Brain aging and speech perception: Effects of background noise and talker variability

Pascale Tremblay [a,b,*], Valérie Brisson [a,b], Isabelle Deschamps [c]

[a] *CERVO Brain Research Center, Québec City, QC, Canada*
[b] *Université Laval, Département de réadaptation, Québec City, QC, Canada*
[c] *Georgian College, Ontario, Canada*

ABSTRACT

Speech perception can be challenging, especially for older adults. Despite the importance of speech perception in social interactions, the mechanisms underlying these difficulties remain unclear and treatment options are scarce. While several studies have suggested that decline within cortical auditory regions may be a hallmark of these difficulties, a growing number of studies have reported decline in regions beyond the auditory processing network, including regions involved in speech processing and executive control, suggesting a potentially diffuse underlying neural disruption, though no consensus exists regarding underlying dysfunctions. To address this issue, we conducted two experiments in which we investigated age differences in speech perception when background noise and talker variability are manipulated, two factors known to be detrimental to speech perception. In Experiment 1, we examined the relationship between speech perception, hearing and auditory attention in 88 healthy participants aged 19 to 87 years. In Experiment 2, we examined cortical thickness and BOLD signal using magnetic resonance imaging (MRI) and related these measures to speech perception performance using a simple mediation approach in 32 participants from Experiment 1. Our results show that, even after accounting for hearing thresholds and two measures of auditory attention, speech perception significantly declined with age. Age-related decline in speech perception in noise was associated with thinner cortex in auditory and speech processing regions (including the superior temporal cortex, ventral premotor cortex and inferior frontal gyrus) as well as in regions involved in executive control (including the dorsal anterior insula, the anterior cingulate cortex and medial frontal cortex). Further, our results show that speech perception performance was associated with reduced brain response in the right superior temporal cortex in older compared to younger adults, and to an increase in response to noise in older adults in the left anterior temporal cortex. Talker variability was not associated with different activation patterns in older compared to younger adults. Together, these results support the notion of a diffuse rather than a focal dysfunction underlying speech perception in noise difficulties in older adults.

1. Introduction

Speech is an extraordinarily complex human behaviour. In natural conversations, speech rate can go up to 6–9 syllables per second (Kent, 2000), that is, 150 to 200 words per minute. Everyday speech recognition requires listeners to not only process the extremely fast speech signal, but also to rapidly adapt to changes in the talker and background conditions. Talkers vary in voice quality, accent, rate of speech, level of underarticulation, and voice intensity level. All these factors can reduce speech intelligibility - the proportion of a talker's output that a listener can readily understand - and make speech comprehension an enormously challenging task. Unsurprisingly, given the demanding nature of speech processing, older adults often report struggling to follow conversations especially in the presence of background noise (Working Group on Speech Understanding and Aging and Committee on Hearing, 1988). The etiology of these age-related difficulties comprehending speech in noise (SPiN) is not completely clear and the factors, other than noise itself, that affect speech processing, such as voice, accents, speech rates and levels of underarticulation, have not been completely identified. It is probably for these reasons that treatments for SPiN difficulties are scarce. Hearing aids do not reduce SPiN difficulties in most users (El-Assal and El-Gharib, 2019; Humes et al., 2002).

Presbycusis -the loss of hearing that occurs with age (Mazelova et al., 2003)- was long considered to be the main predictor of these difficulties. Indeed, numerous studies have shown that hearing decline

contributes to speech perception decline (e.g. Frisina and Frisina, 1997; Gordon-Salant and Fitzgibbons, 1993; Lash et al., 2013). However, a number of studies have shown that SPiN performance in older adults is lower than that of younger adults even when their hearing is within the normal range, as measured using pure tone audiometry (Frisina and Frisina, 1997; Kim et al., 2006; Presacco et al., 2016b; Tun, 1998). Other studies have shown worse SPiN performance in older adults when compared to hearing-matched younger adults (Fullgrabe et al., 2014), and worse SPiN performance in older adults when compared to younger adults after statistically controlling for hearing impairment (Bilodeau-Mercure et al., 2015; Fostick et al., 2013). Together, these studies suggest that SPiN difficulties may not result exclusively from dysfunctions affecting peripheral hearing but that disruptions of central processes may also be involved (Humes et al., 2012). Consistent with this notion, several studies have shown that SPiN difficulties are associated with cerebral decline across several functional networks, including auditory regions, speech processing regions of the dorsal and ventral speech streams, and the cingulo-insular network; this evidence is reviewed below.

Support for the hypothesis of a role for the dorsal stream to SPiN difficulties comes from several studies (Bilodeau-Mercure et al., 2015; Hwang et al., 2007; Salvi et al., 2002; Sheppard et al., 2011; Tremblay et al., 2018; Wong et al., 2010, 2009b). For example, Wong et al. found higher BOLD signal in older adults in the ventral precentral gyrus (PMv/M1) during a word identification task performed in quiet and in noise. Further, a positive correlation was found between signal in this region and performance in older adults, suggesting a compensatory mechanism (Wong et al., 2009b). More recently, Du et al. found a correlation between performance in a forced-choice syllable identification task and BOLD signal in the left posterior inferior frontal gyrus (IFG) in older adults, also suggestive of a compensation (Du et al., 2016). In a recent diffusion MRI study, we found that SPiN performance in aging was associated with the microstructure of the arcuate fasciculus (Tremblay et al., 2018). In addition to dorsal stream regions, ventral stream regions have also been associated with SPiN performance decline. For example, using fMRI, Wong et al. found a compensatory-like increase in activation in the bilateral middle temporal gyrus (MTG) during a word identification task performed in quiet and in noise (Wong et al., 2009a). Manan et al. reported an age difference in activation in the MTG during a word repetition in noise task (Manan et al., 2017). Using structural MRI, Wong et al. found that the volume of the left anterior IFG (pars triangularis) and thickness of the left superior frontal gyrus (SFG) in older but not younger adults predicted performance in an auditory sentence repetition task performed in quiet and noise (Wong et al., 2010). The anterior IFG is connected to the anterior temporal cortex through the uncinate fasciculus (Rauschecker and Scott, 2009), whose role includes lexical selection, sequencing, and higher-order processes (Price, 2010).

Age-related SPiN performance decline has also been associated with aging of primary and associative auditory areas. Abnormal activity patterns within the auditory cortex have been shown using electrophysiology during SPiN tasks in older adults (Presacco et al., 2016a; Presacco et al., 2016b). Others have found increased activity in older adults in the superior temporal gyrus (STG) peaking around 30 ms, indicative of increased acoustical processing (Brodbeck et al., 2018). Manan et al. reported age differences in BOLD signal in the STG during a word repetition in noise task (Manan et al., 2017). In a recent diffusion MRI study from our group, we found that SPiN performance was associated with the microstructure of the middle longitudinal fasciculus or MdLF (Tremblay et al., 2018). The MdLF runs through the temporal lobe connecting the temporal pole to the parietal lobe (Makris et al., 2009, 2013), and could be involved in basic and/or higher-order auditory processing.

Finally, there is also evidence for a role in SPiN performance for the cingulo-opercular network, which is involved in top-down attentional control and monitoring (for a review, see Peelle and Wingfield, 2016).

For instance, Vaden et al. have shown reduced activation in this network in older compared to younger adults, and a positive relationship between the amplitude of the signal in this region and word recognition performance (Vaden et al., 2015). Others have reported a relationship between activation in the anterior insula (AI) and accuracy in a sentence processing task in which rate was manipulated to reduce intelligibility (Peelle et al., 2010). Relatedly, a study from our group found that signal within the AI was lower in older adults during a syllable repetition in noise task and that this decrease was associated with worse performance (Bilodeau-Mercure et al., 2015). These declines could account for the deterioration in attention that occurs in normal aging (Alain and Woods, 1999; Helfer and Freyman, 2008; Mager et al., 2005), and could have a detrimental effect on speech perception (Meister et al., 2013). In sum, while prior studies support the notion that cerebral aging contributes to speech perception decline in aging, they do not provide a unified portrait of the mechanisms involved, given that networks involved in auditory processing, speech processing, and executive control have been associated with these deficits across different studies.

The general objective of the present study was to investigate the neurobiology of speech perception in normal aging using multimodal brain imaging. In addition to examining the effect of background noise on performance and its relationship to brain structure and activity, we examined a second factor, also known to reduce speech intelligibility - talker variability. Talker variability was manipulated because it mimics typical social interaction situations in which more than two people are involved. In such situations, listeners need to rapidly adapt to changes between talkers as well as to changing background listening conditions. Talker variability is known to have a detrimental effect on speech processing, negatively impacting lip reading accuracy (Yakel et al., 2000), and reducing phoneme identification and word recognition and recall, resulting in slower and less accurate responses, in people with hearing loss or cochlear implants (Chang and Fu, 2006; Kaiser et al., 2003; Kirk et al., 1997) as well as in normal hearing individuals (e.g. Assmann et al., 1982; Gilbert et al., 2013; Goldinger et al., 1991; Heald and Nusbaum, 2014; Martin et al., 1989; Nusbaum and Morin, 1992; Strange et al., 1976; Tamati et al., 2013; Wong et al., 2004). The effect of talker variability on speech processing is more detrimental for non-native speakers compared to native speakers (Tamati and Pisoni, 2014). Importantly, adults learn to discriminate non-native speech sounds more accurately when they are trained on a variety of talkers (i.e. high variability training) (Lively et al., 1993; Logan et al., 1991). Speech tasks with high talker variability offer a better assessment of speech capabilities in real-life situations compared to tasks with lower variability. When talker variability is high, speech processing requires *talker normalization,* which is the cognitive process through which a listener is able to access the stored representation of a same word despite extensive variation in the acoustic properties from one talker to the next (Klatt, 1986). Normalization has been associated with increased activity in regions involved in acoustical analysis (superior temporal cortex) and auditory attention (superior parietal region) (Wong et al., 2004). This normalization process may become costlier with age, which may render speech processing slower or more laborious in multi-talker contexts. This is consistent with diminished-resource accounts of cognitive aging which argue that age differences will be exaggerated on tasks that have greater cognitive demands (Reuter-Lorenz and Mikels, 2006). While a number of studies have shown that high talker variability reduces performance, the literature on talker normalization in younger compared to older adults with normal hearing is limited. One study has reported an age-related decline in word recognition performance when talker variability is high (Sommers, 1997), consistent with the notion of an increase in talker normalization cost in aging.

To address our objectives, two experiments were conducted. For Experiment 1 (behavioural), the main hypotheses were that (1) aging would contribute to speech perception decline after statistically controlling for measures of hearing loss and auditory attention and (2) speech perception performance in older adults would be more strongly affected

Table 1

Descriptive statistics (means, standard deviations and ranges) for participant characteristics.

Characteristics	Mean (SD)	Min	Max
Age	55.39 (20.51)	19	86
Education (years)	16.17 (3.71)	8	30
MoCA[a] (/30)	27.58 (1.88)	22	30
GDS[b] (/30)	2.95 (2.91)	0	12
Nb languages[c]	1.51 (0.96)	0	5
Self-reported health (/7)[d]	5.13 (1.07)	2	7
Best ear ePTA [e] (dB HL)	11.10 (9.66)	3.5	34.5
SRT[f]	36.36 (9.92)	15	59

Note. M = Mean. SD = Standard deviation of the mean.

[a] MoCA = Montreal Cognitive Assessment scale. The MOCA is a short cognitive test that is scored on a 30-point scale. Higher scores indicate better cognitive functions. Though the official cut-off score is 26, some researchers have proposed a less strict cut-off of 20/30 (Waldron-Perrine and Axelrod, 2012).

[b] GDS = Geriatric Depression Screening Scale. The GDS includes 30 questions. Each "negative" answer is worth one point; thus, a higher score indicates a more depressed state. For example, question one asks whether the person is globally satisfied with his/her life. A 'no' answer is worth one point, whereas a "yes" answer is worth no point. Participants with scores between 0 and 9 are considered normal, while scores between 10 and 19 indicate a depression, and scores between 20 and 30 indicate a severe depression.

[c] Nb languages = number of spoken languages other than native language (French).

[d] Self-reported health = self-reported general health status on a scale of 0 to 7, with 0 being lowest health level and 7 the maximal one.

[e] Best ear ePTA = For each participant, we calculated the best ear average of 0.25, 0.5, 1, 2, 3, 4 kHz (extended PTA or ePTA).

[f] SRT = Speech Reception Threshold in dB HL. SRT represents the lowest sound intensity level at which participants are able to correctly identify 50% of monosyllabic words.

by background noise and talker variability compared to younger adults. For Experiment 2, the main hypotheses were that (1) brain activity in auditory and speech-related regions during speech perception would differ in older and younger adults reflecting differences in functional cerebral organization, especially when background noise and talker variability are high, and (2) lower speech perception performance in aging would be associated with a decline (functional or structural) within auditory and speech processing areas especially when background noise and talker variability are high.

2. Experiment 1: methods

2.1. Participants

A non-probabilistic sample of 95 adults was recruited to participate in this study through emails, posters, and flyers distributed in the community in Québec City. Participants had normal or corrected-to-normal vision and no self-reported speech, voice, language, swallowing, psychological, neurological, neurodegenerative, or respiratory disorder. Participants were screened for depression using the Geriatric Depression Scale (GDS) (Yesavage et al., 1982). Cognitive level was assessed using the French version of the Montreal Cognitive Assessment scale (MoCA) (Nasreddine et al., 2003). A total of 12 participants were excluded. One was excluded because of failure to comply with task requirements, one because she was unable to comprehend the stimuli, two were excluded because they were considered to have a "severe depression" based on the GDS (19/20), three failed the MoCA (score ≤ 20) (Waldron-Perrine and Axelrod, 2012) and five had abnormal hearing levels. The characteristics of the remaining 83 participants (mean age 55.39 ± 20.51 years [19–86 years]; 40 females, mean of 16.17 ± 3.71 years of education) are reported in Table 1. All participants were native speakers of Canadian

French. 100% of the participants were schooled in French at the elementary and high school levels. English was spoken as a second language by 94% of the participants (78/83). The study was approved by the *Comité d'éthique de la recherche sectoriel en neurosciences et santé mentale, Institut Universitaire en Santé Mentale de Québec* (#360–2014).

