



HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2016 November 28.

Published in final edited form as:

Neuroimage. 2014 February 15; 87: 80–88. doi:10.1016/j.neuroimage.2013.10.064.

Does experience in talking facilitate speech repetition?

Linda I. Shuster^{a,b,*}, Donna R. Moore^a, Gang Chen^c, Dennis M. Ruscello^a, and William F. Wonderlin^d

^a Department of Speech Pathology and Audiology, 805 Allen Hall, West Virginia University, Morgantown, WV 26506, USA

^b Center for Advanced Imaging, 805 Allen Hall, West Virginia University, Morgantown, WV 26506, USA

^c Scientific and Statistical Computing Core, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, USA

^d Department of Biochemistry, 1 Medical Center Drive, P.O. Box 9142, West Virginia University, Morgantown, WV 26506, USA

Abstract

Speech is unique among highly skilled human behaviors in its ease of acquisition by virtually all individuals who have normal hearing and cognitive ability. Vocal imitation is essential for acquiring speech, and it is an important element of social communication. The extent to which age-related changes in cognitive and motor function affect the ability to imitate speech is poorly understood. We analyzed the distributions of response times (RT) for repeating real words and pseudowords during fMRI. The average RT for older and younger participants was not different. In contrast, detailed analysis of RT distributions revealed age-dependent differences that were associated with changes in the time course of the BOLD response and specific patterns of regional activation. RT-dependent activity was observed in the bilateral posterior cingulate, supplementary motor area, and corpus callosum. This approach provides unique insight into the mechanisms associated with changes in speech production with aging.

Introduction

Vocal imitation is one of the earliest abilities that humans display. It is essential for acquiring speech and language, both for the first time and for the acquisition of a second language. Mature speakers routinely and unconsciously imitate various aspects of the speech of others; this has been observed during both structured imitation tasks and in studies of conversation (Fowler et al., 2003; Pardo, 2006). Speech imitation during social interactions has been proposed to serve a variety of functions, and it may drive phonetic changes in language, as well as the acquisition of dialect (*e.g.*, Babel, 2012; Garrett and Johnson, to appear). The

* Corresponding author at: 515 Michigan St. NE, Suite 200, Grand Rapids, MI, USA. shusterl@gvsu.edu (L.I. Shuster), moore.2492@buckeyemail.osu.edu (D.R. Moore), gangchen@mail.nih.gov (G. Chen), Dennis.Ruscello@mail.wvu.edu (D.M. Ruscello), wwonderlin@hsc.wvu.edu (W.F. Wonderlin).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2013.10.064>.

ability to quickly imitate the speech of others can be important. During conversation, for example, a speaker might need to quickly imitate some aspects of their partner's speech to maintain the conversational flow. Unfortunately, the quickness of imitation may decrease with aging. Studies of speech and limb function have revealed that motor responses slow with aging (Fozo and Watson, 1998; Mattay et al., 2002; Rodríguez-Aranda and Jakobsen, 2011). Fozo and Watson demonstrated that the reaction times for producing both simple and complex utterances were slower in older speakers (mean age = 74 yrs) than in young adults. Moreover, imaging studies of motor behavior have shown that older adults seem to recruit more brain regions than younger adults when performing limb movements or speaking (Mattay et al., 2002; Soros et al., 2011).

Although seemingly a trivial skill, speech imitation or repetition putatively involves a variety of underlying cognitive processes, including accurately perceiving the stimulus, holding the stimulus in phonological working memory, constructing a motor plan for producing the utterance and executing the plan. The neurological substrate underlying repetition is proposed to involve a bilateral dorsal stream (Hickok, 2009); however, evidence for this comes predominantly from neuroimaging studies involving young adults. Only one neuroimaging study has investigated the speech repetition performance of older adults (mean age = 71 yrs; Soros et al., 2011). Soros and colleagues used functional magnetic resonance imaging (fMRI) to study older and younger adults while they repeated 'ah' and 'pataka.' They monitored jaw movement as a measure of response latency. Similar to studies of limb movement, older adults demonstrated more areas of activation than younger participants, but there was no difference between older and younger participants with regard to latency and accuracy of response. These data suggest that the older participants were able to achieve the same level of performance as the younger participants, but they required more brain activation to do so. However, the tasks that Soros et al. employed, especially the production of an isolated vowel, are not typical of the everyday use of speech.

In the present study we have brought together three experimental approaches to investigate how changes in the ability of speakers to imitate speech with aging might be related to changes in underlying neural processes. First, we have used an analysis of the distributions of speech response latencies (*i.e.*, reaction time, RT) to obtain a more complete picture of changes in the speed of imitation with aging than can be interpreted from changes in the mean (or median) alone. Second, we recorded overt speech in a magnetic resonance imaging scanner, which enabled us to examine changes in the activity of different regions of the brain that occur in association with word repetition. Third, Yarkoni et al. (2009) estimated the shape of regional response curves using a finite impulse response analysis of fMRI data collected during a variety of behavioral tasks to identify RT-correlated changes in activity in both gray and white matter regions of the brain. We have used an analogous analytical approach for fMRI data collected during a word repetition task to identify regions in which the RT-correlated activity is affected by age and the type of word to be repeated.

An important goal of this study was to determine how the ability of speakers to repeat real words and pseudowords, tasks that are similar to everyday speech, was influenced by age. We included pseudowords because speakers typically encounter and acquire new vocabulary throughout their lifetime, and a speaker's ability to repeat a pseudoword should reflect the

ability to repeat a newly encountered word for which they do not know the meaning. We chose to study middle-aged speakers because there are fewer data for this age group compared to the elderly. Moreover, middle-aged speakers are at risk for diseases such as stroke, but are also under financial pressure to continue to work. Thus, normative data regarding brain function in middle-aged speakers could help in the design of treatments for the recovery of speech post-stroke.

