

Word Retrieval Failures in Old Age: The Relationship between Structure and Function

Meredith A. Shafto, Emmanuel A. Stamatakis, Phyllis P. Tam,
and L. K. Tyler

Abstract

■ A common complaint of normal aging is the increase in word-finding failures such as tip-of-the-tongue states (TOTs). Behavioral research identifies TOTs as phonological retrieval failures, and recent findings [Shafto, M. A., Burke, D. M., Stamatakis, E. A., Tam, P., & Tyler, L. K. On the tip-of-the-tongue: Neural correlates of increased word-finding failures in normal aging. *Journal of Cognitive Neuroscience*, 19, 2060–2070, 2007] link age-related increases in TOTs to atrophy in left insula, a region implicated in phonological production. Here, younger and older adults performed a picture naming task in the fMRI scanner. During successful naming, left insula activity was not affected by age or gray matter integrity. Age differences only emerged during TOTs, with younger but not

older adults generating a “boost” of activity during TOTs compared to successful naming. Older adults also had less activity than younger adults during TOTs compared to “don’t know” responses, and across all participants, less TOT activity was affiliated with lower gray matter density. For older adults, lower levels of activity during TOTs accompanied higher TOT rates, supporting the role of an age-related neural mechanism impacting older more than younger adults. Results support a neural account of word retrieval in old age wherein, despite widespread age-related atrophy, word production processes are not universally impacted by age. However, atrophy undermines older adults’ ability to modulate neural responses needed to overcome retrieval failures. ■

INTRODUCTION

One of the most noticeable cognitive changes that occurs as we age is an increase in word-finding problems; that is, an inability to produce a well-known word despite a strong feeling that recall of the word is imminent (James, 2006; Burke, Locantore, Austin, & Chae, 2004; Cross & Burke, 2004; James & Burke, 2000; Burke, MacKay, Worthley, & Wade, 1991). Recent findings using structural neuroimaging (Shafto, Burke, Stamatakis, Tam, & Tyler, 2007) show that these age-related word retrieval failures are associated with neural atrophy. The aim of the current study is to determine whether age-related increases in word retrieval failures are associated with altered patterns of functional activity, and if so, how they are associated with age-related atrophy.

Normal aging is associated with substantial gray matter atrophy across the brain (Resnick, Pham, Kraut, Zonderman, & Davatzikos, 2003; Sowell et al., 2003; Good et al., 2001), which is linked to age-related cognitive change (e.g., Raz, 2000; Raz, Gunning-Dixon, Head, Dupuis, & Acker, 1998). However, the focus of much of this research has been on a limited set of relationships between behavior and neural structure, especially the neural correlates of executive function and verbal memory, with its emphasis on prefrontal and medial-temporal regions (e.g.,

Raz et al., 1998). For example, the Frontal Lobe Hypothesis (West, 1996) contends that prefrontal cortex is particularly vulnerable to the effects of age, leading to differential decline in functions supported by this region (e.g., inhibition in working memory).

Although links between age-related atrophy and performance have been established in structural imaging studies, the conclusions are less clear-cut in fMRI. Some relationships between functional activity and performance appear analogous to structural imaging findings. For example, age-related changes in activation in prefrontal cortex have been linked to age-related changes in executive function (see Reuter-Lorenz, 2002 for a review). Although some studies report underrecruitment of the neural regions which are activated in young adults (Logan, Sanders, Snyder, Morris, & Buckner, 2002), others find that older adults recruit new regions which are not typically activated in the young (e.g., Morcom, Good, Frackowiak, & Rugg, 2003; Cabeza, Anderson, Locantore, & McIntosh, 2002). Importantly, differences in functional response are also found when behavioral performance is equivalent between age groups (Morcom, Li, & Rugg, 2007; Langenecker & Nielson, 2003) or when age-related performance differences are methodologically controlled (Morcom et al., 2003), demonstrating that fMRI studies may reveal differences in the nature of older adults’ responses, which are not evident from behavioral measures alone.

University of Cambridge, Cambridge, UK

However, matching performance across ages is not possible for most studies examining changes in functional activation in old age; tasks tend to be chosen for aging studies because they involve age-related decrements. Differences in neural activity may reflect these performance differences rather than underlying differences in the cognitive processes or neural activity. If task performance is not controlled for, and if older adults demonstrate both differential neural activity and impaired behavioral performance, changes in activation are difficult to interpret. Similarly, in fMRI studies, the role of structural change, in particular gray matter atrophy, is often not taken into account, so it is not clear whether changes in neural activity are related to reduced gray matter in relevant regions or due to cognitive changes which influence patterns of neural recruitment.

Neural Mechanisms of Word-finding Failures

Normal aging is accompanied by a number of cognitive changes, including notable deficits in specific aspects of language. Although many language abilities are preserved, such as comprehension and vocabulary knowledge (Burke & Shafto, 2008), word-finding failures called tip-of-the-tongue states (TOTs) typically increase (e.g., James, 2006; Burke et al., 1991). TOTs occur when a familiar word temporarily fails to come to mind. Behavioral data indicate that TOTs result from a temporary inability to access the phonology of a word, following successful activation of semantic information (Cross & Burke, 2004; James & Burke, 2000). Moreover, although *partial* phonological information (such as the first sound of the target word or its length in syllables) is sometimes available during a TOT, this is less so for older adults, and older adults take longer to resolve their TOTs than younger adults (Burke et al., 1991).

Shafto et al. (2007) recently demonstrated that increases in phonological retrieval deficits across the lifespan were related to increasing atrophy in left anterior insula. This region is affiliated with postsemantic stages of production, namely, phonological retrieval and articulation, with evidence for this role demonstrated in both patient groups (Cereda, Ghika, Maeder, & Bogousslavsky, 2002; Harasty et al., 2001; Dronkers, 1996) and healthy participants (Blank, Scott, Murphy, Warburton, & Wise, 2002; Wise, Greene, Buechel, & Scott, 1999). Shafto et al. assessed gray matter integrity using voxel-based morphometry (VBM), where gray matter density is measured at each voxel in a segmented gray matter image. Negative correlations of age with voxelwise gray matter density give a measure of the degree and distribution of age-related atrophy. Age-related increases in TOT rates correlated negatively with gray matter integrity in left anterior insula, suggesting that TOT frequency increases as age-related gray matter atrophy worsens in regions critical for phonological retrieval.

In the current experiment, we aimed to determine whether age-related neural atrophy is related to age-related changes in neural activity during TOTs. We avoided the limitations of previous fMRI studies by matching younger and older groups on relevant performance measures: We employed a TOT task modified for use in the fMRI scanner and materials pretested to ensure equivalent TOT rates for younger and older adults. We tested a series of predictions on the basis of our previous finding that atrophy in left insula is affiliated with rising TOT rates. We first predicted that, if left anterior insula is important for phonological retrieval, it should be differentially activated during TOTs, for younger as well as older adults. Second, insula gray matter integrity should be related to patterns of functional activity. Finally, if the main source of decreased gray matter integrity is age-related atrophy, older adults should show a stronger relationship between neural activity and TOT rates compared to younger adults.