2.2. Hearing assessment

Pure tone audiometry was performed using a clinical audiometer (AC40, Interacoustic) for each ear separately, at the following frequencies: 0.25, 0.5, 1, 2, 3, 4, 6, and 8 kHz. For each participant, we calculated the best ear average for the following frequencies: 0.25, 0.5, 1, 2, 3, 4 kHz (extended PTA or ePTA) based on previous studies on age-related hearing decline that showed the importance of including frequencies above the classical 2 KHz (Eckert et al., 2012; Mudar and Husain, 2016). The results of the hearing assessment are summarized in Table 1 and illustrated in Fig. 1. All but 8 participants had normal pure tone thresholds ≤ 25 dB hearing level (HL) from 250 to 4 kHz in at least one ear and interaural asymmetry of ≤ 15 dB HL difference at no more than 2 adjacent frequencies. Eight participants aged 58 or higher had mild-to-moderate peripheral hearing loss (ePTA ≥ 25 dB HL). Because hearing loss is a common consequence of aging, these participants were kept in the analysis. The ePTA was used as a covariate in all analyses. Hearing thresholds are illustrated in Fig. 1.

2.3. Auditory attention

All participants underwent an evaluation of attention using the Advanced Integrated Visual and Auditory Continuous Performance Test (IVA-AE) to determine if speech perception decline is associated with a decline in attention (Sandford and Turner, 1995). Participants were seated in a sound-attenuated room facing a 24-inch computer monitor and wearing high-quality headphones (DT 770 Pro, Beyerdynamic Inc. US). The test was run on a Lenovo ThinkPad W510 computer. Participants were asked to click on a mouse when a visual "3″ or an auditory "5″ was presented. Because of the contribution of sustained attention (i.e., that is the ability to concentrate on a sound source for a certain amount of time) and selective attention (i.e. the ability to focus on a single source, while suppressing others) to SPiN performance (Meister et al., 2013), as well as, more generally, speech processing (e.g. Heald and Nusbaum, 2014), and because age affects both selective (e.g. Alain and Woods, 1999; Helfer and Freyman, 2008; Mager et al., 2005) and sustained attention (Chao and Knight, 1997; Presacco et al., 2016a), two auditory scores from the IVA were used to measure these two attentional processes. Selective attention was measured using the auditory vigilance score, and sustained attention was measured using the focus score. Both scores were used in all subsequent statistical analyses as covariates. Details about the test and results are reported as Supplementary Material 1.

2.4. Speech perception

All stimuli and materials are publicly available via the Scholar Portal Dataverse (https://doi.org/10.5683/SP2/SNW3YO). An auditory sublexical speech discrimination task was used to assess speech perception. In each trial, two syllables were presented, one at a time, at an individually adjusted intensity based on participants' comfort level. The syllables were presented 200 ms apart to minimize working memory demand. Participants were asked to determine if the syllables were identical or different. We used syllables instead of words to avoid semantic and lexical priming effects, which may conceal speech perception difficulties (Samuel, 1981). The presentation of the second syllable was followed by a question mark cueing participant to respond. Participants were asked to answer as quickly as possible using a response box (RB-840 model, Cedrus, USA). The inter-trial interval was 1000 ms. If no response was made, the next trial automatically begun 2000 ms after
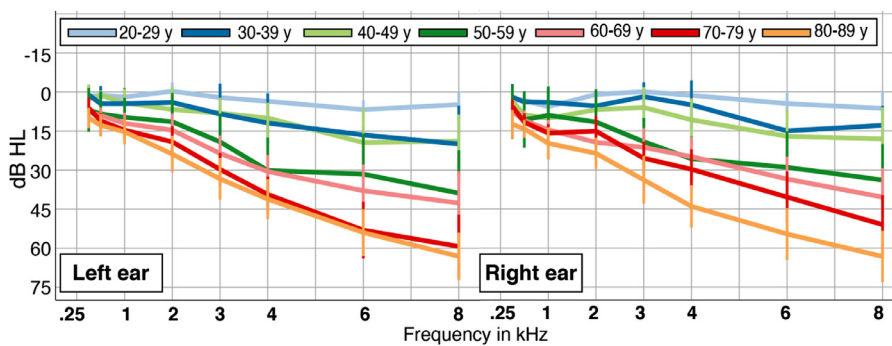
Fig. 1. Hearing thresholds. The line charts present an overview of participants' hearing thresholds for the right ear (right panel) and left ear (left panel) as a function of age and sound frequency tested. Each line in the figure represents the average thresholds for a subgroup of participants. Error bars represent the 95% confidence interval of the mean. As can be seen in the figure, hearing declines with age, especially in the higher frequency range (6 and 8 kHz).

the last stimulus was played. All stimuli were presented using Presentation Software (Neurobehavioral System, USA) through high-quality headphones (DT 770 Pro, Beyerdynamic Inc., USA), while participants were comfortably seated in a soundproof room. The pairing of the responses and button on the response box was counterbalanced across participants. Speech discrimination evaluates sensitivity to the phonetic details of native speech sounds. Unlike identification tasks, it does not require explicit categorical judgment.

The experiment included 720 experimental trials separated into 4 runs of 10 min each, separated by short breaks. Of these, 360 contained identical syllables (e.g., /fe/ vs. /fe/) and 360 contained syllables that differed by one phoneme (e.g., /pe/ vs. /ge/). The pairs were composed of 48 different syllables recorded by three different native adult male French talkers. The syllables had an average duration (mean ± SD) of 350 ± 0.05 ms. The syllables were presented either in the absence of background noise (quiet) or at a signal-noise ratio (dB SNR) of −5 (noise), which was reached by adding pink noise to the recordings, according to the following formula: dB SNR = $10\log_{10}$ (Pressure$_{signal}$/Pressure$_{noise}$), as described by Wong et al. (2008). The pairs were either produced by the same talker (low talker variability) or by different talkers (high talker variability). All stimuli were peak amplitude normalized across all talkers and noise conditions to a mean intensity of 70 dB SPL using Praat. The normalization was conducted after the noise was added to ensure equal amplitude between the Quiet and the Noise conditions. Participants were asked to ignore the talker and to focus on the syllables that were presented.

2.5. Analyses

Speech perception performance was analyzed within the framework of signal detection theory (Macmillan and Creelman, 1990). Specifically, we computed a measure of sensitivity (d-prime or d'), calculated as d' = Z(hit rate) − Z(false alarm rate). In addition to sensitivity, reaction times (RT), measured in milliseconds, were also analyzed to evaluate processing time. The normality of each dependent measure (sensitivity and RT) was assessed before and after removal of outliers defined as values 3 SD above or below the mean (6 values, representing 1.7% of the data, were removed from the response bias data and one from the d' values, representing 0.28% of the data). The RTs were square root transformed to make their distribution normal. There were no outliers. All trials containing an incorrect response were excluded from the analysis of RT.

Linear mixed model (LMM) analyses were conducted using SPSS Version 25 for Mac (IBM), separately for each dependent variable (d', RT), with background Noise level (quiet, noise) and Talker variability (low, high) as a repeated fixed factors, and Age (mean-centered) as a continuous factor. Hearing (mean-centered best ear ePTA) and two mean-centered measures of attention (auditory Vigilance and Focus) were included as between-subject continuous fixed factors. Participants were included as a random factor in the model (random intercept and random slope for Noise level and Talker variability). Model selection was

performed by removing non-significant covariates and non-significant interactions. All effects and interactions with age were kept because they were the focus of our analyses. Predictor collinearity was assessed by calculating the variance inflation factor (VIF) of each predictor in the model. VIF is a measure of collinearity among independent variables. It indexes the proportion of variance in one independent variable that is not explained by the remaining independent variables. All variables in the model had a VIF value lower than 3 (Harrison et al., 2018; Zuur et al., 2010).

3. Results

For each dependent variable (d' and RT), Q-Q plot and histograms were computed, which revealed that the residuals followed a normal distribution. The descriptive statistics for each dependent variable and the detailed results of the LMM analyses are provided in Supplementary Materials 2. The LMM analyses revealed that both dependent variables were modulated by Noise level and Talker variability, with lower sensitivity and longer RT associated with noise and high talker variability (Supplementary Materials 3). In addition, the analyses showed that speech perception performance declined with age. For sensitivity, the analysis revealed no effect of any of the attention factors, so these factors were removed from the analyses. There was a main effect of Age on sensitivity ($\beta = -0.006$, $F_{(1,91)} = 43.413$, $p = < 0.001$), with lower sensitivity with Age, as well as an interaction between Age and Noise level ($\beta = -0.009$, $F_{(1,75)} = 7.555$, $p = .007$). As can be seen in Fig. 2A, this relationship was stronger in the quiet compared to the noise condition. For RT, the analysis revealed a main effect of age ($\beta = 0.229$, $F_{(1,80)} = 10.690$, $p = .002$), with longer RT with Age, as well as an interaction between age and Talker variability ($\beta = 0.069$, $F_{(1,81)} = 11.933$, $p = .001$). As shown in Fig. 2B, the effect of talker variability on RT diminished with Age.

3.1. Interim discussion

Experiment 1 shows that SPiN performance is affected by age, the presence of background noise and by Talker variability. Specifically, while the effect of age on sensitivity was influenced by the presence of background noise, the effect of age on RT was affected by Talker variability. Unlike what was predicted, the impact of Talker variability on RT was greater in younger than in older adults. However, because RTs were strongly affected by age, it is possible that this result reflects a plateau in RT whereby, after a certain age, responses are so slow that they are no longer affected by increasing Talker variability. Because appropriate treatment for speech perception decline with age requires an understanding of the source of these difficulties, next we examined the underlying cerebral mechanisms.

4. Experiment 2

The objective of Experiment 2 was to relate speech perception abilities to brain structure and function using multimodal MRI. The main
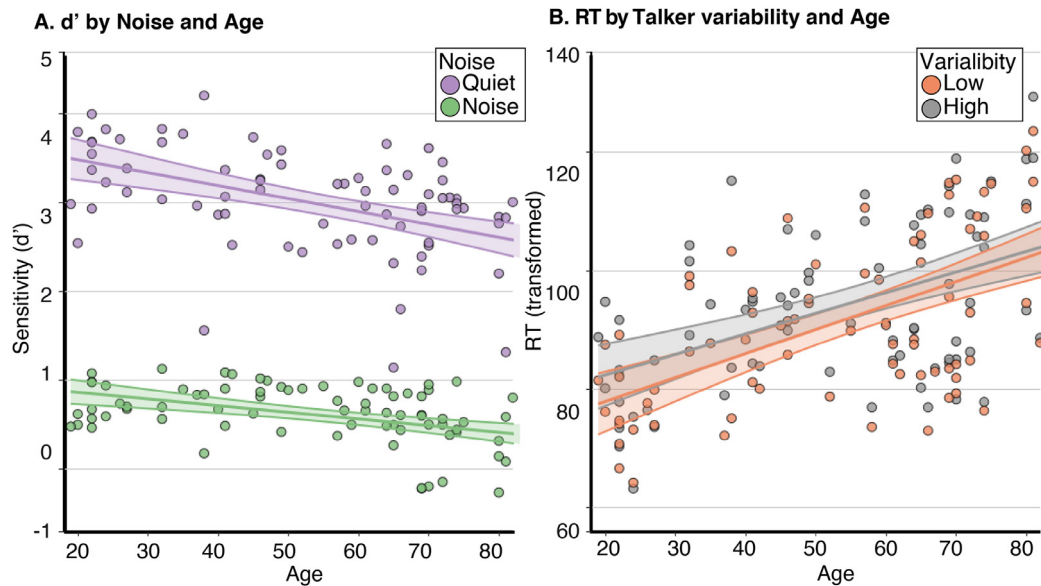
Fig. 2. Behavioural results (Experiment 1). A. The scatter plot shows the relationship between Age and sensitivity (d') values in the quiet (purple) and noise (green) conditions. The 95% confidence interval of the regression line is displayed as shaded confidence bands. B. The scatter plot shows the relationship between Age and RT (square root transformed) in the low (orange) and high (gray) talker variability conditions. The 95% confidence interval of the regression line is displayed as shaded confidence bands. See Supplementary Material 3.2 for a graph illustrating this relationship in untransformed RTs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

hypotheses were that (1) brain activity in auditory and speech-related regions - especially the STG, superior temporal sulcus (STS), PMv and the posterior IFG - during speech discrimination would differ in older compared to younger adults reflecting differences in functional cerebral organization, especially when background noise and talker variability are high, and (2) lower speech perception performance in aging, expressed as longer RT and lower sensitivity (d'), would be associated with functional and/or structural differences within auditory and speech processing areas. To test these hypotheses, a subset of participants from Experiment 1 completed a brain imaging session (Experiment 2).

5. Methods

5.1. Participants

All MRI-compatible participants from Experiment 1 were invited to participate in Experiment 2. 32 right-handed participants from Experiment 1 volunteered to participate in Experiment 2. Participants were divided into a group of 15 younger participants (mean age 29.27 ± 10.12 years [19–46 years]; 5 females, mean of 16.93 ± 2.25 years of education) and a group of 17 older participants (mean age 71.71 ± 5.79 years [65–84 years]; 3 females, mean of 16.12 ± 4.38 years of education). The groups were matched for education ($t_{(30)} = 0.648$, $p = .522$), depression level (GDS) ($t_{(30)} = 0.549$, $p = .587$) and self-reported health ($t_{(30)} = -1.742$, $p = .09$). Handedness was assessed using the Edinburgh Handedness Inventory (Oldfield, 1971). Groups characteristics are reported in Table 2. None of the participants presented with any counter indication to MRI. Experiment 2 was approved by the *Comité d'éthique de la recherche sectoriel en neurosciences et santé mentale, Institut Universitaire en Santé Mentale de Québec* (#360–2014).

