STUDY PROTOCOL

1 Study purpose and rationale

It is not well understood how older adults use context during speech processing, even though this ability could be key to preserving communicative ability in the face of age-related declines in hearing[1]. Older adults are known to experience increased difficulty comprehending speech in real-world situations, especially in noisy environments[2]. However, it has long been known that **older adults can use the linguistic context to compensate for sensory losses**[3–6]. For instance, when perceiving speech in noise, older adults might have trouble distinguishing similar words in isolation (e.g. “keys” vs. “peas”), but when the words are embedded in a meaningful context (“Did you find your car *keys*?”) their recognition performance recovers, often to the level of young adults[7–10]. Yet, **research using the N400, an EEG measure of incremental language processing, has concluded that older adults are *impaired* in their ability to use context, even in speech without noise**[11–17]. These contradictory findings reveal a gap in our understanding of age-related changes in speech perception: behavioral evidence suggests that older adults successfully use context to comprehend speech in noise, but we do not know the mechanism by which they achieve this. Furthermore, the incremental processing deficit likely varies across individuals, as do hearing ability and compensatory cognitive strategies, but we do not yet know how these are related, nor do we have a reliable way to measure different compensatory strategies.

Recent research conducted with young adults (including from PI Brodbeck)[18–22], suggests that incremental processing in natural speech comprehension is far more complex than measured by the N400. While the N400 assesses the use of the sentence context in word perception, listeners also: (1) maintain parallel representations of sublexical and word-form contexts; (2) are sensitive to context-dependent *uncertainty* (entropy) about upcoming speech, suggesting that they (3) actively predict upcoming *words* and *speech sounds*. It is unknown how these different components of speech processing are affected in older adults. Yet, their interplay is likely key to successful speech comprehension. We will address older adults’ use of linguistic context for speech comprehension using behavioral, EEG and fMRI measures. With a large dataset, we will evaluate predictors of successful comprehension, while taking into account individual differences in hearing and cognitive abilities. We capitalize on Brodbeck‘s cutting-edge techniques to model brain responses to continuous speech in *naturalistic* listening conditions, taking advantage of natural variability of narrative speech. This approach ensures that older adults are tested in an engaging context in which they process meaningful speech as in real life.

Aim 1: Test the effect of age on incremental speech processing, in quiet and in noise. We evaluate millisecond-by-millisecond information processing in speech, using EEG responses to narrative speech. We test the central hypothesis that incremental processing is reduced in normal-hearing older adults (60+ years), compared to a control group of young adults. We ask whether these age effects are modulated by the presence of background noise (motivated by the observation that behavioral studies often used speech in noise, while N400 studies used speech in quiet). We also test age effects in non-incremental, alternative processing models.

Aim 2: Determine predictors of individual differences in older adults’ ability to use context to compensate for impaired hearing. We extend the EEG dataset to 200 older adults, adequately powered for individual difference questions (>80% power to detect a medium sized effect[23]). We test hypotheses about different mechanisms of context use to determine which predict individual differences in successful comprehension.

Aim 3. Test whether resilient comprehension depends on compensation from language-internal or -external brain networks. Using fMRI responses to continuous speech, we will test for age effects in the brain networks recruited for speech comprehension in quiet and in noise. We expect that the network related to incremental sentence processing will be attenuated in older adults. We test whether individual differences in comprehension can be predicted from the intactness of this language network, or recruitment of alternative networks.

Given the importance of maintained communication abilities in supporting cognitive and emotional health in aging, **the ability to effectively use context in comprehension may be a protective factor**, and is thus an attractive target for remediation and support. Understanding how older adults use context to compensate for degraded auditory input, and under what conditions such strategies fail, could be instrumental in designing interventions targeting language comprehension. This may be especially relevant if incremental processing *is* impacted, as the N400 studies suggest, but some individuals learn better coping strategies than others, because it may be possible to develop training for adaptive strategies[24]. Furthermore, Aim 3 may reveal potential brain networks as targets for brain stimulation. Neural measures of linguistic processing also have a high potential to be useful measures in audiology applications[19]; for example, verifying successful linguistic processing of speech could be used to assess hearing aid benefit for novel sound processing or noise reduction schemes[25]. Methods developed here can track to what degree speech processing is successfully restored, from auditory to linguistic levels of processing.

2 Description of the population to be studied, inclusion and exclusion criteria

In general, the recruitment targets the healthy general population. For all participants, the following inclusion criteria apply:

- Native speakers of American English.
- No known history of neurological or impairment or disease.
- Free of speech and language disorders (per self-report, and confirmed by short language battery[26]).
- Able to physically travel to the labs.