METHODS

Participants

Participants were 29 healthy adults recruited from the Cambridge community, including 15 younger adults aged 20 to 37 years ($M = 24.60$ years, $SD = 4.45$) and 14 older adults aged 66 to 88 years ($M = 74.50$, $SD = 6.57$). All gave informed consent, and the study was approved by the Suffolk Local Research Ethics Committee. As a requirement for testing in the MRI scanner, no participants had audiometer results indicating severe hearing loss without correction that would impair their ability to follow instructions inside the scanner. All older participants scored 27 or above on the Mini Mental State Exam (Folstein, Folstein, & McHugh, 1975) ($M = 29$, $SD = 0.96$), and there was no age difference in performance on the National Adult Reading Test (Nelson, 1982), $p > .10$. Major exclusion criteria for participation included bilingualism, left-handedness, MR contraindications, neurological or hormonal disorders, recent treatment (within one year) for psychiatric disorders, major head trauma, stroke, or dyslexia.

Materials

Materials consisted of 200 faces of public figures from different categories (e.g., entertainers, politicians, sports figures, authors). Pictures were cropped to include a close up of the face, and resized to the same height (300 pixels) with a range of widths from 218 to 329 pixels. Materials were selected following a pretest in which 10 younger participants aged 23 to 39 years ($M = 28.9$, $SD = 5.1$) and 10 older participants aged 66 to 83 years ($M = 74.5$, $SD = 5.36$), none of whom participated in the fMRI study, viewed 300 pictures of public figures. In the pretest, participants saw each picture on a computer screen for 3 sec, followed by a 1-sec blank screen. Participants, who were

all right handed, made a three-choice button press indicating “know,” “don’t know,” or “TOT,” using the first three fingers of their right hand. The presentation timing was in keeping with fMRI studies of TOTs with younger adults (Kikyo & Sekihara, 2001; Maril, Wagner, & Schacter, 2001). Immediately following this task, they saw the same faces again with the same presentation timing, but this time response choices included the correct name, an incorrect alternative name, and “don’t know.” This second task was used to verify that participants could select the correct name for trials where they had responded know or TOT in the initial task. Materials chosen for use in the scanner task were those which generated at least one TOT for either younger or older adults, and generated a 90% or less don’t know rate for younger and older adults separately. The selected items generated a verified TOT rate of 16%, in keeping with previous behavioral studies (e.g., James & Burke, 2000; Burke et al., 1991), and there were no differences between young and older groups in know, TOT, or don’t know rates (all $ps > .10$). In sum, the pretest helped guarantee that both younger and older adults could perform a TOT task with presentation times and a response mode (i.e., button press) suitable for use in the scanner, and that the materials we chose were likely to elicit TOTs for both younger and older adults.

Procedure

Scanning Session

Before scanning, participants were given a practice session of 20 trials. Each trial involved a 3-sec presentation of a picture followed by a 1-sec blank screen. During the 4-sec trial, participants made their response by button press to indicate “know,” “TOT,” or “don’t know.” Participants were instructed to respond “know” with a button press if they knew and could think of the *name* of the celebrity (familiarity with the face was not sufficient for a “know” response), “TOT” if they were certain they knew the name but could not retrieve it at the moment (again, familiarity was not sufficient; participants had to be certain they knew the person’s name), and “don’t know” if they were certain they did not know the name (even if the face was somewhat familiar). Identical timings and button press responses were then used in the task during the scanning session. During the scanning session, participants had a 10-trial practice session which was followed by 200 trials in two blocks.

Scan Acquisition and Data Analysis

Participants were scanned at the Wolfson Brain Imaging Centre, Cambridge, with a 3-Tesla Bruker scanner using a gradient-echo echo-planar imaging (EPI) sequence (TR = 1100 msec, TE = 27.5 msec, TA = 1100 msec, flip angle 65.49°, matrix size 64 × 64, FOV 24 × 24 cm, in-plane resolution 3.75 × 3.75 mm, 21 oblique slices angled away

from the eyes, 4 mm thick, with body gradients, 101-kHz bandwidth, reconstruction based on a gradient-echo reference scan, trigger pulse length 50 msec). Spoiled gradient recalled T1-weighted scans were acquired for anatomical localization and for comparisons of gray matter integrity between the age groups. For our analyses of gray matter integrity, we undertook the following processing on structural images: T1 images were preprocessed using SPM2 (Wellcome Department of Cognitive Neurology, London, UK) using the optimized VBM protocol (Good et al., 2001). This method was developed to improve segmentation of images into gray matter, white matter, and cerebrospinal fluid. Initially, the images were spatially normalized and segmented into probabilistic gray matter, white matter, and cerebrospinal fluid. Template images specific to our data were constructed from the normalized segmented images and our raw images were normalized again and segmented using these data-specific templates. The final preprocessing step involved smoothing using a 12-mm isotropic Gaussian kernel to make the data conform more closely to the Gaussian field model underlying the statistical procedures used for making inferences, and to render the data more normally distributed (by the central limit theorem; Good et al., 2001).

The fMRI data were preprocessed and analyzed using SPM2 software (Friston et al., 1995) implemented in MATLAB version 6.5 (Mathworks, Natick, MA). Preprocessing involved slice-timing correction to correct differences in image acquisition time between slices. The correction algorithm shifts signal in time to provide an output that represents the same, continuous signal. Within-subject realignment followed slice-timing correction. The process involves realigning all images acquired during a scanning session to the first image in the session (excluding the first five scans to allow for signal stability) to account for head motion. Linear transformations are used at this stage. The process outputs a text file with the amount of translation and rotation required for each image (in x , y , and z directions) in order to bring it to the space of the first image. These translations and rotations in x , y , and z directions were used in the statistical modeling of the data as explained below. The EPI images were then spatially normalized to a standard EPI template based on the MNI reference brain, using a 12-parameter linear affine transformation (translation, rotation, zoom, and shear in x , y , and z directions) and a linear combination of three-dimensional discrete cosine transform basis functions to account for nonlinear differences between target and template. The spatially normalized images were smoothed with an isotropic 8-mm full-width half-maximum Gaussian kernel.