5.2. Task

A shorter version of the speech discrimination task described in Experiment 1 was used as part of the Experiment 2 (with a subset of the same stimuli). The task included 96 experimental trials interleaved with 48 filler trials. A baseline condition (crosshair fixation) was included and interleaved with experimental trials. Within each run, the order of

the conditions and the number and duration of baseline trials were optimized using OPTseq2 (http://surfer.nmr.mgh.harvard.edu/optseq/). 50% of the experimental trials contained identical syllables (e.g., /fe/ vs. /fe/) and 50% contained syllables that differed by one phoneme (e.g., /pe/ vs. /ge/). The syllables were presented in quiet or at a signal-noise ratio (dB SNR) of −5 (noise). All stimuli were presented during the (silent) delay in acquisition, using Presentation Software (Neurobehavioral System, USA) through high-quality MRI-compatible stereo electrostatic earplugs (Nordic Neurolab, Norway), which provide 30 dB of sound attenuation. In each trial, two syllables were presented during a 2700 ms delay in image acquisition (see next section), separated by 200 ms. The presentation of the first sound began 200 ms after the beginning of the delay to allow for echo to resume. The presentation of the second syllable was followed by a question mark cueing participant to respond. Participants were asked to answer as quickly as possible using an MRI compatible response device response box (Celeritas®, Psychology software tools, USA). If no response was made the trial ended after 1000 ms and the new image acquisition begun.

5.3. MRI data acquisition

The data were acquired on a whole-body Philips 3.0 Tesla Achieva TX at the Clinic IRM Québec-Mailloux in Québec City. Structural MR images were acquired with a 3D T1-weighted MPRAGE sequence (TR = 8.2 ms, TE = 3.7 ms, FoV= 250 mm, flip angle = 8°, 256 × 256 matrix, 180 slices/volume, slice thickness = 1 mm, no gap). Single-shot EPI BOLD functional images were acquired using a sparse sampling protocol with parallel imaging (SENSE = 2). Each functional EPI run began with six dummy scans to allow the magnetization to stabilize to a steady state. Two runs of 109 functional images were acquired with a sparse sampling parallel acquisition technique (Gracco et al., 2005) (43 interleaved axial slices, 3 mm isotropic, no gap); TR = 5000 ms; acquisition time = 2300.5 ms, delay in TR = 2700 ms; TE = 30 ms; FOV = 240 × 240 mm; 80 × 80 matrix; Flip angle: 90°. Throughout the procedure, each participant's head was immobilized using a set of cushions and pads. We have recently used a similar protocol to investigate speech perception mechanisms in aging (Bilodeau-Mercure et al.,

Table 2

Descriptive Statistics (Means, Standard Deviations and Ranges) for each group.

Group	Younger adults			Older adults		
Characteristics	Mean (SD)	Min	Max	Mean (SD)	Min	Max
Age	29.27 (10.12)	19	46	71.71 (5.79)	65	84
Handedness[a] (/100)	84.86 (18.09)	50	100	96.06 (8.51)	68	100
Education (years)	16.93 (2.316)	13	21	16.12 (4.38)	10	30
MoCA (/30)	28.67 (1.13)	27	30	27.12 (1.9)	24	30
GDS (/30)	3.6 (3.62)	0	12	2.94 (3.17)	0	8
Nb languages	2.07 (0.70)	1	3	.94 (0.75)	0	3
Self-reported health (/7)	4.83 (1.04)	2	6	5.47 (1.02)	3	7
Best ear ePTA (dB HL)	2.09 (7.02)	−3.33	25.67	16.41 (9.85)	−1.67	43.33
SRT	26.97 (8.25)	15	52	42.44 (8.41)	30	59

Note = See Table 1 for the description of all measures. A. A positive handedness score indicates a right-hand preference while a negative score indicates a left-hand preference. The higher the number the stronger the preference.

2015). Diffusion MRI images were also acquired; these have been published elsewhere and will not be discussed here (Tremblay et al., 2018).

5.4. Analyses

Behavioural data analyses. Statistical analyses focused on two dependent measures: d' and RT. Each dependent variable was assessed for normality and the presence of extreme data. The data were normally distributed and there were no extreme values. Trials containing an incorrect response were excluded from the analysis of RT.

Linear mixed model (LMM) analyses were conducted separately for d' and RT using SPSS Version 25 for Mac (IBM) with REML estimation of covariance parameters, with Noise level (quiet, noise) and Talker variability (low, high) as a repeated fixed categorical factors, and Group (Younger, Older) as a categorical between-subject factor. All categorical factors were dummy coded. Hearing (mean-centered best ear ePTA) and two mean-centered measures of attention (auditory vigilance and focus) were included as covariates. Participants were included as a random factor in the model (random intercept and random slope for Noise level and Talker variability). Model selection was performed by removing non-significant covariates and non-significant interactions. Group effects and interactions were kept because they were the focus of the analyze. Predictor collinearity was assessed by calculating the variance inflation factor (VIF) of each predictor in the model. All variables had a VIF value lower than 3 (Harrison et al., 2018; Zuur et al., 2010). Residuals were visually inspected to assess normality using histograms and Q-Q plots.

Structural MRI data analyses. Reconstruction of participants' cortical surface was carried out with the FreeSurfer Vversion 5 image analysis suite (Dale et al., 1999; Fischl et al., 1999a, 2004). The pipeline includes motion correction, intensity normalization, removal of non-brain tissue (skull and meninges stripping), segmentation of gray and white matter, followed by a triangular tessellation of the gray matter white matter boundary, automated topology correction (Fischl et al., 2001; Segonne et al., 2007) and surface deformation smoothing following intensity gradients to optimally place the gray/white and gray/cerebrospinal fluid borders at the location where the greatest shift in intensity defines the transition to the other tissue class (Dale et al., 1999; Fischl and Dale, 2000). An affine transformation to Talairach space is created in parallel that can be used to report Talairach coordinates (note that the analyses were performed on *native-space* data). Next, surface inflation (Fischl et al., 1999a), and registration to a spherical atlas based on individual cortical folding patterns to match cortical geometry across subjects were performed (Fischl et al., 1999b). Cortical thickness is calculated as the closest distance from the gray/white boundary to the gray/CSF boundary at each vertex on the tessellated surface (Fischl and Dale, 2000). The validity of cortical thickness has been demonstrated (Kabani et al., 2001). Cortical thickness maps were

exported to SUMA (Saad et al., 2004) for group-level surface-based analysis.

To examine whether the relationship between age and speech perception performance is mediated by brain structure within auditory and speech processing areas, a two-step analysis was conducted. First, we identified all regions in which there was an age effect on cortical thickness (older - younger) (AFNI 3dttest++ program). A cluster size of 546 vertices was deemed significant at $p <0.01$ following a Monte Carlo simulation (1000 repetitions) with an individual threshold of $p <0.01$ at the voxel level. Next, we extracted the average regression coefficient (unstandardized beta weight) for each cluster identified through the whole-brain group analyses using a mask of the group result (AFNI 3dROIstats program) and examined the direction of the effects using bar graphs made using BioVinci version 1.1.5 (BioTuring Inc., USA). A series of simple mediation analyses using ordinary least squares path analysis was then conducted separately for all regions showing a group difference. Based on the behavioural results of Experiment 2, the mediation analyses focused on the main effect of Group and on the interaction between Group and Noise. For the interaction effect, a Noise contrast (d' Noise – d' Quiet) was computed and used as the dependent variable. In all analyses, the dichotomous predictor variable was Group (Younger, Older). The mediator variable was cortical thickness. Based on the results of the behavioural data analysis for Experiment 2, hearing (ePTA) was included as a covariate on all analyses involving d', and vigilance was included as a covariate on all analyses involving RT. The mediation models are illustrated in Fig. 3A. A percentile bootstrap confidence interval for the indirect effect was constructed based on 10,000 samples to determine if indirect effects were significantly different from zero. For a description of the analysis and its implementation in SPSS, see Hayes (2013, 2018). Mediation analyses were conducted using the PROCESS 3.0 macro (model #4) for SPSS (Hayes, 2008, 2013). We report the unstandardized beta coefficient, the standard error of the coefficient and the confidence interval for each path (a, b, c' and ab) for the analyses where an indirect effect was found in Table 3.

Functional MRI data analyses. All analyses were conducted using AFNI version 18 (Cox, 1996). First, MRI data were converted to the AFNI format and visually inspected for artefacts. The time series were aligned to the first functional run, motion-corrected, slice timing corrected, corrected for outliers, and mean-normalized. Functional volumes acquired during excessive motion (>1 mm) were excluded from the analyses. Following pre-processing, subject-level regression analyses were conducted. An ordinary least square regression approach was used to analyze subject data. Separate regressors were created for each participant for the correct and incorrect trials (regressors for incorrect trials were not analyzed). Additional regressors were the mean, linear, and quadratic trend components, and the 6-motion parameters (x, y, z and roll, pitch, and yaw). A 2-parameter gamma basis function (AFNI model with a duration of about 12 s 'SPMG2' was used. This regression model was fit to each run separately (separate baseline models for each run).

…

A. Mediation models for cortical thickness ROIs
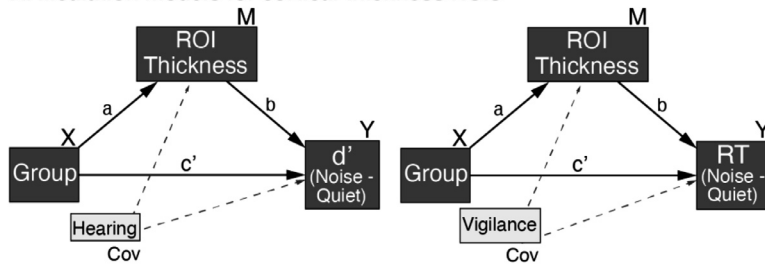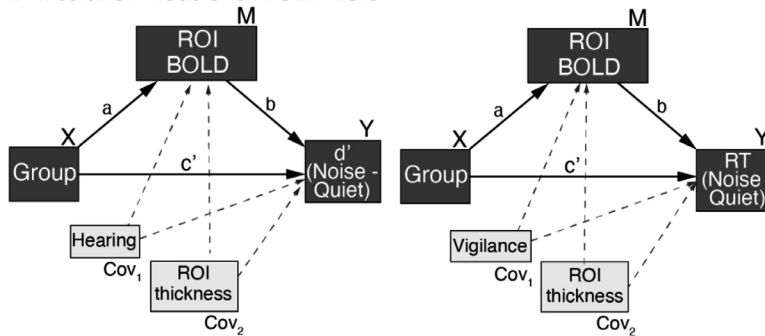

B. Mediation models for BOLD ROIs

Fig. 3. Mediation models. A. Models used for the analysis of cortical thickness. B. Models used for BOLD signal analyses. X = the predictor variable. Y = the dependent variable. M = the mediator variable. Cov = covariate. The a-path represents the effect of Group on either cortical thickness (A) or BOLD signal (B). The a-coefficient tells us that two cases that differ by one unit on X (that is, older adults) differ on the mediator by that amount (the a-coefficient). The b-path represents the effect of the mediator on performance (d' or RT) when X is constant (same group). The c'-path represents the direct effect of Group on d' when the mediator is constant. The ab-path is the indirect effect (mediation). It is the product of the a and b paths. It represents the effect of Group on performance through its effect on the mediator (thickness or BOLD).

Following regression, the outputs from the subject-level regression were converted to the surface space (3dVol2Surf), aligned to the anatomical data and smoothed on the surface to achieve a target smoothing value of 6 mm using a Gaussian FWHM filter. Smoothing on the surface as opposed to the volume ensures that white matter voxels are not included, and that functional data located in anatomically distant locations on the cortical surface are not averaged across sulci (Argall et al., 2006; Jo et al., 2007). The resulting images were visually inspected for quality control. The Destrieux atlas was used for cortical parcellation (Destrieux et al., 2010; Fischl et al., 2004) and to localize changes in BOLD signal in the cortex.

Group-level analyses were conducted in several steps. First, we examined the basic speech perception network (without noise) as a validity check, separately for each group. For each subject, an average of the two conditions without noise (with one and two talkers) was computed and compared against zero using a one-sample t-test (AFNI 3dTtest++). The main focus of the analyses was on the interaction effects between Group and Noise Level, and Group and Talker variability on BOLD signal. The main effect of Group on the BOLD signal was not investigated because unspecific processes such as vascular changes could contribute to any observed group differences (D'Esposito et al., 2003; Garrett et al., 2017). We therefore examined the interaction between Group and Noise level using two-sample t-tests (AFNI 3dTtest++). The same analyses were conducted for Talker variability. Next, we examined the 3-way interaction between Group, Noise and Talker variability. All analyses were corrected for multiple comparisons. The Monte Carlo simulation procedure implemented in Freesurfer (mri_glmfit) was used to identify significant clusters of activated vertices taking into account the number of voxels and the amount of smoothing. The result indicated that, for an individual vertex threshold of $p < 0.05$, corrected for multiple comparisons to achieve a family-wise error (FWE) rate of $p < 0.01$, the appropriate cluster size was ≥ 546 vertices. To examine the pattern of significant BOLD responses, we extracted, for each participant, the average regression coefficient (beta weights) for all the clusters identified through the whole-brain group analyses using a mask of the group result (AFNI 3dROIstats program).

