



Published in final edited form as:

*Neuroimage*. 2022 June ; 253: 119042. doi:10.1016/j.neuroimage.2022.119042.

## Evidence for cortical adjustments to perceptual decision criteria during word recognition in noise

Kenneth I. Vaden Jr<sup>a,\*</sup>, Susan Teubner-Rhodes<sup>a,b</sup>, Jayne B. Ahlstrom<sup>a</sup>, Judy R. Dubno<sup>a</sup>, Mark A. Eckert<sup>a</sup>

<sup>a</sup>Hearing Research Program, Department of Otolaryngology, Head and Neck Surgery, Medical University of South Carolina, 135 Rutledge Ave. MSC 550, Charleston, SC 29455-5500, United States

<sup>b</sup>Department of Psychological Sciences, 226 Thach Hall, Auburn University, AL 36849-9027

### Abstract

Extensive increases in cingulo-opercular frontal activity are typically observed during speech recognition in noise tasks. This elevated activity has been linked to a word recognition benefit on the next trial, termed “adaptive control,” but how this effect might be implemented has been unclear. The established link between perceptual decision making and cingulo-opercular function may provide an explanation for how those regions benefit subsequent word recognition. In this case, processes that support recognition such as raising or lowering the decision criteria for more accurate or faster recognition may be adjusted to optimize performance on the next trial. The current neuroimaging study tested the hypothesis that pre-stimulus cingulo-opercular activity reflects criterion adjustments that determine how much information to collect for word recognition on subsequent trials. Participants included middle-age and older adults ( $N = 30$ ; age =  $58.3 \pm 8.8$  years;  $m \pm sd$ ) with normal hearing or mild sensorineural hearing loss. During a sparse fMRI experiment, words were presented in multitalker babble at +3 dB or +10 dB signal-to-noise ratio (SNR), which participants were instructed to repeat aloud. Word recognition was significantly poorer with increasing participant age and lower SNR compared to higher SNR conditions. A perceptual decision-making model was used to characterize processing differences based on task response latency distributions. The model showed that significantly less sensory evidence was collected (i.e., lower criteria) for lower compared to higher SNR trials. Replicating earlier observations, pre-stimulus cingulo-opercular activity was significantly predictive of correct recognition on a subsequent trial. Individual differences showed that participants with higher criteria also benefitted the most from pre-stimulus activity. Moreover, trial-level criteria changes were significantly linked to higher versus lower pre-stimulus activity. These results suggest

---

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

\*Corresponding author. vaden@musc.edu (K.I. Vaden Jr).

Author contributions

Each of the authors had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. *Conceptualization*: KIV, MAE, JRD, STR; *Methodology*: KIV, JBA, MAE, JRD; *Investigation*: KIV; *Formal Analysis*: KIV; *Resources*: KIV, MAE; *Writing - Original Draft*: KIV; *Writing - Review: & Editing*: KIV, STR, MAE, JBA, JRD; *Visualization*: KIV; *Supervision*: KIV, MAE, JRD; *Funding Acquisition*: KIV, MAE, JRD.

Declaration of Competing Interest

The authors have no conflict of interest to declare.

cingulo-opercular cortex contributes to criteria adjustments to optimize speech recognition task performance.

## Keywords

Adaptive control; Frontal cortex; Speech recognition in noise; Perceptual decision-making; Decision criteria; fMRI

---

## 1. Introduction

Speech recognition in noise is typically poorer for older adults than younger adults, which can reflect a combination of sensory and cognitive-perceptual declines (e.g., Dubno et al., 1984; Humes, 2021; Humes and Christopherson, 1991; Humes and Dubno, 2010; Souza et al., 2007; Wingfield and Tun, 2001). Considerable variability exists, with some older adults performing nearly as well as younger adults. Some of this variance may be explained by differences in the engagement of listening effort or cognitive resources to facilitate speech recognition for some older adults (e.g., Alain et al., 2004; Eckert et al., 2016; Peelle, 2017; Pichora-Fuller et al., 2016; Wingfield, 1996). Neuroimaging evidence suggests that a set of frontal cortex regions are extensively recruited with increasing listening difficulty, but it is unclear how frontal cortex might facilitate speech recognition.

Speech recognition in noise performance is consistently linked to activity in cingulo-opercular regions of frontal cortex in what appears to be a performance monitoring role. Specifically, cingulo-opercular activity typically increases with increasing speech recognition errors, as well as with increasing listening difficulty (Alavash et al., 2019; Eckert et al., 2009; Peelle, 2017; Wild et al., 2012). The robust effects of task errors and difficulty are seen across many task domains, which is suggestive of a domain-general performance monitoring function (Dosenbach et al., 2006) that activates similarly during speech recognition (with or without explicit task feedback). More extensive frontal cortex activity during challenging listening tasks has been shown for older compared to younger adults (Erb and Obleser, 2013; Wingfield and Grossman, 2006), which could reflect task difficulty (Eckert et al., 2016; Erb and Obleser, 2013; Pichora-Fuller et al., 2016). Task performance and difficulty-related increases in cingulo-opercular activity are also believed to lead to changes that can improve performance (Eckert et al., 2016; Erb et al., 2013; Gratton et al., 2016; Vaden et al., 2013), as described next.

In addition to signaling task difficulty or error, cingulo-opercular activity also appears to optimize on-going task performance. For example, higher activity prior to a stimulus presentation (i.e., pre-stimulus activity) is associated with significantly better word recognition in noise (Vaden et al., 2016; Vaden et al., 2015, 2013). That is, word recognition performance is significantly predicted by pre-stimulus cingulo-opercular activity. The performance changes associated with pre-stimulus or poststimulus cingulo-opercular activity suggest these regions are 1) responsive to task performance or uncertainty information and 2) can adjust attention or other resources (Coste et al., 2011; Coste and Kleinschmidt, 2016; Gratton et al., 2016) to optimize performance in challenging tasks as part of an *adaptive control function* (Eckert et al., 2016; Shenhav et al., 2013). Similar effects of pre-stimulus

activity on performance in auditory detection tasks (Coste and Kleinschmidt, 2016; Vaden et al., 2020) suggest that cingulo-opercular effects are not specific to understanding speech in noise. However, the mechanism(s) by which cingulo-opercular cortex provides adaptive control have not been identified.

A growing body of literature indicates that medial frontal cortex can affect response latencies or the decision to initiate a response. A mechanism for speeding or slowing responses may be critical when perception of stimuli is diminished, as in speech recognition in noise, where noisy or ambiguous sensory information interfere with typically rapid perceptual mapping. This framework for understanding the interplay between perception and performance is well-characterized by perceptual decision-making models that focus on response latencies to identify distinct processes (Dutilh et al., 2019; Heathcote and Love, 2012; Matzke and Wagenmakers, 2009), which have been mapped to cortical function, as described later. While there is some variation across models, they all consider: 1) response criteria or response caution where higher criteria reflect slower and more cautious responding; 2) accumulation rate or how quickly information can be perceived where a higher rate reflects faster information processing; and 3) non-decision time or initial sensory function and motor function to produce a response (Anders et al., 2016; Brown and Heathcote, 2005; Donkin et al., 2010; Smith and Ratcliff, 2004). Here, cingulo-opercular function is hypothesized to influence decision criteria through its role in performance monitoring where lower performance would signal for higher criteria when accuracy is important to performance (i.e., as opposed to speeded responding).

Perceptual decision-making provides a promising avenue for investigating speech recognition in noise, because the signal-to-noise ratio (SNR) can influence the speed at which listeners understand words (Meister et al., 2018). Perceptual decision-making models commonly assume that the rate of evidence collection largely depends on stimulus quality (Donkin et al., 2009), which is why recognition would be expected to take longer for words presented in a low compared to high SNR condition. In contrast, decision criteria may be adjusted to optimize the speed and accuracy of task responses. Criteria are raised or lowered to collect more or less evidence during recognition, which affects how quickly recognition can occur in combination with the speed of information collection. Raised criteria require more evidence to be collected prior to recognition (e.g., higher response caution), which may increase accuracy while resulting in slower, more variable response latencies. Although faster responses are strongly associated with lower accuracy (Wickelgren, 1977), lower decision criteria can facilitate faster recognition decisions when there is time pressure (e.g., speed-accuracy tradeoff tasks; Bogacz et al., 2010; Evans et al., 2017; Rae et al., 2014; Van Maanen et al., 2011). Thus, more difficult perceptual conditions with slower information collection may result in lower decision criteria to hasten recognition, or alternatively may result in higher decision criteria to increase accuracy.

The current study used a perceptual decision-making model that derived parameter estimates from task response latencies (speech onsets), using an approach from Anders et al. (2016). The Shifted Wald model combines the computational simplicity and estimate robustness of “measurement models” (e.g., Weibull, Weibull, 1951; gamma, Lukacs, 1955) with the cognitive specificity offered by “process models” of perceptual decision-making (e.g.,

Linear Ballistic Accumulator, Brown and Heathcote, 2008; Drift Diffusion Model, Ratcliff, 1978). A probability distribution function is first fitted to each participant's onset latencies (as with measurement models), then estimates are derived for decision criteria, evidence accumulation, and non-decision time (as with process models; Anders et al., 2016, Anders et al., 2015). Anders et al. (2016) illustrate how decision criteria and evidence accumulation rates are informed by the median and skew of response latencies from trials with correct responses, while non-decision times primarily reflect the minimum response latencies. Longer non-decision times, faster evidence accumulation rates, and higher criteria are reflected in larger number values with the Shifted Wald model (Anders et al., 2016).

Our predictions for perceptual decision-making processes were based on the extant literature showing that both non-decision times and evidence accumulation rates reflect stimulus quality, while decision criteria balance response speed and accuracy in perceptual task performance. Consistent with perceptual decision-making changes in experiments in challenging recognition tasks (e.g., Anders et al., 2016; Hanks and Summerfield, 2017; Mulder et al., 2014), less intelligible words in the +3 dB SNR were predicted to result in slower evidence accumulation rates and longer non-decision times compared to the +10 dB SNR. Lower decision criteria were also expected in the +3 dB SNR condition: when evidence accumulates more slowly, participants may lower their threshold and respond earlier to maintain regular response speed or prevent missing the response interval. Alternatively, if participants prioritize accurate recognition, they may raise criteria in the +3 dB relative to the +10 dB SNR condition to collect more evidence.

Breaking the decision criteria predictions down further, perceptual decision-making models have shown that criteria changes can account for a speed accuracy trade-off (Dutilh et al., 2019). That is, faster responses are often less accurate and slower responses are often more accurate (Wickelgren, 1977). More difficult conditions can lead to compensatory changes in criteria that inversely affect both response latency and accuracy (e.g., Mulder et al., 2014; Hanks and Summerfield, 2017). Poorer SNR conditions were predicted to slow evidence accumulation rates and non-decision times during word recognition. We predicted that time pressure related to slower processes in a poorer SNR could lead to lowered criteria to offset slower recognition. Because the amount of quality sensory information is limited by noise, lowered criteria in poor SNR conditions also might not affect how much evidence accumulates compared to better SNR conditions. An alternative perceptual strategy to optimize accuracy in a poorer SNR condition would raise criteria to gather even more evidence, albeit in exchange for much slower recognition times from the combination of longer non-decision times and prolonged evidence accumulation to reach the raised threshold at a slower rate.

