the composite vector using a cosine similarity measure. This gave a coherence measure for each word in the response, which represented how much the speech produced at that time resembled the typical semantic content of responses to that prompt. A high coherence value indicated that the speech was closely related to the topic being probed. A low coherence value indicated that the participant was speaking about topics that were semantically unrelated to the topic being probed. Examples of high and low coherence responses are shown in Supplementary Table 2.

To obtain a dynamic measure of coherence at each stage of speech production, each 50s speech production periods was divided into blocks of 5s. Coherence in each block was calculated by averaging the coherence values associated with each of the words produced in the block. Reliable coherence estimates could not be obtained for the first block in each response, due to a lack of prior speech to analyse. In these cases, we used the coherence value for the second block instead.



<u>Figure 2: Stages in data analysis</u> 1. Participants were recording while speaking in the scanner. 2. Speech was divided into blocks of 5s, transcribed and analysed using computational linguistic techniques to provide an estimate of coherence for each 5s block. 3. A general linear model was constructed with regressors tracking (a) speech planning periods (b) speech production periods (c) activation correlated with coherence and (d) activation changing over time. 4. The model was fitted to voxel time-series in whole-brain and ROI analyses.

<u>Image acquisition and processing:</u> Images were acquired on a 1.5T GE Signa Horizon scanner with an eight-channel head coil. The EPI sequence included 30 slices covering the whole brain with TE = 40ms, TR = 2200ms, flip angle = 90°, 64 x 64 matrix, reconstructed in-plane resolution = 3.75 x 3.75mm, slice thickness = 4.0mm. 900 volumes were acquired over two runs. T1 and T2 structural images were acquired for co-registration. An MP RAGE sequence was used to acquire a whole-brain image with 256 x 256 matrix, in-plane resolution = 0.9375mm, slice thickness = 1.3mm. A T2 image was acquired with 256 x 256 matrix, in-plane resolution = 0.9375mm, slice thickness = 2mm.

Images were analysed in SPM12. Functional image acquisition during overt speech production can be challenging due to additional head movement. A number of steps were taken to minimise the effects of movement on image quality. First, following spatial realignment, images were processed using the FIACH toolbox which is designed to ameliorate effects of head motion in studies that employ overt speech production (Tierney et al., 2016). The toolbox has two functions. First, it identifies and removes signal spikes in the time-series of individual voxels. Second, it uses principal components analysis to identify noise components across the whole brain in each participant's data. These noise components can be included as covariates of no interest in later analyses (for a similar approach, see Silbert et al., 2014). Here, the first six noise components were included as covariates in all first-level analyses, in addition to the six movement parameters obtained during spatial realignment. Finally, a measure of each participant's propensity to move in the scanner was obtained as follows. The mean scan-to-scan change was calculated for each of the six movement parameters and these were averaged to give a mean scan-to-scan displacement value for each participant (for a similar approach, see Van Dijk, Sabuncu, & Buckner, 2012). This value was included as a nuisance covariate in all second-level analyses.

Functional images were then normalised to MNI space, resampled to 2mm isotropic voxel size and smoothed with a 12mm FWHM Gaussian kernel. Following pre-processing, data were treated with a high-pass filter with a cut-off of 180s and analysed using a general linear model. Four event types were modelled: extended speech planning, extended speech production, automatic speech planning and automatic speech production. Speech planning periods were modelled as 8s blocks. The modelling of speech production periods is illustrated in Figure 2. Each speech production period was modelled as a series of concatenated 5s blocks. This allowed specification of two parametric modulators that coded the characteristics of the speech produced in each 5s block. The first of these was the coherence of the speech produced, calculated as described earlier, and the second was the time point within the 50s speech production period (see Figure 2). Blocks were modelled as boxcars convolved with the canonical hemodynamic response function.

In addition, two supplementary first-level models were estimated to assess the temporal characteristics of coherence-related activation. These comprised an *early* model in which coherence values were shifted back in time by 5s and a *late* model in which they were shifted forwards (e.g., the coherence value for words produced in block 4 of would be

assigned to block 3 in the early model and to block 5 in the late model). These models tested for activation that either preceded or followed the production of high/low coherence speech.

Analyses: Analyses were performed across the whole brain and in anatomically-defined prefrontal ROIs. A voxel-height threshold of $p<0.005$ was adopted for whole-brain analyses, with correction for multiple comparisons performed at the cluster level. The minimum cluster size was determined using a Monte Carlo analysis (Slotnick, Moo, Segal, & Hart, 2003). This modelled the entire smoothed image volume, assumed an individual voxel type-1 error rate of 0.005 and ran 5000 simulations to determine the minimum cluster size associated with a corrected $p<0.05$.

The first set of analyses considered activation for speech planning (extended speech planning minus automatic speech planning) and production (extended speech planning minus automatic speech planning). Subsequent analyses tested for effects of coherence and time on neural activity during speech production. Finally, an exploratory individual differences analysis was performed, to determine whether activity differed between highly coherent and less coherent speakers. For this analysis, the mean level of coherence for each participant's speech was calculated over all of their responses and these values were added as a covariate to the second level analysis.

Prefrontal regions of interest: Anatomical ROIs were specified a priori in the left and right IFG (see Figure 4A), using the Harvard-Oxford probabilistic atlas (Makris et al., 2006). The BA45 region was defined as voxels with >30% probability of falling in the pars triangularis region and BA47 as those with >30% probability of falling in pars orbitalis. The pars orbitalis region includes areas of medial prefrontal cortex that have not been linked with language processing. To exclude these, voxels more medial than x=±30 were removed from the mask.

The Marsbar toolbox (Brett, Anton, Valabregue, & Poline, 2002) was used to extract parameter and contrast estimates for each participant. In addition, to aid visualisation of the parametric effects of coherence and time, the rfxplot package (Gläscher, 2009) was used to estimate effects at five levels of each parameter.

Results

Characteristics of speech production: Participants produced an average of 104 words in response to each prompt (range: 50-153). The mean coherence value for responses was 47.6 (range: 20.9 to 81.4), which is similar to levels observed when speech is produced out of the scanner (Hoffman et al., 2018a).[1] There was considerable variation in coherence across individuals: the least coherent individual had a mean coherence score of 41.5 and the most coherent 55.0. Importantly, the coherence measure showed high test-retest reliability across individuals. Fourteen participants had previously completed a similar speech elicitation task out of the scanner as part of another study (Hoffman et al., 2018a). Their in-scanner and out-of-scanner coherence scores were highly correlated (r = 0.88), indicating that the observed variation represents stable individual differences in the ability to produce coherent speech.

Changes in coherence over the speech production period were also assessed. The mean coherence at each 5s block is plotted in Supplementary Figure 1. There was a strong tendency for speech produced later in a response to be less coherent than early speech (r = -0.93). This likely occurred because deviations from the prescribed topic tended to occur later in the production period, after the most salient information had been provided. To ensure that relationships between coherence and neural activity were not confounded by this factor, time within the speech production period was included as an additional covariate during neuroimaging analyses (as shown in Figure 2).

Activation for speech planning and production: Regions of significant activation for speech planning and production, relative to automatic speech baselines, are shown in Figure 3 (for peak co-ordinates, see Supplementary Table 3). Speech planning activated regions typically associated with semantic processing, including left prefrontal cortex, bilateral anterior temporal cortex (specifically the superior temporal sulcus and middle temporal gyrus) and the angular gyrus. Pronounced medial occipital activation was also observed, presumably related to reading of the written topic prompts. Speech production activated a somewhat similar network of regions, though with a greater left-hemisphere bias. Inferior prefrontal

[1] As described in Method, the coherence measure quantifies the strength of the semantic relationship between the speech produced and the topic being probed. 0 indicates speech that has no semantic relationship with the topic and 100 indicates speech that is identical to it.

activity was centred on BA45 and activation was also observed in dorsolateral prefrontal cortex and the supplementary motor area, the ventrolateral surface of the anterior temporal lobe (ATL) and the left angular gyrus. These regions are similar to those reported in previous studies of extended speech production in young adults (AbdulSabur et al., 2014; Awad et al., 2007; Silbert et al., 2014).