To test the hypothesis that age-related decline in speech discrimination is associated with differences in activity in auditory and speech processing areas, we conducted a series of simple mediation analyses using ordinary least squares path analysis on all regions showing an

interaction effect with Group (either Group by Noise Level; Group by Talker variability or Group by Noise Level by Talker Variability). Noise difference scores (d'/RT $_{Noise}$ − d'/RT $_{Quiet}$) were computed and used as dependent variables (Y). Since no region exhibited a Group by Talker interaction, no mediation was conducted for this contrast. The dichotomous predictor variable was Group (Younger, Older), and the continuous mediator variable was the BOLD signal within the ROI. In all analyses, cortical thickness was used as a covariate to ensure that effects are related only to brain activity. Further, for d', hearing (ePTA) was included as a covariate based on the results of the behavioural data analysis for Experiment 2. For RT, vigilance was used instead of hearing for the same reason. The mediation models are illustrated in Fig. 3B. A percentile bootstrap confidence interval for the indirect effect was constructed based on 10,000 samples to determine if indirect effects were significantly different from zero. We report the unstandardized beta coefficient, the standard error of the coefficient and the confidence interval for each path (a, b, c' and ab) for the analyses where an indirect effect was found in Table 4.

6. Results

6.1. Behavioural results

For each dependent variable (d' and RT), Q-Q plot and histograms were computed, which revealed that the residuals followed a normal distribution. The descriptive statistics for each dependent variable and the detailed results of the LMM analyses are provided in Supplementary Material 4. The LMM analyses revealed that both variables were modulated by Noise level, with lower d' and longer RT. In addition, d' was also modulated by Talker variability, with high variability associated with lower d' (Supplementary Material 5). In terms of Group effects, the analysis revealed a significant main effect of Group on d' ($\beta = -0.014$, $F_{(1,25)} = 4.656$, $p = .041$), with lower sensitivity overall in older compared to younger adults (Fig. 4A). There was also an interaction between Group and Noise level on d' ($\beta = 1.242$, $F_{(1,27)} = 9.953$, $p = .004$). As shown in Fig. 4B, a series of univariate F-tests revealed a Group difference in the quiet condition ($F_{(1,48)} = 11.067$, $p = .002$), with lower d' for older compared to younger adults, but not in the Noise condition ($F_{(1,23)} = 0.096$, $p = .759$). There was no interaction between Group and Talker variability (Fig. 4C), but the analysis revealed a
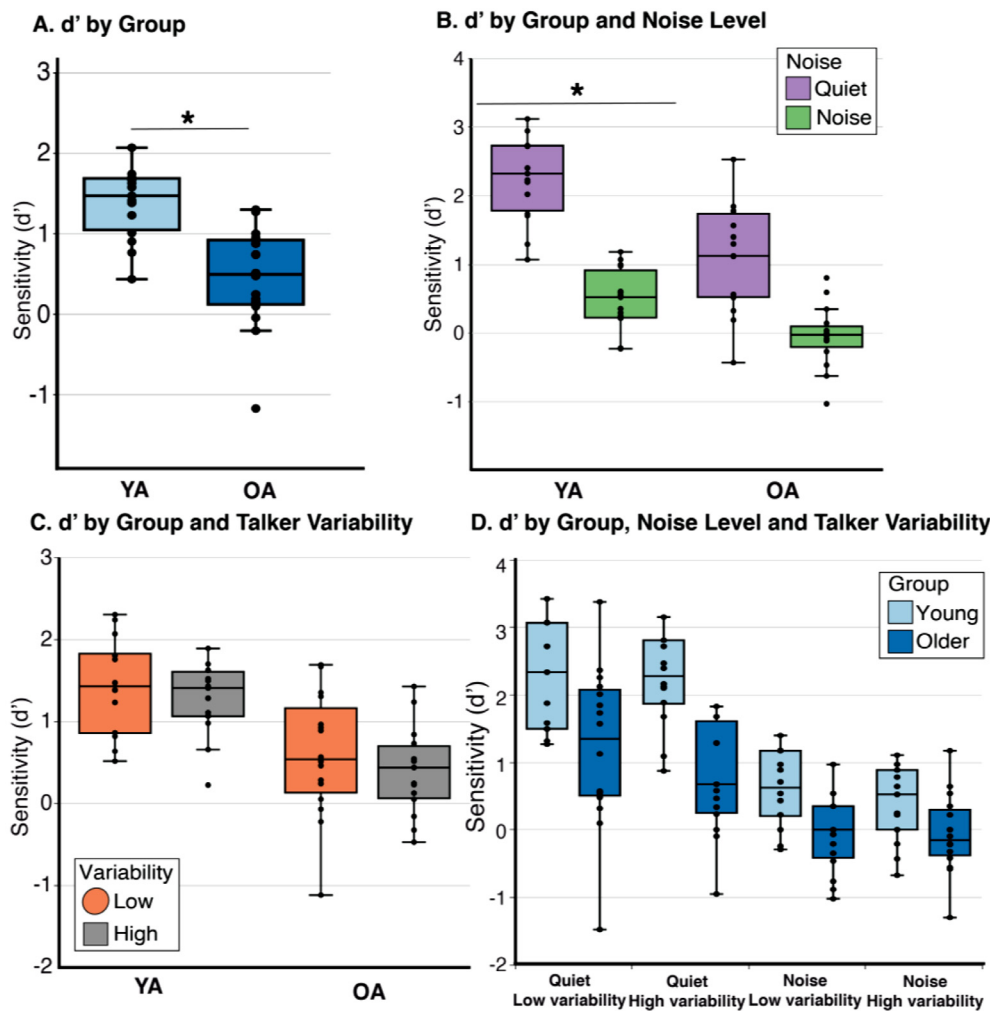
Fig. 4. Behavioural results (Experiment 2). A. The boxplots show the overall sensitivity (d') separately for the younger (YA) and the older groups (OA). B. The boxplots show d' scores separately for the younger and the older groups, for the quiet (purple) and noise (green) conditions. C. The boxplots show d' scores separately for the younger and the older groups, for the low talker variability (orange) and high talker variability (gray) conditions. D. The boxplots show d' scores separately for the younger and the older groups, separately for each condition. In all boxplots, the median is represented by a line traced at the center of the box. The whiskers range from the minimum (Q1 −1.5*interquartile range) and maximum (Q3 + 1.5*interquartile range) values. The black dots represent individual scores. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3-Way interaction between Group, Noise level and Talker variability on d' ($\beta = -0.142$, $F_{(1,20)} = 3.566$, $p = .047$) (Fig. 4D). A series of univariate F-tests revealed that the largest group difference was found in the quiet condition with High talker variability condition ($F_{(1,28)} = 15.265$, $p \leq 0.001$). RTs were not affected by Group.

6.2. Structural MRI results

As shown in Fig. 5 and detailed in Supplementary material 6, the structural analyses revealed group differences (older < younger) in cortical thickness in several areas including the bilateral superior temporal cortex (STC), including the STS, STG, planum temporale (PT), and primary auditory cortex, as well as in the dorsal anterior insula (dAI), precentral and postcentral gyri, and inferior and medial frontal areas.

To address our objective of relating age differences in speech perception performance to brain structure, the analyses focused on relating performance (overall and for the noise contrast) to cortical thickness results through simple mediation analyses. For d', the mediation analyses revealed that Group indirectly influenced overall d' (negatively) through its effect on cortical thickness in several regions: the bilateral STC, bilateral IFG/dAI and the right superior frontal gyrus (SFG). In all analyses, older adults had lower cortical thickness, and participants with thicker cortex had better d'. The bootstrap confidence intervals for the indirect effects based on 10 000 samples were entirely below zero. There was no evidence that Group influenced d' independently of its effect on cortical thickness. The indirect effect was therefore detrimental to performance. The second set of analyses focused on Noise difference

score on d'. These analyses revealed that Group indirectly influenced d' through its effect on cortical thickness in the left posterior cingulate gyrus (pCG). In this analysis, older participants had lower cortical thickness in the pCG. Those with thicker cortex in the pCG showed increased difference in d' (more differentiated performance) as a function of noise. The bootstrap confidence intervals for the indirect effects based on 10 000 samples were entirely above zero. There was, however, evidence that Group influenced the Noise difference score independently of its effect on cortical thickness (i.e. partial mediation). The details of the significant mediations are reported in Table 3.

For RT, the mediation analyses revealed that Group indirectly influenced overall RT through its effect on cortical thickness in the right IFG/dAI. In this ROI, older participants had lower cortical thickness compared to younger participants. Participants with thicker cortex were faster (shorter RT). The bootstrap confidence intervals for the indirect effects based on 10 000 samples were entirely above zero. The indirect effect was therefore detrimental to performance. There was no evidence that Group influenced d' independently of its effect on cortical thickness. The next set of analyses focused on Noise difference on RT (RT $_{Noise}$ − RT $_{Quiet}$). These analyses revealed that Group indirectly influenced RT through its effect on cortical thickness in several regions: the bilateral STC, bilateral IFG/dAI and the left parietooccipital sulcus (POS). In all analyses, older adults had lower cortical thickness. Participants with thicker cortex showed increased difference in RT as a function of noise (more differentiated performance). The bootstrap confidence intervals for the indirect effects based on 10 000 samples were entirely below zero. The details of the significant mediations are reported in Table 4.

A. Right hemisphere
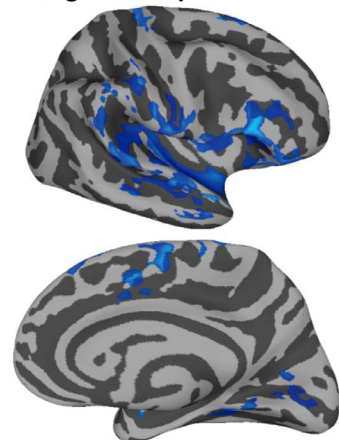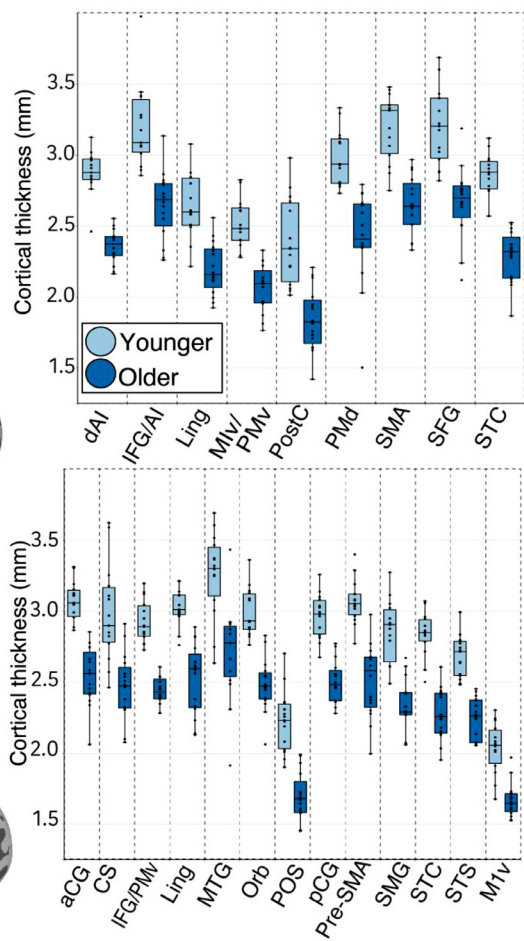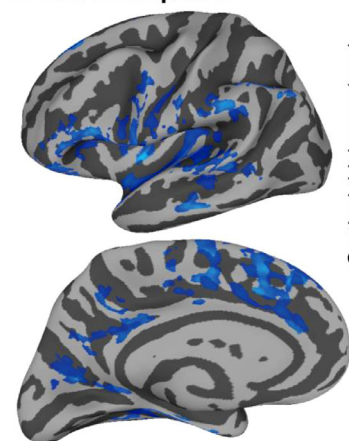

B. Left hemisphere

Fig. 5. Group differences in cortical thickness. On the left panels are the corrected whole-brain maps for Group differences on cortical thickness. Blue indicates lower thickness in older adults compared to younger adults. Group differences are shown on the group lateral and medial average smoothed white matter folded surfaces. The boxplots represent the group differences in cortical thickness within each region identified as part of the whole-brain analysis, separately for the right (A) and left (B) hemispheres. The median is represented by a line traced at the center of the box. The whiskers range from the minimum (Q1 −1.5*interquartile range) and maximum (Q3 + 1.5*interquartile range) values. Each point represents one participant. aCG = anterior cingulate gyrus; CS = calcarine sulcus; dAI = dorsal anterior insula; IFG = inferior frontal gyrus; Ling = lingual gyrus; M1v = ventral primary motor cortex; Orb = orbital gyrus; pCG = posterior cingulate gyrus; POS = parietal-occipital sulcus; PMd = dorsal premotor cortex; PMv = ventral premotor cortex; PostC = postcentral gyrus; pre-SMA = pre-supplementary motor area; SFG = superior frontal gyrus; STC = superior temporal cortex; STS = superior temporal sulcus; MTG = middle temporal gyrus; SMA = supplementary motor area; SMG = supramarginal gyrus. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

6.3. Functional MRI results

We began by examining brain activation patterns in the quiet condition. This analysis revealed the classic bilateral speech processing network including bilateral STC (including primary and associative regions), dAI, PMv, IFG, and medial premotor areas in both younger and older adults (Supplementary Material 7). Here we focus on the interactions between Group and Noise Level, and between Group and Talker variability, but the main effects of Noise and Talker on BOLD are presented in Supplementary Materials 8 and 9.

The following analyses address the hypothesis that group differences in brain activity during speech processing influence speech perception performance. This hypothesis was tested using a combination of whole-brain group comparisons and simple mediation analyses. First, we examined the interaction between Group and Noise level (see Fig. 6 and details are presented in Supplementary Material 8B). Examination of this interaction revealed activation in the left inferior parietal lobule (IPL) and in several parts of the temporal lobe: the bilateral STC, the left anterior temporal cortex and the left inferior temporal cortex. In the bilateral STC, there was a reduced BOLD response to noise in older compared to younger adults, while in all other regions, there was an increased BOLD response to noise, with a tendency for older adults to increase activation from quiet to noise, which was not found in younger adults. The mediation analyses revealed two detrimental indirect effects (i.e. mediations) (left anterior temporal cortex or L ATC and right superior temporal cortex or R STC). For the L ATC, the analysis showed that, relative to the younger adults, the older adults showed a reduced Noise effect on RT (resulting from longer RT in quiet) as a result of the positive effect of Group on BOLD signal in this region (increased response to Noise). For R STC, the analysis shows a reduced effect of noise on RT

for older compared to younger adults, as a result of the negative effect of Group on the BOLD signal (reduced response to noise) in this region. The mediation results are reported in Table 5.