Perceptual decision-making processes have been associated with activity in distinct large-scale cortical networks or systems based on functional neuroimaging studies (Gratton et al., 2016; Hanks and Summerfield, 2017). Evidence accumulation occurs with a build-up of activity within a fronto-parietal network that ceases when object recognition occurs, in contrast with primary sensory cortex activity that remains elevated throughout stimulus presentations (Brosnan et al., 2020; Gratton et al., 2016; Ploran et al., 2011, 2007; c.f., Rowe et al., 2010). More cautious or higher decision criteria occur with increased activity

in cingulo-opercular regions during visual judgement tasks (Cavanagh et al., 2011; Rahnev et al., 2016; Van Maanen et al., 2011). The involvement of cingulo-opercular regions in post-perceptual processes, especially decision criterion adjustments (Desender et al., 2019), is consistent with increased activity during speech recognition in noise tasks (Alain et al., 2018; Eckert et al., 2016), as well as its effects on subsequent task performance (Vaden et al., 2013). We propose that cingulo-opercular activity could also reflect its involvement in decision criterion adjustments to optimize speech recognition. That is, higher pre-stimulus cingulo-opercular activity is predicted to occur with higher decision criteria.

Functional neuroimaging data and perceptual decision-making estimates were used to test the hypothesis that cingulo-opercular cortex adjusts decision criteria, as reflected by pre-stimulus activity, which controls how much evidence is collected for speech recognition in noise. The study was designed to first replicate earlier findings that pre-stimulus cingulo-opercular BOLD contrast (i.e., prior to each presentation) is associated with correct word recognition (Vaden et al., 2015, 2013). We then tested the predictions that 1) larger cingulo-opercular adaptive control effects during word recognition are associated with more cautious decision criteria across listeners (i.e., higher criteria) and 2) elevated pre-stimulus cingulo-opercular activity occurs with trial-level shifts to more cautious decision criteria within listeners, thereby providing evidence for a cingulo-opercular adaptive control function. We characterized perceptual decision-making for speech in noise with middle-aged and older adult participants, to determine the extent to which age-related differences in perceptual decision-making (Dully et al., 2018) could account for speech recognition difficulty (Dubno et al., 1984; Humes and Christopherson, 1991; Johns et al., 2018; Souza et al., 2007).

## 2. Materials and methods

### 2.1. Participants

The current study included 30 middle-aged and older adult participants (22 females, 8 males; age =  $58.3 \pm 8.8$  years;  $M \pm SD$ ), after excluding six participants based on hearing loss screening criteria, one participant based on a technical problem that affected stimuli presentation at the MRI scanner, and another participant based on an incidental radiologic finding. The hearing loss exclusion criteria were designed to minimize audibility differences, based on a modified Articulation Index (Dubno et al., 2008) that predicted equal word recognition scores for participants with pure-tone thresholds below the upper limits (250 Hz to 8000 Hz; Fig. 1) with the stimulus presentation parameters used for the experiment. Pure-tone thresholds were measured for both ears at conventional audiometric frequencies using a Madsen OB922 audiometer and TDH-39 headphones (American National Standards Institute, 2010). Mean pure-tone thresholds from the better ear (250 to 8000 Hz) ranged from 0.6 to 34.4 dB HL ( $M = 12.4 \pm 8.8$ ; Fig. 1) and were significantly correlated with age ( $r_{28} = 0.49$ ,  $p = 0.006$ ).

Participants had to score at least 27 out of 30 on the Mini-Mental Status Examination for inclusion in this study to limit the influence of mild cognitive impairment (Folstein et al., 1983) on the results. All participants were native English speakers, completed between 12 and 20 years of education ( $M = 15.8 \pm 1.8$ ), and exhibited a right-handed distribution of Edinburgh Handedness Inventory scores ( $M = 63.2 \pm 65.3$ ; -100 left-handed and +100

right-handed; Oldfield, 1971). There was no history of neurological or psychiatric events, based on self-report. Informed consent was obtained in compliance with the Institutional Review Board at the Medical University of South Carolina (MUSC), and the experiment was conducted in accordance with the Declaration of Helsinki.

## 2.2. Experiment design

**2.2.1. Stimuli**—During the fMRI experiment, participants were presented with 120 monosyllabic consonant-vowel-consonant words that were recorded by a male speaker (Dirks et al., 2001). Words were presented binaurally in a pseudorandomized order that was fixed across participants. Each word was presented only once during the experiment. Words were presented at 85 or 92 dB SPL in a continuous multitalker babble recording at 82 dB SPL (Kalikow et al., 1977), resulting in +3 or +10 dB SNR. Speech and babble levels were selected to minimize the effects of audibility differences on task performance based on normal hearing to mild hearing loss. None of the subjects reported the levels were uncomfortably loud. The auditory stimuli were presented with MRI-compatible Sensimetrics piezoelectric insert earphones and a sound level meter in the scanner control room was used to calibrate presentation levels.

## 2.3. Procedure

Participants listened to a single word at +3 or +10 dB SNR in multitalker babble, and were instructed to repeat the word aloud or say “nope” if they did not understand the word. A spin echo (EPI) sequence was used to collect BOLD imaging data, with a sparse 8.6 s acquisition that allowed for each word to be presented with limited scanner noise (Vaden et al., 2020; Vaden et al., 2015, 2013). The response interval was visually cued to limit possible speech-related head movement after the inter-scan interval ended (Fig. 2A). Participants viewed a projection screen through a headcoil-mounted periscope, and their responses were recorded using an MRI-compatible microphone system (Magnetic Resonance Technologies). E-Prime software (version 2) was used to present stimuli, record responses, and synchronize trial events to scanner onsets. Task performance feedback was not provided during the experiment. Fig. 2B shows the mixed-block design that included two blocks of word recognition in noise (60 trials per block; 8 min 36 s), habituation to the babble prior to each trial block (17.2 s), and a rest interval at the start, middle, and end of the experiment (60.2 s). Words were presented in blocks of trials for each SNR condition, which included between four and six trials. A mixed-block design was used here because it provides limited predictability about upcoming task demands, which allows participants to optimize decision criteria (for example) through successive adjustments based on response uncertainty or task difficulty on the preceding trials. For that reason, a mixed design may be more sensitive to pre-stimulus adjustments than designs that involve more predictable or less predictable task difficulty.

## 2.4. Word recognition task performance

Responses were scored as correct if the participant repeated a word exactly as it was presented, otherwise responses were scored as incorrect (including “nope”). Two raters in the scanner control room transcribed each response, with 92.2% agreement across trials.



Audio recordings of the participant responses during the experiment were used to resolve discrepancies between the two raters. Because of a recording error for two participants, 13 trials with unresolved score discrepancies were excluded from the analyses (0.3% of the trials). Additionally, there were 46 unintelligible or missing responses excluded from the analyses (1.3% of the trials). Logistic regression analyses were performed on trial-level recognition (0/1) using a generalized linear mixed model (GLMM), which examined the associations between the likelihood for correct word recognition, participant age and SNR condition [R Statistics software, version 3.3.1; R-packages: lme4, version 1.1–17; ImerTest version 3.1–2]. We will use model number 1 to refer to this analysis in the results (Model #1).

Speech onset values were derived from the digital audio recordings of correct participant responses during the fMRI experiment. An edge detection algorithm was used to identify speech onsets in spectrogram data from each recording, based on the largest power increase across frequencies over time. Two raters then examined the spectrogram plots marked with algorithm-based speech onsets to identify errors. Both raters agreed with automatically identified onset values for 91% of the responses, and corrected onsets after inspecting each of the 501 participant recordings that were incorrectly labeled by the algorithm. There was a total of 2164 speech onset latencies from correct recognition trials included in the analyses. Prior to performing the analyses described later in the perceptual decision-making model section, a GLMM-based linear regression analysis (Model #2) was performed on trial-level speech onset times to characterize response latency differences related to participant age and SNR on correct trial response latencies.

Model fit was tested (Hofmann, 1997) to determine whether pure-tone thresholds significantly accounted for word recognition accuracy or speech onsets, based on several measures derived from the audiogram (Bologna et al., 2018): 1) mean pure-tone thresholds across ears; 2) mean pure-tone thresholds from the better ear; 3) mean pure-tone thresholds across 1 and 2 kHz; and 4) interactions between SNR and any of those hearing sensitivity measures. These tests did not show any significant improvement to model fit when a hearing loss measure was included in the model [accuracy: Chi-square < 1.07,  $p = 0.30$ ; latencies: Chi-square < 0.58,  $p = 0.45$ ], which suggests that the hearing loss exclusion criteria and presentation levels limited the influence of audibility differences in our sample. Hearing loss effects were not examined further.

## 2.5. Image acquisition

The fMRI experiment included the collection of 145 T2\*-weighted EPI images (21 min 30 s) with 3 mm isomorphic voxels and whole brain coverage, using a 32-channel head coil on Siemens Tim Trio or Prisma 3 T MRI scanners<sup>1</sup> at the Center for Biomedical Imaging at MUSC [TR = 8600 ms; TE = 35 ms; TA = 1647 ms; 90° flip angle; 36 slices; interleaved and ascending slice order; 3.0 mm thickness; GRAPPA with acceleration factor = 2]. Structural T1-weighted images with 1.0 mm isomorphic voxels were collected using an

---

<sup>1</sup>Imaging data from the first 23 participants were collected on a Siemens Tim Trio. Data from the other 7 participants were collected with an EPI sequence with identical parameters on a Siemens Prisma, after the MRI scanner was upgraded. No significant system upgrade effects were observed when a nuisance covariate was included in group level statistical contrasts.

MP2RAGE sequence [160 slices,  $256 \times 256$  matrix, TR = 5000 ms; TE = 2.84 ms; TI1 = 700 ms; TI 2 = 2500 ms; flip angle 1 =  $4^\circ$ ; flip angle 2 =  $5^\circ$ ; slice gap = 0; GRAPPA with acceleration factor = 3].

## 2.6. Preprocessing neuroimaging data

Preprocessing steps for the fMRI data were performed with the SPM12 software: 1) estimation and correction of head motion and gradient field distortions; 2) coregistration of functional images to native anatomical space; and 3) Gaussian smoothing (FWHM = 8 mm) after spatial normalization with the Advanced Normalization Tools (ANTS version 2.1; Avants and Gee, 2004). Diffeomorphic transformations were estimated with the ANTS software to produce a custom anatomical template for the current study participants. The fMRI data were coregistered to each participant's structural image, then spatially transformed to match the study template. The ANTS software was also used to derive standard coordinates for peak statistic locations, by spatially transforming maps in the anatomical template space to the Montreal Neurological Institute (MNI) template space.