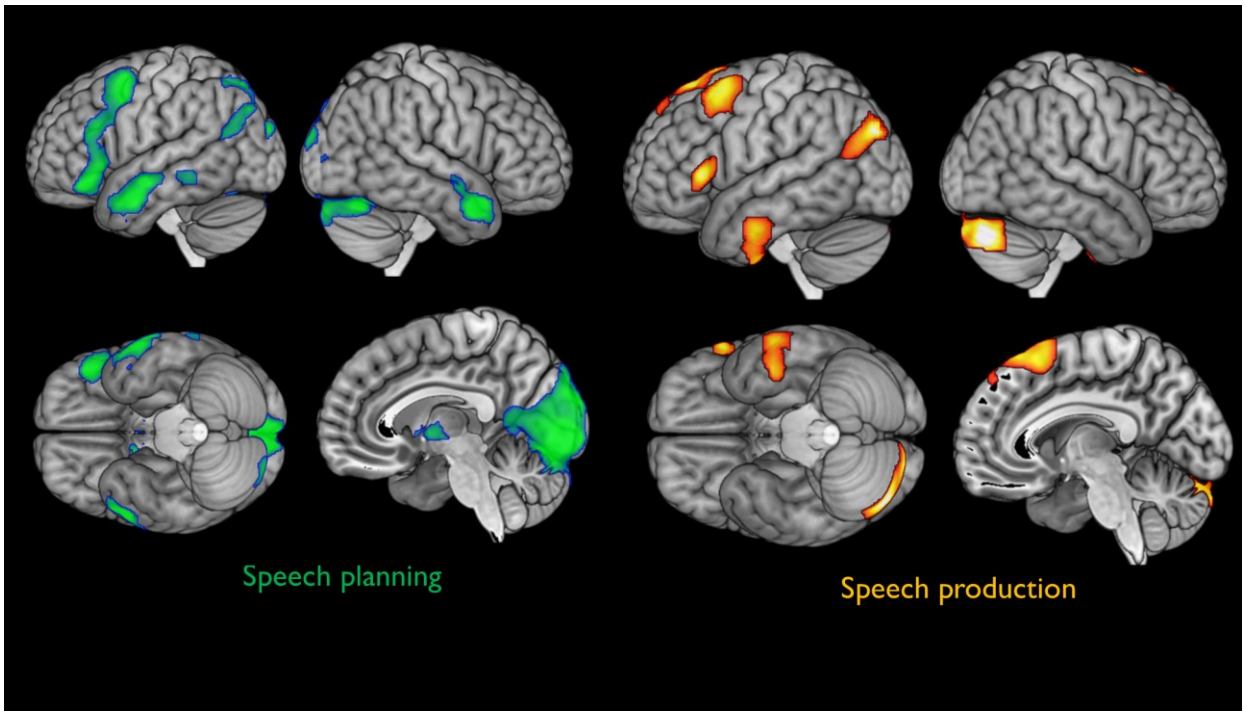


Figure 3: Areas of significant activation for planning and production of extended speech, relative to automatic speech baseline. Images thresholded at p<0.05 cluster-corrected.

Contrast estimates for speech planning and production in the prefrontal ROIs are shown in Figure 4B. There appeared to be a strong left-hemisphere bias in the activation of both BA45 and BA47. To test formally for this effect, a 2 x 2 ANCOVA was performed for each phase of the task, including region (BA45 vs. BA47) and hemisphere (left vs. right) as factors. There were significant effects of hemisphere for planning ($F(1,13) = 40.4$, $p < 0.001$) and production ($F(1,13) = 10.5$, $p = 0.007$), in both cases indicating greater activity in the left hemisphere. There were no main effects of region but there were region x hemisphere interactions for both phases of the task (planning: $F(1,13) = 5.57$, $p = 0.035$; production: $F(1,13) = 5.07$, $p = 0.042$), indicating greater activation of BA45 compared with BA47 in the left hemisphere only.

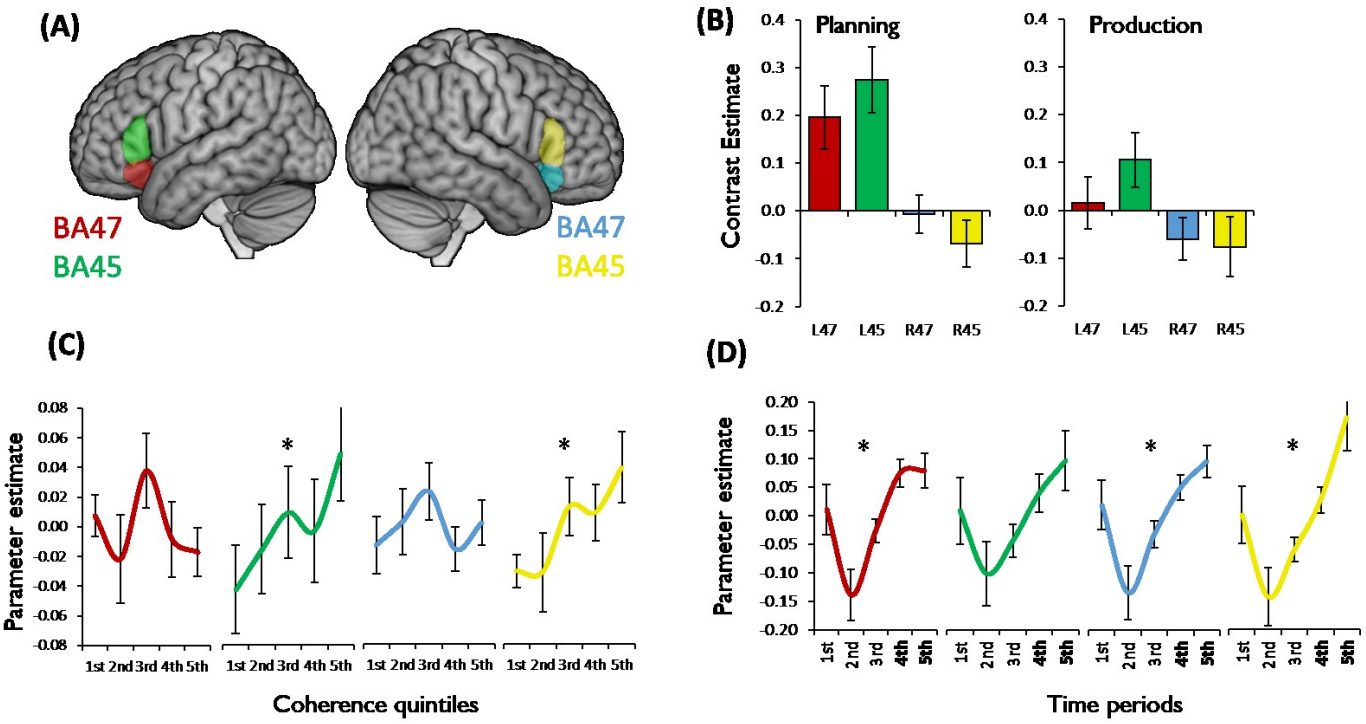


Figure 4: Results in prefrontal regions of interest. (A) Anatomical ROIs were specified for left BA47 (red), left BA45 (green), right BA47 (blue) and right BA45 (yellow). (B) Estimates for contrasts of extended speech over automatic speech. (C) Effects of coherence on activation in each ROI. Speech blocks were divided into five sets based on their coherence (where 1st quintile = least coherent responses and 5th quintile = most coherent). Estimates for each set are relative to the overall speech-related activation for the region. A significant linear effect of coherence was observed in left and right BA45 (indicates $p < 0.05$). (D). Change in activation over time within speech production periods. Speech blocks were divided into five sets based on their temporal position in the speech elicitation period (where 1st = first 10s and 5th = final 10s). All regions except left BA45 showed a linear effect of time on activation (* indicates $p < 0.05$). Bars indicate SEM.*

Activation varying with coherence: The second analysis identified regions whose activity was correlated with the coherence of the speech produced. In the whole-brain analysis, two prefrontal activation clusters showed increased activation when participants produced more coherent speech (see Figure 5; peaks reported in Supplementary Table 4). The right-hemisphere cluster included BA45 and a region of rostrolateral prefrontal cortex (RLPFC; BA10). Left-hemisphere activation was centred on BA45. No regions exhibited negative effects of coherence.

