Project Summary/Abstract

A frequent complaint in older adults is difficulty with speech comprehension, especially in the presence of background noise, such as when at a busy restaurant. A common strategy that listeners use in such situations is to use the linguistic context to fill in gaps in comprehension. Older adults tend to do well with this, at least in laboratory situations: given a meaningful linguistic context, their word recognition is much more like young adults' than when speech is meaningless. An open question is what cognitive mechanisms underly this ability. Based on research in young adults, it is widely believed that prediction plays a central role in language processing – listeners could predict upcoming words from a meaningful context, and thus compensate for noisy input. Yet, a line of research using an electroencephalography (EEG) signature of predictive language processing, the N400, suggests that predictive processing is impaired in older adults. This raises the question of how, then, older adults can successfully use context to facilitate speech recognition. This proposal uses cutting-edge neuroimaging techniques to examine a wider range of predictive speech processing than previous work, as well as non-predictive compensatory mechanisms, in younger and older adults. To maximize relevance to older adults' real-world experience, brain responses are measured while listening to continuous narrative speech, in quiet and in multi-talker babble background noise, approximating conditions in a busy restaurant. EEG is used to model the temporal dynamics of predictive language processing at multiple levels (acoustic-phonetic, sublexical, lexical, sentential). This will test several hypothesis that older adults use predictions during speech processing, but at different levels than young adults. Functional magnetic resonance imaging (fMRI) is used to characterize language networks related to speech processing in quiet and in noise. Whereas previous research has mostly relied on small sample sizes, obscuring individual differences, this work will be based on a large dataset, adequately powered for individual difference analysis, and will collect a range of individual variables that may be related to speech comprehension, such as hearing ability, working memory span and inhibitory functioning. This work will relate individual differences in prediction at multiple levels of the processing hierarchy to language, cognitive, and functional outcomes. Together, this work aims to uncover which cognitive strategies (predictive or non-predictive), neural representations and network are related to successful speech comprehension in older adults. By relating neural measures to comprehension, it will distinguish between successful compensation and maladaptive changes in brain activation. The dataset will be shared publicly, providing a unique testbed for future researchers interested in speech and aging.

Project Narrative

A frequent complaint in older adults is difficulty with speech comprehension, especially in the presence of background noise. Using a high-powered individual difference approach, this project will use neural and behavioral measures to determine successful and unsuccessful adaptation mechanisms. This work aims to discover learnable cognitive strategies for improving speech recognition, and generate biomarkers for successful and harmful adaptations.

INTRODUCTION TO RESUBMISSION

We thank the reviewers of the initial application for their positive comments and constructive suggestions. The reviewers "agreed that the proposed work is highly significant and, if successful, would yield important information about the language processing system in general as well as how language processing changes and shifts in older adults" (panel summary). We have revised this resubmission substantially to address the concerns that were raised. The most substantive changes are described below.

Prediction was given undue attention compared with other potential mechanisms of context use, like integration. We have addressed this both by a more nuanced discussion of existing models, as well as inclusion of new models. First, we note that while brain responses related to surprisal (including the N400) are often interpreted as involving prediction, surprisal effects may also be due to incremental information processing without active prediction of upcoming speech[1,2]. The age difference in the N400 is thus the more striking, as it implies that older adults have a deficit in incrementally integrating information in speech. Incremental processing of information is thought to be critical to successful speech comprehension, because the speech signal evolves rapidly and cannot be stored in its raw form to allow later analysis[3]. The proposed work will: (1) verify whether older adults indeed show a deficit in incremental processing that leads to impaired comprehension (by comparing neural and behavioral data), and (2) use a range of alternative models of incremental and non-incremental processing to test what specific neural mechanisms do explain comprehension performance.

The use of language models trained on web-based corpora biases data analysis towards younger participants. We have added a plan to use historical corpora with texts from 1930 onwards to train new language models that reflect individuals' language experience. In addition to addressing the confound pointed out by the reviewers (younger and older adults are typically assessed using a single corpus), this will also allow us to test whether the language system is tuned to statistics that reflect a given formative period (e.g., young adulthood) or whether it continues to incorporate input throughout adulthood to reflect recent history (e.g., the last two decades). By looking at "generationally-sensitive" corpora, we will be able to contribute to best-practice recommendations for estimating language exposure in diverse populations.