After warping and smoothing fMRI data, additional steps were taken to limit acquisition or motion artifacts in the results. Each voxel time series was corrected for global BOLD signal fluctuations (Macey et al., 2004). A spike detection algorithm was also used to identify BOLD contrast images with intensities exceeding 2.5 standard deviations from the mean across the time series (Vaden et al., 2010), saved as two binary nuisance regressors (possible spikes in  $4.5 \pm 1.1\%$  of the images across subjects). Head position vectors (three rotation and three translation), obtained using SPM12 motion correction, were transformed into a pair of nuisance regressors for head motion based on the Pythagorean Theorem, and first order derivatives were calculated for head motion across BOLD contrast images (Kuchinsky et al., 2012; Wilke, 2012; <http://www.nitrc.org/projects/pythagoras>). Each of the spike detection and head motion vectors were entered as nuisance regressors for the fMRI analyses.

## 2.7. GLM-based fMRI analysis

Multiple regression analyses were performed using a general linear model (GLM, Model #3) in SPM12 software to examine experiment effects related to task performance and SNR manipulations (e.g., Erb et al., 2013; Vaden et al., 2013). The BOLD contrast was modeled by convolution of the hemodynamic response function for each SNR condition (+3 dB SNR and +10 dB SNR trials), which included a parameter for correct or incorrect responses.

## 2.8. GLMM-based fMRI analysis

Regression analyses were performed using a generalized linear mixed model (GLMM) to characterize associations between pre-stimulus BOLD contrast and the likelihood of correct word recognition (e.g., Vaden et al., 2013, Vaden et al., 2015) [R Statistics software, version 3.3.1; R-packages: lme4, version 1.1–17; lmerTest version 3.1–2]. Voxel-level BOLD contrast measurements prior to each word presentation were normalized for each SNR condition ( $M = 0$ ,  $sd = 1$ ). The GLMM equation specified correct or incorrect word recognition ( $W_t$ ) on each trial ( $t$ ) as a function of SNR condition ( $SNR_t$ ) and BOLD contrast prior to the presentation ( $BOLD_{t-1}$ ), as well as an interaction term ( $SNR_t \times BOLD_{t-1}$ ) and random subject effects ( $I|SUB$ ), which can be expressed as: (Model #4)  $W_t = SNR_t +$



$BOLD_{t-1} + SNR_t \times BOLD_{t-1} + (I|SUB) + error$ . Subject-level effects of pre-stimulus BOLD contrast on speech recognition were confirmed by performing an equivalent GLM logistic regression analysis in each voxel:  $W_t = SNR_t + BOLD_{t-1} + SNR_t \times BOLD_{t-1} + error$ , which also simplified the examination of individual differences.

Group level statistic maps from the GLM and GLMM based regression analyses were corrected for multiple comparisons across the whole brain, based on an uncorrected voxel statistic threshold ( $p = 0.001$ ) in combination with a family-wise error corrected cluster extent threshold ( $P_{FWE} = 0.05$ ), using SPM12 and FMRISTAT (Worsley, 2006) respectively.

## 2.9. Perceptual decision-making model

The perceptual decision-making predictions were tested using the Shifted Wald model, described in the Introduction, which fits a probability density distribution to the response latency (speech onset) data from participant responses (Anders et al., 2016; Anders et al., 2015). Simulation tests performed by Anders et al. (2016) showed that this approach can accurately estimate non-decision time ( $\theta$ ), accumulation rate ( $\gamma$ ), and criteria ( $\alpha$ ) parameters with sample sizes similar to the current study. After fitting the Shifted Wald model to speech onset times to estimate decision-making parameters in each SNR condition for each participant, the parameter estimates were entered into a GLMM-based regression analysis to test predictions that SNR conditions affect evidence accumulation, decision criteria, and non-decision times during word recognition (Model #5).

Elevated cingulo-opercular activity prior to word presentations was predicted to lead to higher, more cautious criteria and more accurate word recognition, based on previously observed associations between decision criteria and activity (Forstmann et al., 2010; Mulder et al., 2014). The hypothesis that pre-stimulus cingulo-opercular activity is associated with decision criteria was tested based on two additional analyses. The first test examined individual differences in the Shifted Wald model estimates from Model #5 (above) and cingulo-opercular adaptive control results from Model #4 (above), to determine the extent to which participants with higher, more cautious decision criteria also show larger word recognition benefit from pre-stimulus cingulo-opercular activity (Model #6). First, cingulo-opercular regions of interest were functionally defined based on the significant predictive effect of BOLD on percent correct word recognition from Model #4 that typically demonstrates cingulo-opercular effects. This cingulo-opercular result was found to include three distinct anatomical regions in frontal cortex (see Results), which were analyzed individually as regions of interest. The adaptive control variable for each participant averaged the regression estimates (i.e., betas) from Model #4, within the functionally defined cingulo-opercular regions. Because the betas represented BOLD effects on word recognition accuracy (i.e., correct and incorrect trials), these were statistically independent from Shifted Wald model estimates calculated from response latencies that were exclusively from correct trials.

A second test was performed to characterize the extent to which decision criteria were related to cingulo-opercular function based on BOLD time series (Model #7). The Shifted Wald parameters were estimated within each SNR for onset latencies on trials preceded by high or low BOLD contrast (Vaden et al., 2020; K.I. 2015, 2013). For each participant, high

pre-stimulus cingulo-opercular activity was defined by above-average BOLD contrast, while low activity was below-average BOLD contrast, again from within the cingulo-opercular regions that exhibited pre-stimulus associations with accuracy across the sample from Model #4. A GLMM-based regression was then used to test for significant Shifted Wald model parameter differences across high and low BOLD trials, while controlling for SNR-related effects.

### 3. Results

#### 3.1. Word recognition in noise task performance

Word recognition in noise accuracy was significantly higher in +10 than +3 dB SNR [ $Z = 14.76$ ,  $df = 3524$ ,  $p < 0.001$ ] and significantly poorer with increasing participant age<sup>2</sup> [ $Z = -2.74$ ,  $df = 27.93$ ,  $p = 0.006$ ] (Model #1). Model fit was not significantly improved by including an interaction between participant age and SNR to predict correct word recognition [ $Chi-square = 1.56$ ,  $p = 0.21$ ]. Speech onset latencies were significantly shorter in the +10 than +3 dB SNR [ $Z = -3.78$ ,  $df = 26.63$ ,  $p = 0.0002$ ] and significantly longer with increasing participant age [ $Z = 2.10$ ,  $df = 26.05$ ,  $p = 0.04$ ] (Model #2). The Age  $\times$  SNR interaction was significant for speech onsets [ $Z = -2.29$ ,  $df = 28.84$ ,  $p = 0.02$ ], which reflected significantly longer onsets with increasing age in the +3 dB SNR [ $Z = 2.41$ ,  $df = 26.28$ ,  $p = 0.02$ ] but not in the +10 dB SNR [ $Z = 1.73$ ,  $df = 25.97$ ,  $p = 0.08$ ]. Both task performance measures are plotted for each SNR condition in Fig. 3A–B.

#### 3.2. Perceptual decision-making in different SNR conditions

All three perceptual decision-making parameters of the Shifted Wald model (Anders et al., 2016) were significantly affected by SNR condition (Model #5), as seen in Fig. 3. Decision criteria ( $\alpha$ ) were significantly lower (less cautious responding) with lower SNR [+3 dB SNR:  $\alpha = 39.9 \pm 17.8$ ; +10 dB SNR:  $\alpha = 50.8 \pm 23.0$ ;  $Z = -2.14$ ,  $df = 27$ ,  $p = 0.03$ ]. Significantly lower decision criteria in the +3 dB SNR trials compared to +10 dB SNR trials would allow faster decisions by collecting less evidence before word recognition occurred in the harder SNR condition, consistent with a speed emphasis (Fig. 3C). Meanwhile, higher decision criteria in the +10 dB SNR trials would allow more accurate decisions based on more information than if lower criteria were used (i.e., accuracy emphasis). Information accumulation ( $\gamma$ ) was also significantly slower in the lower SNR [+3 dB SNR:  $\gamma = 0.11 \pm 0.03$ ; +10 dB SNR:  $\gamma = 0.13 \pm 0.04$ ;  $Z = -2.57$ ,  $df = 27$ ,  $p = 0.01$ ], indicating that evidence supporting word recognition was gathered more slowly for noisier speech presentations (Fig. 3D). Non-decision times ( $\theta$ ) were also significantly longer in the lower SNR condition [+3 dB SNR:  $\theta = 213.0 \pm 154.5$ ; +10 dB SNR:  $\theta = 139.5 \pm 156.0$ ;  $Z = 2.32$ ,  $df = 27$ ,  $p = 0.02$ ], suggesting a combination of slower sensory and motor processing for clearer speech (Fig. 3E).

Model testing indicated that neither age effects nor Age  $\times$  SNR interactions significantly improved model fit for the accumulation rate or decision criteria parameters [ $Chi-square = 2.11$ ,  $p = 0.15$ ]. However, there was a significant Age  $\times$  SNR interaction for the non-decision

<sup>2</sup>Degrees of freedom for the GLMM results were calculated using Satterthwaite approximation.

times [ $Z = -2.83$ ,  $df = 26$ ,  $p = 0.004$ ], which reflected longer non-decision times with increasing age on +3 dB SNR trials [ $Z = 1.99$ ,  $df = 26$ ,  $p = 0.05$ ] but not the +10 dB SNR trials [ $p = 0.45$ ]. Because nondecision times represent sensory and motor processes, this could reflect age-related slowing in either process (or both processes) in the more difficult SNR. Together, the Shifted Wald model results indicate that SNR was associated with significant changes in perceptual decision-making processes. Next, we return to associations between Shifted Wald model parameters and neuroimaging measures (Models #6 and #7).

### 3.3. Neuroimaging results related to task events and performance

Results from the GLM-based regression analyses showed significant BOLD contrast effects related to the word recognition task and SNR conditions (Model #3), which are detailed in Table 1 and Fig. 4. Briefly, auditory and cingulo-opercular regions showed significantly elevated BOLD contrast during the word recognition task trials, relative to the implicit baseline. Cingulo-opercular regions and bilateral superior temporal sulci showed significantly higher BOLD contrast for the +3 dB SNR compared to +10 dB SNR. Cingulo-opercular regions also showed significantly higher BOLD contrast for word recognition errors. Smaller, significant clusters with error-related increases in BOLD contrast were also observed in the left caudate, left superior temporal sulcus, and left middle frontal gyrus. These results replicate commonly observed effects of lower intelligibility and word recognition errors for cingulo-opercular activity (Eckert et al., 2009; Erb et al., 2013; Peelle, 2017; Wild et al., 2012; Zekveld et al., 2006).

### 3.4. Cingulo-opercular adaptive control effects on word recognition

Results from the GLMM-based regression analysis (Model #4) showed significant effects of pre-stimulus cingulo-opercular BOLD contrast on word recognition, in addition to the effects of the SNR manipulation on recognition accuracy (Table 1, Fig. 4D). Specifically, prestimulus BOLD contrast in the dorsal cingulate, left frontal operculum, and bilateral anterior insula showed a significant association with subsequent word recognition accuracy. The pre-stimulus BOLD effects on recognition did not interact significantly with SNR, based on nonsignificant cluster sizes [extent = 3 voxels,  $p_{FWE} = 0.98$ ]. These results replicate earlier observations that pre-stimulus activity is associated with increased likelihood for a correct response (Vaden et al., 2020; Vaden et al., 2015; Vaden et al., 2013). Standard estimates (i.e., betas) for the pre-stimulus BOLD effects on word recognition were averaged for each participant within each of the significant cingulo-opercular regions to interrogate reasons for these effects. Participant age was not significantly associated with individual differences in the adaptive control effects, based on the pooled betas [ $p > 0.37$ ].