For the **young healthy control group** (and pilot studies) we recruit using the following criteria:

- Age 18-40. Given differences in language, cognitive, and neural development that continue into adolescent years, a child sample is not appropriate for this study. Adults of middle age (range 41-59) are excluded because they often exhibit mixed characteristics, so that mapping age effects in this population would require an even larger sample.
- Normal hearing.

Older adults:

- Age ≥60. The focus of the study is in understanding individual difference in older adults for which speech in noise hearing complaints are prevalent.
- Normal hearing, or hearing loss that can be corrected during stimulus presentation (approximately < 40-50 dB at any frequency).
- No more than 15 dB between-ear difference hearing loss.

MRI: A subset of participants from the EEG study will be recruited for structural and functional MRI. Additional exclusionary criteria concern compatibility with MRI, including:

- Bodily ferromagnetic materials, including pacemakers, ferromagnetic implants, shrapnel, or tattooed eyeliner.
- Claustrophobia.
- Known or probable pregnancy.
- Mobile enough to get changed and on/off the MRI bed.

3 Sample size (and how sample size was determined)

In Aim 2, the main consideration is the goal of detecting relationships between different measures across individuals. To estimate power, we need to decide what size effect is worth detecting[27]. Because the size of meaningful effect differs substantially across fields, we rely on an estimate based on effects commonly reported in gerontology[23]. For this purpose, a medium size effect is defined as the 50th percentile of effects commonly reported as significant in individual difference research in gerontology, which is $r = 0.2$. To achieve 80% power to detect such an effect, 193 subjects are required. This is the basis for our target sample size of 200.

For a comparison of older and young adults, we estimate 50 young adults. Young adults will have less variability because we do not recruit young adults with hearing loss. In preliminary data, the effect-size of the age difference in the sentence-based TRF peak amplitudes was large ($d \geq 0.99$ in preliminary data, >95% power with group size of 28).

The goal of Aim 3 (fMRI) is comparing individuals with high vs. low **resilient comprehension**. Resilient comprehension is defined as preserved comprehension of meaningful speech relative to acoustic-phonetic hearing ability. This will be operationalized as the residual in a regression model predicting speech recognition with a meaningful context from hearing ability measures for meaningless stimuli. We will recruit participants from the endpoints on this measure to maximize power in the fMRI design[28], aiming for at least 35 older adults in each group (>80% power to detect an effect with $r = 0.2$[28]; younger adults are expected to perform generally well, and we will attempt to match them to resilient comprehenders).

4 Design and detailed description of methodology

The study encompasses three experiments. These may be piloted individually, but for the main sample, we will attempt to recruit the same participants across experiments to allow comparison of measures across experiments (Figure 1). All participants will undergo an initial behavioral assessment. Those that pass the hearing and cognitive screening will be asked to participate in the EEG study. Recruitment will continue until we reach the target sample size. We will selectively recruit EEG participants for the MRI study (see section 3). If we do not find enough MRI participants among EEG participants, we will also recruit participants specifically for the MRI study (who will perform the behavioral assessment and may or may not participate in the EEG study).

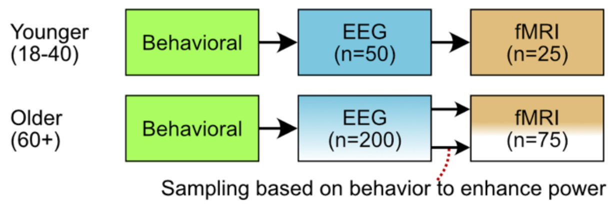

Figure 1. Recruitment strategy, with targeted number of subjects.

4.1 *Auditory stimuli: general parameters*

When background noise is mentioned, this refers to multi-talker babble (unless otherwise noted). This noise is constructed by mixing recordings such as audiobooks from different speakers to produce a relatively uniform background, in which generally no words are recognizable. Backgrounds will be constructed to avoid large fluctuations in amplitude.

The sound pressure level will be fixed for the foreground (audiobook). It will be adjusted before the experiment begins to a level that is comfortable as per self-report from the participant. In hearing tests in Experiment 1, the intensity of the background noise will be adjusted to produce different signal to noise ratios (SNRs) up to -3 dB SNR. This will allow us to determine speech recognition as a function of SNR. These measurements will be used to determine subject-specific SNRs for Experiments 2 and 3, to achieve a fixed percentage correct recognition for 1) meaningless syllables or 2) speech with context (e.g. participants in Experiment 2 will hear speech at an SNR that resulted in 90% of the words being recognized correctly in meaningful sentences in experiment 1).

<u>Hearing loss</u>: For participants with mild hearing loss that do not use hearing aids, stimuli will be amplified in a frequency-specific manner based on their audiogram[29,30]. <u>Hearing aids</u>: For hearing aid users, stimuli will be amplified to the level and characteristics they are accustomed to from their hearing aids.