Next we examined the interaction results for the Group X Talker variability factor. No activation survived correction. The results for the main effect of Talker variability are reported in Supplementary Material 8C and 9B Finally, we examined the 3-way interaction between Group, Talker variability and Noise level. Results are shown in Fig. 7 and detailed in Supplementary Material 8D. The analysis revealed activation within the left ventral postcentral gyrus and IPL, as well as in the right cingulate gyrus, precuneus and parieto-occipital sulcus. In all regions except the precuneus, the BOLD response to noise (Noise – Quiet) was higher in the 2 talkers condition compared to 1 talker in older adults but not in younger adults. In the precuneus, the opposite pattern was found. The mediation analyses did not reveal direct or indirect effect of BOLD response patterns on performance.

7. Discussion

The main objective of this study was to examine age-related decline in speech processing in order to shed new light on underlying mechanisms. Our main hypothesis was that, compared to younger adults, older adults' sensitivity would be worse when noise level and talker variability is high, after statistically controlling for pure tone hearing and auditory sustained and selective attention. This hypothesis was partly verified. In contrast to the effect of Noise, we found very limited evidence of an interaction between age and Talker variability on speech performance. Our results show that age-related decline in SPiN performance is associated with thinner cortex in auditory and speech processing regions (STC, PMv/IFG) as well as in regions involved in executive control (aCG

Table 3

Mediation analyses with d' as the dependent variable, for regions identified through the whole brain cortical thickness analyses.

Dependent variable	ROI	Path	Coef	SE	LLCI	ULCI
A. d' (Overall)	L STC	a-path	−0.448	0.072	−0.595	−0.302
		b-path	1.531	0.577	0.348	2.715
		c'	0.236	0.339	−0.459	0.931
		ab	−0.687	0.199	−1.094	−0.303
	L IFG	a-path	−0.413	0.053	−0.522	−0.303
		b-path	1.890	0.788	0.273	3.507
		c'	0.330	0.394	−0.479	1.138
		ab	−0.780	0.217	−1.208	−0.347
	L aSFG	a-path	−0.557	0.083	−0.727	−0.386
		b-path	1.032	0.519	−0.034	2.097
		c'	0.124	0.369	−0.633	0.881
		ab	−0.574	0.263	−1.089	−0.039
	R STC	a-path	−0.434	0.071	−0.580	−0.289
		b-path	1.169	0.612	−0.087	2.425
		c'	0.057	0.352	−0.664	0.779
		ab	−0.508	0.242	−1.039	−0.074
	R AI/IFG	a-path	−0.438	0.060	−0.560	−0.317
		b-path	1.409	0.731	−0.090	2.909
		c'	0.167	0.394	−0.642	0.976
		ab	−0.618	0.286	−1.308	−0.150
	R SFG	a-path	−0.400	0.118	−0.642	−0.158
		b-path	0.783	0.362	0.040	1.526
		c'	−0.138	0.269	−0.689	0.414
		ab	−0.313	0.154	−0.630	−0.035
B. d' (Noise - Quiet)	L pCG	a-path	−0.339	0.064	−0.469	−0.208
		b-path	1.907	0.936	0.013	−3.827
		c'	0.933	0.447	0.016	1.850
		ab	−0.646	0.311	−1.309	−0.082

Note. L = left hemisphere; R = right hemisphere; STC = superior temporal cortex; STS = superior temporal sulcus; PMv = ventral premotor cortex; AI = anterior insula; IFG = inferior frontal gyrus; aSFG = anterior superior frontal gyrus; pCG = posterior cingulate cortex. The significance of the effects was tested using 95% confidence internals. All significant effects are color-coded (green). LLCI = lower limit of the confidence interval, ULCI = upper limit of the confidence interval. In all analyses, the a-path represents the effect of Group on cortical thickness. The a-coefficient tells us that two cases that differ by one unit on X (that is, older adults) differ on the mediator (cortical thickness) by that amount (the a-coefficient). The b-path represents the effect of cortical thickness on d' (sensitivity) when X is constant (same group). The c'-path represents the direct effect of Group on d' when the mediator (cortical thickness) is constant. The ab-path is the indirect effect (mediation). It is the product of the a and b paths. It represents the effect of Group on d' through its effect on cortical thickness. A. In these analyses the dependent variable is the overall d'. Negative indirect effects indicate that d' is lower in the older adults because of the negative effect of Group on cortical thickness. B. The last analysis has a Noise difference d' score as the dependent variable. A negative indirect effect indicates that the impact of noise is reduced in the older adults because of the negative effect of Group on cortical thickness. This is explained by a reduced sensitivity for older adults in the quiet condition. This is thus a detrimental effect.

Table 4

Mediation analyses with RT as the dependent variable, for regions identified through the whole brain cortical thickness analyses.

Dependent variable	ROI	Path	Coef	SE	LLCI	ULCI
RT (Overall)	R AI/IFG	a-path	−0.496	0.125	−0.751	−0.241
		b-path	−0.161	0.082	−0.330	0.007
		c'	−0.071	0.068	−0.210	0.068
		ab	0.080	0.046	0.004	0.185
RT (Noise - Quiet)	L STC	a-path	−0.512	0.078	−0.671	−0.353
		b-path	0.306	0.076	0.151	0.461
		c'	0.151	0.050	0.049	0.252
		ab	−0.157	0.044	−0.249	−0.074
	L IFG	a-path	−0.500	0.057	−0.616	−0.384
		b-path	0.273	0.120	0.026	0.520
		c'	0.130	0.070	−0.014	0.274
		ab	−0.136	0.051	−0.233	−0.030
	L ParOc	a-path	−0.472	0.099	−0.675	−0.269
		b-path	0.144	0.070	0.001	0.288
		c'	0.062	0.049	−0.039	0.163
		ab	−0.068	0.032	−0.130	−0.004
	R STC	a-path	−0.511	0.079	−0.673	−0.348
		b-path	0.231	0.083	0.061	0.400
		c'	0.112	0.055	−0.001	0.224
		ab	−0.118	0.041	−0.190	−0.027

Note. L = left hemisphere; R = right hemisphere; AI = anterior insula; IFG = inferior frontal gyrus; pCG = posterior cingulate cortex; ParOc = parieto-occipital sulcus. ULCI = upper limit of the confidence interval.

is worse when talker variability is high in adults with normal hearing (e.g. Assmann et al., 1982; Heald and Nusbaum, 2014; Martin et al., 1989; Mullennix et al., 1989; Nusbaum and Morin, 1992; Nygaard et al., 1995; Strange et al., 1976; Summerfield, 1981; Wong et al., 2004) and in those with hearing impairments (Kirk et al., 1997). Here, we show that talker variability slowed syllable discrimination and made it less accurate, even though participants were instructed to ignore talkers and to focus on the phonetic content of the syllables. This finding supports the notion that during speech perception, in the presence of multiple talkers, phonetic categorization cannot take place without talker normalization/recalibration -the process by which the listener adapts to talker-specific acoustic space. In a study by Wong et al. (2004), talker normalization was associated with increased activity in the STC and the superior parietal cortex during a word identification task with high talker variability. In contrast, in the present study, talker variability did not modulate activity in the superior parietal cortex, but it was associated with increased activation in the ventral part of the posterior IFG and dAI bilaterally, and in left IPL including the sulcus of Jensen and the angular gyrus. Activation in the IFG has been shown to be influenced by selective auditory attention during the processing of voice and other sounds (Hausfeld et al., 2018), and it interacts with temporal areas during speech perception (Hausfeld et al., 2018; Mesgarani and Chang, 2012; Sabri et al., 2008). Interestingly, one study showed that the left IFG was activated prior to the STG, suggesting an involvement in top-down processes (Sohoglu et al., 2012). The IFG would make predictions of the speech input and send them to sensory areas to clarify speech, especially when speech is degraded. It is thus not surprising that BOLD response in this region was increased in the high talker variability condition in our experiment. Indeed, in the high Talker variability condition, participants had to ensure that the focus of their attention was on the phonological information, not on the voice or talker. These seemingly contrasting findings from those of Wong et al. (2004) may be related to the vastly different tasks used between studies; an auditory word identification task vs. an auditory syllable discrimination task. Given the nature of our task (discrimination at the sublexical level), the regions involved in acoustical analysis were highly active in both the low and the high talker variability conditions, which could explain the lack of a difference in the STC. A word identification task does not require extensive sublexical acoustical analysis, which could explain the change in activation in this region when a word identification task in-
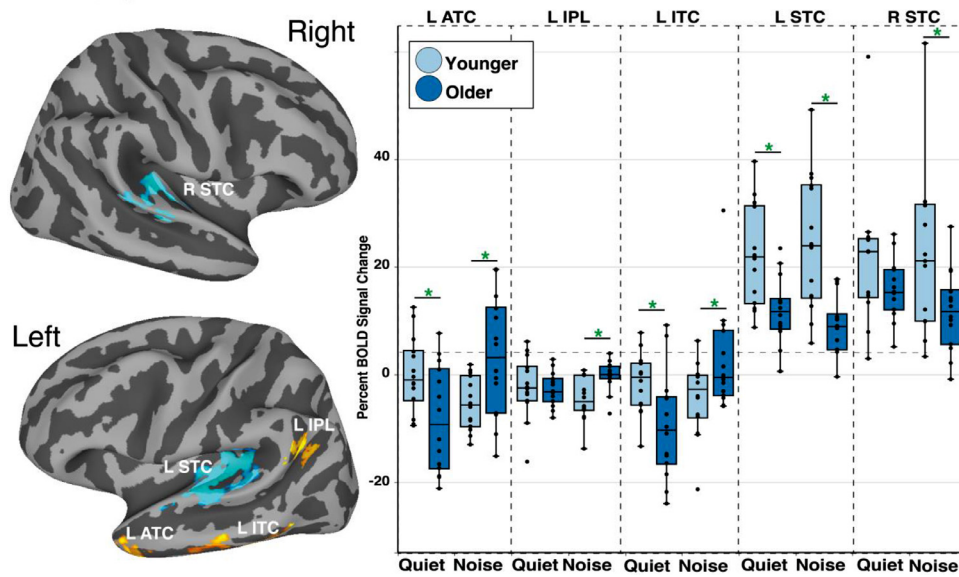
and medial frontal cortex, pCG, dAI and IFG). Moreover, age-related decline in SPiN performance was also associated with differences in BOLD response patterns to noise in the bilateral temporal cortex. Talker variability was not associated with different functional activation patterns in older compared to younger adults. We also found no evidence of compensation in older adults. Together, these results, which we detail below, support our hypothesis of a diffuse rather than a focal neurobiological dysfunction underlying SPiN performance decline in aging; and further show that performance decline is more strongly tied to brain structure than brain function.

7.1. Talker normalization in older adults

As expected, talker variability had a detrimental influence on performance in both experiments (Supplementary Material 3B and 5B), replicating a number of previous studies showing that speech identification

Fig. 6. Group X Noise interaction on BOLD signal. A. Corrected whole-brain maps for the Group by noise interaction. L = Left hemisphere. R = Right hemisphere. ATC = anterior temporal cortex; IPL = inferior parietal lobule; ITC = inferior temporal cortex; STC = superior temporal cortex. In the group map, blue indicates a reduced BOLD response to noise in older compared to younger adults (bilateral STC), while yellow indicates a stronger BOLD response to noise in older compared to younger adults (L IPL, LATC, LITC). These patterns are decomposed in the boxplots. B. In each plot, the y-axis represents the BOLD signal in the noise and quiet conditions, displayed separately for each group. Asterisks represent a significant group difference. In the boxplots, each point represents one participant. The median is represented by a horizontal line across the box. The whiskers range from the minimum (Q1 −1.5*interquartile range) and maximum (Q3 + 1.5*interquartile range) values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5

Mediation results for the regions identified through the Group by Noise interaction on BOLD (whole brain analyses).