A control analysis that modified Model #4 by including only trials that immediately followed correct word recognition showed significant associations between pre-stimulus BOLD and correct word recognition within each of the cingulo-opercular regions (Table 1), which suggests these effects do not depend on error-related activity from preceding trials (Vaden et al., 2015; Vaden et al., 2013).

### 3.4. Individual differences in adaptive control and decision criteria

Evidence that decision criteria were adjusted to optimize task performance came from individual differences in the decision criteria and prestimulus BOLD effects on word recognition (Fig. 5A–B, Model #6). BOLD effects on subsequent word recognition (betas) pooled within the dorsal cingulate region were significantly higher for participants with more cautious decision criteria [both SNR:  $r = 0.47$ ,  $df = 26$ ,  $p = 0.01$ ; +3 dB SNR:  $r = 0.54$ ,  $df = 26$ ,  $p = 0.003$ ; +10 dB SNR:  $r = 0.24$ ,  $df = 26$ , ns]. The cingulate betas were not significantly related to individual differences in accumulation rate [ $p = 0.30$ ] nor non-decision time [ $p = 0.65$ ]. Betas from the other cingulo-opercular regions were not associated with decision criteria estimates for each subject [ $p > 0.89$ ]. Thus, participants who responded more cautiously exhibited stronger effects of dorsal cingulate activity on word recognition.

### 3.5. Pre-stimulus activity and ongoing decision criteria changes

Trial-level changes in decision criteria were also tested for each cingulo-opercular region of interest based on the Shifted Wald model, which was fitted to response latencies for high or low pre-stimulus BOLD contrast trials for each SNR condition (Fig. 5C; Model #7). Decision criteria were significantly higher on trials with high compared to low prestimulus BOLD contrast in the left insula/frontal operculum [change in  $\alpha = 9.32 \pm 15.6$ ;  $Z = 2.72$ ,  $df = 82$ ,  $p = 0.008$ ] and in the right insula [change in  $\alpha = 7.16 \pm 19.5$ ;  $Z = 2.01$ ,  $df = 82$ ,  $p = 0.048$ ]. Pre-stimulus BOLD contrast in the dorsal cingulate was not significantly associated with trial-level decision criteria changes [ $p = 0.84$ ]. The significant associations between pre-stimulus BOLD and decision criteria suggest that higher pre-stimulus cingulo-opercular activity shifted decision criteria, so that more evidence was collected before word recognition. In other words, cingulo-opercular activity appears to benefit speech recognition accuracy in background noise by raising decision criteria.

## 4. Discussion

Adaptive control is characterized by elevated pre-stimulus activity in cingulo-opercular regions that is associated with subsequent improvements in task performance. The results of this study support the premise that a shift in decision criteria is one mechanism underlying cingulo-opercular adaptive control during speech recognition in noise. Lower decision criteria were observed during word recognition in the poorer SNR, which was also associated with slower evidence accumulation and longer non-decision times. Those perceptual decision-making changes have been observed for other types of auditory and visual stimuli that were degraded or presented in challenging task conditions (e.g., Anders et al., 2016; Ben-David et al., 2014; Bode et al., 2018; Hanks and Summerfield, 2017; Tillman et al., 2017). The finding that decision criteria were lower for trials with poorer speech intelligibility suggests that recognition occurred more quickly to offset effects of slower evidence accumulation and non-decision times. Participants may notice slow-downs in performance when the SNR decreases and lower their threshold to maintain an acceptable response rate. This is also consistent with a strategy that emphasizes response speed in the poorer SNR and response accuracy in the better SNR, where there may be more evidence to collect.

We also replicated findings that pre-stimulus cingulo-opercular activity is associated with correct word recognition in noise (Vaden et al., 2015, Vaden et al., 2013). Our novel findings linked pre-stimulus cingulo-opercular activity to raised decision criteria, based on a Shifted Wald model fitted to response latencies (Anders et al., 2016). Dorsal cingulate activity appeared to predict word recognition accuracy and more cautious responding for correct responses. More cautious responses allow for more information to be collected while listeners choose the most likely response, which was more likely to occur on trials with increased pre-stimulus anterior insula and opercular activity.

Elevated cingulo-opercular activity is commonly observed during difficult speech recognition tasks (Adank, 2012; Zekveld et al., 2012), which is thought to reflect performance monitoring in difficult task conditions with recognition errors (Erb et al., 2013; Vaden et al., 2013). This response appears to occur when there is utility in adjusting behavior to optimize performance (Eckert et al., 2016), in part because cingulo-opercular activity drops off when a task is easy or when performance is impossible (e.g., Reuter-Lorenz and Cappell, 2008; Reuter-Lorenz and Lustig, 2005). We previously observed that cingulo-opercular activity benefits task performance for speech recognition in noise (Vaden et al., 2013, 2016; Vaden et al., 2015), memory encoding for speech understood in noise (Vaden et al., 2017), and gap detection in noise (Vaden et al., 2020). Similarly, pre-stimulus cingulo-opercular BOLD contrast was significantly associated with correct word recognition in the present study.

With respect to the implementation of adaptive control during difficult speech recognition tasks, the results of the current study suggest that the pre-stimulus cingulo-opercular adaptive control effect includes decision criteria adjustments. The word recognition benefit from elevated pre-stimulus cingulo-opercular activity was larger for participants with higher compared to lower decision criteria. That is, participants whose dorsal cingulate activity predicted correct recognition on the next trial also showed evidence of more cautious responding (only for +3 dB SNR when tested separately), based on criteria differences derived from response latencies. The specificity of the current findings to anterior insula-operculum or dorsal cingulate may reflect that cingulo-opercular regions are functionally heterogeneous (Menon and D'Esposito, 2022) and appear sensitive to a wide-range of task and reward-related information (Horst and Laubach, 2012). Indeed, even in the current study, cingulo-opercular BOLD activity showed spatial and temporal differences in relation to SNR conditions, word recognition errors, and subsequent word recognition. Dorsal cingulate activity has been shown to affect response caution (Bogacz et al., 2010; Mulder et al., 2014), which supports a hypothesized braking function through interactions with the basal ganglia (Cavanagh et al., 2011). In the current study, this putative braking function could explain how participants implement slower and more cautious responding to emphasize response accuracy (Van Maanen et al., 2011).

Individual differences in adaptive control and response caution effects were not seen in bilateral anterior insula and frontal operculum, but the magnitude of increased pre-stimulus activity in both regions was significantly associated with more cautious responding on a trial-to-trial basis. While increased activity for cautious response selection and response inhibition (Aron et al., 2003; Forstmann et al., 2008; Hughes et al., 2013) would be expected

during a high criterion trial when long reaction times are observed (Binder et al., 2004), elevated pre-stimulus activity in insula/opercula regions may provide feedback to the dorsal cingulate to inform value-based calculations and alter decision criteria. Depth electrode recordings have shown that anterior insula computes prediction error (expected versus actual outcome) and appears to communicate this error to dorsal cingulate (Billeke et al., 2020). Thus, anterior insula, and perhaps the frontal operculum, appear to contribute to trial-to-trial confidence-weighting (Braun et al., 2018) or expectation about the relative difficulty of the next trial so that changes in response caution could be implemented (Urai et al., 2017) to optimize accuracy. The current results suggest that the extent of change in response caution depends on the extent of the prediction error coded in the anterior insula.

The current study links pre-stimulus activity changes to decision criteria during speech recognition for the first time, suggesting that cingulo-opercular cortex can adjust decision-making, in addition to attention and behaviors, to optimize task performance (Eckert et al., 2016). Our finding accords with mid-frontal EEG activity measures that were related to response confidence, as well as subsequent criterion changes during a visual judgment task (Desender et al., 2019). Sensitivity to pre-stimulus cingulo-opercular effects on decision criteria may have been enhanced from the block design in the current study, which presented four to six consecutive trials in each SNR. This design allowed participants to recognize the change in SNR and modify decision criteria. For example, a consistently better SNR, which provided a shorter nondecision time and a faster accumulation rate, allowed for more flexible decision criteria. Participants were able to raise their criteria to optimize performance in the better SNR condition or capitalize on the relatively easier condition to optimize overall word recognition. An event-related design with unpredictable SNR conditions and a limited response interval could limit listeners' ability to alter their decision thresholds. One drawback to the mixed-block design in the current study was that the small number of correct recognition trials following a different SNR did not allow effects related to preceding SNR conditions (e.g., same or different) to be characterized. Manipulating the predictability of task demands by varying the duration for a given SNR, instructing participants to favor a speeded response over accuracy (Van Maanen et al., 2011), or varying response intervals may allow the flexibility of decision thresholds to be evaluated in future studies.

Age-related differences in cingulo-opercular activation have been observed in difficult speech recognition tasks with younger and older adult participants (Erb and Obleser, 2013; Harris et al., 2009; Vaden et al., 2015). Even though cingulo-opercular adaptive control was significantly related to speech recognition for middle-aged and older adults, this trial-level effect did not explain age-related differences in speech recognition in this study. Non-decision times in the poorer SNR were significantly longer with increasing age in the current study, which is consistently observed across numerous studies (Ben-David et al., 2014; Ratcliff et al., 2004). However, longer non-decision times are thought to reflect a combination of sensory and motor processes (Anders et al., 2016) and did not explain the association between speech recognition and participant age in this cross-sectional study. The limited age-related differences in perceptual decision-making, despite age-related speech recognition differences, suggest that these processes can facilitate performance independent of age-related effects on speech recognition to some degree. Our results



indicated that middle-aged and older adults did not maintain cautious decision criteria across SNR conditions, although age-related criteria differences may be limited to specific task conditions that emphasize response speed (e.g., Ratcliff et al., 2004). Decision criteria may be adjusted to stabilize performance at an optimal level for a listener based on their motivation, uncertainty, and value from correctly understanding words in noise (Desender et al., 2019; Eckert et al., 2016; Shenhav et al., 2013; Yee et al., 2021).

There were fewer correct response latencies fitted with the Shifted Wald model for older participants due to their poorer scores. Response latencies from incorrect trials are typically excluded, because of distributional differences and under-sampling (Anders et al., 2016; Anders et al., 2015). This issue may have limited statistical sensitivity to age-related differences in decision criteria or accumulation rates. Despite this limitation, significant SNR-related effects and cingulo-opercular effects were observed for perceptual decision-making processes modeled using this approach. We statistically controlled for SNR and performance effects on pre-stimulus activity, however we did not identify what drove the pre-stimulus activity changes. For example, it is possible that participants were uncertain about their correct response because no feedback was provided, and this produced elevated cingulo-opercular activity that predicted changes in criteria and word recognition. Although it is unclear what factors influenced pre-stimulus cingulo-opercular activity (e.g., response uncertainty; Desender et al., 2019), potential SNR and performance influences were statistically controlled for while characterizing adaptive control effects for understanding speech.