There was a minor concern that the proposal covered too many areas, including ones in which the team lacked expertise. We have clarified that the focus of this proposal is on relating neural mechanisms (Brodbeck, Myers) to comprehension and hearing ability (Hinchey, Anderson) in older adults (Anderson, Myers). However, given the large number of alternative hypotheses in the literature, we err on the side of over-including <u>standardized</u> individual difference measures. We also failed to highlight relevant expertise in using these measures in our team: Co-I Myers and consultant Anderson have expertise in using working memory[4–6], processing speed[4–6], and inhibitory funtioning[5–7] measures. Myers also has published expertise with brain morphology measures[8,9].

It was unclear how the samples of younger and older adults with different sample sizes would be analyzed, as the two samples were not matched. To clarify our approach we have now separated the former Aim 1 into two aims: Aim 1 is to compare older adults to younger adults, whereas Aim 2 is studying individual differences among 200 older adults. For Aim 1, one-to-one matching between younger and older adults is virtually impossible, because these populations differ inherently in many respects. We take two approaches: (1) We characterize general age-related differences (and replicate studies that themselves did not perform detailed matching[10–14]) by comparing 50 older adults with a 50 younger adults. We match 50 older adults, a subset from the larger sample collected for Aim 2, to the younger adults by approximating the prevalence of clinical hearing loss (which we expect to be low in younger adults, leading to similar differences as in the published literature). These population differences are relevant even if they may eventually be explained by other individual differences. (2) We expect a sample of younger adults to be more homogeneous (e.g., on average better hearing). However, regression models (e.g., predicting comprehension from hearing loss) allow testing for an interaction with age. In effect, we ask whether the regression fit to older adults can also account for younger adults' data. This will allow us to use the full dataset, and will allow us to determine in what respects younger adults are like older adults unencumbered by sensory deficits, and in what respects they differ more fundamentally.

To address a gap in the **characterization of individual differences in the peripheral auditory system** we have added extended pure tone threshold measurements to detect extended high frequency hearing loss[15,16], and a click auditory brainstem response. To improve measurement of **working memory**, we now use three tests. We have also clarified several aspects of the Research Strategy that reviewers pointed out were underspecified, including the processing speed measure (NIH Toolbox), signal to noise ratios for speech in noise stimuli (see *Approach*).

SPECIFIC AIMS

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[17]. Older adults are known to experience increased difficulty comprehending speech in real-world situations, especially in noisy environments[18]. However, it has long been known that **older adults can use the linguistic context to compensate for sensory losses**[19–22]. 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[23–26]. 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**[10–14,27,28]. 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)[29–33], 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[34]). 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[35]. 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[30]; for example, verifying successful linguistic processing of speech could be used to assess hearing aid benefit for novel sound processing or noise reduction schemes[36]. Methods developed here can track to what degree speech processing is successfully restored, from auditory to linguistic levels of processing.

SIGNIFICANCE

Intact communication abilities are key to quality of life in aging. Hearing acuity invariably declines with age, even in older adults classified as normal-hearing[17], posing a challenge for speech perception, particularly in noisy environments[18]. As a group, **older adults can use linguistic context to compensate for sensory losses**[19–26]. Context can facilitate speech perception because it provides a top-down constraint to help interpret ambiguous acoustic input[37,38]. Older adults may, thus, compensate for afferent auditory deficits by relying more heavily on context[39]. This is especially plausible given a wide range of evidence for preserved[40–45] or even increased[46–51] offline language knowledge in older adults, as well as evidence that older adults rely more on context than young adults during reading[52–54], non-linguistic perception[55] and motor function[56].

The use of context in speech processing is thought to be highly incremental[3]. The speech signal evolves rapidly in time, and limitations of listeners' memory systems create a pressure to compress and recode the input as rapidly as possible[3]. Empirically, new language input is quickly used to update internal representations at multiple levels[31,57–60]. A widely studied neural signature of incremental processing is the N400 event-related potential[61]. The N400 is a response thought to occur after every word[62], reflecting the effect of this word on an evolving semantic representation[63]. In young adults, its amplitude is systematically reduced when a word is predictable from its context[61] (e.g., reduced response to <u>bar</u> in "We got drunk in the local <u>bar</u>", compared to "Tom asked about the <u>bar</u>"). This suggests that processing is incremental, i.e., (1) semantic updates that could be anticipated have already been done, and (2) semantic updates possible due to the new word are done immediately. In older adults, this N400 difference of predictability is often reduced and/or delayed in time. This has been found in spoken sentences[10,64–67], sentence reading[12–14,68–70] and reading of primed vs unprimed words[71–73], and **implies that older adults process language less incrementally than younger adults**[74].