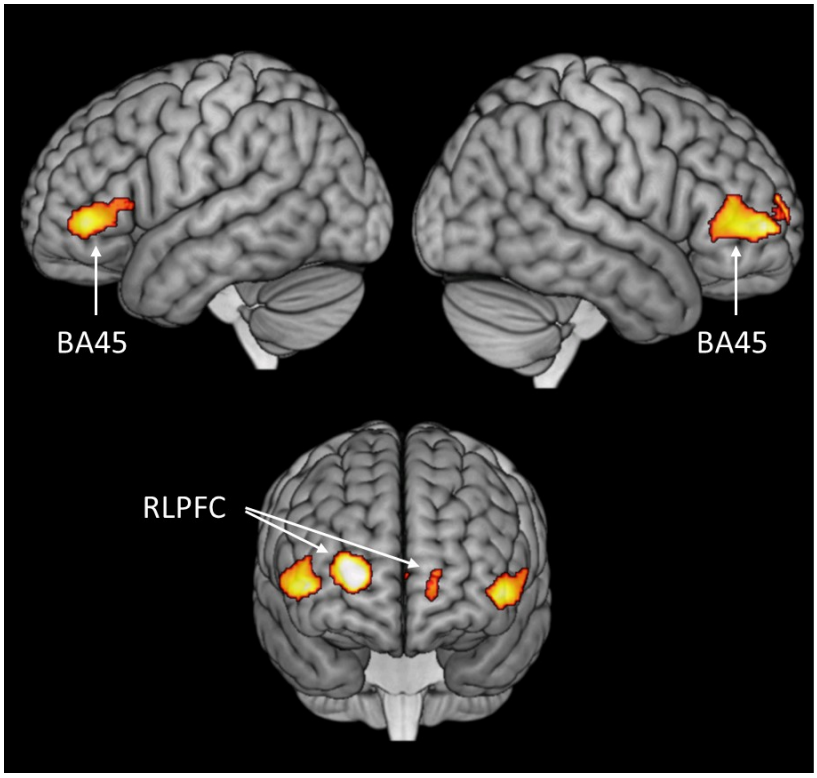


Figure 5: Regions showing greater activation when participants produced more coherent speech. Images thresholded at p<0.05 cluster-corrected.

Effects of coherence in the BA45 and BA47 ROIs are illustrated in Figure 4C. Left and right BA45 showed a clear linear increase in activity as utterances became more coherent, while no such effect was present in BA47. The effects of coherence in the prefrontal ROIs were analysed with a 2 x 2 (hemisphere x region) ANCOVA. This revealed a main effect of region ($F(1,13) = 7.34$, $p = 0.018$), confirming that the effect of coherence was significantly larger in BA45 compared with BA47. There were no between-hemisphere differences in the effect of coherence ($F(1,13) = 0.62$, $p = 0.44$).

To explore the timing of coherence-related activity, two further models were estimated for the neuroimaging data. These were an *early* model in which coherence values associated with speech were shifted backward in time by 5s and a *late* model in which they were shifted forward by 5s. Significant effects of coherence in these models are shown in Supplementary Figure 2. The late model principally revealed activation in left BA45, indicating that activation associated with highly coherent speech in this area persisted after the production of the speech itself. However, the early model revealed larger coherence-related activations in RLPFC bilaterally and in right BA45. This indicates that these regions increased

activity immediately *prior* to the production of highly coherent speech, suggesting that they may play an earlier role in shaping the content of upcoming utterances.

Activation varying over time within speech production periods: The third analysis identified regions whose activity increased over time during a speech production act (e.g., regions whose activity was greater at the end of a 50s production period, relative to its start). Results of the whole-brain analysis are shown in Supplementary Figure 3 (activation peaks in Supplementary Table 4). A number of regions showed increasing activation as each speech period progressed, including dorsomedial prefrontal cortex, right IFG, bilateral ATL and the right angular gyrus. Reduced activity over time was observed in primary auditory regions, perhaps due to a slowing of speech production later in the period.

Change over time in the IFG ROIs are plotted in Figure 4D. All regions showed broadly similar patterns throughout the speech window, with high levels of activation during the first 10s of production which then dropped before increasing monotonically over the remainder of the period. One possible interpretation of this pattern is that it reflects high early executive demands when participants initiate their speech act, after which demands subside before gradually building over time as the continued production of speech becomes more challenging. A 2 x 2 (hemisphere x region) ANCOVA performed on the effects of time in IFG revealed no main effects, although there was an interaction between hemisphere and region ($F(1,13) = 7.37$, $p = 0.018$). This appeared to arise because there was a more pronounced effect on time in the right hemisphere than in the left, but only within BA45.

Differences in activation between more and less coherent individuals: Finally, an exploratory analysis was conducted to identify differences in activation between more and less coherent speakers. The mean coherence level for each participant was included as a covariate in second-level analyses of speech planning and speech production, results of which are shown in Figure 6 (activation peaks in Supplementary Table 5). Thus, while the previous analysis identified activity that varied with coherence *within* individuals, this analysis aimed to identify activity that differed *across* speakers as a function of their coherence. Individual differences were observed during both the planning and production phases of speech. More coherent participants tended to show greater activation in the right IFG during planning of speech (including portions of BA47 and BA45). People who spoke less coherently exhibited greater

activity in left ATL during speech planning. This pattern was partially reversed during the speech production period, where a region of left ATL was associated with greater activation in *more* coherent speakers. No areas showed a negative association with coherence during production.