## Conclusion

Perceptual decision-making, and specifically an elevated decision criterion, is one explanation for the cingulo-opercular adaptive control effect on word recognition. More cautious responding enables listeners to optimize their speech understanding despite the cost of taking more time to respond (Drugowitsch et al., 2012). Taking more time to understand can be costly in conversation, when a listener could fall behind or lose the thread - common complaints for older adults, especially with hearing loss. Cingulo-opercular regions exhibit structural declines with increasing age and poor health (Eckert et al., 2010; Fjell et al., 2009; Hahn et al., 2019) and less adaptive control is observed with increasing age (Vaden et al., 2015). More research is needed to determine the extent to which decision criteria flexibility, and perceptual decisionmaking more broadly, change with increasing age and have detrimental consequences for the ability to adapt to changing listening conditions.

## Acknowledgments

We thank the study participants. This work was supported by an Emerging Research Grant from the Hearing Health Foundation and by NIH/NIDCD (P50 DC000422), MUSC Center for Biomedical Imaging, the National Center for Advancing Translational Sciences of the National Institutes of Health under Grant Number (UL1 TR001450). This investigation was conducted in a facility constructed with support from the NIH/NCRR Research Facilities Improvement Program (C06 RR014516).

## Data and code availability statement

Data from this study can be requested from the corresponding author and accessed with appropriate IRB and institutional approvals. R code for analyzing response latencies with the Shifted Wald Model is available online from Anders et al. (2016).

## References

- Adank P, 2012. The neural bases of difficult speech comprehension and speech production: two activation likelihood estimation (ALE) meta-analyses. *Brain Lang.* 122, 42–54. doi:10.1016/j.bandl.2012.04.014. [PubMed: 22633697]
- Alain C, Du Y, Bernstein LJ, Barten T, Banai K, 2018. Listening under difficult conditions: an activation likelihood estimation meta-analysis. *Hum. Brain Mapp* 39, 2695–2709. doi:10.1002/hbm.24031. [PubMed: 29536592]
- Alain C, McDonald KL, Ostroff JM, Schneider B, 2004. Aging: A switch from automatic to controlled processing of sounds? *Psychol. Aging* 19, 125–133. doi:10.1037/0882-7974.19.1.125. [PubMed: 15065936]
- Alavash M, Tune S, Obleser J, 2019. Modular reconfiguration of an auditory-control brain network supports adaptive listening behavior. *Proc. Natl. Acad. Sci. USA* 116, 660–669. doi:10.1101/409797. [PubMed: 30587584]
- American National Standards Institute, 2010. Specification for audiometers. ANSI S3.6-2010. American National Standards Institute, New York.
- Anders R, Alario FX, Van Maanen L, 2016. The shifted Wald distribution for response time data analysis. *Psychol. Methods* 21, 309–327. doi:10.1037/met0000066. [PubMed: 26867155]
- Anders R, Ries S, Van Maanen L, Alario FX, 2015. Evidence accumulation as a model for lexical selection. *Cogn. Psychol* 82, 57–73. doi:10.1016/j.cogpsych.2015.07.002. [PubMed: 26375509]
- Aron AR, Fletcher PC, Bullmore ET, Sahakian BJ, Robbins TW, 2003. Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nat. Neurosci* 6, 115–116. doi:10.1038/nrn1003. [PubMed: 12536210]
- Avants BB, Gee JC, 2004. Geodesic estimation for large deformation anatomical shape averaging and interpolation. *Neuroimage* 23 (Suppl 1), S139–S150. doi: 10.1016/j.neuroimage.2004.07.010. [PubMed: 15501083]
- Ben-David BM, Eidels A, Donkin C, 2014. Effects of aging and distractors on detection of redundant visual targets and capacity: do older adults integrate visual targets differently than younger adults? *PLoS ONE* 1–29. doi:10.1371/journal.pone.0113551.
- Billeke P, Ossandon T, Perrone-Bertolotti M, Kahane P, Bastin J, Jerbi K, Lachaux JP, Fuentelba P, 2020. Human anterior insula encodes performance feedback and relays prediction error to the medial prefrontal cortex. *Cereb. Cortex* 30, 4011–4025. doi: 10.1093/CERCOR/BHAA017. [PubMed: 32108230]
- Binder JR, Liebenthal E, Possing ET, Medler DA, Ward BD, 2004. Neural correlates of sensory and decision processes in auditory object identification. *Nat. Neurosci* 7, 295–301. [PubMed: 14966525]
- Bode S, Bennett D, Sewell DK, Paton B, Egan GF, Smith PL, Murawski C, 2018. Dissociating neural variability related to stimulus quality and response times in perceptual decision-making. *Neuropsychologia* 111, 190–200. doi: 10.1016/j.neuropsychologia.2018.01.040. [PubMed: 29408524]
- Bogacz R, Wagenmakers E, Forstmann BU, Nieuwenhuis S, 2010. The neural basis of the speed - accuracy tradeoff. *Trends Neurosci.* 33, 10–16. doi: 10.1016/j.tins.2009.09.002. [PubMed: 19819033]
- Bologna WJ, Vaden KI, Ahlstrom JB, Dubno JR, 2018. Age effects on perceptual organization of speech: Contributions of glimpsing, phonemic restoration, and speech segregation. *J. Acoust. Soc. Am* 144, 267–281. doi:10.1121/1.5044397. [PubMed: 30075693]

- Braun A, Urai AE, Donner TH, 2018. Adaptive history biases result from confidence-weighted accumulation of past choices. *J. Neurosci* 38, 2418–2429. doi: 10.1523/JNEUROSCI.2189-17.2017. [PubMed: 29371318]
- Brosnan MB, Sabaroein K, Silk T, Gene S, Newman DP, Loughnane GM, Fornito A, O’Connell RG, Bellgrove MA, 2020. Evidence accumulation during perceptual decisions in humans varies as a function of dorsal frontoparietal organization. *Nat. Hum. Behav* 4, 844–855. [PubMed: 32313233]
- Brown S, Heathcote A, 2005. A ballistic model of choice response time. *Psychol. Rev* 112, 117–128. doi:10.1037/0033-295X.112.1.117. [PubMed: 15631590]
- Brown SD, Heathcote A, 2008. The simplest complete model of choice response time: Linear ballistic accumulation. *Cogn. Psychol* 57, 153–178. doi: 10.1016/j.cogpsych.2007.12.002. [PubMed: 18243170]
- Cavanagh JF, Wiecki TV, Cohen MX, Figueroa CM, Samanta J, Sherman SJ, Frank MJ, 2011. Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nat. Neurosci* 14, 1462–1469. doi:10.1038/nn.2925. [PubMed: 21946325]
- Coste CP, Kleinschmidt A, 2016. Cingulo-opercular network activity maintains alertness. *Neuroimage* 128, 264–272. doi:10.1016/j.neuroimage.2016.01.026. [PubMed: 26801604]
- Coste CP, Sadaghiani S, Friston KJ, Kleinschmidt A, 2011. Ongoing brain activity fluctuations directly account for intertrial and indirectly for intersubject variability in stroop task performance. *Cereb. Cortex* 21, 2612–2619. doi:10.1093/cercor/bhr050. [PubMed: 21471558]
- Desender K, Boldt A, Verguts T, Donner TH, 2019. Confidence predicts speed-accuracy tradeoff for subsequent decisions. *Elife* 8, 1–25. doi:10.7554/eLife.43499.
- Dirks DD, Takayanagi S, Moshfegh A, Noffsinger PD, Fausti SA, 2001. Examination of the neighborhood activation theory in normal and hearing-impaired listeners. *Ear Hear.* 22,1–13. [PubMed: 11271971]
- Donkin C, Averell L, Brown S, Heathcote A, 2009. Getting more from accuracy and response time data: methods for fitting the linear ballistic accumulator. *Behav. Res. Methods* 41, 1095–1110. doi:10.3758/BRM.41.4.1095. [PubMed: 19897817]
- Donkin C, Brown S, Heathcote A, Wagenmakers E-J, 2010. Diffusion versus linear ballistic accumulation: different models but the same conclusions about psychological processes? *Psychon. Bull. Rev* 18, 61–69.
- Dosenbach NUF, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE, 2006. A core system for the implementation of task sets. *Neuron* 50, 799–812. doi: 10.1016/j.neuron.2006.04.031. [PubMed: 16731517]
- Drugowitsch J, Moreno-Bote RN, Churchland AK, Shadlen MN, Pouget A, 2012. The cost of accumulating evidence in perceptual decision making. *J. Neurosci* 32, 3612–3628. doi:10.1523/JNEUROSCI.4010-11.2012. [PubMed: 22423085]
- Dubno JR, Dirks DD, Morgan DE, 1984. Effects of age and mild hearing loss on speech recognition in noise. *J. Acoust. Soc. Am* 76, 87–96. [PubMed: 6747116]
- Dubno JR, Lee F-S, Matthews LJ, Ahlstrom JB, Horwitz AR, Mills JH, 2008. Longitudinal changes in speech recognition in older persons. *J. Acoust. Soc. Am* 123, 462–475. [PubMed: 18177174]
- Dully J, McGovern DP, O’Connell RG, 2018. The impact of natural aging on computational and neural indices of perceptual decision making: A review. *Behav. Brain Res* 355, 48–55. doi:10.1016/j.bbr.2018.02.001. [PubMed: 29432793]
- Dutilh G, Annis J, Brown SD, Cassey P, Evans NJ, Grasman RPPP, Hawkins GE, Heathcote A, Holmes WR, Kryptos A, Kupitz CN, Leite FP, Lerche V, Lin Y, Logan GD, Palmeri TJ, Starns JJ, Trueblood JS, Maanen L. Van, Ravenzwaaij D. Van, Vandekerckhove J, 2019. The quality of response time data inference: A blinded, collaborative assessment of the validity of cognitive models. *Psychon. Bull. Rev* 26, 1051–1069. [PubMed: 29450793]
- Eckert MA, Keren NI, Roberts DR, Calhoun VD, Harris KC, 2010. Age-related changes in processing speed: Unique contributions of cerebellar and prefrontal cortex. *Front. Hum. Neurosci* 4, 1–14. doi:10.3389/neuro.09.010.2010. [PubMed: 20204154]
- Eckert MA, Menon V, Walczak A, Ahlstrom JB, Denslow S, Horwitz A, Dubno JR, 2009. At the heart of the ventral attention system: The right anterior insula. *Hum. Brain Mapp* 30, 2530–2541. doi:10.1002/hbm.20688. [PubMed: 19072895]