Following preprocessing, data for each subject were modeled with the general linear model. Stick functions representing the onset of the participant’s keypress responses were convolved with the canonical hemodynamic response function and these were entered into a

regression analysis. Regressors were constructed based on participants' button press choice into "know," "don't know," and "TOT" responses, and these response sets were used for all relevant contrasts. For each modeled session, we included six confounding covariates (translations and rotations in x , y , and z directions produced at the realignment stage) to reduce the probability of obtaining false positives that could be attributed to residual movement-related artifacts. Parameter estimate images from each subject were combined into a group random effects analysis. We rendered statistical parametric maps at $p < .001$, voxel-level uncorrected, and report clusters that survived random field corrected p value of $< .05$. Cluster statistics correction for multiple comparisons was carried out for the entire brain, unless otherwise stated. MNI coordinates are reported. In order to identify anatomical regions within clusters and cluster maxima the MNI coordinates were converted to Talairach equivalent coordinates (Brett, Christoff, Cusack, & Lancaster, 2001). To confirm that our naming procedure produced correct names, we also superimposed activations on a mean T1-weighted image (mean of all participants' spatially normalized T1 images) and manually inspected activations using the Mai, Assheuer, and Paxinos (2004) atlas of the human brain and cross-referenced names to the AAL atlas (Tzourio-Mazoyer et al., 2002).

Based on our previous finding relating left insula to phonological retrieval (Shafto et al., 2007), we tested specific hypotheses about the relation of activity to age and performance in contrasts restricted to left or right insula. We identified insular cortex using the WFU Pick Atlas (Maldjian, Laurienti, Kraft, & Burdette, 2003), and used the MarsBaR toolbox (<http://marsbar.sourceforge.net/>) to construct binary masks of left and right insula. We used these masks to identify right or left insula clusters active in group analyses for every contrast. We then calculated the average parameter estimates within active clusters, and extracted these values for each participant. This procedure provided a measure of mean left and right insula activity for each contrast for each participant, and these activity measures were then related to performance scores or age. We used these data to carry out comparisons between age groups where the analyses within insula had yielded significant clusters for each age group separately. A similar process was used for examining the relationship of gray matter to performance and activity, using the same clusters of active voxels in insular cortex. However, instead of extracting mean activity measures, we extracted mean probabilistic gray matter density.

RESULTS

Behavioral Results

The proportion of know, TOT, and don't know responses were analyzed in a 3 (response type: know, TOT, or don't know) \times 2 (age: young or older) mixed ANOVA. Results

are summarized in Figure 1A and demonstrated a main effect of response type [$F(1, 27) = 44.31$, $MSE = 0.022$, $p < .001$], with fewer TOT responses than either know [$t(28) = 6.74$, $p < .001$] or don't know responses [$t(28) = 5.93$, $p < .001$], but no difference in the proportion of know compared to don't know responses ($t < 1.0$). There was no effect of age, and no interaction of age with response type. Overall TOT rate per subject (14.6%) was in keeping with the verified TOT rate measured in behavioral pretesting (see Materials), confirming that the fMRI task provides a reliable measure of TOT rate.

The same pattern held for response times in a 3 (response type: know, TOT, or don't know) \times 2 (age: young or older) mixed ANOVA. There was a similar main effect of response type [$F(1, 27) = 199.69$, $MSE = 68,678.312$, $p < .001$], where TOT responses were slower than either know [$t(28) = 14.10$, $p < .001$] or don't know [$t(28) = 12.00$, $p < .001$] responses, but there was no difference in the speed of know and don't know responses ($t < 1.0$). There was no effect of age, and no interaction of age with response type (see Figure 1B).

Thus, TOTs were slower and less frequent than know or don't know responses, but there was no effect of age

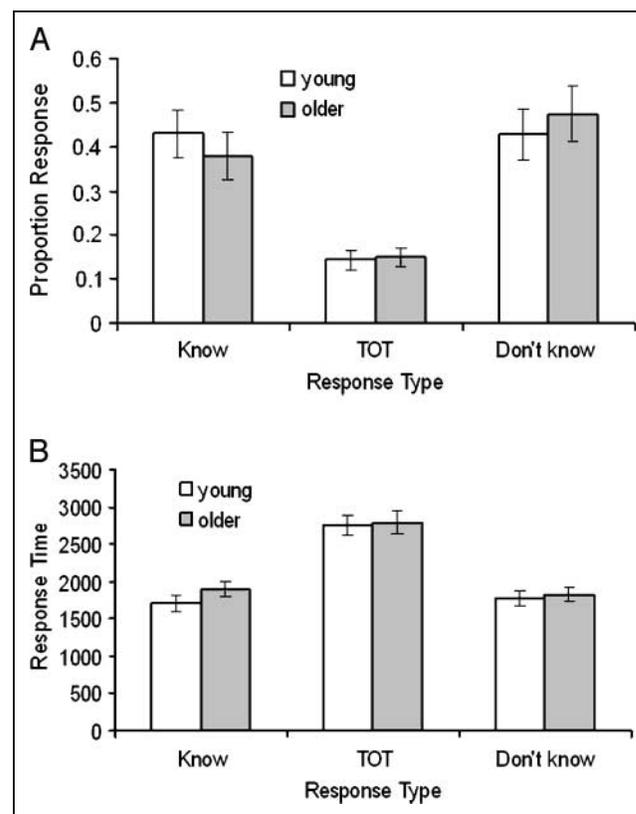


Figure 1. Behavioral performance by response type and age group. (A) Proportion of responses in each response type and (B) mean response times indicate that TOTs were less frequent than knows or don't knows, and were responded to more slowly. There were no effects of age in either proportion response or response times. Error bars show standard error.

on response times or frequency of any response type. This verifies the effectiveness of our materials development, wherein we selected items that would elicit TOTs for younger and older adults. The fact that older and younger adults were matched on their behavioral performance allows us to examine age-related functional differences without confounding factors such as different TOT rates or response times.

Imaging Results

There were two main contrasts of interest: (1) A know contrast (know > don't know), which revealed the broad network of regions involved in producing the name of familiar celebrities in response to their faces, having removed activation that can be attributed to visual processing; and (2) a TOT contrast (TOT > don't know), which revealed a network of similar regions, active during TOTs. We also directly compared know and TOT activity in two contrasts (TOT > know and know > TOT) in order to highlight the differences between successful and unsuccessful production.