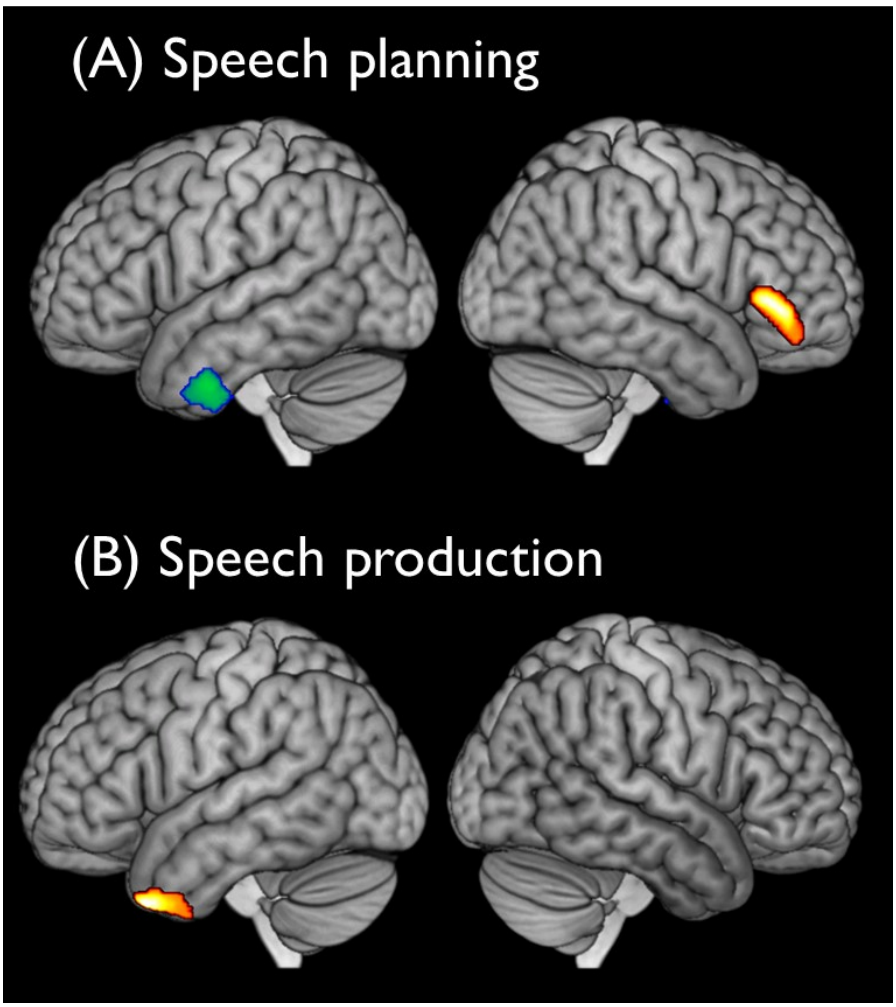


Figure 6: Regions where activation during speech planning and production was correlated with individual differences in coherence. Red-orange shows regions with greater activation in highly coherent participants and blue-green shows regions with greater activation in less coherent participants. Images thresholded at p<0.05 cluster-corrected.

Discussion

The ability to maintain coherence while speaking is critical for effective communication but declines in later life. Here, fMRI was used in healthy older people to investigate the neural correlates of coherent speech production. When speaking for 50s

periods about specified topics, older people activated a left-lateralised network of frontal, anterior temporal and inferior parietal regions, similar to that observed in previous studies with young people (AbdulSabur et al., 2014; Awad et al., 2007; Silbert et al., 2014). In a major advance on previous studies, formal analyses were performed on speech produced in the scanner to determine its coherence (i.e., the degree to which utterances were meaningfully related to the subject probed). This allowed discovery of areas in which increased neural activity was correlated with the production of highly coherent utterances. When participants spoke with high coherence, they showed greater activation in two distinct areas of prefrontal cortex: bilateral IFG (BA45) and RLPFC. Both regions are associated with goal-directed executive control of behaviour. BA45 has been implicated specifically in the selection of task-relevant information from semantic knowledge (Badre & Wagner, 2007; Thompson-Schill et al., 1997) while RLPFC contributes to the planning of complex behavioural sequences (Fuster, 2004; Koechlin & Summerfield, 2007). These findings have major implications for how understanding how people maintain coherence and focus during speech production and, more broadly, for the role of left and right IFG in controlled cognition in later life.

The most notable result in the study is that, as predicted, activity in left and right BA45 was positively correlated with coherence during speech production. In other words, this region was more engaged in processing when participants produced utterances that were closely related to the current topic, and less engaged when they deviated off-topic. These results suggest that BA45 plays a central role in regulating the selection of topic-relevant information during speech production (i.e., semantic selection). This conclusion is consistent with evidence for left BA45 involvement in semantic selection across a range of experimental tasks (Badre & Wagner, 2007; Gold et al., 2006; Noonan et al., 2013; Whitney, Kirk, O'Sullivan, Lambon Ralph, & Jefferies, 2011). For example, this region shows increased activation in comprehension when participants are asked to attend to specific semantic features and avoid competition from irrelevant semantic associations (Badre et al., 2005; Thompson-Schill et al., 1997). In semantic single-word production tasks, BA45 activity is greater for stimuli that prime multiple, competing possible responses (Kan & Thompson-Schill, 2004; Thompson-Schill et al., 1997). The present study demonstrates for the first time that BA45 is also involved in selecting from activated semantic knowledge during extended periods of natural speech.

Why is selection so important for coherent communication? A conversational cue, such as “what’s your favourite season?”, may automatically cause a wide range of general

semantic knowledge to become activated. 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 information which is directly relevant at the current time, while suppressing aspects of knowledge that have been activated but are less pertinent. Behavioural work supports the idea the effective semantic selection is critical for maintaining coherence in speech. In a recent study, we found that individuals who highly skilled at semantic selection (assessed using a forced-choice "feature selection" task; Thompson-Schill et al., 1997) produced more coherent speech (Hoffman et al., 2018a). This effect was independent of the breadth of participants' semantic knowledge (assessed with vocabulary tests) and of non-semantic, domain-general executive function. This study suggests an important role for semantic processes in regulating speech. It is important to note, however, that the knowledge that drives speech production is not solely semantic in nature – memories of specific events and experiences also shape and influence the content of our utterances. Nor is the involvement of BA45 in competition resolution limited to the semantic domain: competition in episodic and working memory tasks also increase activation in this region (Badre & Wagner, 2005; Dobbins & Wagner, 2005; Jonides & Nee, 2006). The neuroimaging literature therefore indicates a general role for BA45 in selecting the most relevant knowledge retrieved from memory, irrespective of the specific type of information.

In contrast, the more anterior portion of IFG, BA47, was not related to speech coherence (and differed significantly from BA45 in this respect). Although left BA47 is also associated with the control and regulation of semantic knowledge, it is thought to serve a different function to BA45 (Badre & Wagner, 2007). It is typically described as supporting "controlled retrieval" of semantic knowledge when stimulus-driven activation of knowledge is insufficient to complete the task at hand. Under these circumstances, BA47 is thought to support a goal-directed controlled search through the semantic store for less salient information (Badre & Wagner, 2002; Hoffman et al., 2018b). The present results suggest that increased engagement of BA47 does not result in more coherent speech. This may be because the less salient knowledge accessed through controlled retrieval is less centrally related to the topic under discussion, so carries a higher risk of deviation off-topic. Although BA47 showed no relationship with coherence, it did display increasing activity over time during speech production periods (along with BA45). This suggests that controlled retrieval is more engaged during the later stages of a response, once the participant has exhausted the most

salient knowledge about the topic at hand and is forced to search for other relevant information.

In our older participants, the association of BA45 activity with high-coherence speech output was present in both hemispheres. In contrast, the majority of studies in young people indicate a dominant role for the *left* IFG in controlled semantic processing. In younger people, IFG activation during verbal semantic processing is strongly left-lateralised, especially for speech production tasks (Berlingeri, Danelli, Bottini, Sberna, & Paulesu, 2013; Binder, Desai, Graves, & Conant, 2009; Persson et al., 2004; Rice, Lambon Ralph, & Hoffman, 2015). When right IFG activation is observed during semantic processing, it is less extensive and typically found under conditions of high demand (Krieger-Redwood, Teige, Davey, Hymers, & Jefferies, 2015; Noonan et al., 2013). Older people show a more bilateral pattern of frontal recruitment during semantic processing (Hoffman & Morcom, 2018; Logan et al., 2002; Persson et al., 2004), which is in line with reduced lateralisation in frontal activity across a variety of domains (Cabeza, 2002). The consequences of this age-related reduction in laterality continue to be debated. Some researchers hold that increased recruitment of right IFG serves to compensate for decreased efficiency of processing in left prefrontal regions (Cabeza, Anderson, Locantore, & McIntosh, 2002; Grady, 2012). Others hold that additional right-hemisphere activation is a symptom of reduced specificity in brain activity in old age and does not benefit task performance (Li, Lindenberger, & Sikström, 2001; Logan et al., 2002).