- Eckert MA, Teubner-Rhodes S, Vaden KI, 2016. Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. *Ear Hear* 37, 101S–110S. doi:10.1097/AUD.0000000000000300. [PubMed: 27355759]
- Erb J, Henry MJ, Eisner F, Obleser J, 2013. The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *J. Neurosci* 33, 10688–10697. [PubMed: 23804092]
- Erb J, Obleser J, 2013. Upregulation of cognitive control networks in older adults' speech comprehension. *Front. Syst. Neurosci* 7. doi:10.3389/fnsys.2013.00116, 116.1–13. [PubMed: 24399939]
- Evans NJ, Rae B, Bushmakin M, Rubin M, Brown SD, 2017. Need for closure is associated with urgency in perceptual decision-making. *Mem. Cognit* 45, 1193–1205. doi:10.3758/s13421-017-0718-z.
- Fjell AM, Westlye LT, Amlie I, Espeseth T, Reinvang I, Raz N, Agartz I, Salat DH, Greve DN, Fischl B, Dale AM, Walhovd KB, 2009. High consistency of regional cortical thinning in aging across multiple samples. *Cereb. Cortex* 19, 2001–2012. doi:10.1093/cercor/bhn232. [PubMed: 19150922]
- Folstein MF, Robins LN, Helzer JE, 1983. The mini-mental state examination. *Arch. Gen. Psychiatry* 40, 812. doi:10.1001/archpsyc.1983.01790060110016. [PubMed: 6860082]
- Forstmann BU, Anwander A, Schafer A, Neumann J, Brown S, Wagenmakers EJ, Bogacz R, Turner R, 2010. Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proc. Nat. Acad. Sci. U.S.A* 107, 15916–15920. doi:10.1073/pnas.1004932107.
- Forstmann BU, Jahfari S, Scholte HS, Wolfensteller U, van den Wildenberg WPM, Ridderinkhof KR, 2008. Function and structure of the right inferior frontal cortex predict individual differences in response inhibition: A model-based approach. *J. Neurosci* 28, 9790–9796. [PubMed: 18815263]
- Gratton C, Neta M, Sun H, Ploran EJ, Schlaggar BL, Wheeler ME, Petersen SE, Nelson SM, 2016. Distinct stages of moment-to-moment processing in the cinguloopercular and frontoparietal networks. *Cereb. Cortex* 27, 2403–2417. doi: 10.1093/cercor/bhw092.
- Hahn S, Lotze M, Domin M, Schmidt S, 2019. The association of health-related quality of life and cerebral gray matter volume in the context of aging: A voxel-based morphometry study with a general population sample. *Neuroimage* 191, 470–480. doi:10.1016/j.neuroimage.2019.02.035. [PubMed: 30790673]
- Hanks TD, Summerfield C, 2017. Perceptual decision making in rodents, monkeys, and humans. *Neuron* 93, 15–31. doi:10.1016/j.neuron.2016.12.003. [PubMed: 28056343]
- Harris KC, Dubno JR, Keren NI, Ahlstrom JB, Eckert MA, 2009. Speech recognition in younger and older adults: A dependency on low-level auditory cortex. *J. Neurosci* 29, 6078–6087. doi: 10.1523/JNEUROSCI.0412-09.2009. [PubMed: 19439585]
- Heathcote A, Love J, 2012. Linear deterministic accumulator models of simple choice. *Front. Psychol* 3, 1–19. doi:10.3389/fpsyg.2012.00292. [PubMed: 22279440]
- Hofmann DA, 1997. An overview of the logic and rationale of hierarchical linear models. *J. Manag* 23, 723–744. doi: 10.1177/014920639702300602.
- Horst NK, Laubach M, 2012. Working with memory: Evidence for a role for the medial prefrontal cortex in performance monitoring during spatial delayed alternation. *J. Neurophysiol* 108, 3276–3288. doi:10.1152/jn.01192.2011. [PubMed: 23019007]
- Hughes ME, Johnston PJ, Fulham WR, Budd TW, Michie PT, 2013. Stop-signal task difficulty and the right inferior frontal gyrus. *Behav. Brain Res* 256, 205–213. doi:10.1016/j.bbr.2013.08.026. [PubMed: 23973765]
- Humes LE, 2021. Factors underlying individual differences in speech-recognition threshold (SRT) in noise among older adults. *Front. Aging Neurosci* 13. doi:10.3389/FNAGI.2021.702739.
- Humes LE, Christopherson L, 1991. Speech identification difficulties of hearing-impaired elderly persons: The contributions of auditory processing deficits. *J. Speech. Hear. Res* 34, 686–693. [PubMed: 2072694]
- Humes LE, Dubno JR, 2010. Factors affecting speech understanding in older adults. In: *Springer Handbook of Auditory Research*. Springer, New York, NY, pp. 211–257.
- Johns AR, Myers EB, Skoe E, 2018. Sensory and cognitive contributions to age - related changes in spoken word recognition. *Lang. Linguist. Compass* 1–25. doi:10.1111/lnc3.12272.

- Kalikow DN, Stevens KN, Elliott LL, 1977. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J. Acoust. Soc. Am* 61, 1337–1351. [PubMed: 881487]
- Kuchinsky SE, Vaden KI, Keren NI, Harris KC, Ahlstrom JB, Dubno JR, Eckert MA, 2012. Word intelligibility and age predict visual cortex activity during word listening. *Cereb. Cortex* 22, 1360–1371. doi:10.1093/cercor/bhr211. [PubMed: 21862447]
- Lukacs E, 1955. A characterization of the gamma distribution. *Ann. Math. Stat* 26, 319–324.
- Macey PM, Macey KE, Kumar R, Harper RM, 2004. A method for removal of global effects from fMRI time series. *Neuroimage* 22, 360–366. doi: 10.1016/j.neuroimage.2003.12.042. [PubMed: 15110027]
- Matzke D, Wagenmakers EJ, 2009. Psychological interpretation of the ex-gaussian and shifted wald parameters: A diffusion model analysis. *Psychon. Bull. Rev* 16, 798–817. doi:10.3758/PBR.16.5.798. [PubMed: 19815782]
- Meister H, Rahlmann S, Lemke U, Besser J, 2018. Verbal response times as a potential indicator of cognitive load during conventional speech audiometry with matrix sentences. *Trends Hear.* 22, 1–11. doi:10.1177/2331216518793255.
- Menon V, D’Esposito M, 2022. The role of PFC networks in cognitive control and executive function. *Neuropsychopharmacology* 47, 90–103. doi:10.1038/s41386-021-01152-w. [PubMed: 34408276]
- Mulder MJ, van Maanen L, Forstmann BU, 2014. Perceptual decision neurosciences - a model-based review. *Neuroscience* 277, 872–884. doi:10.1016/j.neuroscience.2014.07.031. [PubMed: 25080159]
- Oldfield RC, 1971. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9, 97–113. [PubMed: 5146491]
- Peelle J, 2017. Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear Hear.* 39, 204–214.
- Pichora-Fuller MK, Kramer SE, Eckert MA, Edwards B, Hornsby BWY, Humes LE, Lemke U, Lunner T, Matthen M, Mackersie CL, Naylor G, Phillips NA, Richter M, Rudner M, Sommers MS, Tremblay KL, Wingfield A, 2016. Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear Hear.* 37, 5S–27S. doi:10.1097/AUD.0000000000000312. [PubMed: 27355771]
- Ploran EJ, Nelson SM, Velanova K, Donaldson DI, Petersen SE, Wheeler ME, 2007. Evidence accumulation and the moment of recognition: Dissociating perceptual recognition processes using fMRI. *J. Neurosci* 27, 11912–11924. doi:10.1523/JNEUROSCI.3522-07.2007. [PubMed: 17978031]
- Ploran EJ, Tremel JJ, Nelson SM, Wheeler ME, 2011. High quality but limited quantity perceptual evidence produces neural accumulation in frontal and parietal cortex. *Cereb. Cortex* 21, 2650–2662. doi:10.1093/cercor/bhr055. [PubMed: 21498405]
- Rae B, Heathcote A, Donkin C, Averell L, Brown S, 2014. The hare and the tortoise: Emphasizing speed can change the evidence used to make decisions. *J. Exp. Psychol.: Learn. Mem. Cognit* 40, 1226–1243. doi:10.1037/a0036801.
- Rahnev D, Nee DE, Riddle J, Larson AS, D’Esposito M, 2016. Causal evidence for frontal cortex organization for perceptual decision making. *Proc. Natl. Acad. Sci. USA* 113, 6059–6064. doi:10.1073/pnas.1522551113. [PubMed: 27162349]
- Ratcliff R, 1978. A theory of memory retrieval. *Psychol. Rev* 85, 59–108. doi: 10.1037/0033-295X.85.2.59.
- Ratcliff R, Thapar A, Mckoon G, 2004. A diffusion model analysis of the effects of aging on recognition memory. *J. Mem. Lang* 50, 408–424. doi:10.1016/j.jml.2003.11.002.
- Reuter-Lorenz PA, Cappell KA, 2008. Neurocognitive aging and the compensation hypothesis. *Curr. Dir. Psychol. Sci* 17, 177–182. doi: 10.1111/j.1467-8721.2008.00570.x.
- Reuter-Lorenz PA, Lustig C, 2005. Brain aging: Reorganizing discoveries about the aging mind. *Curr. Opin. Neurobiol* 15, 245–251. [PubMed: 15831410]
- Rowe JB, Hughes L, Nimmo-Smith I, 2010. Action selection: A race model for selected and non-selected actions distinguishes the contribution of premotor and prefrontal areas. *Neuroimage* 51, 888–896. doi:10.1016/j.neuroimage.2010.02.045. [PubMed: 20188184]