4.2 *Experiment 1: Hearing health and cognitive evaluation*

This experiment will collect hearing health and cognitive measures which may be relevant for speech recognition ability. This will be achieved using a combination of tests: 1) Standard audiological tests, which provide general measures of hearing health. 2) Audiological tests that are modified to make measurements more comparable for the subsequent EEG and fMRI experiments. These are primarily speech recognition in noise tests, for which a custom background noise will be substituted to provide a common type of background noise across experiments. 3) Cognitive tests commonly used in research or clinical practice, to measure cognitive abilities which have been related to speech comprehension ability in older adults in previous research (e.g., working memory capacity and inhibitory ability).

4.3 Experiment 2: EEG

The goal of this experiment is to measure cognitive processes during listening to naturalistic speech, and to test hypotheses about how these cognitive processes relate to successful comprehension.

Paradigm: EEG is recorded while participants listen to 2-3 minute long segments from an audiobook (*The Big Sleep* by Raymond Chandler, or *The Old Man and the Sea* by Ernest Hemingway). Segment length varies to interrupt segments at meaningful positions. Some of the segments are mixed with background noise (see Section 4.1). In between segments they answer comprehension questions on the content of the preceding segment, to encourage and monitor continued attention. A small set of segments may be acoustically distorted to sound like muffled speech for a language network localizer[31]. The total recording time will be less than 2 hours. After the experiment, subjects will be asked for a subjective judgement of difficulty and listening effort[32] and familiarity with the audiobook, author and genre.

Analysis: The primary approach is to implement neural and cognitive models of speech processing, and to test how well these cognitive models fit the EEG data. Models include models of auditory processing[33] and linguistic predictive processing[20]. For example, how much of the variability in the EEG data can sublexical, lexical and sentence level processing models explain? A primary hypothesis is that sentence-level predictive speech processing (measured with EEG) is related to better comprehension of speech with context, when controlling for peripheral hearing ability (measured in the behavioral experiment). The dataset may also provide further individual difference measures relevant in aging and cognitive functioning, e.g. 1/f scale of EEG background activity[34]. A primary goal is to test hypotheses about neurophysiological correlates of successful speech processing, which will be answered through regression models predicting comprehension performance (in the Behavioral experiment and the EEG experiment) from neurophysiological indices. A primary hypothesis is that the quality of predictive processing, measured in neurophysiological responses[20], predicts intact comprehension over and above non-linguistic hearing ability.

4.4 Experiment 3: MRI

The goal of this experiment is to determine brain networks which contribute to resilient comprehension. For example: is better comprehension associated with increased activity in the classical language network? Or by recruitment of a general-domain executive function network?

Paradigm: Participants will be wearing active noise canceling headphones to minimize interference form the scanner noise. The fMRI paradigm is similar to the EEG experiment, except that segments will be approximately 40 s long to optimize data collection for a block design analysis[35]. Total duration will be less than 90 minutes.

Structural scans: Structural MRI will be acquired to assess overall structural brain health (primarily white matter hyperintensities).

Analysis: Functional brain activity will be assessed based on fit with cognitive models (as in the EEG experiment) as well as through block design analysis (e.g. speech in quiet vs speech in noise). In the primary analysis, functional brain activity and structural brain health will be used to predict resilient comprehension. Secondary hypotheses concerning relationships between other variables will also be tested.

5 Definition of end-point(s)

The primary end-point of the study will be cognitive and neural determinants of resilient comprehension, i.e., the ability to use context to compensate for speech signal degradation during speech recognition (see definition in 3). This will be determined through the interrelationship among peripheral hearing, cognitive abilities, neural information processing strategies (temporal dynamics and anatomical localization), and speech recognition performance. A central hypothesis is that predictive speech processing is an important determinant of resilient comprehension.

A secondary endpoint is the trajectory of speech processing strategies as a function of age.

6 Measurements and measurement instruments

6.1 *Experiment 1: Hearing health and cognitive evaluation*

Tests include standard tests to assess cognitive and hearing health, measures of speech recognition in noise, and cognitive abilities that may predict speech comprehension. Cognitive tests will be administered as pen-and-paper tests or through computerized versions. Verbal responses will be scored by an experimenter. NIH Toolbox tests are administered with the official iPad version of the toolbox (https://www.nihtoolbox.org/).

Hearing Aid Characterization: This is a standard part of hearing aid fitting procedures. Participants will be asked to temporarily remove their hearing aids. Hearing aids will be placed in a Audioscan Verifit2 test box. A sound will be played and output of the hearing aids recorded. This measurement allows estimating general amplification characteristics of the hearing aids, which will then be used to simulate the effect of the hearing aid when presenting stimuli through earphones.