In addition to investigating age-dependent changes in word repetition, we also sought to identify regions of the brain in which the level of activity measured by fMRI covaried with response latency. Variability in the RTs for behavioral tasks typically deviates from a Gaussian distribution around an average value, with the addition of a variable degree of skewing of the distribution towards longer RTs. Many investigators have fitted the positively-skewed distributions of RTs with multiparameter functions (*e.g.*, ex-Gaussian, Weibull, Gamma, ex-Wald) from which putative relationships between the parameters of these functions (*i.e.*, the shape of the distributions) and various motor and cognitive processes underlying a behavioral task have been proposed (reviewed in Luce, 1986; Van Zandt, 2000; Van Zandt, 2002). However, these relationships are complicated. Therefore, rather than choosing among alternative mathematical functions based on an *a priori* assumption about an underlying model for word repetition, we have used the ex-Gaussian function, which is a convolution of an exponential function and a Gaussian function. The ex-Gaussian function is very robust in fitting distributions of RTs, and it can provide insight into changes in skewness and central tendency, albeit in a model-independent manner (Van Zandt, 2002). We also performed an analysis of covariance of the fMRI data to identify regions of the brain in which the RT-correlated component of the activity was significantly affected by age and/or word type. By comparing these results with the effects of age and/or word type on the parameters of an ex-Gaussian function fitted to the distributions of RTs, we could potentially identify regions of the brain in which the level of activity identified by fMRI analysis might be related to specific features of the distributions of RTs. Although this approach does not enable us to demonstrate causal relationships between brain activity and the distribution of RTs, it does represent an important step towards identifying regions of the brain that might play a role in influencing word repetition.

Material and methods

Participants

Participants were 23 healthy adults, 11 young adults (19.11–28.6 yrs., $M = 22.8$, $sd = 2.4$, 1 male) and 12 middle-aged (48.11–68 yrs., $M = 56.5$, $sd = 7.1$, 5 males) with no history of speech, language or neurological problems and who passed a hearing screening at 20 dB for the frequencies 125–4000 Hz. They were all native speakers of English. The younger adults were undergraduate or graduate students. The middle-aged adults' education ranged from 16 to 20 years ($M = 17.8$, $sd = 1.8$) and all were employed in full-time jobs. There was one left-handed participant (middle-aged) as determined by the Edinburgh (Oldfield, 1971). All provided written consent under a protocol approved by the West Virginia University (WVU) Institutional Review Board. Two of the MA participants were unable to repeat all of the words accurately (including the real words) because they had difficulty hearing them, even

though they had passed the hearing screening and the stimulus intensity was increased to the maximum. Therefore, the data from these two participants were excluded from the analysis.

Stimuli

The stimuli were 30 four-syllable real words and 30 four-syllable pseudowords that were created by re-arranging the syllables of the real words. The pseudowords were created so as to have the same initial syllable frequency (Celex database Release 2, 1995) and phonotactic probability (Vitevitch and Luce, 1999) as the real words. They were edited to have the same RMS intensity and preceded by the carrier phrase “Do say_____”. The total stimulus length (carrier phrase + - word/pseudoword) was 1750 ms. There were also 30 pink noise stimuli of the same duration and RMS intensity as the word stimuli and preceded by the phrase “Don't say _____,” which served as a no-speech control condition.

Experimental design

Participants repeated all of the words/pseudowords aloud one time prior to entering the scanner. They repeated only the word or pseudoword portion of the stimulus (not the carrier phrase). Their responses were recorded on one channel of a digital recorder and the stimuli they heard were simultaneously recorded on the second channel. For the fMRI, we used an event-related design as in Shuster (2009) and Shuster and Lemieux (2005). The fixed interstimulus interval was 12 s. The stimuli were presented in random order in two sessions under the control of the Presentation® software (Neurobehavioral Systems). Forty-five stimuli were presented during each run, 15 of each of the three types (noise, real word, pseudoword) and the duration of each run was 9 min, 20 s. Participants repeated the words/pseudowords aloud immediately upon hearing them, and they were asked to not move their mouths in any way (*e.g.*, lick lips, swallow) in the period immediately following the presentation of the noise stimulus, as well as immediately after the repetition. Again, participants' responses and the stimuli that were presented were recorded digitally. The speech was recorded using an MR compatible microphone (Optoacoustics) mounted to the head coil, and the stimuli were presented over MR compatible Stax ear buds. Participants wore ear muffs over the ear buds. Prior to beginning the actual experimental run, stimuli were presented to each participant while the spiral in/out pulse sequence was running, and the volume was adjusted so that the stimuli were clearly audible. Because the ear muffs fit tightly into the head coil, it was difficult for participants to move their heads; however, in order to minimize head movement as much as possible, a piece of tape was placed across participants' foreheads and attached to each side of the head coil. They were told that if they felt a tug on the tape, this meant they were moving their heads and they should try to avoid this.

Analysis of accuracy of responses

Pseudoword productions that turned the pseudoword into a real word or that were not four syllables were excluded from the analysis, as were any incorrect real word productions, with one exception. In order to create 30 real word and pseudoword stimuli that were matched for initial syllable frequency and phonotactic probability, we had to use both ‘laboratory’ and ‘lavatory.’ If a participant produced ‘lavatory’ in place of ‘laboratory,’ or vice versa, these productions were retained in the analysis. This occurred twice. One middle-aged participant

produced 'lavatory' when the stimulus was 'laboratory,' and one younger participant produced 'laboratory' when the stimulus was 'lavatory.' All participants repeated all of the words. However, for pseudoword repetition, three of the 300 total utterances produced by the 10 middle-aged participants who were included in the analysis and two of the 330 utterances produced by the 11 younger participants were omitted from the analysis, based on the error criteria described above. For real word repetition, one utterance was omitted for one of the middle-aged participants and none were omitted for the younger participants. Repetitions were orthographically transcribed.

Analysis of reaction times

The reaction time (RT) for each utterance was measured from the onset time of the stimulus carrier phrase to the onset time of the participant's response. The RT and the duration of each word were measured from speech spectrograms using a digital cursor in Cool Edit Pro. In order to more clearly visualize the speech responses, the participants' utterances were filtered using an FFT filter created using the noise profile of the scanner noise. Nine subjects were randomly selected and RTs for one scanner run for each subject were independently re-measured by a second investigator (43% of the total utterances). Reliability was determined using a Pearson's correlation.