- Shenhav A, Botvinick MM, Cohen JD, 2013. The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron* 79, 217–240. doi:10.1016/j.neuron.2013.07.007.
- Smith PL, Ratcliff R, 2004. Psychology and neurobiology of simple decisions. *Trend Neurosci* 27, 161–168. doi:10.1016/j.tins.2004.01.006. [PubMed: 15036882]
- Souza PE, Boike KT, Witherell K, Tremblay KL, 2007. Prediction of speech recognition from audibility in older listeners with hearing loss: Effects of age, amplification, and background noise. *J. Am. Acad. Audiol* 18, 54–65. [PubMed: 17252958]
- Tillman G, Benders T, Brown SD, Ravenzwaaij D.van, 2017. An evidence accumulation model of acoustic cue weighting in vowel perception. *J. Phon* 61, 1–12. doi:10.1016/j.wocn.2016.12.001.
- Urai AE, Braun A, Donner TH, 2017. Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nat. Commun* 8. doi: 10.1038/ncomms14637. [PubMed: 28364116]
- Vaden KI, Eckert MA, Dubno JR, Harris KC, 2020. Cingulo-opercular adaptive control for younger and older adults during a challenging gap detection task. *J. Neurosci. Res* 98, 680–691. doi:10.1002/jnr.24506. [PubMed: 31385349]
- Vaden KI, Kuchinsky SE, Ahlstrom JB, Dubno JR, Eckert MA, 2015. Cortical activity predicts which older adults recognize speech in noise and when. *J. Neurosci* 35, 3929–3937. doi:10.1523/JNEUROSCI.2908-14.2015. [PubMed: 25740521]
- Vaden KI, Kuchinsky SE, Ahlstrom JB, Teubner-Rhodes S, Dubno JR, Eckert MA, 2016. Cingulo-opercular function during word recognition in noise for older adults with hearing loss. *Exp. Aging Res* 42, 67–82. doi:10.1080/0361073X.2016.1108784. [PubMed: 26683042]
- Vaden KI, Kuchinsky SE, Cute SL, Ahlstrom JB, Dubno JR, Eckert MA, 2013. The cingulo-opercular network provides word-recognition benefit. *J. Neurosci* 33, 18979–18986. doi: 10.1523/JNEUROSCI.1417-13.2013. [PubMed: 24285902]
- Vaden KI, Muftuler LT, Hickok G, 2010. Phonological repetition-suppression in bilateral superior temporal sulci. *Neuroimage* 49, 1018–1023. doi: 10.1016/j.neuroimage.2009.07.063. [PubMed: 19651222]
- Vaden KI, Teubner-Rhodes S, Ahlstrom JB, Dubno JR, Eckert MA, 2017. Cingulo-opercular activity affects incidental memory encoding for speech in noise. *Neuroimage* 157, 381–387. doi:10.1016/j.neuroimage.2017.06.028.
- Van Maanen L, Brown SD, Eichele T, Wagenmakers EJ, Ho T, Serences J, Forstmann BU, 2011. Neural correlates of trial-to-trial fluctuations in response caution. *J. Neurosci* 31, 17488–17495. doi:10.1523/JNEUROSCI.2924-11.2011. [PubMed: 22131410]
- Weibull W, 1951. A statistical distribution function of wide applicability. *J. Appl. Mech* Vol. 18, 293–297.
- Wickelgren WA, 1977. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol.* 41, 67–85. doi:10.1016/0001-6918(77)90012-9.
- Wild CJ, Yusuf A, Wilson DE, Peelle JE, Davis MH, Johnsrude IS, 2012. Effortful listening: The processing of degraded speech depends critically on attention. *J. Neurosci* 32, 14010–14021. doi:10.1523/JNEUROSCI.1528-12.2012. [PubMed: 23035108]
- Wilke M, 2012. An alternative approach towards assessing and accounting for individual motion in fMRI timeseries. *Neuroimage* 59, 2062–2072. doi: 10.1016/j.neuroimage.2011.10.043. [PubMed: 22036679]
- Wingfield A, 1996. Cognitive factors in auditory performance: Context, speed of processing, and constraints of memory. *J. Am. Acad. Audiol* 7, 175–182. [PubMed: 8780990]
- Wingfield A, Grossman M, 2006. Language and the aging brain: Patterns of neural compensation revealed by functional brain imaging. *J. Neurophysiol* 96, 2830–2839. [PubMed: 17110737]
- Wingfield A, Tun PA, 2001. Spoken language comprehension in older adults: Interactions between sensory and cognitive change in normal aging. *Semin. Hear* 22, 287–301.
- Worsley KJ, 2006. FMRISTAT: A general statistical analysis for fMRI data.
- Yee DM, Crawford JL, Lamichhane B, Braver TS, 2021. Dorsal anterior cingulate cortex encodes the integrated incentive motivational value of cognitive task performance. *J. Neurosci* 41, 3707–3720. doi:10.1523/JNEUROSCI.2550-20.2021. [PubMed: 33707296]



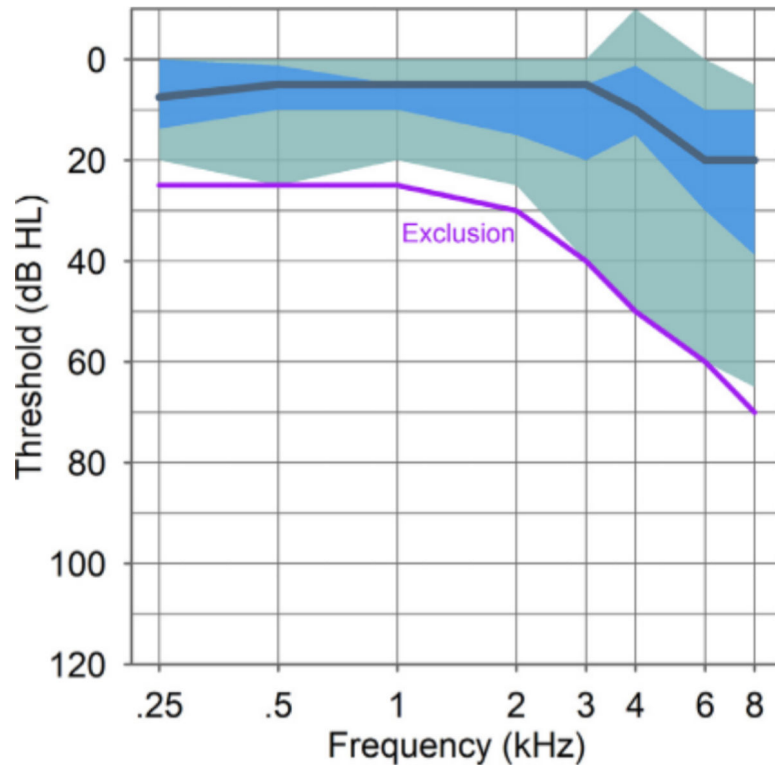
- Zekveld AA, Rudner M, Johnsrude IS, Heslenfeld DJ, Ronnberg J, 2012. Behavioral and fMRI evidence that cognitive ability modulates the effect of semantic context on speech intelligibility. *Brain Lang.* 122, 103–113. doi:10.1016/j.bandl.2012.05.006. [PubMed: 22728131]
- Zekveld AA, Heslenfeld DJ, Festen JM, Schoonhoven R, 2006. Top-down and bottom-up processes in speech comprehension. *Neuroimage* 32, 1826–1836. [PubMed: 16781167]

Author Manuscript

Author Manuscript

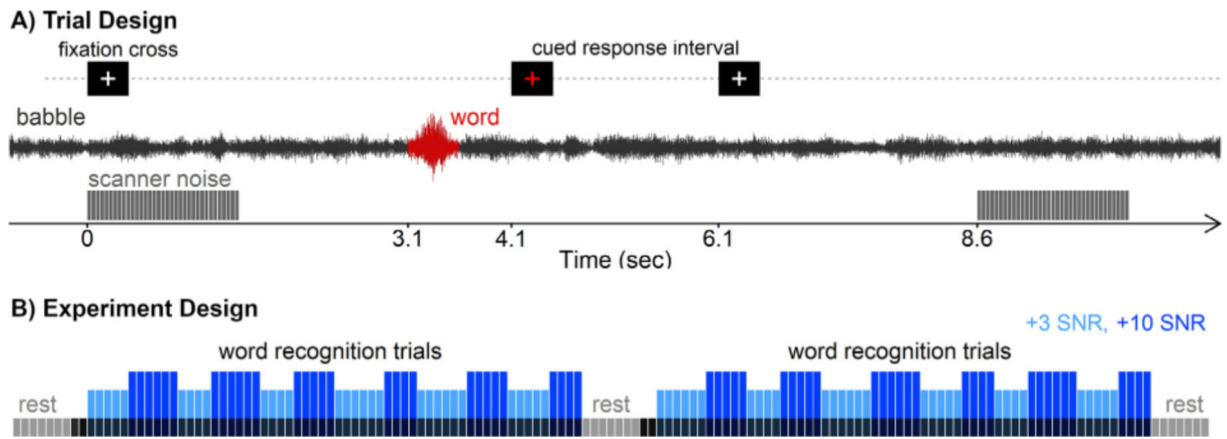
Author Manuscript

Author Manuscript

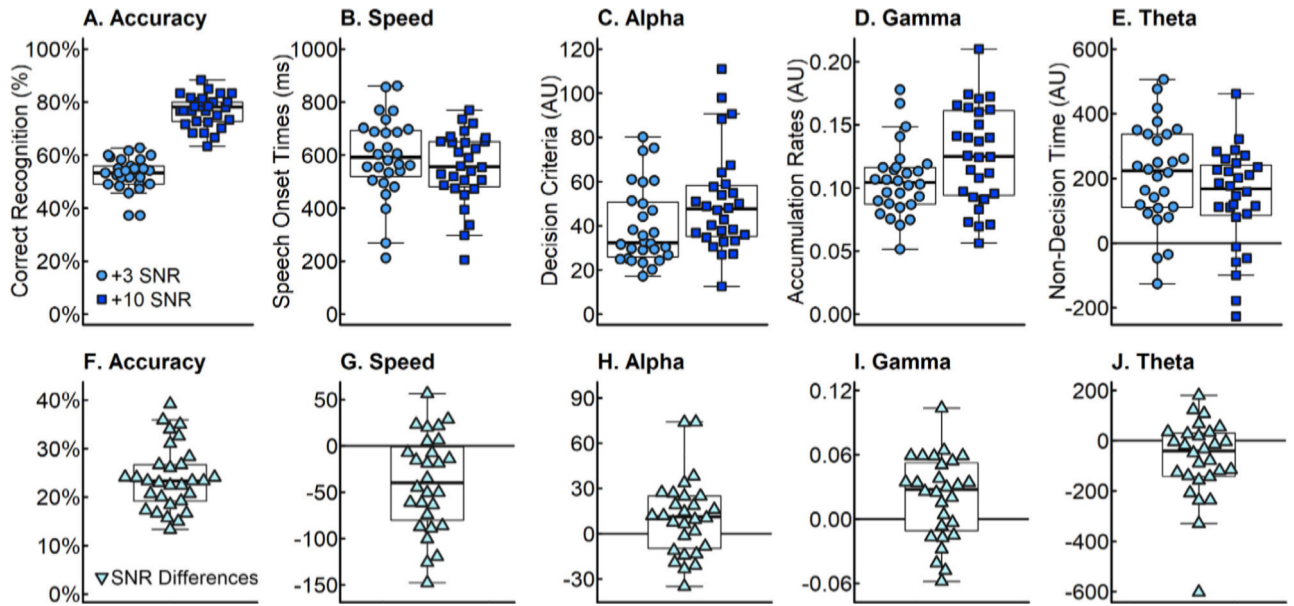


**Fig. 1.**

Better ear pure-tone thresholds for each participant are shown with 25% quantiles (shaded regions). The median threshold at each frequency for the study participants is indicated with a dark gray line. The purple line shows the exclusion cutoff threshold at each frequency based on a modified Articulation Index calculation (Dubno et al., 2008), which minimized effects related to audibility differences (described in the text).