ROI	Path	d' Coef	SE	LLCI	ULCI	RT Coef	SE	LLCI	ULCI
L	a-path	−4.475	2.952	−10.567	1.618	−3.690	2.885	−9.644	2.264
STC	b-path	**−0.070**	.0257	−0.124	−0.017	0.002	0.003	−0.004	0.008
	c'	−0.149	0.389	−0.953	0.655	0.077	0.041	−0.008	0.162
	ab	0.315	0.201	−0.028	0.762	−0.007	0.011	−0.034	0.009
L	a-path	8.461	4.250	−0.312	17.234	**16.562**	5.168	5.896	27.228
ATC	b-path	−0.006	0.018	−0.044	0.031	**−0.004**	0.002	−0.007	−0.001
	c'	0.314	0.408	−0.530	1.157	0.077	0.044	−0.015	0.169
	ab	−0.054	0.175	−0.393	0.359	**−0.061**	0.027	−0.113	−0.008
L	a-path	8.284	4.162	−0.307	16.875	**9.410**	4.545	0.029	18.790
ITC	b-path	0.012	0.019	−0.027	0.051	−0.001	0.002	−0.005	0.004
	c'	0.230	0.413	−0.623	1.084	0.021	0.051	−0.084	0.126
	ab	0.100	0.172	−0.220	0.480	−0.007	.0233	−0.055	0.039
L	a-path	3.152	1.719	−0.396	6.700	**4.648**	1.708	1.123	8.174
IPL	b-path	0.000	0.044	−0.092	0.091	0.001	0.005	−0.010	0.011
	c'	0.238	0.399	−0.586	1.063	−0.018	0.048	−0.117	0.081
	ab	−0.001	0.183	−0.418	0.362	0.003	0.028	−0.036	0.077
R	a-path	**−12.737**	4.665	−22.365	−3.110	**−13.784**	5.128	−24.367	−3.201
STC	b-path	−0.030	0.021	−0.074	0.014	**0.005**	0.002	0.001	0.010
	c'	−0.087	0.556	−1.237	1.063	0.132	0.065	−0.002	0.266
	ab	0.384	0.458	−0.521	1.335	**−0.071**	0.047	−0.181	−0.003

Note. STC = superior temporal cortex; ATC = anterior temporal cortex; ITC = Left inferior temporal cortex; IPL = inferior parietal lobule. All significant effects are color-coded (green). LLCI = lower limit of the confidence interval, ULCI = upper limit of the confidence interval. The a-path represents the effect of Group on BOLD signal (Noise – quiet). A positive a-path here represents an increase in BOLD signal in the noise condition. The b-path represents the effect of BOLD signal on performance when X is constant (same group). Remember that the dependent variable is also a contrast (Noise – quiet), and that d' and RT behave differently. A positive effect means that the difference in d' between the Noise and the Quiet is reduced. This is explained by the reduced sensitivity for older adults in the Quiet condition. This is thus a detrimental effect. For RT, a negative effect means that the difference in RT between the Noise and the Quiet is reduced. The c'-path represents the direct effect of Group on performance when the mediator (BOLD signal) is constant. The ab-path is the indirect effect (mediation). It represents the effect of Group on performance through its effect on BOLD signal.

volves multiple talkers. Increased activation in the dAI and IPL suggests up-regulation within the (ventral) bottom-up attentional system in the presence of multiple talkers (Corbetta et al., 2008). Even though talker identity was irrelevant to the task, talker information needed to be processed in order for the syllable discrimination to take place, thus there was an increase in the amount of sensory information to process in the high talker condition.

Contrary to Sommers et al. (1997), who reported that performance of older adults in a word recognition task is more strongly affected by talker variability than that of younger adults, here we found very limited evidence of an interaction between age and talker variability on speech performance. It is possible that our high talker variability condition was not variable enough for age differences to emerge. Indeed, Sommers et al. used 10 different talkers including male and female talkers, whereas we only had three different talkers and they were all men, such that it is possible that participants learned the talker spaces throughout the experiment. An alternative interpretation, given the relatively low performance level in the present task, is that it was too difficult thus masking potentially more subtle differences between groups related to talkers.

Additional studies are needed to determine the extent to which talker normalization process is maintained throughout the lifespan and its
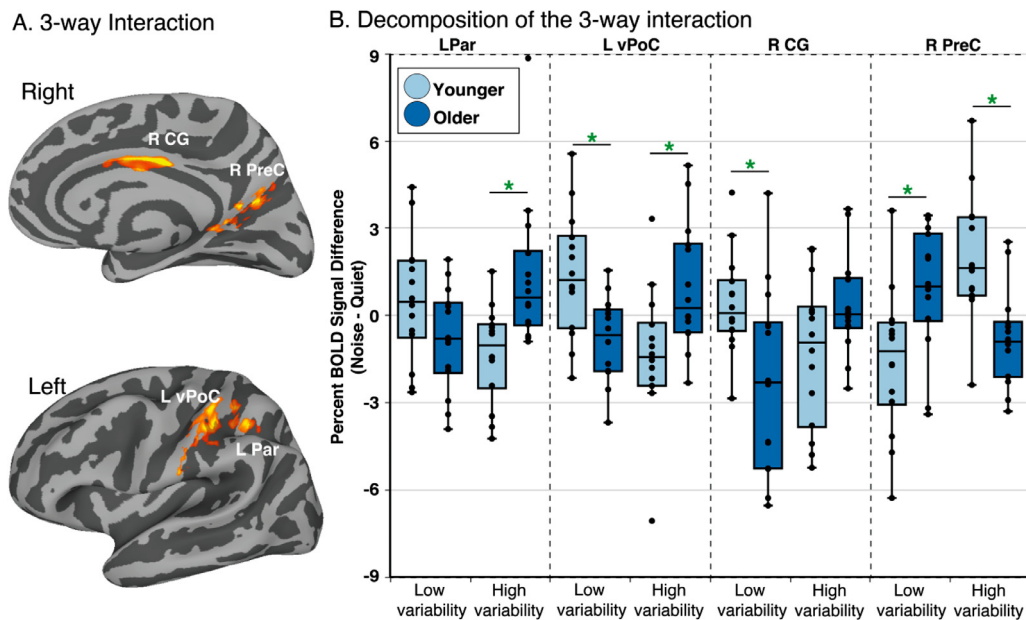
Fig. 7. Group by Noise by Talker interaction on the BOLD signal. A. Corrected whole-brain group maps for the 3-way interaction. B. Decomposition of the 3-way interaction separately for each ROI. In each plot, the y-axis represents the difference in BOLD signal between the noise and quiet conditions (Noise – Quiet), displayed separately for each group, and for the low and high talker variability conditions. L = Left hemisphere. R = Right hemisphere. Par = parietal cortex; vPoC = ventral postcentral gyrus and sulcus; CG = cingulate gyrus; PreC = precuneus. Asterisks represent a significant group difference. In the boxplots, each point represents one participant. The median is represented by a horizontal line across the box. The whiskers range from the minimum (Q1 −1.5*interquartile range) and maximum (Q3 + 1.5*interquartile range) values.

potential contribution to communication difficulties in older adults. Despite the lack of robust age differences in the effect of talker variability, we did find group differences in the brain response to the high talker variability condition. In the bilateral STC, we found that lower signal in this region in older adults was associated with slower RT and lower sensitivity (d'). Moreover, lower cortical thickness in the left STC in older adults was associated with lower sensitivity. These results are consistent with the finding by Wong et al. (2004) of an increased role for this region when speech is produced by multiple talkers, probably reflecting the processing of talker-related spectral information necessary for talker normalization.

7.2. Thinner cortex in auditory and speech processing regions is associated with poorer spin performance in aging

As was expected, age-related decline in cortical thickness, a measure of the depth of the cortical mantle (Fischl and Dale, 2000; Fischl et al., 1999a), was diffuse. Importantly, in several speech-related regions such as the STC, the IFG, PMv and M1v, this decline was associated with worse SPiN performance in older adults. The supratemporal cortex, as a whole, is involved in prelexical processing of auditory speech (Price, 2010). Several models of speech processing assign the bilateral STG a role in basic, spectrotemporal (i.e., acoustic) analysis of auditory stimuli including speech (Binder et al., 2000; Hickok and Poeppel, 2007). A recent study with stroke patients showed that accuracy in phoneme discrimination correlates with damage to the STG (Kim et al., 2019). The task that was used in the present study — an auditory sublexical speech discrimination task with three talkers — involved both acoustical (spectrotemporal) analysis as well as phonetic and phonological processing. In each trial, successful comparison of the two syllables required phonological processing to segment relevant from non-relevant signals, including voice and phonetic differences, to determine if the syllables were different. Hence, the finding that lower thickness affected SPiN performance in this region is not surprising. Importantly, in our study, decline was not restricted to primary auditory cortices but extended to much of the bilateral STC including the PT and the

lateral STG. This finding extends previous reports of abnormal activity patterns within the auditory cortex in older adults (Presacco et al., 2016a, 2016b), as well as reports of early abnormal processing in the STG (Brodbeck et al., 2018; Manan et al., 2017).

Another important finding of the present study is that the structure of the left PMv was associated with speech perception performance. The Directions Into Velocities of Articulators (DIVA) model of speech production proposes that the PMv, which is part of the dorsal speech stream, contains speech motor programs (Guenther et al., 2006). Modern accounts of the motor theory of speech perception posit that the speech motor system generates internal models of the predicted sensory consequences of speech gestures under consideration, and that these predictions are matched with acoustic representations to constrain perception (for a review, see e.g. McGettigan and Tremblay, 2018). A recent model further suggests that interactions between auditory and motor processes are pervasive and occur at acoustical and phonological processing levels (Liebenthal and Mottonen, 2017). The finding that lower thickness in the left PMv, which could reflect less specific or less distinct speech representations, is associated with lower sensitivity is consistent with the notion of a role for this region in speech perception (e.g. D'Ausilio et al., 2009; Grabski et al., 2013; Meister et al., 2007; Sato et al., 2009; Watkins and Paus, 2004; Watkins et al., 2003; Wilson et al., 2004).

7.3. SPiN performance in aging and bold signal in the temporal lobe

Age differences in the effect of Noise on the BOLD signal were limited to the bilateral temporal lobes and in the left IPL. Specifically, examination of the patterns of BOLD response in these regions, controlling for cortical thickness, shows that, in the bilateral STC, the BOLD response was overall lower for older adults compared to the younger ones, and that it decreased with noise, while in younger adults, the signal was stable across noise conditions. In the right STC, this pattern of activity was associated with worse (slower) performance. In the other regions, older adults showed increased activation from the Quiet to the Noise condition, which was not found in younger adults. This pattern was only

associated to speech performance in the left anterior temporal lobe (L ATC), where it was associated with worse (slower) performance. Hence, in the present study, we found no evidence of neural compensation, operationalized as a difference in the pattern of brain response in older compared to younger adults that is associated with a positive outcome for older adults. This kind of difference can take the form of a similar performance level for younger and older adults at the group level, or it can be reflected as a positive outcome at the subject level when those exhibiting such pattern show better performance compared to those who do not (Cabeza et al., 2002; Grady, 2012). We found no such patterns in the present study.

Whether SPiN comprehension in older adults engage compensatory-like mechanisms is unclear. Indeed, the patterns of BOLD-behaviour relationship during SPiN comprehension are heterogeneous in the literature. Wong et al. have reported compensation-like activity during a picture-word matching task performed in the presence of noise in the right MTG and in the right frontal lobe, in an area encompassing the MFG and the precentral gyrus (Wong et al., 2009a). In a more recent study, Du et al. found increased activation in the left the IFG pars opercularis and in the left precentral gyrus that was associated with better SPiN performance (Du et al., 2016). Others have found reduced beneficial activity in the cingulo-opercular network (Vaden et al., 2015) and in the left dAI during SPiN tasks (Bilodeau-Mercure et al., 2015). Manan et al. also failed to find evidence of compensation during a word repetition in noise task (Manan et al., 2015). Similarly, a recent study of suprasegmental speech processing did not report compensatory-like activity in older adults (Keller et al., 2019). In that study, an increased in the recruitment of right-hemisphere regions was found (i.e., a reduced left lateralization) but it was not associated with better performance. Taken together, these results, past and present, suggest that individuals may approach SPiN tasks differently. Some compensate, some do not. But even in those who compensate, the literature shows no robust pattern across studies. This suggests that individuals may compensate through different mechanisms, possibly as a reflection of the resources at their disposal; a hypothesis that will require further investigation. Another possibility is that differences across studies are related to the heterogeneous nature of the SPiN tasks used in the literature, which engage partly distinct processes (acoustical, phonological, lexical, attention, working memory). These two possibilities need to be addressed simultaneously, by assessing SPiN capabilities with multiple SPiN tasks involving different levels of processing to shed light on the capacity for the brain to maintain SPiN performance by engaging additional neural resources.

7.4. Decline in executive network also influences performance

Our results show that the structure of the cingulate gyrus declines with age, consistent with prior studies (Good et al., 2001; Mann et al., 2011; Resnick et al., 2003). Importantly, a relationship between the structure of the anterior (aCC) and posterior cingulate cortex (pCC) and speech performance in aging was found. The pCC is a central node of the default mode network (DMN) (Buckner et al., 2008), a set of regions that are more active during at rest than during tasks (Raichle et al., 2001). There is also evidence suggesting that the pCC is involved in regulating the focus of attention (e.g., Leech et al., 2011; Leech and Sharp, 2014). The aCC is believed to be involved in signaling conflicts in information processing to trigger compensatory adjustments in cognitive control (Botvinick et al., 2004). The implication of cognitive control networks in our challenging speech discrimination task is thus unsurprising and is consistent with a number of prior studies. For example, Liebenthal and colleagues reported activation in the aCC and pCC task during a speech discrimination task (Liebenthal et al., 2005). Harris et al. reported age-related changes in dACC activation in the form of an increase in activation in older adults as compared to young adults in a word recognition task; this change was not related to performance, but it was suggestive of an upregulation of error monitoring in older

adults (Harris et al., 2009). More recently, we found compensatory-like activation in the right pCC in older adults during a sublexical speech production task, suggesting that allocation of more cognitive resources, such as increased focused attention, was contributing to maintaining articulation in older age during demanding tasks in a general manner (not speech specific) (Tremblay et al., 2017). The present finding that thinner cortex in the cingulate cortex was associated with worse sensitivity in older adults suggests a contribution of general cognitive control networks to SPiN performance. Together, these results demonstrate that age differences in the structure and functioning of executive networks is a contributing factor to SPiN performance decline in aging.

8. Conclusions

Listening to speech, especially in a noisy environment, can be a daunting task, especially for older adults. Speech is a fast, continuous and intrinsically noisy signal that varies enormously in its acoustical realization from one talker to another. Successful speech perception requires extensive acoustical and phonological analyses to decode the incoming signal, normalize it to talkers' characteristics, and ultimately comprehend its meaning. The cause of the well-known age-related decline in speech perception in noise is not clear. The current study demonstrates that sensitivity to speech sounds declines with age, especially in the presence of background noise. Talker variability, in contrast, did not affect the performance of older adults more strongly than that of younger adults, suggesting preserved (at least to some extent) talker normalization capabilities. Further, we show evidence of detrimental effects of cortical thinning in auditory, speech processing and executive control on speech perception performance. Finally, our results show relatively circumscribed regions of the temporal cortex in which downregulation was associated with worse speech performance in older adults after controlling for cortical thickness, suggesting that disruptions within auditory or speech processing can be harmful to speech perception in the elderly. Together, these results suggest that diffuse underlying dysfunctions, mostly structural, could be causing age-related SPiN performance decline. Further studies need to address the issue of the heterogeneity in the brain response to adverse listening conditions, to understand why some individuals can set off neural compensations, and others seemingly cannot. This information is critical for the development of rehabilitation interventions to enhance SPiN performance in the elderly, and, in turn, their quality of life, allowing them to actively participate in their communities for more years.