Montreal Cognitive Assessment (MOCA): A widely used pen and paper test to screen for mild cognitive impairment.[36]

Adult Language Battery: Participants will complete three tasks that have been found to identify history of developmental language disorder in adults with good sensitivity and specificity.[26]

- **Clinical Evaluation of Language Fundamentals 4 (CELF-4) – Word Definitions subtest**: Participants are asked to define 24 words. Each word is presented in a sentence to provide semantic context. (e.g., The word is neat, as in "Grandma said, 'You keep your room very neat'").
- **Spelling Test**: Participants are asked to spell 15 words. Words are read to the participants as a single word, used within a sentence, and then again as a single word.
- **Modified Token Test**: An array of colored geometric shapes is laid out in front of the participant. The participant is asked to follow a sequence of instructions (e.g. "Touch the large white circle and the small green square.").

Hearing case history: Participants will be asked standardized questions for a clinical hearing case history in a structured interview (e.g., work in a noisy environment). This is a shortened version of a clinically used audiology intake form.

Speech, Spatial and Qualities of Hearing Scale (SSQ): Questionnaire on perceived hearing and communication difficulties in everyday life[37].

Pure tone audiometry: Standardized audiological measurement. test in which participants are asked to detect pure tones.

Audible Contrast Threshold (ACT) test: A newer test that predicts aided speech-in-noise ability without language stimuli. The subject listens to modulated noise along with intermittent siren sounds. Each time they hear the siren sound, they press the response button.

Distortion Product Otoacoustic Emissions (DPOAE) Testing: DPOAE testing is a safe, non-invasive method of obtaining objective data about the degree of function in the cochlear outer hair cells with no potential risks to participants. Participants will be seated comfortably in a quiet space. A sterilized earphone is placed in the outer one-third of their ear canal. A combined two-tone stimulus sweep is played at simultaneous levels of 55 dB SPL and 65 dB SPL across a frequency range of 1000 Hz to 8000 Hz at minimum of 3 points per octave for a period of approximately 1 minute. A second run of tonal sweeps is performed to identify repeatability of the results. This process is then repeated on the second ear.

Auditory brainstem response: Participants listen to auditory clicks while EEG is measured to assess signal transmission in the auditory brainstem. The system additionally allows non-invasively measuring electrocochleography in parallel with the ABR through earplugs which double as in-ear electrodes (Etymotic ER3-26A). These are simply foam insert earphones wrapped in gold foil, and do not pose any additional risk to participants compared to traditional foam insert earphones.

Speech recognition in noise: measure the effect of background noise on different components of speech recognition:

- **Acoustic speech recognition**: The California Syllables Test (CaST)[38,39] procedure is used to measure recognition of consonant-vowel-consonant (CVC) syllables in the presence of background noise. Participants listen to CVC syllables that were constructed by randomly combining consonants and vowels. After each token, the noise is interrupted, and they repeat what they heard to the experimenter. CVC syllables that form words with strong emotional content are avoided.

- **Speech recognition with context**: The influence of linguistic context will be measured with sentence recognition[6]. In each trial, participants hear a sentence played with background noise, the noise is paused, and they repeat the sentence to the experimenter. Sentences will be drawn from the widely used HINT[40], QuickSIN[41], and R-SPIN[42] tests, which provide a low to high range of contextual constraint.

Working memory: three tests of working memory (including multiple tests allows deriving a common factor, which is more stable than any single test):

- **Reading Span Test**: This is a test of linguistic working memory[43]. In each trial, participants are asked to read several sentences aloud, one sentence at a time. After a variable number of sentences (2-6) a cue prompts participants to verbally reproduce the last word of every sentence they have read in this trial (in any order).We will use the standard computerized version of the test[44].
- **Letter-monitoring task**: another test of linguistic working memory[45]. Participants see a sequence of letters on a screen, one letter at a time. They are instructed that whenever three successive letters represent a word in the English language, they are to press a button. Words in the test are: BAD, BAG, BOG, BUG, DIN, DOG, DOT, GAP, MAD, MAN, MAP, MIT, MUD, MUG, NET, NOT, PAN, PIG, PIT, POP, POT, PUT, TAN, TOP.
- **NIH Toolbox List Sorting Working Memory Test**[46]: “A series of stimuli is presented on the computer screen visually (object) and orally (spoken name), 1 at a time. Participants are instructed to repeat the stimuli to the examiner in order of size, from smallest to largest.”[47] For example, dog, horse, elephant. The task takes approximately 7 minutes.