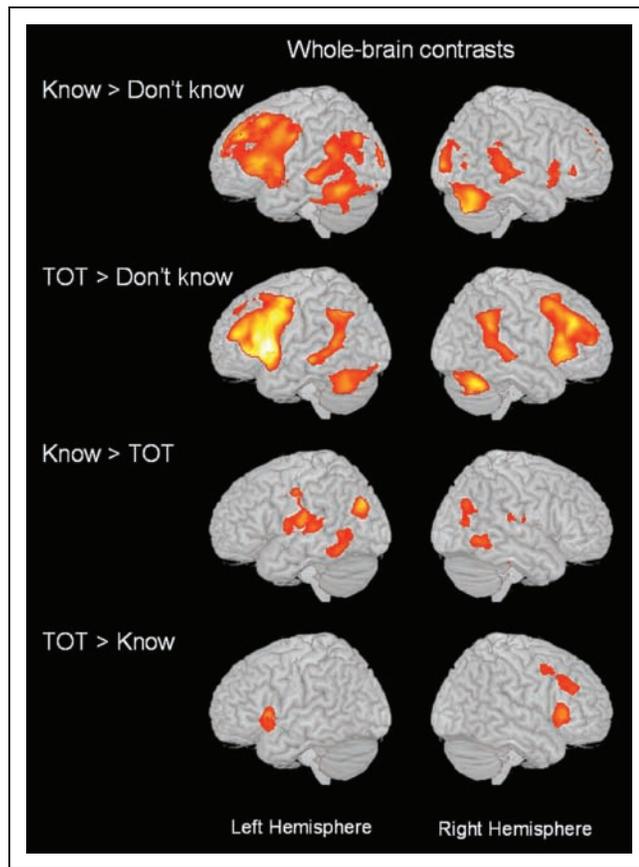


Figure 2. Activity related to know and TOT contrasts and the direct comparison of know and TOT conditions. Whole-brain analyses show significant clusters for all contrasts, thresholded at $p < .001$. Activity from insula is rendered on the surface of the brain.

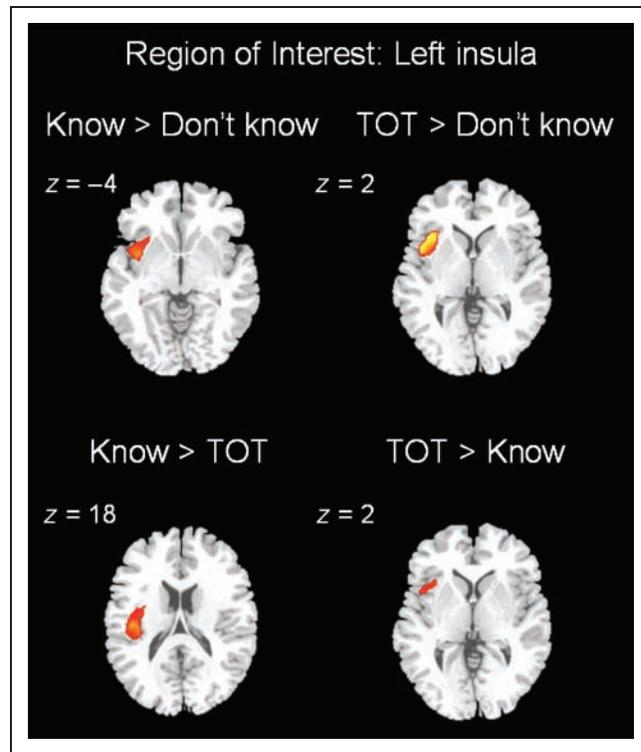


Figure 4. Left insula activity related to know and TOT contrasts, and the direct comparison of know and TOT conditions. Analyses within left insula show significant clusters for all contrasts, thresholded at $p < .001$.

Whole-brain Analyses

The whole-brain know contrast for all participants revealed a widespread network of active regions (see Figure 2). These included a number of left hemisphere regions identified as being involved in picture naming (Indefrey & Levelt, 2004; Murtha, Chertkow, Beauregard, & Evans, 1999), including middle occipital cortex, fusiform gyrus (FG), middle and superior temporal cortex, middle and inferior frontal cortex (LIFC), and supplementary motor area (SMA). There was homologous right hemisphere activity for most of these active regions, with the notable exception of activity in precentral gyrus, SMA, middle frontal, LIFC, and insular cortex, which was significant in the left hemisphere only. The whole-brain analysis for the TOT contrast revealed a network of active regions which were largely a subset of the regions active in the know contrast. These included bilateral middle temporal gyrus, middle frontal and IFC, and SMA (see Figure 2).

The direct contrast of TOT and know responses revealed four clusters of activity for TOT greater than know responses (see Figure 2). The most significant cluster involved LIFC ($x = -52, y = 16, z = 2, Z\text{-score} = 4.67; BA 47$), left anterior insula, and left temporal pole (see Figure 2). The inclusion of left insula indicates the additional phonological retrieval difficulty in the TOT condition, and is in keeping with our previous finding that left insula gray matter integrity correlates with TOT

rates (Shafto et al., 2007). Other significant clusters included a right lateralized cluster including right insula and RIFC ($x = 52, y = 24, z = 4$, Z-score = 4.50; BA 45), a cluster in right middle frontal cortex ($x = 32, y = 34, z = 28$, Z-score = 4.04; BA 9), and a cluster centered in right middle cingulate gyrus ($x = 12, y = 14, z = 48$, Z-score = 4.23; BA 32), extending superiorly to bilateral SMA, and anteriorly to anterior cingulate cortex (ACC). These regions are consistent with previous findings from fMRI studies of TOTs with younger adults (Maril, Simons, Weaver, & Schacter, 2005; Kikyo & Sekihara, 2001). There were also a number of regions more active for know than TOT responses, which were primarily a subset of the regions active in the know contrast, including bilateral middle occipital cortex, FG, and middle and superior temporal cortex.

Left Insula Analyses

Because of our recent findings (Shafto et al., 2007) and previous evidence for the role of the left insula in phonological production (Blank et al., 2002; Cereda et al., 2002; Harasty et al., 2001; Wise et al., 1999; Dronkers, 1996), we used a series of analyses in left insula, anatomically defined using the WFU Pick Atlas (Maldjian et al., 2003), to test our predictions about phonological retrieval during know and TOT responses.

These analyses revealed a significant cluster of activity within left insula during know responses (peak voxel, $x = -38, y = 4, z = -4$, Z-score = 4.60), which was significant for both younger and older adults separately. An independent *t* test revealed that for the know contrast, there was no age difference in mean insula activity (see Figure 3A). There was also significant left insula activity during TOT responses (peak voxel, $x = -40, y = 20, z = 2$, Z-score = 6.01), which was significant for both younger and older adults separately. However, younger adults had greater activity than older adults (see Figure 3B) [$t(27) = 1.77, p < .05$, one-tailed].

Following these contrasts, the direct comparison of TOT and know conditions confirmed the age difference observed in TOT activity: There was a cluster in left insula more active during TOTs than know responses (peak voxel, $x = -46, y = 12, z = 2$, Z-score = 3.91), but this was only significant for younger adults when the age groups were examined separately. The same was true of a cluster of activity in posterior insula with more activity during know responses than TOTs (peak voxel, $x = -42, y = -20, z = 18$, Z-score = 4.72), although older adults had a marginally significant cluster (cluster-level $p = .067$, corrected for multiple comparisons). As can be seen from Figure 4, this cluster is posterior to the activity of the other contrasts, and may not reflect the same processes as those seen in the other contrasts (see Discussion).