Although the present study was not designed to adjudicate on these issues directly, our results are more consistent with the compensatory view of right prefrontal activation. Overall, IFG activation during speech planning and production was left-lateralised in our participants. This suggests that, even in older people, the left hemisphere is dominant in guiding production of extended speech. However, the profile of right IFG regions with respect to coherence and time was very similar to that of left IFG: right BA45 displayed increased activity when participant produced highly coherent utterances and right BA45 and BA47 both showed increasing activation over each speech production period. These results suggest that right IFG was engaged in the regulation of speech content, albeit at lower overall levels of activity than left IFG. In addition, the individual differences analysis revealed that the more coherent individuals in the study exhibited greater right IFG activation during the planning period before beginning to speak. This result should be interpreted with caution, given the limited sample size. It does, however, suggest that greater right IFG activity preceding speech

is associated with a more positive behavioural outcome. This association of right prefrontal engagement with improved behavioural performance in older people has been reported previously in other domains but not in speech production (Grady, McIntosh, & Craik, 2005; Lee, Grady, Habak, Wilson, & Moscovitch, 2011; Rypma & D'Esposito, 2000).

Finally, the only region aside from IFG to show a relationship with coherence was RLPFC. This frontopolar region was more active during production of highly coherent speech, which is in line with the idea that RLPFC is involved in the planning and sequencing of complex, goal-directed behaviours (Desrochers, Chatham, & Badre, 2015; Fuster, 2004; Gilbert, 2011; Koechlin & Summerfield, 2007; Momennejad & Haynes, 2012). In particular, Badre and Nee (2017) have proposed that RLPFC controls the sequencing and execution of sub-tasks in service of an over-arching goal, especially when this process is informed by learned knowledge structures (schemas) coded by the ventromedial prefrontal cortex. This organised sequencing of information is critical for producing a well-structured and coherent narrative. Indeed, many of the prompts used to elicit speech here encouraged participants to describe well-learned behavioural schemas (e.g., “What do people do when getting ready for work in the morning?”). The most coherent responses were those that accurately narrated the schema content without deviating to other topics. Interestingly, the time-shifting analysis indicated that RLPFC showed increased activation in the 5s prior to the production of highly coherent utterances, as well as during production itself. This suggests that this area contributes to coherence prior to production, which is again consistent with hierarchical models of cognitive control which assign a high-level planning role to this area (Fuster, 2004; Koechlin & Summerfield, 2007). In contrast, the later effects in left BA45 are more consistent with a downstream filtering role, to prevent irrelevant information from intruding into the narrative stream.

In summary, this first investigation of the neural correlates of coherent speech has revealed that multiple prefrontal control systems contribute to coherence. It suggests that timing of these contributions differ, with the most anterior RLPFC region making an earlier contribution than the left IFG. These data will allow future investigations to probe in more detail the precise roles of these systems in speech regulation and their differential recruitment in young vs. older people. This will provide further insights into how complex verbal interactions are planned and regulated.

References

AbdulSabur, N. Y., Xu, Y., Liu, S., Chow, H. M., Baxter, M., Carson, J., & Braun, A. R. (2014). Neural correlates and network connectivity underlying narrative production and comprehension: A combined fMRI and PET study. *Cortex, 57*, 107-127.

Arbuckle, T. Y., & Gold, D. P. (1993). Aging, inhibition, and verbosity. *J Gerontol, 48*(5), P225-232.

Arbuckle, T. Y., Nohara-LeClair, M., & Pushkar, D. (2000). Effect of off-target verbosity on communication efficiency in a referential communication task. *Psychol Aging, 15*(1), 65.

Awad, M., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. (2007). A common system for the comprehension and production of narrative speech. *Journal of Neuroscience, 27*(43), 11455-11464. doi:10.1523/JNEUROSCI.5257-06.2007

Badre, D., & Nee, D. E. (2017). Frontal cortex and the hierarchical control of behavior. *Trends in Cognitive Sciences*.

Badre, D., Poldrack, R. A., Pare-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47*(6), 907-918.

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

Badre, D., & Wagner, A. D. (2005). Frontal lobe mechanisms that resolve proactive interference. *Cerebral Cortex, 15*(12), 2003-2012.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883-2901.

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage, 37*(1), 90-101.

Berlingeri, M., Danelli, L., Bottini, G., Sberna, M., & Paulesu, E. (2013). Reassessing the HAROLD model: is the hemispheric asymmetry reduction in older adults a special case of compensatory-related utilisation of neural circuits? *Experimental Brain Research, 224*(3), 393-410.

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

Blank, S. C., Scott, S. K., Murphy, K., Warburton, E., & Wise, R. J. (2002). Speech production: Wernicke, Broca and beyond. *Brain, 125*(8), 1829-1838.

Braun, A. R., Guillemin, A., Hosey, L., & Varga, M. (2001). The neural organization of discourse: An H215O-PET study of narrative production in English and American sign language. *Brain, 124*(10), 2028-2044.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). *Region of interest analysis using an SPM toolbox*. Paper presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: the HAROLD model. *Psychol Aging, 17*(1), 85.

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.

Cusack, R., Cumming, N., Bor, D., Norris, D., & Lyzenga, J. (2005). Automated post-hoc noise cancellation tool for audio recordings acquired in an MRI scanner. *Human Brain Mapping, 24*(4), 299-304.

Desrochers, T. M., Chatham, C. H., & Badre, D. (2015). The necessity of rostrolateral prefrontal cortex for higher-level sequential behavior. *Neuron, 87*(6), 1357-1368.

Dobbins, I. G., & Wagner, A. D. (2005). Domain-general and domain-sensitive prefrontal mechanisms for recollecting events and detecting novelty. *Cerebral Cortex, 15*(11), 1768-1778.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198.

Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with latent semantic analysis. *Discourse Processes, 25*(2-3), 285-307. doi:10.1080/01638539809545029

Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in Cognitive Sciences, 8*(4), 143-145.

Gilbert, S. J. (2011). Decoding the content of delayed intentions. *Journal of Neuroscience, 31*(8), 2888-2894.

Gläscher, J. (2009). Visualization of group inference data in functional neuroimaging. *Neuroinformatics, 7*(1), 73-82.

Glosser, G., & Deser, T. (1992). A comparison of changes in macrolinguistic and microlinguistic aspects of discourse production in normal aging. *J Gerontol, 47*(4), P266-272.

Gold, B. T., Balota, D. A., Jones, S. J., Powell, D. K., Smith, C. D., & Andersen, A. H. (2006). Dissociation of automatic and strategic lexical-semantics: functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *Journal of Neuroscience, 26*(24), 6523-6532. doi:10.1523/JNEUROSCI.0808-06.2006

Gold, D., Andres, D., Arbuckle, T., & Schwartzman, A. (1988). Measurement and correlates of verbosity in elderly people. *Journal of gerontology, 43*(2), P27-P33.

Gracco, V. L., Tremblay, P., & Pike, B. (2005). Imaging speech production using fMRI. *Neuroimage, 26*(1), 294-301.

Grady, C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience, 13*(7), 491-505.

Grady, C. L., McIntosh, A. R., & Craik, F. I. (2005). Task-related activity in prefrontal cortex and its relation to recognition memory performance in young and old adults. *Neuropsychologia, 43*(10), 1466-1481.

Hoffman, P. (2018). An individual differences approach to semantic cognition: Divergent effects of age on representation, retrieval and selection. *Scientific Reports, 8*(1), 8145.

Hoffman, P., Loginova, E., & Russell, A. (2018a). Poor coherence in older people's speech is explained by impaired semantic and executive processes. *eLife, 7*, e38907. doi:10.7554/eLife.38907

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

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

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

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience, 139*(1), 181-193.

Juncos-Rabadan, O., Pereiro, A. X., & Rodriguez, M. S. (2005). Narrative speech in aging: quantity, information content, and cohesion. *Brain Lang, 95*(3), 423-434. doi:10.1016/j.bandl.2005.04.001

Kan, I. P., & Thompson-Schill, S. L. (2004). Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cognitive, Affective and Behavioral Neuroscience, 4*, 43-57.