In addition to comparing mean RTs, we also analyzed the distributions of RTs by fitting each distribution with a Gaussian and an ex-Gaussian function. The ex-Gaussian function is a convolution of Gaussian and exponential functions with parameters including the mean (μ) and standard deviation (σ) of the Gaussian component and the time constant (τ) of the exponential component (Lacouture and Cousineau, 2008). Compared to a symmetric Gaussian distribution, the addition of the exponential component improves the fit to asymmetric distributions that have a tail extending to longer reaction times. The shape of the ex-Gaussian function can range from a normal, bell-shaped curve (when τ is small) to a highly positively skewed curve (when τ is large). The ex-Gaussian function can be useful as an initial step in the analysis of RT distributions because it does not make any assumptions about an underlying model, compared to alternative models for RT distributions. Best-fit values of μ , σ , and τ were estimated using the Continuous Maximum Likelihood (CML) method implemented either in the QMPE/CML software package (v2.18; Heathcote et al., 2004) or in a software routine written by one of the authors (Wonderlin). Reaction time data were pooled across participants within each combination of age group and word type, which yielded a robust estimation of the parameters using CML because of the large number of RTs in each data set. The standard error of each parameter estimate was calculated using two methods. First, the standard error was calculated from the Hessian matrix by an internal routine in QMPE/CML. Second, a bootstrapping method (Efron, 1979) was used in which each data set was fitted 100 times, with one RT randomly selected and removed from the data set (with replacement) for each iteration. The standard deviation of the 100 estimates of the parameter provides an estimate of the standard error of the parameter. Both methods yielded similar values for the standard errors, with the largest discrepancy being less than 10% of the parameter value. Standard errors calculated by QMPE/CML were used in all figures. Differences between data sets in the estimated values of a parameter were statistically tested using the likelihood ratio test of a $\chi^2(1)$ distribution with likelihood ratio

$L_{\text{ratio}} = 2 * (L_{\text{nested}} - L_{\text{full}})$, where the log-likelihood of the full model (L_{full}) was compared to that of a nested model (L_{nested}) in which the parameter being tested was shared during the fitting of the two data sets.

fMRI data acquisition, processing, and analysis

Imaging data were acquired on the GE Signa 3 T MR scanner in the WVU Center for Advanced Imaging. The functional imaging parameters were as follow: spiral in/out; axial plane; TR = 2 s; FOV = 240; matrix = 64×64 ; slice = 5 mm. We also collected high resolution anatomical images with the following parameters: SPGR; FOV = 250; matrix 256×256 ; slice = 1.8 mm; 124 slices. These data were analyzed using the AFNI software suite (Cox, 1996). For each participant, after reconstruction (Glover and Law, 2001), the images were slice timing corrected, warped to the ICBM 452 template, and motion corrected, and the time series were scaled by the temporal mean so that the response estimates can be interpreted in percent signal change. The data were then entered into a regression analysis (3dDeconvolve). The motion parameters that were output during the motion correction process were used as regressors of no interest in the regression analysis. Speech motion artifact was controlled by censoring the time points during which the speech response occurred, which was consistently between 0 and 2.1 s after the onset of the stimulus. Censoring means that those time points were not used in the statistical model. RT was used as an explanatory variable at the trial level to determine the extent to which the brain activation was proportional to the RT for each response. To accomplish this, the RT for each response was entered into the model using the amplitude modulation (also known as parametric modulation) approach. This yields two sets of regressors, one associated with the average RT and the other in proportion to the RT (RT-modulated or RT-correlated component). The relationship between the two effects can be visualized as the intercept and slope when the x and y axes are the RT and BOLD responses respectively (with RT centered around the average value). This approach is not unusual in fMRI analysis, but in the majority of studies, the RT-correlated component is viewed as a nuisance variable and ignored. The first component (average RT) shows the typical BOLD response and has a unit of percent signal change. The second component (RT-correlated) indicates the response change in percent signal change (increase or decrease) when RT increases by one time unit (second).

We did not want to make any *a priori* assumptions regarding the shape of the hemodynamic response function. Therefore, we estimated and captured the shape with six tent basis functions that equally spanned 10 s beginning 2 s after the stimulus onset (also known as finite impulse response (FIR) method). The amplitudes for the first and last basis functions were assumed to be 0; therefore, there were effectively four basis functions. The flexibility of the tent basis function offers the advantage of more accurately fitting the time course of the BOLD response compared to the standard approach commonly used in fMRI investigations, in which the shape of the BOLD response is pre-specified, *e.g.*, using a gamma variate function. The analysis at the individual subject level resulted in four effect estimates per word type that characterized the response shape associated with average RT, and another set of four effect estimates per word type that captured the trial-to-trial variability and indicated the RT-correlated effects. We analyzed the RT-correlated effects at the group level with a three-way ANOVA (3dMVM, Chen et al., 2013): one between-

subjects factor (two levels: young and middle-aged), and two within-subjects factors, word type (two levels: pseudo and real) and time points for the RT-correlated effects (four levels). Multiple comparisons correction for family-wise errors was accomplished using Monte Carlo simulations through 3dClustSim, and all of the results presented below have a corrected p of 0.05 at the cluster level.

Results

Behavioral results

Independent measurement of the RTs by two investigators revealed excellent reliability ($r = .98$). In a preliminary analysis, we used the conventional approach of comparing the mean values of RT for each combination of group and word type. Using a linear mixed-effects modeling approach, we found that there was no significant difference in the mean RT for the repetition of either real words or pseudowords by younger versus middle-aged participants, and the mean RT for both age groups was significantly shorter for repeating real words than repeating pseudowords (middle-aged: $t(9) = 4.01$, $p = .003$, two-tailed; younger: $t(10) = 2.98$, $p = .014$, two-tailed; Fig. 1e).

There are good reasons for extending our analysis to include possible differences in the shape of the RT distributions, rather than limiting the analysis to only comparing measures of central tendency. First, the RT distribution often deviates from a Gaussian distribution, and the arithmetic mean can be biased by these deviations in the distribution. Second, the cognitive processes underlying a behavior such as speech are too complex to tease apart by a simple comparison of mean RTs, and many studies have provided evidence that the analysis of RT distributions has the potential to provide additional information regarding the cognitive and motor processes underlying complex behaviors (discussed in Balota and Spieler, 1999; Luce, 1986; Van Zandt, 2002). Therefore, we also analyzed the RT distributions to determine if we could obtain more insightful data for comparison with brain activation patterns during word repetition.