Inhibitory functioning:

- **Speeded Classification Task**[7]: In each trial, participants hear a word, which may vary on two dimensions: The initial phoneme may be /p/ or /b/, and the voice may be male or female. Participants respond to each word by indicating either the first phoneme or the speaker gender (the task varies by block). In the control condition, the irrelevant dimension is held constant. In the test condition, the relevant and irrelevant dimensions both vary randomly. The increases in response latency in the test condition measures the listeners' ability to inhibit variation in the irrelevant dimension.
- **Auditory Stroop task**[7]: The stimuli consist of three words (*mother*, *father*, and *person*) produced by four different talkers (two male and two female). On each trial, participants hear a token and respond as quickly and as accurately as possible with whether the word was spoken by a man or by a woman. Inhibition is measured by the influence of the semantic content of the word on response latency and error rates.
- **Stroop Color Naming Task**[48,49]: On each trial, participants see a string of letters and are asked to name the color of those letters as quickly and accurately as possible. On baseline trials, the letters form a meaningless sequence (“XXXX”). On experimental trials, the letters form a color name that is inconsistent with the color (e.g., “BLUE” written in red letters). Trials on which the color name is inconsistent with the color measure the interference from the written word on color naming.

NIH Toolbox Pattern Comparison Processing Speed Test[50]: In each trial, participants see two pictures side-by-side and are asked to identify, as quickly as possible, whether the two pictures are the same or different by button-press[47].

Edinburgh Handedness Inventory: Participants indicate which hand they prefer to use on various common tasks[51].

Some of these tasks may indicate potentially unknown medical conditions. If this is the case, we will inform participants, stressing that our tests are not clinical results but that they may want to follow up with a relevant specialist. If hearing tests indicate hearing loss (for participants not already wearing hearing aids) we will inform subjects that they may benefit from a hearing aid. For MOCA scores suggesting possible mild cognitive impairment (score < 26) we will recommend that they follow up with their family doctor. The study team will offer to contact their family doctor with this information. If they wish us to do this, we will ask them to fill in the required information on a form. This form will be destroyed after the family doctor’s practice has been informed.

6.2 Experiment 2: EEG

Participants will sit in a comfortable chair while the experimenter places the electrode cap, attaches electrodes on the cap, and applies contact gel to the individual electrodes to ensure good conductance. Participants will be seated while stimuli are presented through earphones and on a computer screen. After the experiment, technicians or the experimenter will help participants remove any gel from their hair and skin. Participants also have an opportunity to wash their hair using a salon-style sink to remove gel. Towels and shampoo will be provided.

6.3 Experiment 3: MRI

The behavioral task will be the same as in Experiment 2.

MRI Procedures. For the MRI study, study staff will review an MRI safety screening form with the potential participant via email or phone. If the participant indicates a condition that would prevent him/her from participating in MRI scans, (e.g. a participant might have a pacemaker), the participant will not undergo MRI scanning. If the participant indicates that he/she has undergone a surgical procedure, we will request more details of this procedure from the participant and his/her family to ensure that no metallic objects or devices remain in the participant. If there is any doubt as to the participant's safety in the MRI environment, he/she will not be included in the MRI study.

After consent and safety screening, participants will enter the scan suite, will lie supine on the scanner bed, will wear hearing protection (active noise-cancelling headphones). We use cushions to position the participant comfortably and provide a squeeze bulb if the participant needs to urgently communicate with the study staff. Participants will be requested to remain as still as possible during imaging. During functional scans, participants will respond to comprehension questions using a response device. Scan sessions will last no more than 90 minutes, typically closer to 60 minutes.

After the experiment, subjects will be asked for a subjective judgement of difficulty and listening effort[32] and familiarity with the audiobook, author and genre.

7 Data analysis plan

7.1 Handling of identifying information

Linking data across experiments: Participants' name and date of birth will be recorded on the consent form at the beginning of each experiment. This record will be kept to minimize the risk of mixing up participants across experiment, because linking individuals across experiments is critical for the study goals. Each participant will be assigned an arbitrary code, and only this code will be used to link data from the same participant. The only place where name and date of birth will be recorded together with the participant's code is the consent form, and these paper forms will be stored in a locked cabinet in the PI's office.

Participants' names and contact information will be temporarily kept in a McMaster institutional email (Microsoft Teams) account for scheduling purposes, and may be retained as a list of participants that agree to be contacted for future studies, or request information about the outcome of the study. These lists will *not* include participant codes.

Questionnaires: Some questionnaires may result in identifying information (primarily Hearing Case History). These questionnaires will be stored in a locked cabinet in the PI's office. Only summary variables without identifying information will be electronically recorded and used for data analysis (e.g. binary variable for history of noise exposure at work). Date of birth will be recorded on the paper form but will be re-coded and electronically recorded as "age at the time of the experiment".