Kemper, S., Schmalzried, R., Hoffman, L., & Herman, R. (2010). Aging and the vulnerability of speech to dual task demands. *Psychol Aging, 25*(4), 949.

Kintsch, W., & Vandijk, T. A. (1978). Toward a Model of Text Comprehension and Production. *Psychological Review, 85*(5), 363-394. doi:Doi 10.1037//0033-295x.85.5.363

Kintz, S., Fergadiotis, G., & Wright, H. H. (2016). Aging effects on discourse production. In H. H. Wright (Ed.), *Cognition, Language and Aging* (pp. 81-106). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends Cogn Sci, 11*(6), 229-235. doi:10.1016/j.tics.2007.04.005

Krieger-Redwood, K., Teige, C., Davey, J., Hymers, M., & Jefferies, E. (2015). Conceptual control across modalities: graded specialisation for pictures and words in inferior frontal and posterior temporal cortex. *Neuropsychologia, 76*, 92-107.

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

Lee, Y., Grady, C. L., Habak, C., Wilson, H. R., & Moscovitch, M. (2011). Face processing changes in normal aging revealed by fMRI adaptation. *Journal of Cognitive Neuroscience, 23*(11), 3433-3447.

Li, S.-C., Lindenberger, U., & Sikström, S. (2001). Aging cognition: from neuromodulation to representation. *Trends in Cognitive Sciences, 5*(11), 479-486.

Logan, J. M., Sanders, A. L., Snyder, A. Z., Morris, J. C., & Buckner, R. L. (2002). Under-recruitment and nonselective recruitment: dissociable neural mechanisms associated with aging. *Neuron, 33*(5), 827-840.

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

Makris, N., Goldstein, J. M., Kennedy, D., Hodge, S. M., Caviness, V. S., Faraone, S. V., . . . Seidman, L. J. (2006). Decreased volume of left and total anterior insular lobule in schizophrenia. *Schizophr Res, 83*(2), 155-171.

Marini, A., & Andreetta, S. (2016). Age-related effects on language production: A combined psycholinguistic and neurocognitive perspective. In H. H. Wright (Ed.), *Cognition, Language and Aging* (pp. 55-79). Amsterdam/Philadelphia: John Benjamins Publishing Company.

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

Momennejad, I., & Haynes, J.-D. (2012). Human anterior prefrontal cortex encodes the 'what'and 'when'of future intentions. *Neuroimage, 61*(1), 139-148.

Morcom, A. M., & Johnson, W. (2015). Neural reorganization and compensation in aging. *Journal of Cognitive Neuroscience, 27*, 1275-1285.

Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013). Going beyond Inferior Prefrontal Involvement in Semantic Control: Evidence for the Additional Contribution of Dorsal Angular Gyrus and Posterior Middle Temporal Cortex. *Journal of Cognitive Neuroscience, 25*(11), 1824-1850.

North, A. J., Ulatowska, H. K., Macaluso-Haynes, S., & Bell, H. (1986). Discourse performance in older adults. *The International Journal of Aging and Human Development, 23*(4), 267-283.

Persson, J., Sylvester, C. Y., Nelson, J. K., Welsh, K. M., Jonides, J., & Reuter-Lorenz, P. A. (2004). Selection requirements during verb generation: differential recruitment in older and younger adults. *Neuroimage, 23*(4), 1382-1390. doi:10.1016/j.neuroimage.2004.08.004

Pushkar, D., Basevitz, P., Arbuckle, T., Nohara-LeClair, M., Lapidus, S., & Peled, M. (2000). Social behavior and off-target verbosity in elderly people. *Psychol Aging, 15*(2), 361.

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Current Directions in Psychological Science, 17*(3), 177-182.

Rice, G. E., Lambon Ralph, M. A., & Hoffman, P. (2015). The roles of left versus right anterior temporal lobes in conceptual knowledge: an ALE meta-analysis of 97 functional neuroimaging studies. *Cerebral Cortex, 25*(11), 4374-4391.

Rypma, B., & D'Esposito, M. (2000). Isolating the neural mechanisms of age-related changes in human working memory. *Nature Neuroscience, 3*(5), 509.

Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Sciences of the United States of America, 111*(43), E4687-E4696. doi:DOI 10.1073/pnas.1323812111

Slotnick, S. D., Moo, L. R., Segal, J. B., & Hart, J. (2003). Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Cognitive Brain Research, 17*, 75-82.

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proc Natl Acad Sci U S A, 107*(32), 14425-14430. doi:10.1073/pnas.1008662107

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America, 94*, 14792-14797.

Tierney, T. M., Weiss-Croft, L. J., Centeno, M., Shamshiri, E. A., Perani, S., Baldeweg, T., . . . Carmichael, D. W. (2016). FIACH: A biophysical model for automatic retrospective noise control in fMRI. *Neuroimage, 124*, 1009-1020.

Van Dijk, K. R., Sabuncu, M. R., & Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage, 59*(1), 431-438.

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2011). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cerebral Cortex, 21*(5), 1066-1075.

Wright, H. H., Koutsoftas, A. D., Capilouto, G. J., & Fergadiotis, G. (2014). Global coherence in younger and older adults: Influence of cognitive processes and discourse type. *Aging, Neuropsychology, and Cognition, 21*(2), 174-196.

Supplementary Table 1: Demographic information and mean test scores

Measure	Score
Age	77.5 (8.4)
Sex M:F	6:9
Years of education	15.0 (2.9)
MMSE /30	29.5 (0.7)
Category fluency (items per category)	22.3 (3.9)
Letter fluency (items per category)	17.2 (6.7)
Trails A errors	0.1 (0.4)
Trails B errors	1.0 (1.7)
Trails A time (s)	38.5 (13.6)
Trails B time (s)	80.2 (40.4)
Trails ratio score (B time / A time)	2.0 (0.5)
Lexical decision (% correct; chance = 50%)	94.2 (5.2)
Synonym judgement (% correct; chance = 25%)	72.6 (15.2)

Standard deviations are shown in parentheses.

Supplementary Table 2: Examples of highly coherent and less coherent responses to the prompt: What do people usually do when getting ready for work in the morning?

Participant	Block	Speech	Coherence
Participant A	1	When people get ready to go to work in the morning, you would	0.43
	2	normally have set an alarm clock to go off at a time that would … them in enough	0.43
	3	time to get ready. They would have time for	0.63
	4	a shower, perhaps a shave, teeth cleaning,	0.70
	5	hair drying, breakfast. Making the breakfast	0.69
	6	first of all, of course, hopefully with toast or scrambled eggs or	0.50
	7	something to start off the day, and a cup of tea or coffee	0.50
	8	just to settle you down, ready for the day. Many people have to travel	0.33
	9	to work so they would have to have appropriate time to do so, whether by train	0.18
	10	coach, car, bicycle or walking	0.48
Participant B	1	Well it's a matter of getting up, getting out	0.77
	2	of bed and getting dressed. Washing, shaving, toiletry.	0.77
	3	Downstairs, breakfast and then	0.80
	4	make plans for the day, like going to work.	0.45
	5	Because being aged ninety two and having	0.05
	6	retired twenty eight years ago, it's a long time since I was personally….	0.10
	7	I very much enjoy the life I have	0.15
	8	without the demands of the call or the urgency	0.11
	9	of going through to a work situation. I certainly	0.19
	10	enjoy not having to drive and finish the morning	0.53

Supplementary Table 3: Activation peaks for speech planning and production, relative to baseline