Credit authorship contribution statement

Pascale Tremblay: Conceptualization, Funding acquisition, Methodology, Investigation, Project administration, Formal analysis, Visualization, Supervision, Resources, Project administration, Writing - original draft, Data curation. **Valérie Brisson:** Writing - review & editing, Data curation. **Isabelle Deschamps:** Formal analysis, Writing - review & editing.

Acknowledgments

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2020.117675.

References

Alain, C., Woods, D.L., 1999. Age-related changes in processing auditory stimuli during visual attention: evidence for deficits in inhibitory control and sensory memory. Psychol. Aging 14 (3), 507–519. doi:10.1037//0882-7974.14.3.507.

Argall, B.D., Saad, Z.S., Beauchamp, M.S., 2006. Simplified intersubject averaging on the cortical surface using SUMA. Hum. Brain Mapp. 27 (1), 14–27.

Assmann, P.F., Nearey, T.M., Hogan, J.T., 1982. Vowel identification: orthographic, perceptual, and acoustic aspects. J. Acoust. Soc. Am. 71 (4), 975–989. doi:10.1121/1.387579.

Bilodeau-Mercure, M., Lortie, C.L., Sato, M., Guitton, M.J., Tremblay, P., 2015. The neurobiology of speech perception decline in aging. Brain Struct. Funct. 220 (2), 979–997. doi:10.1007/s00429-013-0695-3.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S.F., Springer, J.A., Kaufman, J.N., Possing, E.T, 2000. Human temporal lobe activation by speech and nonspeech sounds. Cereb. Cortex 10 (5), 512–528. doi:10.1093/cercor/10.5.512.

Botvinick, M.M., Cohen, J.D., Carter, C.S., 2004. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn. Sci. 8 (12), 539–546.

Brodbeck, C., Presacco, A., Anderson, S., Simon, J.Z., 2018. Over-representation of speech in older adults originates from early response in higher order auditory cortex. Acta Acust. United Acust. 104 (5), 774–777. doi:10.3813/AAA.919221.

Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L., 2008. The brain's default network: anatomy, function, and relevance to disease. Ann. N. Y. Acad. Sci. 1124, 1–38. doi:10.1196/annals.1440.011.

Cabeza, R., Anderson, N.D., Locantore, J.K., McIntosh, A.R., 2002. Aging gracefully: compensatory brain activity in high-performing older adults. Neuroimage 17 (3), 1394–1402.

Chang, Y.P., Fu, Q.J., 2006. Effects of talker variability on vowel recognition in cochlear implants. J. Speech Lang. Hear. Res. 49 (6), 1331–1341. doi:10.1044/1092-4388(2006/095).

Chao, L.L., Knight, R.T., 1997. Prefrontal deficits in attention and inhibitory control with aging. Cereb. Cortex 7 (1), 63–69. doi:10.1093/cercor/7.1.63.

Corbetta, M., Patel, G., Shulman, G.L., 2008. The reorienting system of the human brain: from environment to theory of mind. Neuron 58 (3), 306–324. doi:10.1016/j.neuron.2008.04.017.

Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29 (3), 162–173.

D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., Fadiga, L., 2009. The motor somatotopy of speech perception. Curr. Biol. 19, 381–385.

D'Esposito, M., Deouell, L.Y., Gazzaley, A., 2003. Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging. Nat. Rev. Neurosci. 4 (11), 863–872. doi:10.1038/nrn1246.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based Analysis: I. Segmentation and surface reconstruction. Neuroimage 9 (2), 179–194.

Destrieux, C., Fischl, B., Dale, A., Halgren, E., 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53 (1), 1–15. doi:10.1016/j.neuroimage.2010.06.010.

Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2016. Increased activity in frontal motor cortex compensates impaired speech perception in older adults. Nat. Commun. 7, 12241. doi:10.1038/ncomms12241.

Eckert, M.A., Cute, S.L., Vaden Jr., K.I., Kuchinsky, S.E., Dubno, J.R, 2012. Auditory cortex signs of age-related hearing loss. J. Assoc. Res. Otolaryngol. 13 (5), 703–713. doi:10.1007/s10162-012-0332-5.

El-Assal, H.A., El-Gharib, A.M., 2019. Encoding of speech in noise in adults using hearing aids: effect of noise reduction algorithm. Hearing Balance Commun. doi:https://doi.org/10.1080/21695717.2019.1667687.

Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. U. S. A. 97 (20), 11050–11055. doi:10.1073/pnas.200033797.

Fischl, B., Liu, A., Dale, A.M., 2001. Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. IEEE Trans. Med. Imaging 20 (1), 70–80.

Fischl, B., Sereno, M.I., Dale, A.M., 1999a. Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. Neuroimage 9 (2), 195–207.

Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M., 1999b. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum. Brain Mapp. 8 (4), 272–284.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., …, Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cerebral Cortex 14 (1), 11–22. doi:10.1093/cercor/bhg087.

Fostick, L., Ben-Artzi, E., Babkoff, H., 2013. Aging and speech perception: beyond hearing threshold and cognitive ability. J. Basic Clin. Physiol. Pharmacol. 24 (3), 175. doi:10.1515/jbcpp-2013-0048.

Frisina, D.R., Frisina, R.D., 1997. Speech recognition in noise and presbycusis: relations to possible neural mechanisms. Hear. Res. 106 (1–2), 95–104. doi:10.1016/S0378-5955(97)00006-3.

Fullgrabe, C., Moore, B.C., Stone, M.A., 2014. Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. Front. Aging Neurosci. 6, 347. doi:10.3389/fnagi.2014.00347.

Garrett, D.D., Lindenberger, U., Hoge, R.D., Gauthier, C.J., 2017. Age differences in brain signal variability are robust to multiple vascular controls. Sci. Rep. 7 (1), 10149. doi:10.1038/s41598-017-09752-7.

Gilbert, J.L., Tamati, T.N., Pisoni, D.B., 2013. Development, reliability, and validity of PRESTO: a new high-variability sentence recognition test. J. Am. Acad. Audiol. 24 (1), 26–36. doi:10.3766/jaaa.24.1.4.

Goldinger, S.D., Pisoni, D.B., Logan, J.S., 1991. On the nature of talker variability effects on recall of spoken word lists. J. Exp. Psychol. Learn. Mem. Cogn. 17 (1), 152–162. doi:10.1037//0278-7393.17.1.152.

Good, C.D., Johnsrude, I.S., Ashburner, J., Henson, R.N., Friston, K.J., Frackowiak, R.S., 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage 14 (1 Pt 1), 21–36. doi:10.1006/nimg.2001.0786.

Gordon-Salant, S., Fitzgibbons, P.J., 1993. Temporal factors and speech recognition performance in young and elderly listeners. J. Speech, Language, Hear. Res. 36 (6), 1276–1285. doi:10.1044/jshr.3606.1276.

Grabski, K., Tremblay, P., Gracco, V.L., Girin, L., Sato, M., 2013. A mediating role of the auditory dorsal pathway in selective adaptation to speech: a statedependent transcranial magnetic stimulation study. Brain Res. 1515, 55–65. doi:10.1016/j.brainres.2013.03.024.

Gracco, V.L., Tremblay, P., Pike, B., 2005. Imaging speech production using fMRI. Neuroimage 26 (1), 294–301. doi:10.1016/j.neuroimage.2005.01.033.

Grady, C., 2012. The cognitive neuroscience of ageing. Nat. Rev. Neurosci. 13 (7), 491–505. doi:10.1038/nrn3256.

Guenther, F.H., Ghosh, S.S., Tourville, J.A., 2006. Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang. 96 (3), 280–301. doi:10.1016/j.bandl.2005.06.001.

Harris, K.C., Dubno, J.R., Keren, N.I., Ahlstrom, J.B., Eckert, M.A., 2009. Speech recognition in younger and older adults: a dependency on low-level auditory cortex. J. Neurosci. 29 (19), 6078–6087. doi:10.1523/JNEUROSCI.0412-09.2009.

Harrison, X.A., Donaldson, L., Correa-Cano, M.E., Evans, J., Fisher, D.N., Goodwin, C.E.D., …, Inger, R., 2018. A brief introduction to mixed effects modelling and multi-model inference in ecology. Peer J. 6, e4794. doi:10.7717/peerj.4794.

Hausfeld, L., Riecke, L., Formisano, E., 2018. Acoustic and higher-level representations of naturalistic auditory scenes in human auditory and frontal cortex. Neuroimage 173, 472–483 doi:https://doi.org/10.1016/j.neuroimage.2018.02.065.

Hayes, A.F., 2008. SPSS Macro for Multiple Mediation Retrieved from http://www.comm.ohio-state.edu/ahayes/ .

Hayes, A.F., 2013. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. The Guilford Press.

Hayes, A.F., 2018. Introduction to Mediation, Moderation, and Conditional Process Analysis. A Regression-Based Approach, 2nd Ed. The Guilford Press, New York. London.

Heald, S.L., Nusbaum, H.C., 2014. Speech perception as an active cognitive process. Front. Syst. Neurosci. 8, 35. doi:10.3389/fnsys.2014.00035.

Helfer, K.S., Freyman, R.L., 2008. Aging and speech-on-speech masking. Ear Hear. 29 (1), 87–98. doi:10.1097/AUD.0b013e31815d638b.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8 (5), 393–402. doi:10.1038/nrn2113.

Humes, L.E., Dubno, J.R., Gordon-Salant, S., Lister, J.J., Cacace, A.T., Cruickshanks, K.J., …, Wingfield, A., 2012. Central presbycusis: a review and evaluation of the evidence. J. Am. Acad. Audiol. 23 (8), 635–666. doi:10.3766/jaaa.23.8.5.

Humes, L.E., Wilson, D.L., Barlow, N.N., Garner, C., 2002. Changes in hearing-aid benefit following 1 or 2 years of hearing-aid use by older adults. J. Speech Lang. Hear. Res. 45 (4), 772–782. doi:10.1044/1092-4388(2002/062.

Hwang, J.H., Li, C.W., Wu, C.W., Chen, J.H., Liu, T.C., 2007. Aging effects on the activation of the auditory cortex during binaural speech listening in white noise: an fMRI study. Audiol. Neurootol. 12 (5), 285–294. doi:10.1159/000103209.

Jo, H.J., Lee, J.M., Kim, J.H., Shin, Y.W., Kim, I.Y., Kwon, J.S., Kim, S.I., 2007. Spatial accuracy of fMRI activation influenced by volume- and surface-based spatial smoothing techniques. Neuroimage 34 (2), 550–564. doi:10.1016/j.neuroimage.2006.09.047.

Kabani, N., Le Goualher, G., MacDonald, D., Evans, A.C., 2001. Measurement of cortical thickness using an automated 3-D algorithm: a validation study. Neuroimage 13 (2), 375–380. doi:10.1006/nimg.2000.0652.

Kaiser, A.R., Kirk, K.I., Lachs, L., Pisoni, D.B., 2003. Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. J. Speech Lang. Hear. Res. 46 (2), 390–404. doi:10.1044/1092-4388(2003/032.

Keller, M., Neuschwander, P., Meyer, M., 2019. When right becomes less right: neural dedifferentiation during suprasegmental speech processing in the aging brain. Neuroimage 189, 886–895. doi:10.1016/j.neuroimage.2019.01.050.

Kent, R.D., 2000. Research on speech motor control and its disorders: a review and prospective. J. Commun. Disord. 33 (5), 391–427 quiz 428. doi:S0021-9924(00)00023-X [pii].

Kim, K., Adams, L., Keator, L.M., Sheppard, S.M., Breining, B.L., Rorden, C., …, Hillis, A.E., 2019. Neural processing critical for distinguishing between speech sounds. Brain Lang. 197, 104677. doi:10.1016/j.bandl.2019.104677.

Kim, S., Frisina, R.D., Mapes, F.M., Hickman, E.D., Frisina, D.R., 2006. Effect of age on binaural speech intelligibility in normal hearing adults. Speech Commun. 48 (6), 591–597. doi:10.1016/j.specom.2005.09.004.

Kirk, K.I., Pisoni, D.B., Miyamoto, R.C., 1997. Effects of stimulus variability on speech perception in listeners with hearing impairment. J. Speech Lang. Hear. Res. 40 (6), 1395–1405. doi:10.1044/jslhr.4006.1395.

Klatt, D., 1986. The problem of variability in speech recognition and in models of speech perception. In: Perkell, J., Klatt, D. (Eds.), Invariance and Variability in Speech Processes. Lawrence Erlbaum Associates, Hillsdale.

Lash, A., Rogers, C.S., Zoller, A., Wingfield, A., 2013. Expectation and entropy in spoken word recognition: effects of age and hearing acuity. Exp. Aging Res. 39 (3), 235–253. doi:10.1080/0361073X.2013.779175.

Leech, R., Kamourieh, S., Beckmann, C.F., Sharp, D.J., 2011. Fractionating the default mode network: distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. J. Neurosci. 31 (9), 3217–3224. doi:10.1523/JNEUROSCI.5626-10.2011.

Leech, R., Sharp, D.J., 2014. The role of the posterior cingulate cortex in cognition and disease. Brain 137 (Pt 1), 12–32. doi:10.1093/brain/awt162.

Liebenthal, E., Binder, J.R., Spitzer, S.M., Possing, E.T., Medler, D.A., 2005. Neural substrates of phonemic perception. Cereb. Cortex 15 (10), 1621–1631. doi:10.1093/cercor/bhi040.