MRI: Structural MRI contain identifying information because they can be used to reconstruct facial features. This is commonly solved by removing those facial features form the recording ("defacing"). We will store original MRIs in a secure location (physical hard drive in PI's office and encrypted backup on McMaster institutional OneDrive). For data analysis purposes we will work only with defaced images only (using FreeSurfer "mideface").

7.2 Analysis using de-identified data

Data used for analysis will not contain identifying information (see Data Collection Form). This data will be stored on computers used for data analysis and backed up on local hard drives.

Individual difference variables will be derived from measures of cognitive health, hearing, and speech recognition tests. Measures characterizing brain health and individual fits with different cognitive models (information processing strategies, for example, predictive processing based on a large sentence context vs a local sublexical context[20]) will be derived from neural data (EEG/MRI). Data will also be used for model development (e.g., which of two cognitive model fits data better overall) before testing individual differences in those models. The primary goal is to analyze the relationship between Individual difference variables, i.e., determine which variables predict speech recognition ability. Secondary hypotheses concern main effects of age on information processing, after controlling for hearing ability.

De-identified data may be shared with other researchers or uploaded to an online repository and made publicly available to maximize the potential research benefits from the data.

8 How subjects will be recruited, including advertisements/publicity

Participants will be recruited via internet by flier and social media post. The McMaster Institute for Research on Aging (MIRA) will assist in advertising the study by distributing the flier, including in community email bulletin, on Voice Canada, and at community aging events.

Potential participants will be asked to confirm they meet inclusion criteria by email or phone (see Section 2). Participants will receive a copy of the Consent Form by email or paper mail before visiting the lab.

Participants who gave consent (on the consent form for the EEG study) will be selectively re-contacted for the fMRI study. To determine eligibility for fMRI, initial screening questions from the CIAMI MRI Screening Form will be administered by phone or email.

As incentives participants will be paid $20 per hour for experiments. Older adults will be paid $30/hour for fMRI to account for potentially increased discomfort in the scanner, and to increase retention of the same subjects. Participants will also receive a reimbursement for travel and parking (either $10, or, for older adults only, up to $80 with cab receipt, as older adults may prefer to take a cab).

9 References

1 Morrell CH, Gordon-Salant S, Pearson JD, Brant LJ, Fozard JL. Age- and gender-specific reference ranges for hearing level and longitudinal changes in hearing level. *J Acoust Soc Am* 1996;**100**:1949–67. https://doi.org/10.1121/1.417906.

2 Schneider BA, Daneman M, Pichora-Fuller MK. Listening in aging adults: From discourse comprehension to psychoacoustics. *Can J Exp Psychol Can Psychol Expérimentale* 2002;**56**:139–52. https://doi.org/10.1037/h0087392.

3 Pichora-Fuller MK, Schneider BA, Daneman M. How young and old adults listen to and remember speech in noise. *J Acoust Soc Am* 1995;**97**:593–608. https://doi.org/10.1121/1.412282.

4 Pichora-Fuller MK. Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *Int J Audiol* 2008;**47**:S72–82. https://doi.org/10.1080/14992020802307404.

5 Sheldon S, Pichora-Fuller MK, Schneider BA. Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *J Acoust Soc Am* 2008;**123**:489–99. https://doi.org/10.1121/1.2783762.

6 Woods DL, Arbogast T, Doss Z, Younus M, Herron TJ, Yund EW. Aided and Unaided Speech Perception by Older Hearing Impaired Listeners. *PLoS ONE* 2015;**10**:1–29. https://doi.org/10.1371/journal.pone.0114922.

7 Sommers MS, Danielson SM. Inhibitory Processes and Spoken Word Recognition in Young and Older Adults: The Interaction of Lexical Competition and Semantic Context. *Psychol Aging* 1999;**14**:15. https://doi.org/10.1037/0882-7974.14.3.458.

8 Dubno JR, Ahlstrom JB, Horwitz AR. Use of context by young and aged adults with normal hearing. *J Acoust Soc Am* 2000;**107**:538–46. https://doi.org/10.1121/1.428322.

9 Stewart R, Wingfield A. Hearing Loss and Cognitive Effort in Older Adults' Report Accuracy for Verbal Materials. *J Am Acad Audiol* 2009;**20**:147–54. https://doi.org/10.3766/jaaa.20.2.7.