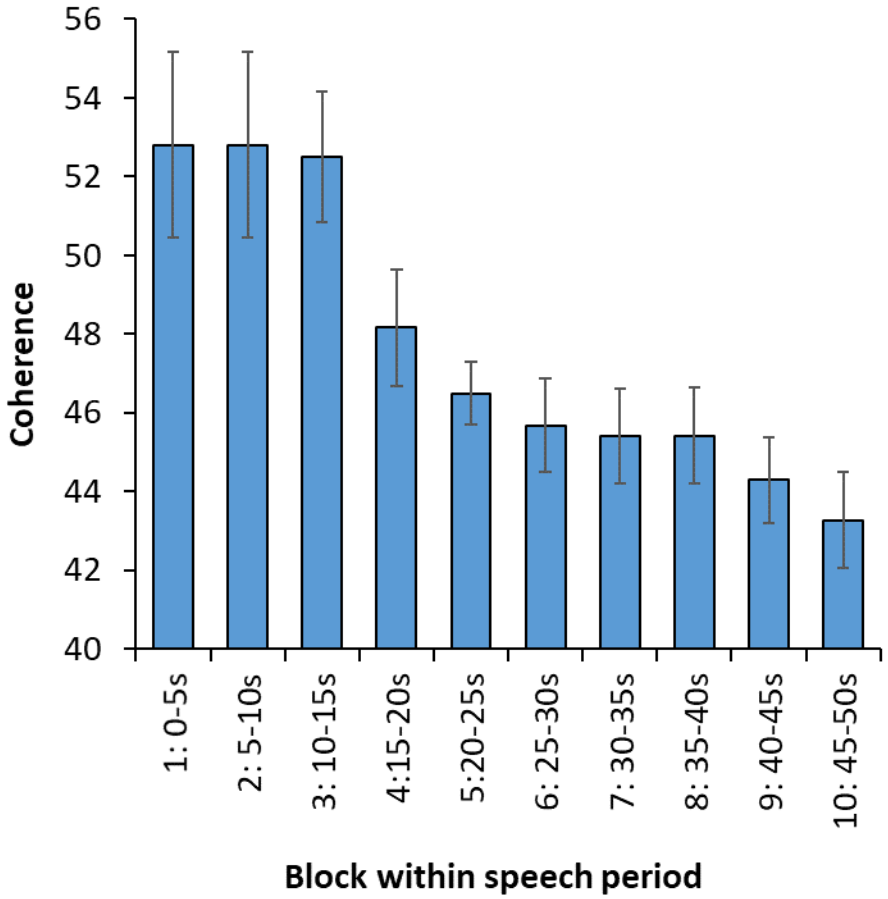
Contrast	Location	Extent (mm^3)	*t*	MNI co-ords x	y	z
Speech planning	Occipital pole	95392	10.45	2	-94	-4
	R lingual gyrus		6.87	16	-78	-14
	R cerebellum		6.01	42	-66	-26
	L superior parietal		4.45	-24	-78	54
	L angular gyrus		4.27	-40	-82	38
	R temporal pole	3544	7.41	58	12	-24
	R anterior STS		4.67	68	-2	-10
	L anterior MTG	11256	6.76	-68	-8	-14
	L anterior MTG		6.53	-54	0	-26
	L mid MTG		3.62	-56	-34	-6
	L pars orbitalis	23544	5.83	-44	34	-14
	L middle frontal gyrus		5.56	-44	6	58
	L pars triangularis		5.43	-46	24	4
	Substantia nigra	11336	5.51	-8	-22	-14
	R globus pallidus		5.36	16	2	-4
	L globus pallidus		5.26	-16	4	6
	L SMA	1264	4.03	-6	10	58
Speech production	R cerebellum	22112	8.99	36	-78	-34
	R cerebellum		5.18	20	-92	-30
	R cerebellum		4.83	16	-78	-24
	L lateral occipital	7192	6.60	-44	-84	32
	L angular gyrus		4.20	-42	-64	18
	L superior frontal gyrus	7760	5.26	-8	32	62
	L SMA		4.99	-8	18	60
	L superior frontal gyrus		3.70	-10	50	48
	L middle frontal gyrus	6800	5.24	-40	14	58
	L middle frontal gyrus		5.15	-36	10	64
	L pars triangularis	2376	4.99	-52	24	6
	L anterior MTG	7000	4.91	-60	-8	-28
	L anterior ITG		4.55	-44	-8	-50
	L anterior ITG		4.27	-46	-8	-40

Supplementary Table 4: Activation peaks associated with coherence and time during speech production

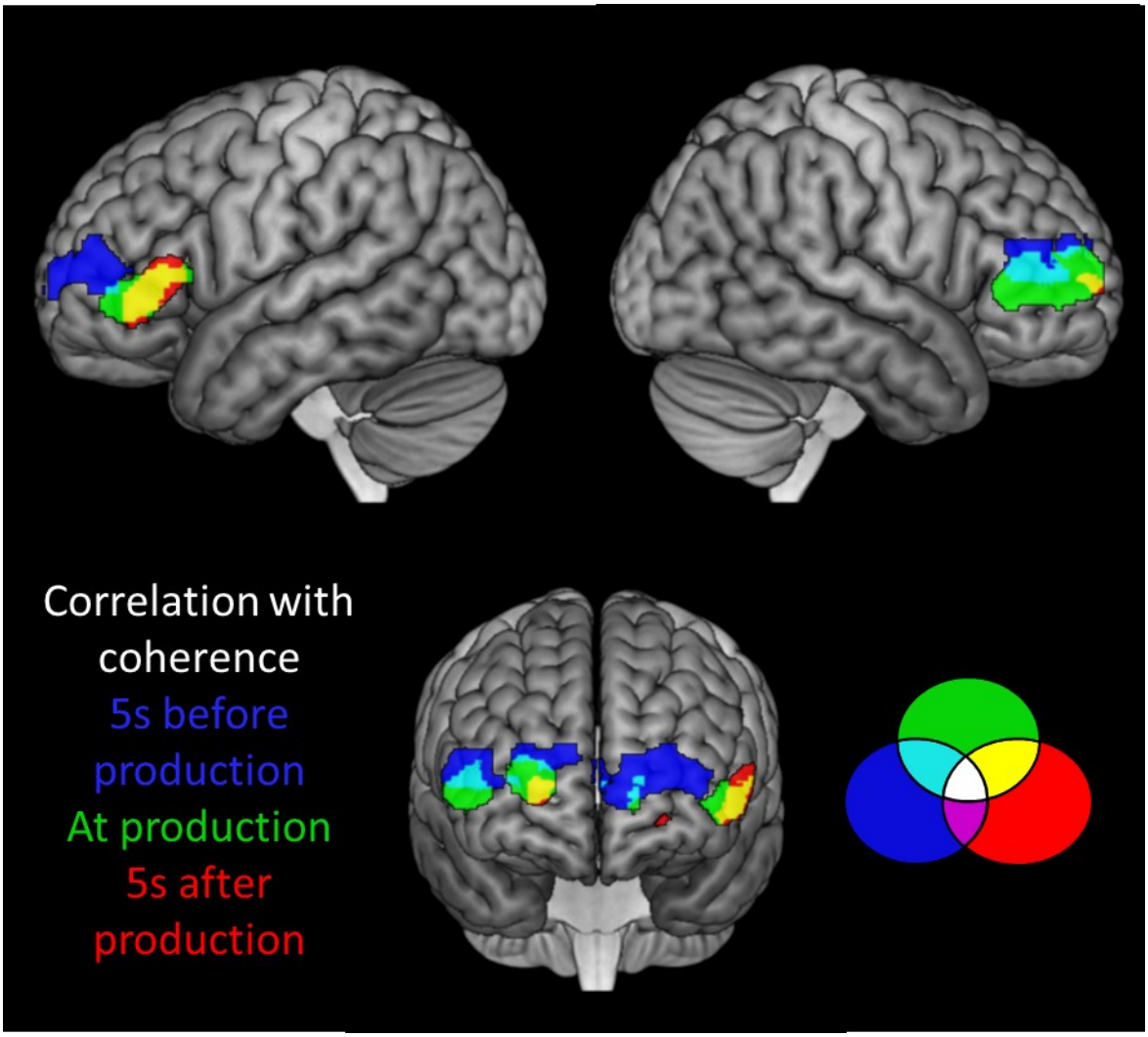
Effect	Location	Extent (mm^3)	*t*	MNI co-ords x	y	z
Positive effect of coherence	R frontopolar cortex	11720	7.20	22	58	6
	R pars triangularis		5.49	42	30	4
	R anterior MFG		5.13	48	56	2
	L anterior cingulate	11656	5.23	-16	40	10
	L pars triangularis		4.86	-44	40	-2
	L anterior insula		4.60	-24	30	6
Negative effect of coherence	No peaks					
Increase over time	L lateral occipital	10944	8.27	-42	-90	-16
	L occipital pole		7.41	-26	-104	-6
	R pars orbitalis	55048	5.69	44	38	-6
	Ventromedial PFC		4.91	4	38	-6
	R pars orbitalis		4.63	30	22	-16
	R temporal pole		3.96	36	10	-32
	L parahippocampal gyrus		3.90	-14	-4	-30
	R middle frontal gyrus	3680	5.09	28	6	32
	R precentral gyrus		4.52	34	0	30
	R precentral gyrus		3.70	26	-6	40
	R angular gyrus	1040	3.60	60	-46	40
	R supramarginal gyrus		3.22	58	-34	48
Decrease over time	L posterior STS	3264	4.04	-58	-32	-2
	R cerebellum	2136	4.03	32	-86	-30
	R cerebellum		3.84	50	-66	-28
	R superior temporal gyrus	1040	3.63	66	2	-6
	R superior temporal gyrus		3.52	58	-8	-2
	R lingual gyrus	1360	3.40	14	-54	-2
	R lingual gyrus		3.35	4	-48	0
	Posterior cingulate		3.07	-4	-46	2