Analysis of the distributions of RTs revealed a very different view of the performance by middle-aged versus younger participants (Fig. 1). An ex-Gaussian function produced a significantly better fit than a Gaussian function for all four distributions (Likelihood Ratio test), indicating that all distributions were significantly skewed towards longer RTs. However, the RT distributions for the repetition of real words and pseudowords by middle-aged participants (Figs. 1a,c) were much more skewed to longer RTs than the corresponding RT distributions for repetitions by younger participants (Figs. 1b,d). This age-dependent difference was evident in a significantly larger value of the τ parameter of the ex-Gaussian function (which corresponds to a longer right tail) for middle-aged participants repeating both real words ($p < .0001$) and pseudowords ($p < .0001$) (Fig. 1h). In contrast, the mean (μ) of the Gaussian component was significantly shorter for middle-aged participants repeating real words ($p < .0001$), but not pseudowords (Fig. 1f), demonstrating an age-by-word type interaction. The μ parameter was significantly longer for middle-aged subjects repeating pseudowords, compared to real words ($p < .0001$; Fig. 1f).

The deconvolution of RT distributions into Gaussian and exponential components revealed differences between younger and middle-aged participants that were not evident in a comparison of the overall mean RTs. This discrepancy can be accounted for by the fact that the arithmetic mean of each distribution is equal to the sum of μ and τ . If two experimental groups have the same mean, but a different τ , then the change in μ must be equal in size but opposite in direction to the change in τ . Thus, the large increase in τ for middle-aged participants repeating both word types produced a corresponding decrease in μ for both word types. This accounts for the differences between the patterns of the overall mean values in Fig. 1e and the μ values in Fig. 1f. It is important to note that μ is a relatively unbiased estimator of the mean of the Gaussian component of the distribution, whereas the overall mean of the distribution can be strongly biased by the positive skew in the data.

fMRI results

The goal of the fMRI analysis was to identify regions of the brain in which the level of activity was correlated (either positively or negatively) with RT. Two primary findings from the analysis of the RT distributions, described above, led us to focus on particular aspects of the fMRI data analysis. The first finding from the analysis of the RT distribution was the interaction between age and word type for the μ parameter, in which the middle-aged participants were faster than younger participants for repeating real words, but not for repeating pseudowords. Analysis of the RT-correlated component identified by ANOVA resulted in a significant group by word type by time interaction (the two groups had significantly different shapes between the two word types) in the bilateral posterior cingulate, with a peak in the left hemisphere (Talairach = -14, 36, 36). Fig. 2a shows the area with significant RT-correlated activation, and Fig. 2b shows the time course of RT-correlated activation for each age group and each word type for that area. It is important to note here that the plots shown in Fig. 2b (and later in Fig. 3b) are not hemodynamic response functions (HRFs) in the traditional sense. They are plots of the RT-correlated component of the HRF at each time point, which is a small portion of the total HRF. The largest difference between age groups for the repetition of real words was observed at 4 s after the stimulus onset. There was a large negative correlation between RT and activation for younger subjects (*i.e.*, greater activation associated with shorter RTs), but RT and activation were not correlated in middle-aged subjects at 4 s after the stimulus onset. This large difference between age groups in the correlation between RT and activity was not observed at 4 s after the stimulus onset for the repetition of pseudowords. A negative correlation between RT and activation at early timepoints was also reported by Yarkoni et al (2009).

The second finding of interest from the analysis of RT distributions was the difference between young and middle-aged subjects with regard to the τ parameter, for which there was no significant difference between the repetition of real words and pseudowords for either group. Analysis of the RT-correlated component of the BOLD response identified two regions for which there was a group by time interaction (the two age groups differed significantly in correlation pattern across time, independent of word type), but not a group by word type by time interaction (the two groups were not significantly different in the RT-correlation pattern between the two word types). These areas for the significant two-way

interaction effects were the bilateral supplementary motor area (SMA), with a peak in the right hemisphere (Talairach = 1, -9, 56) and the corpus callosum (CC; Talairach = -1, -14, 24). Fig. 3a shows the regions with significant RT-correlated activation and Fig. 3b shows the corresponding time courses of RT-correlated activity for those regions. The overall correlation pattern across time is similar between SMA and CC. The largest difference between the time courses of the middle aged *versus* the younger participants was evident at 4 s after the stimulus onset, at which point the younger participants had a large negative correlation, while the middle-aged participants had a moderate-sized positive correlation. Thus, 4 s after the stimulus onset was the time point at which the largest correlation effects of either age or word type were observed. This is roughly the time at which a BOLD response reaches its plateau.

Discussion

Our analysis of the RT distributions for the repetition of real words and pseudowords led to the identification of two significant effects of age. First, the mean (μ) parameter of the Gaussian component of the ex-Gaussian distribution was significantly shorter for middle-aged participants repeating real words than for younger participants repeating real words, but there was no significant difference between the age groups for the repetition of pseudowords. A simple explanation for this interaction of age and word type is that it represents an effect of practice. The longer experience of middle-aged participants in producing real words shortened their response times, but this practice effect did not occur with pseudowords for which they had no experience. In contrast, we did not observe an effect of age on the repetition of real words when we compared mean RTs, which are biased towards longer values in middle-aged participants by the large positive skew in the distributions.

One might predict that because of increased expertise over more years of speaking, the older participants could be faster at producing both real words *and* pseudowords. However, they were faster (in terms of the μ parameter of the RT distribution) only at producing the real words. One reason for this might be the nature of the pseudowords. There are a variety of methods for creating pseudowords. In the past, we started with Turkish words, but produced them using American English consonants and vowels (Shuster, 2009). As in the current study, those Turkish-derived pseudowords were matched to the real American English words with regard to phonotactic probability and initial syllable frequency of occurrence. In the current study, we rearranged the syllables of the four syllable real American English words to create the pseudowords, again matching for phonotactic probability and initial syllable frequency of occurrence. An unanticipated consequence of this method was that some of the participants (both middle-aged and younger) occasionally did recognize some of the rearranged pseudowords, even before they had heard the real word (*e.g.*, 'gayluhadder,'/geɪl æ /for 'alligator'). Therefore, expertise may have hurt the older participants for these particular pseudowords, because it may have taken more effort for them to inhibit the real word counterpart, if it came to mind after hearing the stimulus, than it did for the younger speakers with fewer years of speaking experience.

A second effect of age was the greater extent to which the distribution of RTs for repeating both real words and pseudowords was skewed to longer RTs (increased τ in the ex-Gaussian distribution) in the middle-aged participants than in the young participants. An increase in τ in the distributions of RTs has been observed with aging for other behavioral tasks (McAuley et al., 2006; Spieler et al., 1996).