10 Goy H, Pelletier M, Coletta M, Pichora-Fuller MK. The Effects of Semantic Context and the Type and Amount of Acoustic Distortion on Lexical Decision by Younger and Older Adults. *J Speech Lang Hear Res* 2013;**56**:1715–32. https://doi.org/10.1044/1092-4388(2013/12-0053).
11 Federmeier KD, McLennan DB, Ochoa ED, Kutas M. The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology* 2002;**39**:133–46. https://doi.org/10.1111/1469-8986.3920133.
12 Federmeier KD, Kutas M, Schul R. Age-related and individual differences in the use of prediction during language comprehension. *Brain Lang* 2010;**115**:149–61. https://doi.org/10.1016/j.bandl.2010.07.006.
13 Wlotko EW, Federmeier KD. Age-related changes in the impact of contextual strength on multiple aspects of sentence comprehension. *Psychophysiology* 2012;**49**:770–85. https://doi.org/10.1111/j.1469-8986.2012.01366.x.
14 Dave S, Brothers TA, Traxler MJ, Ferreira F, Henderson JM, Swaab TY. Electrophysiological evidence for preserved primacy of lexical prediction in aging. *Neuropsychologia* 2018;**117**:135–47. https://doi.org/10.1016/j.neuropsychologia.2018.05.023.
15 Payne BR, Federmeier KD. Contextual constraints on lexico-semantic processing in aging: Evidence from single-word event-related brain potentials. *Brain Res* 2018;**1687**:117–28. https://doi.org/10.1016/j.brainres.2018.02.021.
16 Wlotko EW, Lee C-L, Federmeier KD. Language of the Aging Brain: Event-Related Potential Studies of Comprehension in Older Adults. *Lang Linguist Compass* 2010;**4**:623–38. https://doi.org/10.1111/j.1749-818X.2010.00224.x.
17 Payne BR, Silcox JW. Aging, context processing, and comprehension. In: Federmeier KD, editor. *Psychol. Learn. Motiv.*, vol. 71. Academic Press; 2019. p. 215–64.
18 Brodbeck C, Hong LE, Simon JZ. Rapid Transformation from Auditory to Linguistic Representations of Continuous Speech. *Curr Biol* 2018;**28**:3976-3983.e5. https://doi.org/10.1016/j.cub.2018.10.042.
19 Gillis M, Vanthornhout J, Simon JZ, Francart T, Brodbeck C. Neural markers of speech comprehension: measuring EEG tracking of linguistic speech representations, controlling the speech acoustics. *J Neurosci* 2021;**41**:10316–29. https://doi.org/10.1523/JNEUROSCI.0812-21.2021.
20 Brodbeck C, Bhattasali S, Cruz Heredia AA, Resnik P, Simon JZ, Lau E. Parallel processing in speech perception with local and global representations of linguistic context. *eLife* 2022;**11**:e72056. https://doi.org/10.7554/eLife.72056.
21 Gaston P, Brodbeck C, Phillips C, Lau E. Auditory Word Comprehension Is Less Incremental in Isolated Words. *Neurobiol Lang* 2023;**4**:1–24. https://doi.org/10.1162/nol_a_00084.
22 Xie Z, Brodbeck C, Chandrasekaran B. Cortical Tracking of Continuous Speech Under Bimodal Divided Attention. *Neurobiol Lang* 2023;**4**:1–26. https://doi.org/10.1162/nol_a_00100.
23 Brydges CR. Effect Size Guidelines, Sample Size Calculations, and Statistical Power in Gerontology. *Innov Aging* 2019;**3**:igz036. https://doi.org/10.1093/geroni/igz036.
24 Huettig F, Pickering MJ. Literacy Advantages Beyond Reading: Prediction of Spoken Language. *Trends Cogn Sci* 2019;**23**:464–75. https://doi.org/10.1016/j.tics.2019.03.008.
25 Alickovic E, Lunner T, Wendt D, Fiedler L, Hietkamp R, Ng EHN, *et al.* Neural Representation Enhanced for Speech and Reduced for Background Noise With a Hearing Aid Noise Reduction Scheme During a Selective Attention Task. *Front Neurosci* 2020;**14**:846. https://doi.org/10.3389/fnins.2020.00846.
26 Fidler LJ, Plante E, Vance R. Identification of Adults With Developmental Language Impairments. *Am J Speech Lang Pathol* 2011;**20**:2–13. https://doi.org/10.1044/1058-0360(2010/09-0096).
27 Cohen J. A power primer. *Psychol Bull* 1992;**112**:155–9. https://doi.org/10.1037/0033-2909.112.1.155.
28 de Haas B. How to Enhance the Power to Detect Brain–Behavior Correlations With Limited Resources. *Front Hum Neurosci* 2018;**12**:.
29 Byrne D, Dillon H, Ching T, Katsch R, Keidser G. NAL-NL1 Procedure for Fitting Nonlinear Hearing Aids: Characteristics and Comparisons with Other Procedures. *J Am Acad Audiol* 2001;**12**:37–51. https://doi.org/10.1055/s-0041-1741117.
30 Decruy L, Vanthornhout J, Francart T. Hearing impairment is associated with enhanced neural tracking of the speech envelope. *Hear Res* 2020;**393**:107961. https://doi.org/10.1016/j.heares.2020.107961.
31 Scott TL, Gallée J, Fedorenko E. A new fun and robust version of an fMRI localizer for the frontotemporal language system. *Cogn Neurosci* 2017;**8**:167–76. https://doi.org/10.1080/17588928.2016.1201466.
32 Peelle JE. Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior. *Ear Hear* 2018;**39**:204–14.