Supplementary Table 5: Peak co-ordinates for analysis of individual differences in coherence

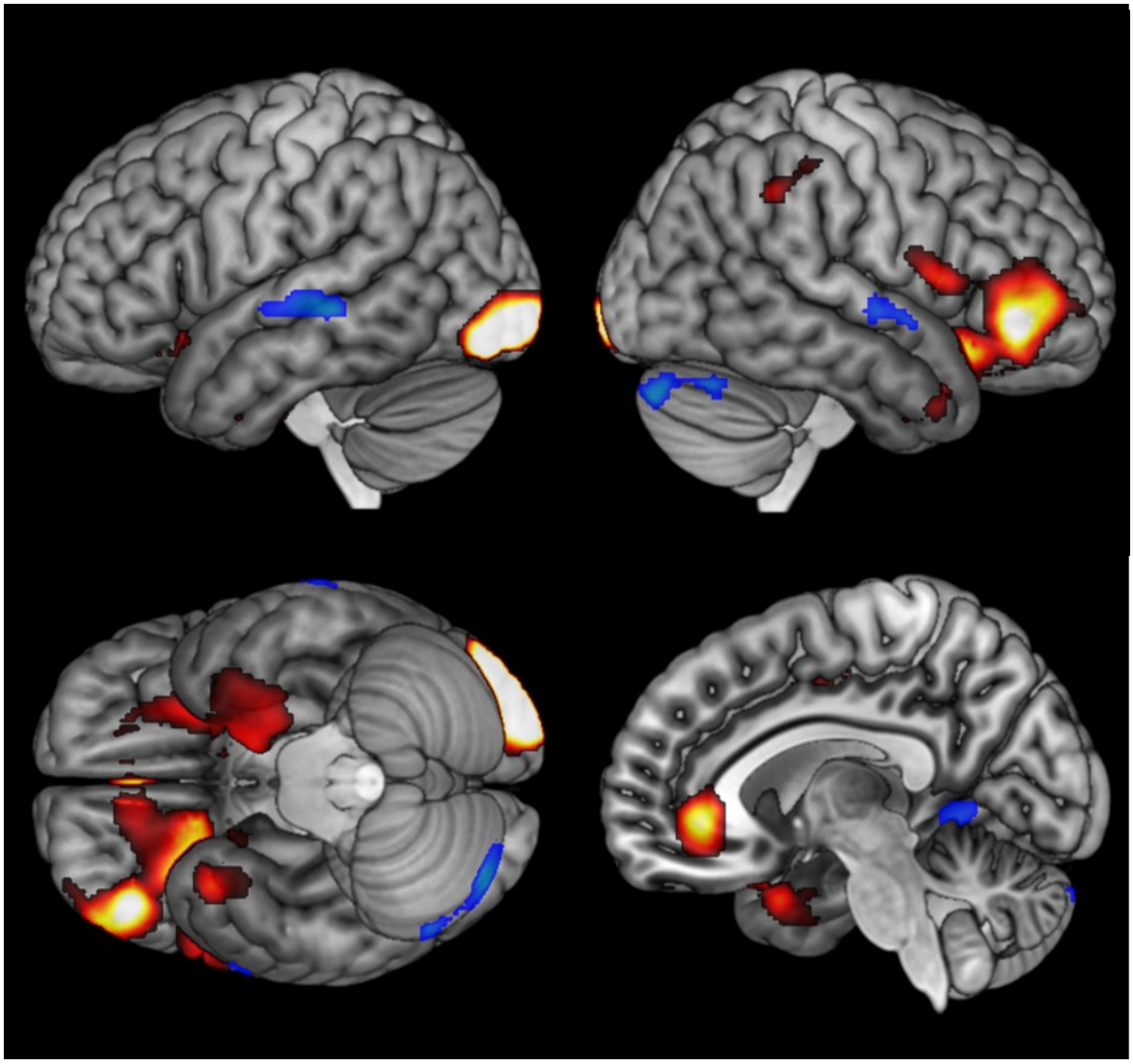
Effect	Location	Extent (mm^3)	*t*	MNI co-ords x	y	z
Greater activation in more coherent speakers during Planning	R pars triangularis	5433	8.51	54	34	2
	R pars orbitalis		4.82	40	40	-6
Greater activation in less coherent speakers during Planning	L anterior ITG	1360	4.67	-52	-6	-38
	L anterior ITG		3.26	-52	-14	-42
Greater activation in more coherent speakers during Production	L temporal pole	2192	5.93	-38	18	-46
	L anterior ITG		3.39	-48	-2	-50
Greater activation in less coherent speakers during Production	No peaks					



Supplementary Figure 1: Average coherence of speech produced in each 5s block of the speech production periods. Bars indicate SEM.



Supplementary Figure 2: Regions showing greater activation when participants produced more coherent speech at time of production (green; standard model), 5s before the speech was produced (blue; early model) and 5s after the speech was produced (red; late model). Images thresholded at p<0.05 cluster-corrected.



<u>Supplementary Figure 3: Regions showing increasing (hot colours) or decreasing (cool colours) activation over time within speech production periods.</u> Images thresholded at p<0.05 cluster-corrected.

Supplementary Materials

List of prompts used to elicit speech:

Which is your favourite season and why?
What do you like or dislike about Christmas?
Why do people come to Scotland on holiday?
What sort of things usually happen at a wedding?
Describe a typical visit to a restaurant
What sort of things do you have to do to look after a dog?
What do people usually do when getting ready for work in the morning?
Describe the steps you would take if going somewhere by train
Describe how you would make a cup of tea or coffee
How would you prepare to go on holiday?
What would it have been like to live in the Middle Ages?
What would it be like to live in Antarctica?
What do the police do when a crime has been committed?
What happens when a storm is forecast in the UK?
What are the advantages and disadvantages of going to university?
What sort of things does the Queen do on a typical day?
What happens during a general election in the UK?
Why are some people concerned about climate change?
Do you think it's a good idea to send people to live on Mars?
Do you think the internet has improved people's lives?