We cannot attribute these effects of age and word type on μ and τ to changes in specific neural processes based on behavioral data alone. The ex-Gaussian function is often viewed as a model-independent function (e.g., Luce, 1986; Van Zandt, 2002), and the μ , σ , and τ parameters are not linked by theory to specific underlying neural processes. Some investigators have proposed a general interpretation in which the Gaussian component (defined by μ and σ) is thought to reflect a more automatic (non-analytic) component (with added noise) of a response, whereas the exponential component (defined by τ) might reflect a less automatic (analytic) cognitive processing component (e.g., working memory, attention) (Balota and Spieler, 1999; Heathcote et al., 1991; Hohle, 1965; Luce, 1986). This interpretation is based on evidence from a large variety of behavioral tasks designed to examine these putative relationships. However, discrepant results in some studies (e.g., Heathcote et al., 1991) have demonstrated that this simple correspondence between the components of the ex-Gaussian function and motor *versus* cognitive processes is not universal.

Although we cannot directly attribute the effects of age and word type on μ and τ to changes in specific neural processes, we can use these effects to guide our analysis of the fMRI data and potentially identify regions of the brain in which changes in activity are correlated with age- and word-type-dependent changes in RT. The RT-correlated component of the BOLD response is very small (a few percent signal change per second) compared to the RT-independent component, so we used parametric modulation analysis to separate the RT-correlated component of the BOLD responses, and it is within this RT-correlated component that we tested for effects of age and word type. This combined approach for analyzing RT distributions and fMRI data was, therefore, model independent, and it made no *a priori* assumptions about how μ and τ might be related to either motor function or cognitive processing.

Analysis of the fMRI data revealed that a significant interaction between age and word type was evident in the RT-correlated activity in the bilateral posterior cingulate. A recent review paper by Leech and Sharp (2013) reveals that the posterior cingulate plays an important role in cognition. Although its precise functions are not universally agreed upon, it has been proposed that it plays a role in arousal and awareness, as well as in attention (both internally and externally directed). Several studies have shown that normal aging brings about changes in both the connectivity and function of this region. However, the participants in these studies have been older than those in our study, and some studies have employed tasks, such as the n-back task, that would not tap into the expertise of the older participants (e.g., Prakash et al., 2012; Sambataro et al., 2010). Leech and Sharp proposed a model of posterior cingulate functioning and suggest that it is involved in the retrieval of episodic and semantic memories. Studies have shown that the posterior cingulate is involved in memory during communication (e.g., Awad et al., 2007; de Zubicaray et al., 2000). It is not surprising that

the memory network would be involved in the comprehension and production of speech, because communication is not learned and used in a vacuum, but in the context of everyday experience and, therefore, it is entwined in our episodic memories. Age-related changes with regard to the speed with which one can access memory stores could be either positive or negative. Years of practice in talking might lead to quicker access to these representations for the middle-aged participants, especially for such an easy task as repeating real words. Thus, the observation of a significant interaction between age and word type in both the μ parameter of the distribution of RTs, from which we inferred a practice effect (discussed above), and in the activity in the posterior cingulate region is intriguing. Yarkoni et al. (2009) predicted that “there should be a *negative* correlation between RT and the BOLD response in regions associated with deployment of task-related resources.” To the extent that activity in the posterior cingulate is required for accessing memory resources during speech, it is plausible that the more effortful requirements for younger, less-experienced speakers could produce a negative correlation between RT and the BOLD response, as predicted by Yarkoni et al., but the negative correlation would not be observed in mature adults who are more practiced in speaking and more efficient in accessing memory resources.

Analysis of the fMRI data also revealed a significant main effect of age on RT-correlated activity in the SMA and corpus callosum. The SMA plays a role in movement, including movement during speech, although its precise role is still not agreed upon (*e.g.*, Alario et al., 2006; Brown et al., 2009; Indefrey and Levelt, 2004). Activation in the SMA is almost always reported in functional imaging studies of speech production (*e.g.*, Bohland and Guenther, 2006; Riecker et al., 2008; Shuster, 2009; Shuster and Lemieux, 2005). Moreover, lesions to either the right or left SMA result in various problems with speech production, such as an inability to initiate spontaneous speech (Krainik et al., 2004); thus, it is not surprising that activity in SMA would be correlated with RT. The pattern of RT correlation shown in Fig. 3b reveals a negative correlation between the BOLD and RT at 4 s for the Y participants, and a similar negative correlation between BOLD and RT at 8 s for the MA participants. This suggests that the MA participants were slower to deploy motor resources to the task than were the Y participants. This is consistent with previous findings suggesting that psychomotor slowing occurs with age. A role of the SMA in the initiation of speech is particularly interesting because age produced a significant increase in the τ parameter of the distribution of RTs, in addition to affecting activity in the SMA. The increase in τ was identical for the repetition of real words and pseudowords, indicating that the neural mechanism underlying the increase in τ with age is likely to be common to the repetition of both word types. An important feature of this neural mechanism is that it is associated with delays in the onset of word repetition that are exponentially distributed (as evident in the positive skew of the RTs), rather than delays that are normally distributed around a mean response time. This condition could be met easily if the exponential distribution of response latencies results from time-invariant random variability in the latency to the initiation of speech (*i.e.*, the response latencies correspond to waiting times in a Markov process). Further study will be required to determine if age-dependent changes in the influence of the SMA on speech initiation could produce an increase in response latencies that are exponentially distributed.

Our finding of RT-correlated activity in the CC is consistent with a report by Yarkoni et al. (2009), who found that both gray and white matter showed RT-correlated variability. Reliable activation in white matter (including corpus callosum) has been demonstrated in a variety of additional studies (Chow et al., 2007; Fraser et al., 2012; Gawryluk et al., 2009; Mazerolle et al., 2008, 2010; Tettamanti et al., 2002; Weber et al., 2005). The area of activation in our study, in the posterior body/isthmus of the CC, is in a similar location to that reported by Mazerolle and colleagues. Using electroencephalography, Wohlert (1993) investigated the readiness potential (RP) generated prior to two oral motor movement tasks and one speech task. She found that the RP prior to the speech task (which required the greatest number of sequenced muscle elements) was the largest. Moreover, of the three recording sites, right side, left side and vertex, the RP was largest over the vertex. She suggested that this reflected the activation of the bilateral SMA, most likely as a result of extensive transcallosal connections. Thus, the pattern of RT-correlated brain activation that we observed could reflect a portion of the speech network involving the bilateral SMA and its reciprocal connections. However, there are other possible reasons for observing “activation” in white matter, including smoothing with gray matter and pulsatile artifacts. Therefore, we must be cautious in interpreting our finding of activation in CC.