Taken together, these results indicate that left insula activity is not affected by age during normal retrieval (know responses), but older adults have less activity than

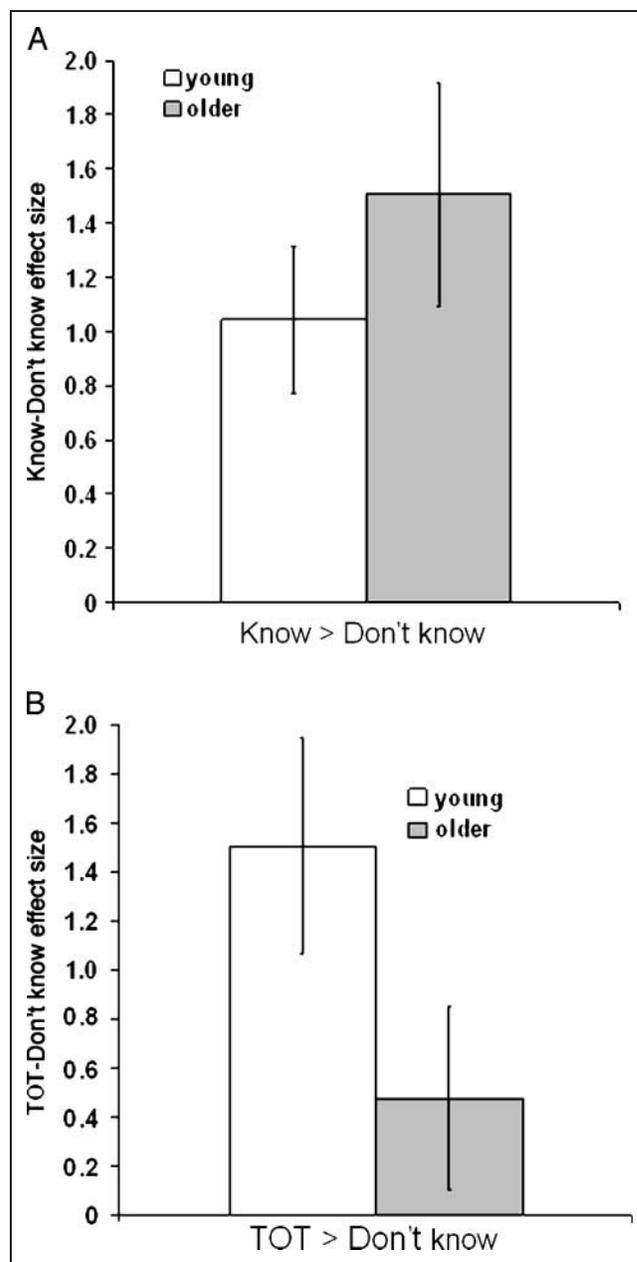


Figure 3. Comparisons of mean activity in left insula across age groups show that (A) there is no age difference for know activity, but (B) younger adults have more TOT activity than older adults.

younger adults during TOT responses. This interaction between response type and age was also found in the relationship between activity and performance: Both younger and older adults show a significant negative correlation between know rates and know activity in left insula ($r = -.70, p < .01$; $r = -.81, p < .001$, respectively; see Figure 5A) and these correlations were not significantly different (by a Fisher's Z-test). However, only older adults showed a significant correlation between TOT activity and performance, with participants who generated greater activity during a TOT overall having fewer TOTs (Figure 5B; $r = -.79, p < .001$).

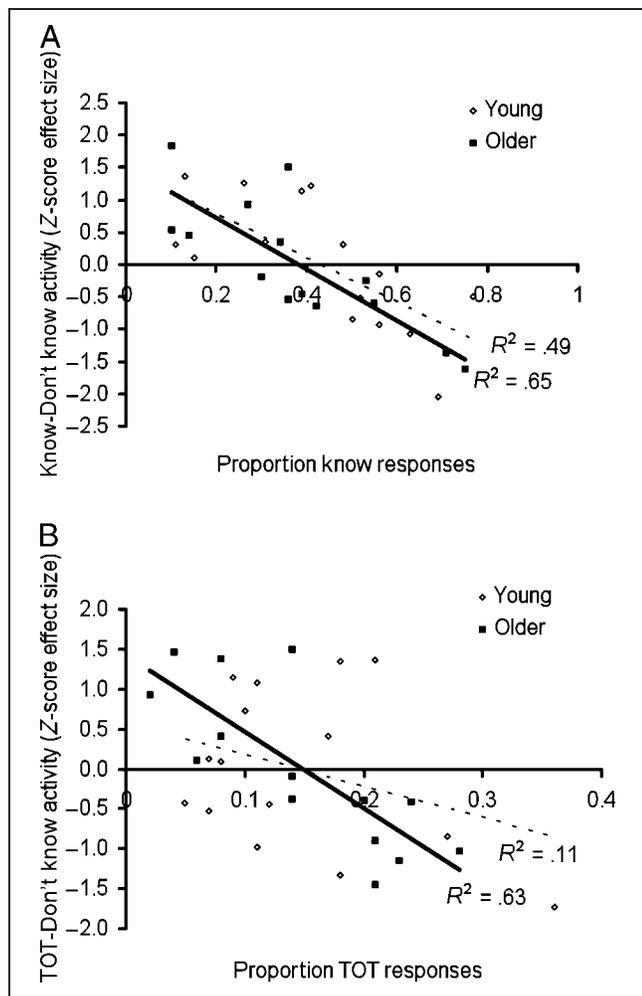


Figure 5. Correlations between performance and activity in left insula. (A) Both younger and older adults had significant negative correlations between proportion know responses and mean activity across voxels in left insula which were significant for the know contrast ($r = -.70$, $p < .01$; $r = -.81$, $p < .001$, respectively). (B) Only older adults had significant negative correlation between proportion TOT responses and mean activity across voxels in left insula which were significant for the TOT contrast ($r = -.79$, $p < .001$).

This last finding is in keeping with the hypothesis that older adults should show a stronger relationship between neural activity and TOT rates compared to younger adults because of lower gray matter integrity in older adults. To verify that lower gray matter integrity is associated with lowered activity during TOTs, we used the previously defined clusters in left insula from the know and TOT contrasts to extract mean gray matter density for these voxels using a VBM analysis as in our previous study (Shafto et al., 2007). Negative correlations of gray matter with age across all subjects confirmed the presence of age-related atrophy in the clusters active for know responses ($r = -.69$, $p < .001$) and TOTs ($r = -.69$, $p < .001$). In order to test whether levels of gray matter are related to activity, we performed a median split on the extracted gray matter measures, and for each contrast, assigned participants to either high or low gray matter density.

For each contrast, this division occurred largely along age group boundaries with low gray matter density groups older than the high gray matter density groups (lowest $t = 4.98$, $p < .001$). Know activity did not differ as a function of high versus low gray matter density, but participants with high gray matter density had higher TOT activity than participants with low gray matter density [$F(1, 27) = 7.66$, $MSE = 2.15$, $p < .05$; see Figure 6]. When this comparison was repeated with age group as a covariate, the effect of gray matter density was reduced to a marginal effect [$F(1, 26) = 3.93$, $MSE = 2.23$, $p = .06$]. Although this marginal effect indicates the potential role of individual gray matter variation independent of age, it is clear that *age-related* changes to gray matter play a dominant role in influencing TOT-related activity.