Liebenthal, E., Mottonen, R., 2017. An interactive model of auditory-motor speech perception. Brain Lang. doi:10.1016/j.bandl.2017.12.004.

Lively, S.E., Logan, J.S., Pisoni, D.B., 1993. Training Japanese listeners to identify English /r/ and /l/. II: the role of phonetic environment and talker variability in learning new perceptual categories. J. Acoust. Soc. Am. 94 (3 Pt 1), 1242–1255. doi:10.1121/1.408177.

Logan, J.S., Lively, S.E., Pisoni, D.B., 1991. Training Japanese listeners to identify English /r/ and /l/: a first report. J. Acoust. Soc. Am. 89 (2), 874–886. doi:10.1121/1.1894649.

Macmillan, N.A., Creelman, C.D., 1990. Response bias: characteristics of detection theory, threshold theory, and" nonparametric" indexes. Psychol. Bull. 107 (3), 401.

Mager, R., Falkenstein, M., Stormer, R., Brand, S., Muller-Spahn, F., Bullinger, A.H., 2005. Auditory distraction in young and middle-aged adults: a behavioural and event-related potential study. J. Neural Transm. (Vienna) 112 (9), 1165–1176. doi:10.1007/s00702-004-0258-0.

Makris, N., Papadimitriou, G.M., Kaiser, J.R., Sorg, S., Kennedy, D.N., Pandya, D.N., 2009. Delineation of the middle longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. Cereb. Cortex 19 (4), 777–785. doi:10.1093/cercor/bhn124.

Makris, N., Preti, M.G., Wassermann, D., Rathi, Y., Papadimitriou, G.M., Yergatian, C., …, Kubicki, M., 2013. Human middle longitudinal fascicle: segregation and behavioral-clinical implications of two distinct fiber connections linking temporal pole and superior temporal gyrus with the angular gyrus or superior parietal lobule using multi-tensor tractography. Brain Imag. Behav. 7 (3), 335–352. doi:10.1007/s11682-013-9235-2.

Manan, H.A., Franz, E.A., Yusoff, A.N., Mukari, S.Z., 2015. The effects of aging on the brain activation pattern during a speech perception task: an fMRI study. Aging Clin. Exp. Res. 27 (1), 27–36. doi:10.1007/s40520-014-0240-0.

Manan, H.A., Yusoff, A.N., Franz, E.A., Mukari, S.Z.-M.S., 2017. Effects of Aging and Background Babble Noise on Speech Perception Processing: an fMRI Study. Neurophysiology 49 (6), 441–452.

Mann, S.L., Hazlett, E.A., Byne, W., Hof, P.R., Buchsbaum, M.S., Cohen, B.H., …, Chu, K.-.W., 2011. Anterior and posterior cingulate cortex volume in healthy adults: effects of aging and gender differences. Brain Res. 1401, 18–29. doi:10.1016/j.brainres.2011.05.050.

Martin, C.S., Mullennix, J.W., Pisoni, D.B., Summers, W.V., 1989. Effects of talker variability on recall of spoken word lists. J. Exp. Psychol. Learn. Mem. Cogn. 15 (4), 676–684. doi:10.1037//0278-7393.15.4.676.

Mazelova, J., Popelar, J., Syka, J., 2003. Auditory function in presbycusis: peripheral vs. central changes. Exp. Gerontol. 38 (1–2), 87–94.

McGettigan, C., Tremblay, P., 2018. Links between perception and production: examining the roles of motor and premotor cortices in understanding speech. In: Gaskell, G., Rueschemeyer, S.A. (Eds.), Oxford Handbook of Psycholinguistics. Oxford University Press, Oxford.

Meister, H., Schreitmuller, S., Grugel, L., Ortmann, M., Beutner, D., Walger, M., Meister, I.G., 2013. Cognitive resources related to speech recognition with a competing talker in young and older listeners. Neuroscience 232, 74–82. doi:10.1016/j.neuroscience.2012.12.006.

Meister, I.G., Wilson, S.M., Deblieck, C., Wu, A.D., Iacoboni, M., 2007. The essential role of premotor cortex in speech perception. Curr. Biol. 17, 1692–1696.

Mesgarani, N., Chang, E.F., 2012. Selective cortical representation of attended speaker in multi-talker speech perception. Nature 485 (7397), 233–236. doi:10.1038/nature11020.

Mudar, R.A., Husain, F.T., 2016. Neural alterations in acquired age-related hearing loss. Front. Psychol. 7, 828. doi:10.3389/fpsyg.2016.00828.

Mullennix, J.W., Pisoni, D.B., Martin, C.S., 1989. Some effects of talker variability on spoken word recognition. J. Acoust. Soc. Am. 85 (1), 365–378. doi:10.1121/1.397688.

Nasreddine, Z.S., Chertkow, H., Phillips, N., Bergman, H., Whitehead, V., 2003. Sensitivity and specificity of the montreal cognitive assessment (MoCA) for detection of mild cognitive deficits. Can. J. Neurol. Sci. 30 (30).

Nusbaum, H.C., Morin, T.M., 1992. Paying attention to differences among talkers. In: Tohkura, Y., Sagisaka, Y., Vatikiotis-Bateson, E. (Eds.), Speech perception, production, and Linguistic Structure. Ohmasha Publishing, Tokyo, pp. 113–134.

Nygaard, L.C., Sommers, M.S., Pisoni, D.B., 1995. Effects of stimulus variability on perception and representation of spoken words in memory. Percept. Psychophys. 57 (7), 989–1001.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9 (1), 97–113.

Peelle, J.E., Troiani, V., Wingfield, A., Grossman, M., 2010. Neural processing during older adults' comprehension of spoken sentences: age differences in resource allocation and connectivity. Cereb. Cortex 20 (4), 773–782. doi:10.1093/cercor/bhp142.

Peelle, J.E., Wingfield, A., 2016. The neural consequences of age-related hearing loss. Trends Neurosci. 39 (7), 486–497. doi:10.1016/j.tins.2016.05.001.

Presacco, A., Simon, J.Z., Anderson, S., 2016a. Effect of informational content of noise on speech representation in the aging midbrain and cortex. J. Neurophysiol. 116 (5), 2356–2367. doi:10.1152/jn.00373.2016.

Presacco, A., Simon, J.Z., Anderson, S., 2016b. Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. J. Neurophysiol. 116 (5), 2346–2355. doi:10.1152/jn.00372.2016.

Price, C., 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann. N.Y. Acad. Sci. 1191, 62–88.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. Proc. Natl. Acad. Sci. U. S. A. 98 (2), 676–682. doi:10.1073/pnas.98.2.676.

Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat. Neurosci. 12 (6), 718–724 doi:nn.2331 [pii] 10.1038/nn.2331.

Resnick, S.M., Pham, D.L., Kraut, M.A., Zonderman, A.B., Davatzikos, C., 2003. Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. J. Neurosci. 23 (8), 3295–3301.

Reuter-Lorenz, P., Mikels, J., 2006. The aging mind and brain: implications of enduring plasticity for behavioral and cultural change. Lifespan Devel. Brain 255–276.

Saad, Z., Reynolds, R., Argall, B., Japee, S., Cox, R., 2004. SUMA: an interface for surface-based intra-and inter-subject analysis with AFNI. IEEE Int. Symp. Biomed. Imaging 1510–1513.

Sabri, M., Binder, J.R., Desai, R., Medler, D.A., Leitl, M.D., Liebenthal, E., 2008. Attentional and linguistic interactions in speech perception. Neuroimage 39 (3), 1444–1456. doi:10.1016/j.neuroimage.2007.09.052.

Salvi, R.J., Lockwood, A.H., Frisina, R.D., Coad, M.L., Wack, D.S., Frisina, D.R., 2002. PET imaging of the normal human auditory system: responses to speech in quiet and in background noise. Hear. Res. (1–2), 96–106.

Samuel, A.G., 1981. Phonemic restoration: insights from a new methodology. J. Exp. Psychol. Gen. 110 (4), 474–494.

Sandford, J.A., Turner, A., 1995. Manual For the Integrated Visual and Auditory Continuous Performance Test. Braintrain, Richmond, VA.

Sato, M., Tremblay, P., Gracco, V.L., 2009. A mediating role of the premotor cortex in phoneme segmentation. Brain Lang. 111 (1), 1–7. doi:10.1016/j.bandl.2009.03.002.

Segonne, F., Pacheco, J., Fischl, B., 2007. Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. IEEE Trans. Med. Imaging 26 (4), 518–529.

Sheppard, J.P., Wang, J.P., Wong, P.C., 2011. Large-scale cortical functional organization and speech perception across the lifespan. PLoS ONE 6 (1), e16510. doi:10.1371/journal.pone.0016510.

Sohoglu, E., Peelle, J.E., Carlyon, R.P., Davis, M.H., 2012. Predictive top-down integration of prior knowledge during speech perception. J. Neurosci. 32 (25), 8443–8453. doi:10.1523/jneurosci.5069-11.2012.

Sommers, M.S., 1997. Stimulus variability and spoken word recognition. II. The effects of age and hearing impairment. J. Acoust. Soc. Am. 101 (4), 2278–2288.

Strange, W., Verbrugge, R.R., Shankweiler, D.P., Edman, T.R., 1976. Consonant environment specifies vowel identity. J. Acoust. Soc. Am. 60 (1), 213–224. doi:10.1121/1.381066.

Summerfield, A.Q., 1981. On articulation rate and perceptual constency in phonetic perception. J. Exp. Psychol. Hum. Percept. Perform 7, 1074–1095.

Tamati, T.N., Gilbert, J.L., Pisoni, D.B., 2013. Some factors underlying individual differences in speech recognition on PRESTO: a first report. J. Am. Acad. Audiol. 24 (7), 616–634. doi:10.3766/jaaa.24.7.10.

Tamati, T.N., Pisoni, D.B., 2014. Non-native listeners' recognition of high-variability speech using PRESTO. J. Am. Acad. Audiol. 25 (9), 869–892. doi:10.3766/jaaa.25.9.9.

Tremblay, P., Perron, M., Deschamps, I., Kennedy-Higgins, D., Houde, J.C., Dick, A.S., Descoteaux, M., 2018. The role of the arcuate and middle longitudinal fasciculi in speech perception in noise in adulthood. Hum. Brain Mapp. doi:10.1002/hbm.24367.

Tremblay, P., Sato, M., Deschamps, I., 2017. Age differences in the motor control of speech: an fMRI study of healthy aging. Hum. Brain Mapp. 38 (5), 2751–2771. doi:10.1002/hbm.23558.

Tun, P.A., 1998. Fast noisy speech: age differences in processing rapid speech with background noise. Psychol. Aging 13 (3), 424.

Vaden Jr., K.I., Kuchinsky, S.E., Ahlstrom, J.B., Dubno, J.R., Eckert, M.A, 2015. Cortical activity predicts which older adults recognize speech in noise and when. J. Neurosci. 35 (9), 3929–3937. doi:10.1523/JNEUROSCI.2908-14.2015.

Waldron-Perrine, B., Axelrod, B.N., 2012. Determining an appropriate cutting score for indication of impairment on the montreal cognitive assessment. Int. J. Geriatr. Psychiatry 27 (11), 1189–1194. doi:10.1002/gps.3768.

Watkins, K., Paus, T., 2004. Modulation of motor excitability during speech perception: the role of Broca's area. J. Cogn. Neurosci. 16 (6), 978–987. doi:10.1162/0898929041502616.

Watkins, K., Strafella, A., Paus, T., 2003. Seeing and hearing speech excites the motor system involved in speech production. Neuropsychologia 41 (8), 989–994 doi:S0028393202003160 [pii].

Wilson, S., Saygin, A., Sereno, M., Iacoboni, M., 2004. Listening to speech activates motor areas involved in speech production. Nat. Neurosci. 7 (7), 701–702.

Wong, P.C., Ettinger, M., Sheppard, J.P., Gunasekera, G.M., Dhar, S., 2010. Neuroanatomical characteristics and speech perception in noise in older adults. Ear Hear. 31 (4), 471–479. doi:10.1097/AUD.0b013e3181d709c2.

Wong, P.C., Jin, J.X., Gunasekera, G.M., Abel, R., Lee, E.R., Dhar, S., 2009a. Aging and cortical mechanisms of speech perception in noise. Neuropsychologia 47 (3), 693–703 doi:S0028-3932(08)00462-4 [pii] 10.1016/j.neuropsychologia.2008.11.032.

Wong, P.C., Nusbaum, H.C., Small, S.L., 2004. Neural bases of talker normalization. J. Cogn. Neurosci. 16 (7), 1173–1184. doi:10.1162/0898929041920522.

Wong, P.C.M., Jin, J.X., Gunasekera, G.M., Abel, R., Lee, E.R., Dhar, S, 2009b. Aging and cortical mechanisms of speech perception in noise. Neuropsychologia 47 (3), 693–703 doi:https://doi.org/10.1016/j.neuropsychologia.2008.11.032.

Wong, P.C.M., Uppunda, A.K., Parrish, T.B., Dhar, S, 2008. Cortical mechanisms of speech perception in noise. J. Speech, Language, Hear. Res. 51 (4), 1026–1041. doi:10.1044/1092-4388(2008/075.

Working Group on Speech Understanding and Aging and Committee on Hearing, B., and Biomechanics, 1988. Speech understanding and aging. J. Acoust. Soc. Am. 83 (3), 859–895.

Yakel, D.A., Rosenblum, L.D., Fortier, M.A., 2000. Effects of talker variability on speechreading. Percept. Psychophys 62 (7), 1405–1412. doi:10.3758/bf03212142.

Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M., Leirer, V.O., 1982. Development and validation of a geriatric depression screening scale: a preliminary report. J. Psychiatr. Res. 17 (1), 37–49.

Zuur, A.F., Leno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1, 3–14. doi:10.1111/j.2041-210X.2009.00001.x.