It is unlikely that the patterns of brain activation we observed in the study were due to stimulus-correlated movement. First, it was difficult for participants to move their heads because the ear muffs fit so snugly into the head coil. Second, we placed a piece of tape across their heads to help them monitor for head movement. Third, we used the output from the motion correction as regressors of no interest in the regression analysis. Fourth, the time points during which the speaking occurred were not included in the regression analysis. Fifth, examination of the movement parameters that are output from the AFNI motion correction program (3dvolreg) did not reveal excessive head movement (see supplementary information). Sixth, we used a spiral in/out pulse sequence, which has been shown to be less sensitive to motion than other pulse sequences (*e.g.*, Glover and Lee, 1995). Lastly, at the group level, we compared the analysis of the data from the six effect estimates per word type that characterized the response shape associated with average RT (the effect estimates that were not RT correlated) with the effect estimates from the non-speech control condition. This analysis revealed significant activation in the speech network in the speaking as compared to the non-speaking condition, showing that our experimental paradigm was robust. These data and a graph of the movement parameters are presented in the supplementary information.

We did not observe a slowing in the average speed to repeat real words by the middle-aged participants as compared to the young adults, which is inconsistent with the findings of previous investigations that revealed motor slowing with age (Fozo and Watson, 1998; Mattay et al., 2002; Rodríguez-Aranda and Jakobsen, 2011). One reason might be that our middle-aged participants were younger. Another reason might be differences in the experimental tasks that were used. For example, Fozo and Watson used a light cue to prompt participants to respond with the same word or short sentence. In our task, auditory presentation of the word to be repeated may have primed the response, leading to slightly faster average RTs for the middle-aged speakers for the real words.

In our description of the fMRI data, we noted that 4 s after the stimulus onset was the time point at which the largest effects of either age or word type were observed. We acknowledge that at the temporal resolution of fMRI, we cannot clearly determine which component of the response can be attributed to the perception of the stimulus and which can be attributed to the repetition because these events occur on a scale of milliseconds. It is well-established in the speech perception literature that listeners frequently can recognize words before they hear the entire word, and some of the participants did initiate responses prior to the end of the stimulus.

Conclusions

We have demonstrated that our novel approach of using information obtained from the analysis of RT distributions to guide the analysis and interpretation of RT-correlated variability in patterns of brain activation has great potential. This approach has allowed us to associate specific mechanisms that underlie speech (memory and motor processing) with activity in different brain regions. In particular, our approach has provided an opportunity to hypothesize two relationships relevant to word repetition: (1) an association between RT-correlated activity in the posterior cingulate gyrus and a practice effect for the repetition of real words, and (2) an age-dependent increase in the variability of RT (an increase in τ) related to RT-correlated activity in the SMA. Rigorous testing of these hypotheses will require further study with an increased number of stimuli and different manipulations of the repetition task, with the potential of providing greater insight into the roles of different brain regions in specific aspects of speaking. Finally, our results demonstrate that the common practice in fMRI analysis of using RT simply as a regressor of no interest or even a nuisance variable could lead to the loss of valuable insight that can be otherwise gained by the alternate approach of exploring RT-correlated activity in the brain.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research reported in this publication was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under grant number R15DC011136 and by the Department of Radiology at West Virginia University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Alario F-X, Chainay H, Lehericy S, Cohen L. The role of supplementary motor area (SMA) in word production. *Brain Res.* 2006; 1076:129–143. [PubMed: 16480694]
- Awad M, Warren JE, Scott SK, Turkheimer FE, Wise RJ. A common system for the comprehension and production of narrative speech. *J. Neurosci.* 2007; 27:11455–11464. [PubMed: 17959788]
- Babel M. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *J. Phon.* 2012; 40:177–189.
- Balota DA, Spieler DH. Word frequency, repetition, and lexicality effects in word recognition tasks: beyond measures of central tendency. *J. Exp. Psychol. Gen.* 1999; 128:32–55. [PubMed: 10100390]