Contralateral Recruitment: Right Insula

Although our analyses have focused so far on left insula, bilateral insula activity was present for the whole-brain contrasts (see Figure 2). To assess the contribution of the right insula, we repeated the main contrasts within right insula, as anatomically defined using the WFU Pick Atlas (Maldjian et al., 2003). Analyses revealed significant right insula activity during know responses ($x = 36$, $y = 10$, $z = -6$, Z-score = 4.08) and TOTs ($x = 34$, $y = 24$, $z = 2$, Z-score = 5.02), as well as for the direct comparisons of TOTs and know response, with more activity for TOTs than know responses in anterior right insula ($x = 44$, $y = 16$, $z = -2$, Z-score = 3.73), and a posterior cluster with more activity for know responses than TOTs ($x = 40$, $y = -8$, $z = 20$, Z-score = 4.18).

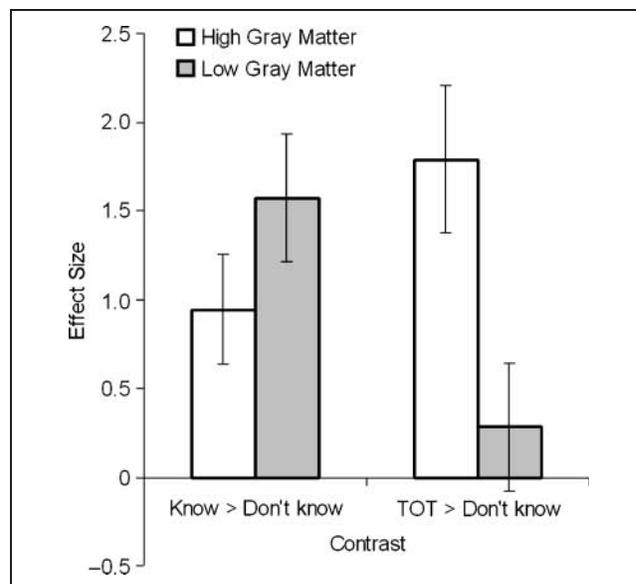


Figure 6. Activity by level of gray matter density. Mean activity in active clusters for know and TOT contrasts is shown divided into groups with low versus high gray matter in these same clusters. t tests revealed a difference for TOT activity but not know activity between participants with low or high levels of gray matter.

Follow-up analyses revealed that although both younger and older adults activated right insula during TOTs (with no age difference in activity), only older adults activated right insula during know responses. This pattern is different from that in left insula, and suggests that younger adults only recruit right insula for difficult retrieval (TOT responses), whereas older adults recruit right insula for normal retrieval (know responses). However, as with the left insula activity, only younger adults showed differential activity between know and TOT conditions, with significant TOT > know and know > TOT clusters for younger but not older adults. Additionally, as in left insula, older but not younger adults demonstrated a correlation between TOT rate and TOT activity ($r = -.54$, $p < .05$), although there were no correlations between know rate and know activity. Finally, there was no relationship for any of the contrasts between activity and gray matter level (high vs. low using a median split) in right insula. To test the possibility that right insula activity reflects recruitment in response to left hemisphere atrophy, we repeated this analysis dividing participants into groups based on *left* hemisphere gray matter level, but again there were no significant effects on right insula activity.

DISCUSSION

Our results support two main predictions regarding word retrieval failures across the lifespan: First, TOTs involve retrieval failures at the phonological level, and second, the success of phonological retrieval depends on the integrity of neural regions including left insula. In keeping with our predictions, we found that (a) left insula was differentially active during TOTs for younger and older adults, (b) the degree of activity during TOTs was related to the neural integrity of left insula, and (c) older adults' performance is more closely linked to neural measures (activity) than that of younger adults.

Our results argue against a general age-related decline in neural activity or cognitive functions, and instead provide evidence which identifies specific functional impacts of age-related atrophy. The age-related changes we find are specific to certain aspects of word production—namely, phonological rather than semantic retrieval, and difficult rather than typical production. Although age-related atrophy affected neural regions involved in both know and TOT responses, gray matter integrity was only related to activity during TOTs.

TOTs as Phonological Retrieval Deficits

Successfully retrieving a name (know response) compared to not knowing the name for a face (don't know response) activated a wide network of regions, consistent with previous overt naming studies using MEG and PET (e.g., Indefrey & Levelt, 2004; Murtha et al., 1999). Thus,

we assume that at least a subset of the regions active for the know contrast reflect the phonological retrieval stage of production, and this was supported by activity in left anterior insula. Because TOTs, like know responses, reflect attempts to retrieve familiar words, we assume that the direct contrast between TOT and know responses reflects the locus of retrieval failure, where know and TOT responses diverge. The presence of more left anterior insula activity for TOTs compared to know responses is consistent with a model which localizes the retrieval failure during TOTs to the phonological level.

Role of Posterior Insula in Phonological Retrieval

The cluster in left insula more active for know than TOT responses was clearly in a more posterior area than the other left insula clusters (see Figure 4), and it is not entirely clear from our results what this posterior activity reflects. It did not correlate significantly with performance (TOT or know rates) or differ significantly across high and low gray matter groups. However, it did correlate with the anterior insula activity for the TOT > know contrast ($r = .48$, $p < .01$), while not correlating with know activity ($p > .10$). These correlations suggest the posterior insula activity may reflect some aspect of TOT processing, and the *positive* direction of the correlation suggests that activity for TOT > know and know > TOT contrasts is not simply trading off.

The different roles of anterior and posterior insula are not well understood with regard to stages of lexical production, but the notion that insular regions are functionally distinct is not new. Anterior and posterior insula have different cytoarchitecture and different connectivity (Bamiou, Musiek, & Luxon, 2003; Augustine, 1985). Anterior insula is implicated in postsemantic processing at a variety of levels including general postsemantic processes (Gitelman, Nobre, Sonty, Parrish, & Mesulam, 2005), articulatory planning (Wise et al., 1999; Dronkers, 1996), and motor planning (Ackermann & Riecker, 2004). Posterior insula, on the other hand, is implicated in more basic auditory processing (e.g., Bamiou et al., 2003). The distinction is not always clear cut, with both regions having been implicated in “automatic” aspects of language (Raichle, 1994), with integrating information from different modalities (Noesselt, Shah, & Jancke, 2003; Mesulam & Mufson, 1985), and with modulating frontal regions during language tasks (Noesselt et al., 2003; Paulesu et al., 1996). In light of these mixed results, it is difficult to conclusively interpret the posterior insula activity, in particular, because the current study was not designed to identify different subprocesses involved in phonological retrieval. However, the positive correlation between anterior and posterior insula activity suggests that a more complete explanation of word retrieval failures should incorporate the interaction between early auditory and later linguistic aspects of phonological retrieval in word production.