33 Brodbeck C, Jiao A, Hong LE, Simon JZ. Neural speech restoration at the cocktail party: Auditory cortex recovers masked speech of both attended and ignored speakers. *PLOS Biol* 2020;**18**:e3000883. https://doi.org/10.1371/journal.pbio.3000883.
34 Voytek B, Kramer MA, Case J, Lepage KQ, Tempesta ZR, Knight RT, *et al.* Age-Related Changes in 1/f Neural Electrophysiological Noise. *J Neurosci* 2015;**35**:13257–65. https://doi.org/10.1523/JNEUROSCI.2332-14.2015.
35 Henson R. Chapter 15 - Efficient experimental design for fMRI. In: Friston K, Ashburner J, Kiebel S, Nichols T, Penny W, editors. *Stat. Parametr. Mapp.* Elsevier; 2007.
36 Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, *et al.* The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc* 2005;**53**:695–9. https://doi.org/10.1111/j.1532-5415.2005.53221.x.
37 Gatehouse S, Noble W. The Speech, Spatial and Qualities of Hearing Scale (SSQ). *Int J Audiol* 2004;**43**:85–99. https://doi.org/10.1080/14992020400050014.
38 Woods DL, Yund EW, Herron TJ. Measuring consonant identification in nonsense syllables, words, and sentences. *J Rehabil Res Dev* 2010;**47**:243. https://doi.org/10.1682/JRRD.2009.04.0040.
39 Woods DavidL, Yund EW, Herron TJ, Cruadhlaoich MAIU. Consonant identification in consonant-vowel-consonant syllables in speech-spectrum noise. *J Acoust Soc Am* 2010;**127**:1609–23. https://doi.org/10.1121/1.3293005.
40 Nilsson M, Soli SD, Sullivan JA. Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am* 1994;**95**:1085–99. https://doi.org/10.1121/1.408469.
41 Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 2004;**116**:2395–405. https://doi.org/10.1121/1.1784440.
42 Bilger RC, Nuetzel JM, Rabinowitz WM, Rzeczkowski C. Standardization of a Test of Speech Perception in Noise. *J Speech Lang Hear Res* 1984;**27**:32–48. https://doi.org/10.1044/jshr.2701.32.
43 Daneman M, Carpenter PA. Individual differences in working memory and reading. *J Verbal Learn Verbal Behav* 1980;**19**:450–66. https://doi.org/10.1016/S0022-5371(80)90312-6.
44 van den Noort M, Bosch P, Haverkort M, Hugdahl K. A Standard Computerized Version of the Reading Span Test in Different Languages. *Eur J Psychol Assess* 2008;**24**:35–42. https://doi.org/10.1027/1015-5759.24.1.35.
45 Gatehouse S, Naylor G, Elberling C. Benefits from hearing aids in relation to the interaction between the user and the environment. *Int J Audiol* 2003;**42**:77–85. https://doi.org/10.3109/14992020309074627.
46 Tulsky DS, Carlozzi N, Chiaravalloti ND, Beaumont JL, Kisala PA, Mungas D, *et al.* NIH Toolbox Cognition Battery (NIHTB-CB): List Sorting Test to Measure Working Memory. *J Int Neuropsychol Soc* 2014;**20**:599–610. https://doi.org/10.1017/S135561771400040X.
47 Weintraub S, Dikmen SS, Heaton RK, Tulsky DS, Zelazo PD, Bauer PJ, *et al.* Cognition assessment using the NIH Toolbox. *Neurology* 2013;**80**:. https://doi.org/10.1212/WNL.0b013e3182872ded.
48 Stroop JR. Studies of interference in serial verbal reactions. *J Exp Psychol* 1935;**18**:643–62. https://doi.org/10.1037/h0054651.
49 Knight S, Heinrich A. Different Measures of Auditory and Visual Stroop Interference and Their Relationship to Speech Intelligibility in Noise. *Front Psychol* 2017;**8**:. https://doi.org/10.3389/fpsyg.2017.00230.
50 Carlozzi NE, Beaumont JL, Tulsky DS, Gershon RC. The NIH Toolbox Pattern Comparison Processing Speed Test: Normative Data. *Arch Clin Neuropsychol* 2015;**30**:359–68. https://doi.org/10.1093/arclin/acv031.
51 Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 1971;**9**:97–113. https://doi.org/10.1016/0028-3932(71)90067-4.