- Bohland JW, Guenther FH. An fMRI investigation of syllable sequence production. *NeuroImage*. 2006; 32:821–841. [PubMed: 16730195]
- Brown S, Laird AR, Pfordresher PQ, Thelen SM, Turkeltaub P, Liotti M. The somatotopy of speech: phonation and articulation in the human motor cortex. *Brain Cogn*. 2009; 70:31–41. [PubMed: 19162389]
- Chen G, Saad ZS, Britton JC, Pine DS, Cox RW. Linear mixed-effects modeling approach to fMRI group analysis. *NeuroImage*. 2013; 73:176–190. [PubMed: 23376789]
- Chow LS, Cook GG, Whitby E, Paley MN. Investigation of axonal magnetic fields in the human corpus callosum using visual stimulation based on MR signal modulation. *J. Magn. Reson. Imaging*. 2007; 26:265–273. [PubMed: 17654726]
- Cox R. Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res*. 1996; 29:162–173. [PubMed: 8812068]
- de Zubicaray GI, Zelaya FO, Andrew C, Williams SCR, Bullmore ET. Cerebral regions associated with verbal response initiation, suppression and strategy use. *Neuropsychologia*. 2000; 38:1292–1304. [PubMed: 10865105]
- Efron B. Computers and the theory of statistics: thinking the unthinkable. *SIAM Rev*. 1979; 21:460–480.
- Fowler CA, Brown JM, Sabadini L, Weihing J. Rapid access to speech gestures in perception: evidence from choice and simple response time tasks. *J. Mem. Lang*. 2003; 49(3):396–413. [PubMed: 20622982]
- Fozo MS, Watson BC. Task complexity effect on vocal reaction time in aged speakers. *J. Voice*. 1998; 12(4):404–414. [PubMed: 9988027]
- Fraser LM, Stevens MT, Beyea SD, D'Arcy RCN. White versus gray matter: fMRI hemodynamic responses show similar characteristics, but differ in peak amplitude. *BMC Neurosci*. 2012; 13:91. [PubMed: 22852798]
- Garrett, A.; Johnson, K. Phonetic bias in sound change. In: Yu, ACL., editor. *Origins of Sound Change: Approaches to Phonologization*. Oxford University Press; Oxford: 2013.
- Gawryluk JG, Brewer KD, Beyea SD, D'Arcy RCN. Optimizing the detection of white matter fMRI using asymmetric spin echo spiral. *NeuroImage*. 2009; 45:83–88. [PubMed: 19084071]
- Glover GH, Law CS. Spiral in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. *Magn. Res. Med*. 2001; 46:515–522.
- Glover GH, Lee AT. Motion artifacts in fMRI: comparison of 2DFT with PR and spiral scan methods. *Magn. Reson. Med*. 1995; 33:624–635. [PubMed: 7596266]
- Heathcote A, Popiel SJ, Mewhort DJ. Analysis of response time distributions: an example using the Stroop task. *Psychol. Bull*. 1991; 109(2):340–347.
- Heathcote A, Brown S, Cousineau D. QMPE: estimating Lognormal, Wald, and Weibull RT distributions with a parameter-dependent lower bound. *Behav. Res. Methods Instrum. Comput*. 2004; 36(2):277–290. [PubMed: 15354694]
- Hickok G. The functional neuroanatomy of language. *Phys. Life Rev*. 2009; 6:121–143. [PubMed: 20161054]
- Hohle RH. Inferred components of reaction times as functions of foreperiod duration. *J. Exp. Psychol*. 1965; 69(4):382–386. [PubMed: 14286308]
- Indefrey P, Levelt WJM. The spatial and temporal signatures of word production components. *Cognition*. 2004; 92:101–144. [PubMed: 15037128]
- Krainik A, Duffau H, Capelle L, Cornu P, Boch AL, Mangin JF, Le Bihan D, Marsault C, Chiras J, Lehericy S. Role of the healthy hemisphere in recovery after resection of the supplementary motor area. *Neurology*. 2004; 62(8):1323–1332. [PubMed: 15111669]
- Lacouture Y, Cousineau D. How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times. *Tutor. Quant. Methods Psychol*. 2008; 4(1):35–45.
- Leech, R.; Sharp, DJ. The role of the posterior cingulate cortex in health and disease. *Brain Advance Access*. 2013. <http://dx.doi.org/10.1093/brain/awt162> (published July 18, 2013)
- Luce, RD. *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press; New York, NY: 1986.

- Mattay VS, Fera F, Tessitore A, Hariri AR, Das S, Callicott JS, Weinberger DR. Neurophysiological correlates of age-related changes in human motor function. *Neurology*. 2002; 58(4):630–635. [PubMed: 11865144]
- Mazerolle EL, D'Arcy RCN, Beyea SD. Detecting fMRI activation in white matter: interhemispheric transfer across the corpus callosum. *BMC Neurosci*. 2008; 9:84. [PubMed: 18789154]
- Mazerolle EL, Beyea SD, Gawryluk JR, Brewer KD, Bowen CV, D'Arcy RCN. Confirming white matter fMRI activation in the corpus callosum: co-localization with DTI tractography. *NeuroImage*. 2010; 50:616–621. [PubMed: 20053383]
- McAuley T, Yap M, Christ SE, White DA. Revisiting inhibitory control across the life span: insights from the ex-Gaussian distribution. *Dev. Neuropsychol*. 2006; 29(3):447–458. [PubMed: 16671861]
- Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*. 1971; 9:97–113. [PubMed: 5146491]
- Pardo JS. On phonetic convergence during conversational interaction. *J. Acous. Soc. Am*. 2006; 119:2382–2393.
- Prakash RS, Heo S, Voss MW, Patterson B, Kramer AF. Age-related differences in cortical recruitment and suppression: implications for cognitive performance. *Behav. Brain Res*. 2012; 230:192–200. [PubMed: 22348896]
- Riecker A, Brendel B, Ziegler W, Erb M, Ackermann H. The influence of syllable onset complexity and syllable frequency on speech motor control. *Brain Lang*. 2008; 107(2):102–113. [PubMed: 18294683]
- Rodríguez-Aranda C, Jakobsen M. Differential contribution of cognitive and psychomotor functions to the age-related slowing of speech production. *J. Int. Neuropsychol. Soc*. 2011; 17(5):807–821. [PubMed: 21729398]
- Sambataro F, Murty VP, Callicott JH, Tan HY, Das S, Weinberger DR, et al. Age-related alterations in default mode network: impact on working memory performance. *Neurobiol. Aging*. 2010; 31:839–852. [PubMed: 18674847]
- Shuster LI. The effect of sublexical and lexical frequency on speech production: an fMRI investigation. *Brain Lang*. 2009; 111:66–72. [PubMed: 19616296]
- Shuster LI, Lemieux SK. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain Lang*. 2005; 93:20–31. [PubMed: 15766765]
- Soros P, Bose A, Sokoloff L, Graham SJ, Stuss DT. Age-related changes in the functional neuroanatomy of overt speech production. *Neurobiol. Aging*. 2011; 32:1505–1513. [PubMed: 19782435]
- Spieler DH, Balota DA, Faust ME. Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *J. Exp. Psychol. Hum*. 1996; 22(2):461–479.
- Tettamanti M, Paulesu E, Scifo P, Maravita A, Fazio F, Perani D, Marzi CA. Interhemispheric transmission of visuomotor information in humans: fMRI evidence. *J. Neurophys*. 2002; 88:1051–1058.
- Vitevitch MS, Luce PA. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *J. Mem. Lang*. 1999; 40:374–408.
- Van Zandt T. How to fit a response time distribution. *Psychon. Bull. Rev*. 2000; 7(3):424–465. [PubMed: 11082851]
- Van Zandt, T. Analysis of response time distributions. In: Pashler, H.; Wixted, J., editors. *Steven's Handbook of Experimental Psychology*. 3rd. Vol. 4. John Wiley & Sons, Inc.; New York, NY: 2002. p. 461–516.
- Weber B, Treyer V, Oberholzer N, Jaermann T, Boesiger P, Brugger P, Regard M, Buck A, Savazzi S, Marzi CA. Attention and interhemispheric transfer: a behavioral and fMRI study. *J. Cogn. Neurosci*. 2005; 17(1):113–123. [PubMed: 15701243]
- Wohlert AB. Event-related brain potentials preceding speech and non-speech oral movements of varying complexity. *J. Speech Hear. Res*. 1993; 36:897–905. [PubMed: 8246478]
- Yarkoni T, Barch DM, Gray JR, Conturo TE, Braver TS. BOLD correlates of trial-by-trial reaction time variability in gray and white matter: a multi-study fMRI analysis. *PLoS ONE*. 2009; 4(1):e4257–e4257. [PubMed: 19165335]