The Role of Right Insula

Although right insula is active during retrieval failures for all participants, it was active during successful retrieval only for older adults. Younger adults' TOT-specific activation of the right hemisphere could reflect the need for additional resources during difficult retrieval. If so, older adults' activation of right insula during both successful and unsuccessful responses could indicate that older adults require additional resources during normal retrieval. However, if this is the case, it is difficult to explain why older adults do not have higher TOT-related activity in right insula compared to younger adults, or a significant correlation between know rate and know activity. If right hemisphere activity is compensatory, our findings suggest that age-related atrophy of critical left hemisphere regions is a likely mechanism. However, there was no relationship between right insula activity and gray matter integrity of either left or right insula. Taken together, our results suggest that right insula does not contribute critically to normal word production, but is active during difficult retrieval (TOT responses). This activity may reflect recruitment of additional resources in aid of phonological retrieval, or it may reflect some unique contribution of right insula. Age differences may reflect increased right hemisphere recruitment, although the current results do not show a consistent interaction of retrieval difficulty and age.

Age Differences in Word Retrieval

In the insula-specific analyses, older adults demonstrated attenuated activity in left insula during TOTs compared to younger adults. A negative correlation across all participants ($r = -.51, p < .01$) between TOT rate and TOT activity in left insula indicates that increased activity in response to retrieval difficulty can help overcome retrieval failures. Although we selected stimuli to match performance across age in the current study, older adults were less able to generate additional activity during TOTs compared to know responses, which may provide insight on the typically higher rate of TOTs in older adults (Shafto et al., 2007). Despite an overall reduction in TOT activity, in older adults there is a stronger relationship between activity and TOT rate than in younger adults, implying that TOT rates are more strongly determined by neural factors in older than in younger adults. However, activity during normal retrieval revealed similar effects for young and older adults, and both age groups demonstrated significant negative correlations between activity and know rate. Thus, activity during word production is not universally affected in old age, and both age groups showed a similar relationship between retrieval success and neural activity.

A Neural Model of TOTs: Atrophy and Activity

Taken together, our results suggest that neural atrophy underpins the increase in TOT rates typically seen in older

populations, as well as the age-related differences in TOT-related activity observed in the current study. Although participants with reduced gray matter levels had lower TOT-related activity, lower gray matter did not significantly impact on know activity, possibly because know responses involve less difficult retrieval (as also evidenced by faster response times for know responses compared to TOTs). By the same token, more severe atrophy, such as that seen in age-related pathologies like Alzheimer's disease, may lead to more severe phonological retrieval impairments (Harasty et al., 2001). Phonemic paraphasias are also characteristic of patients with left insula damage following stroke (Cereda et al., 2002). Although patients may have production problems for common words, TOTs in healthy adults tend to be for low-frequency words (e.g., Burke et al., 1991) which are vulnerable to retrieval failure. Thus, normal aging may reflect an intermediary stage, where only words that are particularly vulnerable to retrieval failure (i.e., low-frequency words) are affected by age-related atrophy.

This view places know and TOT responses on a continuum of retrieval success. The question remains as to whether the increase in activity during TOTs over and above know responses merely reflects a progression along this continuum, or reflects a qualitative "boost" initiated when retrieval has failed. Other findings support a qualitative difference between TOT and know responses. For example, although activity during successful retrieval correlates negatively with better performance (increasing know rates), activity during TOTs correlates positively with better performance (lower TOT rates). That is, although overall higher activity is affiliated with the less successful TOT response, *within* the category of TOTs, more activity is associated with better performance. Additionally, know activity does not correlate with TOT rate (i.e., there is no evidence that people with higher activity have both low know rates and high TOT rates). Moreover, know and TOT response rates do not correlate with each other. These findings suggest a qualitative difference between the neural characteristics of TOTs and know responses.

TOTs and the Language System: Beyond Phonological Retrieval

In the whole-brain analyses, there were a number of overlapping regions active for TOT and know responses, implying that differential activity during TOTs reflects, at least in part, the modulation of the language production system as a result of the demands of difficult production. This conclusion is supported by the greater activity for TOT compared to know or don't know responses not only in left insula, but also in ACC and LIFC, regions involved in modulation of the language system in response to task demands (Stamatakis, Marslen-Wilson, Tyler, & Fletcher, 2005; Tyler, Stamatakis, Post, Randall, & Marslen-Wilson, 2005). From this perspective, activity during TOTs reflects modulation of the existing language production network rather than recruitment of additional processes.

An alternative interpretation comes from previous fMRI studies of TOTs using younger participants (Maril et al., 2001, 2005) where ACC, in conjunction with right prefrontal cortex, is claimed to form a cognitive control network responsible for monitoring and mediating post-TOT retrieval processes aimed at resolving the TOT. In the present study, TOTs generated right frontal activity which was notably absent during successful retrieval, indicating recruitment of processes during TOTs which may not be part of normal language production. Modulation of normal language production and the recruitment of a non-linguistic cognitive control network may both contribute to activity during TOTs, and may reflect the onset of TOTs or the resolution processes that follow. A particularly important possibility is that linguistic and nonlinguistic processes interact, with partial phonological information available during TOTs providing critical input for use by the cognitive control network (Maril et al., 2001). Examining this potential is beyond the scope of the current article, but future research could test these alternatives by directly manipulating phonological availability, as previously done in behavioral studies (e.g., James & Burke, 2000).

Conclusions

The aim of this article was to determine how age-related increases in both atrophy and word-finding failures relate to neural activity. Our findings revealed that younger and older adults employ a similar network of regions during both successful retrieval (know responses) and during difficult retrieval (TOT responses). However, younger adults show a differential increase in activity in key regions during difficult retrieval, which was attenuated in the older participants. This age difference is related to decreased gray matter density in the same regions, and in conjunction with other recent results, suggests that age-related atrophy affects older adults' ability to successfully modulate neural activity when word retrieval is difficult. Taken together, these results support a specific, phonological locus for word retrieval failures which increase in old age due to the deleterious effect of age-related atrophy on the flexible functioning of the word production network.

Acknowledgments

This research was supported by an MRC Programme grant, a Newton Trust grant, a grant from the Strategic Promotion of Ageing Research Capacity to L. K. T., and a Junior Research Fellowship from Christ Church College, Oxford, to M. A. S. We thank Deborah Burke for assistance with this research and many stimulating discussions.

Reprint requests should be sent to Dr. Meredith A. Shafto, Department of Experimental Psychology, University of Cambridge, Cambridge, CB2 3EB, UK, or via e-mail: mshafto@cs1.psychol.cam.ac.uk.