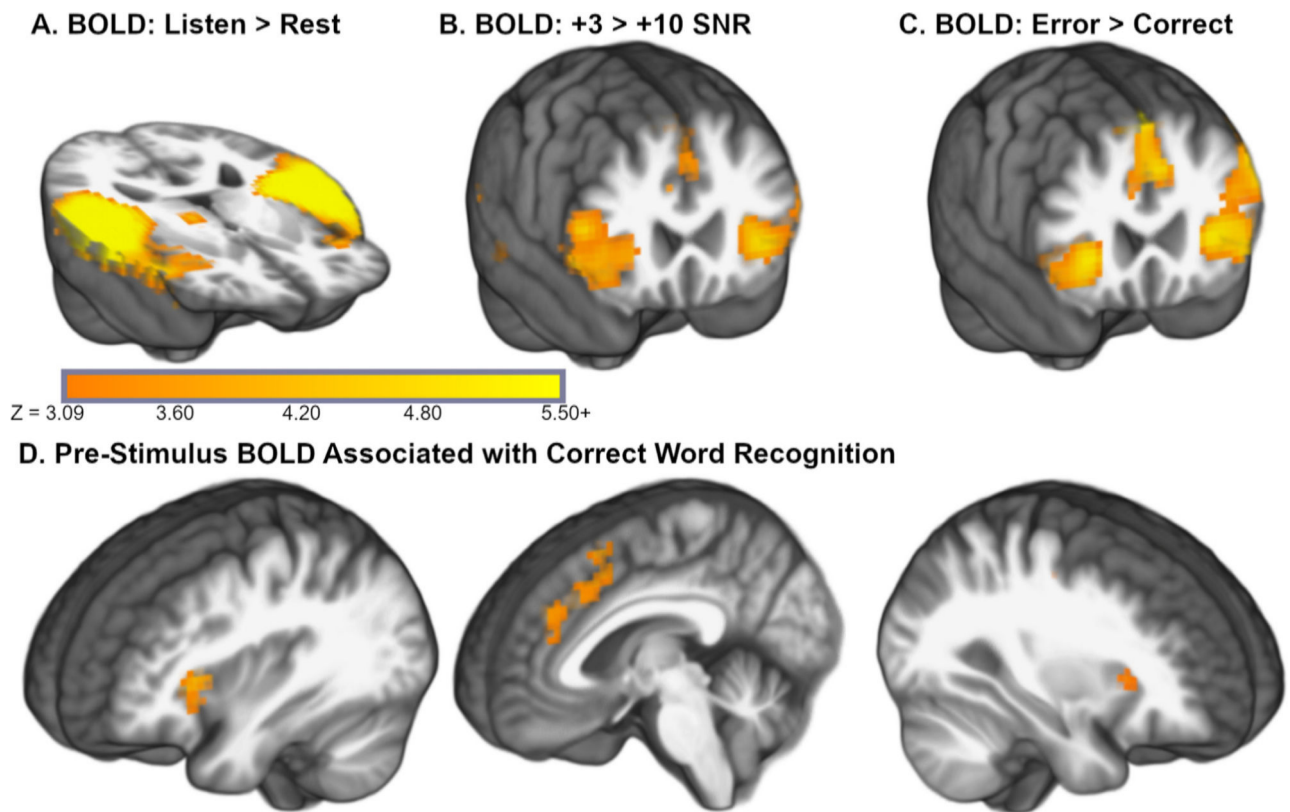
**Fig. 2.**

A) During the sparse acquisition fMRI experiment, each word was presented between scans at +3 or +10 dB SNR in multitalker babble. The crosshair onscreen changed color from white to red to visually cue the response interval (4.1 to 6.1 s). B) A mixed block design was used, with 60-word recognition trials in the +3 dB SNR and 60 trials in the +10 dB SNR, which were presented in alternating sets of four to six trials. There were no cues presented for SNR condition or task feedback. There were 7 TRs of rest before and after each block of task trials. A continuous multitalker babble was presented throughout the speech recognition task blocks and preceded each block by two TRs (black).

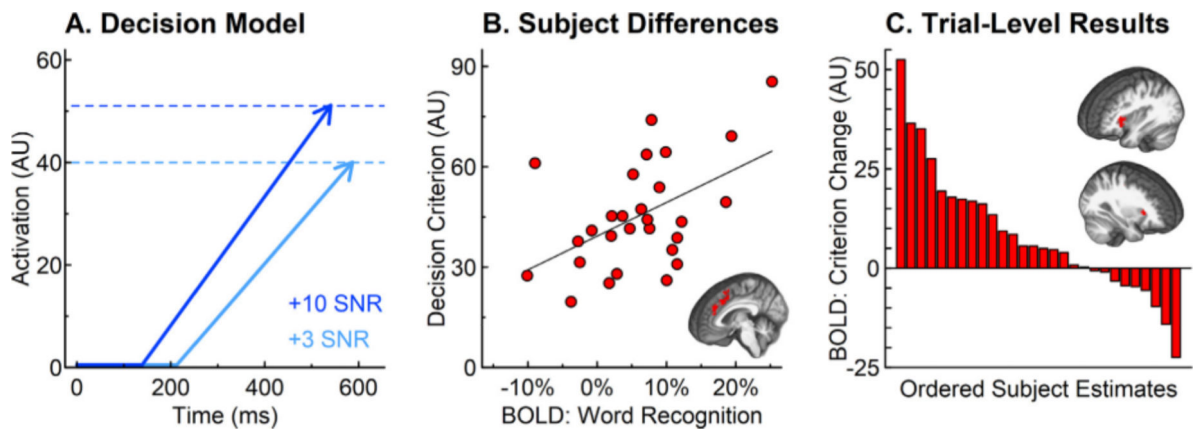


**Fig. 3.**

During the speech recognition in noise task, performance and perceptual decision-making showed significant SNR-related differences. Relative to the +10 dB SNR condition, the +3 dB SNR was associated with A ) lower percent correct word recognition (Model #1), B) longer average speech onset times (Model #2), C) lower decision criteria (alpha), D) slower evidence accumulation (gamma), and E) longer non-decision times (theta; Model #5). Participant-level SNR differences in F-J are shown below the corresponding values in A-E, calculated by subtracting values in the +3 dB SNR from the +10 dB SNR. The Shifted Wald model parameter estimates are plotted in arbitrary units (AU) on the y-axes.

**Fig. 4.**

A) The word recognition task significantly increased activity throughout superior temporal regions and cingulo-opercular regions of cortex. B) Lower SNR was associated with higher BOLD contrast in cingulo-opercular regions, as well as posterior superior temporal sulci. C) On trials with recognition errors, there was significantly increased BOLD contrast in cingulo-opercular regions and left dorsolateral prefrontal cortex (Model #3). D) Elevated cingulo-opercular BOLD contrast prior to a word presentation was also associated with increased likelihood for correct word recognition in noise (Model #4). The colored voxels in each subplot show significant BOLD effects. A voxel statistic threshold of  $Z = 3.09$ , uncorrected  $p = 0.001$  was used in combination with a Family-Wise Error corrected  $p < 0.05$  extent threshold, such that significant clusters included more than 20 voxels.



**Fig. 5.**

A) The Shifted Wald model showed that participants used significantly higher decision criteria in combination with faster accumulation rates and shorter non-decision times in the +10 dB SNR compared to the +3 dB SNR. Model parameter estimates were averaged for each SNR across participants for this illustration, with activation shown in model-based arbitrary units on the y-axis and time in milliseconds on the x-axis (Model #5). B) Participants with higher, more cautious decision criteria across SNR conditions demonstrated greater word recognition benefit from pre-stimulus BOLD contrast (x-axis: BOLD-related increase in percent word recognition; y-axis: decision criteria values in arbitrary units). The dorsal cingulate region (red) demonstrated a significant association between larger prestimulus activity effect sizes and higher criteria. Fitted regression estimates indicated that on average correct recognition was  $6.0 \pm 8.0\%$  higher for trials with high pre-stimulus dorsal cingulate activity (top 25% normalized BOLD contrast) compared to low pre-stimulus activity (bottom 25%; Vaden et al., 2013; Model #6). C) Trials with high pre-stimulus BOLD contrast in the left frontal operculum and bilateral insula (shown in red) yielded significantly more cautious criteria, compared to trials with low pre-stimulus BOLD contrast. Subject-level criteria differences for high versus low pre-stimulus BOLD are shown in descending order (y-axis: estimated criteria differences, arbitrary units; x-axis: individual participants; Model #7).



Table 1

BOLD changes during the word recognition task.

Statistical contrasts and cluster regions	Peak Z	Extent	MNI
<b>Task dependent BOLD: listening &gt; rest</b>			
R. posterior to middle superior temporal gyrus, R. precentral gyrus, R. inferior frontal sulcus, R. insula and frontal operculum, R. anterior middle frontal gyrus	8.47	2540	57, -9, 2
L. posterior to middle superior temporal gyrus, L. precentral gyrus, L. precentral sulcus, L. insula and frontal operculum	8.14	1563	-56, -18, 6
Dorsal middle cingulate, paracingulate, medial superior frontal gyrus	4.94	324	2, 16, 40
R. inferior occipital gyrus	4.74	122	28, -85, -6
L. inferior occipital gyrus	4.51	93	-27, -94, -9
R. inferior temporal gyrus, R. superior posterior cerebellum	4.30	40	26, -60, -22
Anterior cerebellum, primary fissure	3.76	35	-11, -59, -19
<b>Speech in poorer SNR (BOLD: +3 &gt; +10 dB SNR)</b>			
R. inferior frontal gyrus, R. anterior insula	5.10	396	48, 38, -5
L. anterior insula, L. inferior frontal gyrus, L. anterior inferior frontal sulcus	5.02	278	-39, 22, -5
L. posterior to middle superior temporal sulcus	4.78	98	-66, -30, 7
Dorsal paracingulate cortex, medial superior frontal gyrus	4.51	170	8, 19, 63
R. middle superior temporal sulcus	4.20	47	57, -22, -3
R. posterior superior temporal sulcus	3.88	49	65, -39, 16
<b>Recognition errors (BOLD: incorrect &gt; correct)</b>			
Dorsal paracingulate cortex, medial superior frontal gyrus	5.97	448	1, 16, 59
R. insula, R. frontal operculum	5.22	174	34, 24, -1
L. insula, L. frontal operculum, L. inferior frontal sulcus, L. middle frontal gyrus, L. precentral sulcus, L. precentral gyrus, L. anterior inferior frontal sulcus	5.16	708	-35, 21, 2
L. caudate	4.02	25	-14, 8, 4
L. posterior superior temporal sulcus	3.63	22	-49, -37, -1
L. anterior middle frontal gyrus	3.57	23	-25, 65, 23
<b>BOLD prior to correct speech recognition</b>			
L. superior temporal sulcus	4.61	70 [35]	-54, -33, 1
L. anterior insula, L. frontal operculum	4.48	63 [12]	-35, 21, 6
R. central sulcus	4.18	50 [48]	42, -16, 46
Dorsal paracingulate	4.16	138 [123]	2, 36, 34
R. precentral sulcus	3.94	31 [26]	43, 6, 45

<b>Statistical contrasts and cluster regions</b>	<b>Peak Z</b>	<b>Extent</b>	<b>MNI</b>
R. anterior insula, R. frontal operculum	3.76	29 [3]	38, 22, 6

*Notes:* MNI: Montreal Neurological Institute coordinates; L: left, R: right, otherwise bilateral. Significant results were based on a  $pFWE = 0.05$  cluster extent in combination with a cluster-defining threshold of  $Z = 3.09$ ,  $pLNC = 0.001$ . Cluster extent is presented as the number of voxels per cluster (voxel dimensions =  $3 \text{ mm}^3$ ). Numbers in square brackets indicate significant voxel counts within each cluster, when the GLMM analysis included post-correct response trials only; the voxel statistic threshold for this post-hoc test was  $Z > 2.32$ ,  $p = 0.01$ . Anatomical labels describe the entire cluster extent, not just the peak coordinate location.