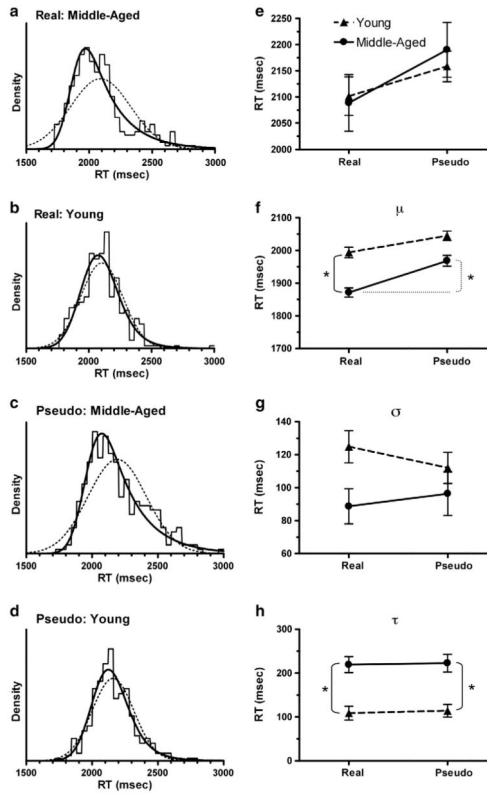


Fig. 1.

Distributions of RTs grouped by age and type of word. (a, b, c, d). The histograms were calculated with a bin width of 40 ms. The solid curve is the best fit of an ex-Gauss function, and the dotted curve is the best fit of a Gaussian function. (e) Mean RTs calculated by age and type of word. (f, g, h). Best-fit estimates of the parameters μ , σ , and τ for the ex-Gauss function fitted to the RT distributions shown in panels a–d. Each error bar is an estimate of the standard error of the parameter value, calculated from the Hessian matrix in QMPE/CML. Significant differences (Likelihood Ratio test, $p < .0001$) are marked by a bracket and asterisk.

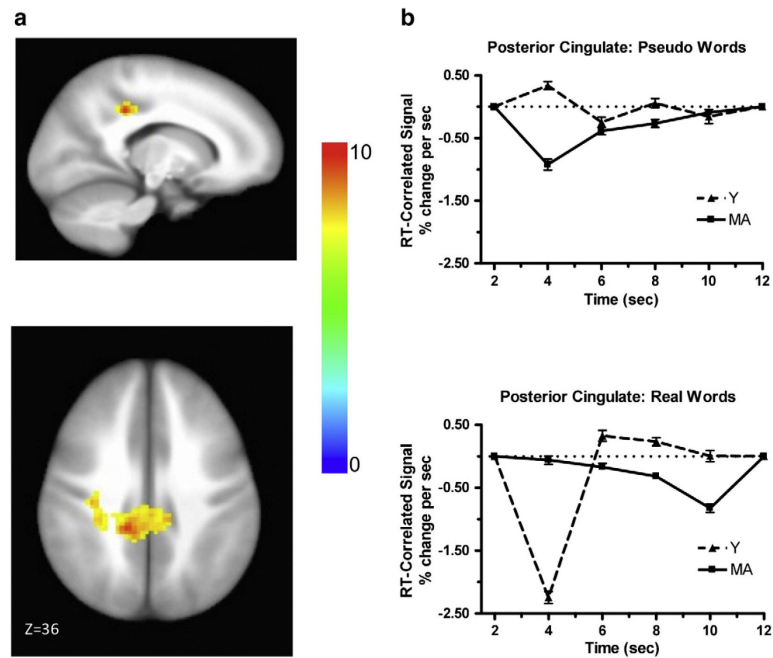


Fig. 2. Brain activation associated with word type by age by time interaction. (a) Activation in posterior cingulate. The color bar codes for the F values. (b) Estimated time courses of the RT-correlated signal for the repetition of pseudowords (upper panel) and real words (lower panel) for each age group.

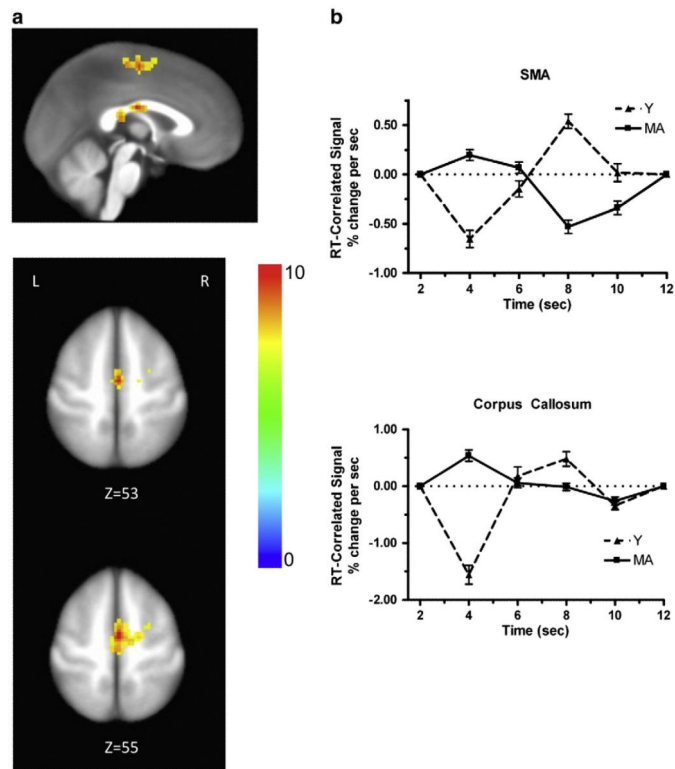


Fig. 3. Brain activation associated with age by time interaction. (a) Activation in SMA and corpus callosum. The color bar codes for the F values. (b) Estimated time courses of the RT-correlated signal for each age group for SMA (upper panel) and corpus callosum (lower panel). The responses to real words and pseudowords were combined within each age group for this analysis.