REFERENCES

- Ackermann, H., & Riecker, A. (2004). The contribution of the insula to motor aspects of speech production: A review and a hypothesis. *Brain and Language*, *89*, 320–328.
- Augustine, J. R. (1985). Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Research Reviews*, *22*, 229–244.
- Bamiou, D. E., Musiek, F. E., & Luxon, L. M. (2003). The insula (Island of Reil) and its role in auditory processing literature review. *Brain Research Reviews*, *42*, 143–154.
- Blank, S. C., Scott, S. K., Murphy, K., Warburton, E., & Wise, R. J. S. (2002). Speech production: Wernicke, Broca and beyond. *Brain*, *125*, 1829–1838.
- Brett, M., Christoff, K., Cusack, R., & Lancaster, J. (2001). Using the Talairach atlas with the MNI template. *Neuroimage*, *13*, S85.
- Burke, D. M., Locantore, J. K., Austin, A. A., & Chae, B. (2004). Cherry pit primes Brad Pitt: Homophone priming effects on young and older adults' production of proper names. *Psychological Science*, *15*, 164–170.
- Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip-of-the-tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language*, *30*, 542–579.
- Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 373–443). New York: Psychology Press.
- Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *Neuroimage*, *17*, 1394–1402.
- Cereda, C., Ghika, J., Maeder, P., & Bogousslavsky, J. (2002). Strokes restricted to the insular cortex. *Neurology*, *59*, 1950–1955.
- Cross, E. S., & Burke, D. M. (2004). Do alternative names block young and older adults' retrieval of proper names? *Brain and Language*, *89*, 174–181.
- Dronkers, N. (1996). A new brain region for coordinating speech articulation. *Nature*, *384*, 159–161.
- 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.
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-B., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*, 189–210.
- Gitelman, D. R., Nobre, A. C., Sonty, S., Parrish, T. B., & Mesulam, M. M. (2005). Language network specializations: An analysis with parallel task designs and functional magnetic resonance imaging. *Neuroimage*, *26*, 975–985.
- Good, C. D., Johnsrude, I., Ashburner, J., Henson, R. N. A., Friston, K. J., & Frackowiak, R. S. J. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage*, *14*, 21–36.
- Harasty, J. A., Halliday, G. M., Xuereb, J., Croot, K., Bennett, H., & Hodges, J. R. (2001). Cortical degeneration associated with phonologic and semantic language impairments in AD. *Neurology*, *56*, 944–950.
- Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*, 101–144.
- James, L. E. (2006). Specific effects of aging on proper name retrieval: Now you see them, now you don't. *Journal of Gerontology B: Psychological Sciences and Social Sciences*, *61*, 180–183.

- James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1378–1391.
- Kikyo, H., & Sekihara, K. (2001). Temporal characterization of memory retrieval processes: An fMRI study of the “tip of the tongue” phenomenon. *European Journal of Neuroscience*, *14*, 887–892.
- Langenecker, S. A., & Nielson, K. A. (2003). Frontal recruitment during response inhibition in older adults replicated with fMRI. *Neuroimage*, *20*, 1384–1392.
- 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*, 827–840.
- Mai, J. K., Assheuer, J., & Paxinos, G. (2004). *Atlas of the human brain* (2nd ed.). San Diego: Elsevier Academic.
- Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage*, *19*, 1233–1239.
- Maril, A., Simons, J. S., Weaver, J. J., & Schacter, D. L. (2005). Graded recall success: An event-related fMRI comparison of tip of the tongue and feeling of knowing. *Neuroimage*, *24*, 1130–1138.
- Maril, A., Wagner, A. D., & Schacter, D. L. (2001). On the tip of the tongue: An event-related fMRI study of semantic retrieval failure and cognitive conflict. *Neuron*, *31*, 653–660.
- Mesulam, M. M., & Mufson, E. J. (1985). The insula of Reil in man and monkey. Architectonics, connectivity, and function. In E. G. Jones & A. Peters (Eds.), *Cerebral cortex* (Vol. 4, pp. 179–226). New York: Plenum Press.
- Morcom, A. M., Good, C. D., Frackowiak, R. S. J., & Rugg, M. D. (2003). Age effects on the neural correlates of successful memory encoding. *Brain*, *126*, 213–229.
- Morcom, A. M., Li, J., & Rugg, M. D. (2007). Age effects on the neural correlates of episodic retrieval: Increased cortical recruitment with matched performance. *Cerebral Cortex*, *17*, 2491–2506.
- Murtha, S., Chertkow, H., Beaugard, M., & Evans, A. (1999). The neural substrate of picture naming. *Journal of Cognitive Neuroscience*, *11*, 399–423.
- Nelson, H. E. (1982). *The National Adult Reading Test (NART): Test manual*. London: The National Hospital for Nervous Diseases.
- Noesselt, T., Shah, N. J., & Jancke, L. (2003). Top-down and bottom-up modulation of language related areas—An fMRI study. *BMC Neuroscience*, *4*, 13.
- Paulesu, E., Frith, U., Snowling, M., Gallagher, A., Morton, J., Frackowiak, R. S. J., et al. (1996). Is developmental dyslexia a disconnection syndrome? Evidence from PET scanning. *Brain*, *119*, 143–157.
- Raichle, M. E. (1994). Images of the human mind. *Neuropsychopharmacology*, *10*, S28–S33.
- Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of aging and cognition II* (pp. 1–90). Mahwah, NJ: Erlbaum.
- Raz, N., Gunning-Dixon, F. M., Head, D., Dupuis, J. H., & Acker, J. D. (1998). Neuroanatomical correlates of cognitive aging: Evidence from structural magnetic resonance imaging. *Neuropsychology*, *12*, 95–114.
- Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: A shrinking brain. *Journal of Neuroscience*, *23*, 3295–3301.
- Reuter-Lorenz, P. (2002). New visions of the aging mind and brain. *Trends in Cognitive Sciences*, *6*, 394–400.
- Shafto, M. A., Burke, D. M., Stamatakis, E. A., Tam, P., & Tyler, L. K. (2007). On the tip-of-the-tongue: Neural correlates of increased word-finding failures in normal aging. *Journal of Cognitive Neuroscience*, *19*, 2060–2070.
- Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience*, *6*, 309–315.
- Stamatakis, E. A., Marslen-Wilson, W. D., Tyler, L. K., & Fletcher, P. C. (2005). Cingulate control of fronto-temporal integration reflects linguistic demands: A three-way interaction in functional connectivity. *Neuroimage*, *28*, 115–121.
- Tyler, L. K., Stamatakis, E. A., Post, B., Randall, B., & Marslen-Wilson, W. D. (2005). Temporal and frontal systems in speech comprehension: An fMRI study of past tense processing. *Neuropsychologia*, *43*, 1963–1974.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, *15*, 273–289.
- West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*, 272–292.
- Wise, R. J. S., Greene, J., Buechel, C., & Scott, S. K. (1999). Brain regions involved in articulation. *Lancet*, *353*, 1057–1061.