This reveals a gap in our understanding of how older adults use context: Do older adults generally process speech less incrementally, and instead rely on an alternative mechanism to successfully use context?[75] Or are these group-level findings obscuring individual differences, such that older adults with more preserved incremental processing are the ones driving the finding of successful context use? More generally, context use strategies may differ across individuals[11,76,77], as does speech in noise recognition ability[78–81]. Individual differences might also explain some heterogeneity among studies with smaller numbers of participants, as not all studies found the expected N400 effect difference between younger and older adults[77]. This question has high public health relevance for older adults, because of the importance of maintained communication abilities in supporting cognitive and emotional health. **The ability to effectively use context in comprehension may thus be a protective factor in aging**, and an attractive target for remediation and support. In this proposal, we use neural signatures of different mechanisms for context use to determine the mechanisms underlying successful comprehension as participants listen to speech in quiet and in noise. Using natural language materials ensures that we are better positioned to translate this work to real-world listening conditions.

Aim 1: The effect of age on incremental speech processing

Aim 1 is to better understand age effects in the neural mechanisms of speech processing, including (1) a more refined neuro-cognitive model of incremental speech processing, and (2) novel neuro-cognitive models of non-incremental speech processing. The N400 literature referenced above has focused on incremental processing of words in isolated sentences. However, recent **research conducted only with young adults, including research by PI Brodbeck, suggests that incremental processing in natural speech comprehension is more complex**: (1) Incremental processing is not restricted to the word level, but occurs even at the level of speech sounds (e.g., upon hearing "Massachuse–", listeners anticipate that the next sound will be "t")[29–31,82–85]. (2) Incremental processing also involves probabilistic (graded) prediction of upcoming speech sounds and words[2,31], although theoretical accounts differ on whether this is an essential component of speech processing[82,86–89], or not[90]. Importantly, predictive processing may aid comprehension[35], and may be amenable to training[91,92]. Predictive processing can be dissociated in brain responses from other forms of incremental processing through measures of uncertainty (i.e., *entropy*; e.g., "bant–" will likely continue as be "banter", whereas "barr–" could continue as "barrel", "barracks", "barren", …)[2,29,31,32], and may engage specific brain regions[93]. (3) Brodbeck's recent work also shows that young listeners maintain not just a single incremental model of speech input, but instead maintain parallel models that represent different levels of context: *sublexical context* (sequences of speech sounds independent of how they form words), *word-form context* (speech sounds used to identify the current word, independent of preceding words) and *sentence context* (combining speech sounds used to identify the current word with predictions about what words are more likely given the preceding sentence context)[31]. It is unknown how these different components of incremental processing are affected in older adults. Yet, their interplay may be key to successful speech comprehension.

Alternatively, older adults may shift to a less incremental mode of language processing, relying on a different mechanism instead. Most well-established neural correlates of linguistic speech processing relate to incremental processing (although this may also reflect a bias in the field towards studying younger populations, underlining the need for more data on older adults). Given the lack of well-established measures, we will evaluate two more exploratory directions. (1) Brain activity associated with syntactic structure, including open nodes and parser actions (e.g., brain activity increases over the course of syntactic phrases, and drops after phrase ends)[94–96]. This may reflect a mechanism that accumulates information until phrases can be built, rather than using information incrementally. Indeed, our pilot data suggest that older adults have an *increased* response associated with such cumulative phrase processing (Figure 4). (2) One hypothesis holds that instead of processing context incrementally, older adults draw on context selectively when they encounter distorted acoustic input[75]. Thus, older adults should exhibit increased brain activity in response to ambiguous segments in the acoustic input. We will use time-dependent estimates of acoustic ambiguity in speech in noise to test for such a mechanism.

We will capitalize on Brodbeck's cutting-edge techniques to model neural responses relating to these mechanisms in continuous speech, as it unfolds in *naturalistic* listening conditions. The temporal response function (TRF) method models EEG signals as brain responses to different, concurrent speech features (Figure 1). <u>Our preliminary data show that, with this method, it is feasible to simultaneously characterize age-related changes in speech processing at multiple levels, from subcortical to sentence-level processing</u> (Figure 2). The use of continuous, narrative speech will ensure that older adults are tested in an engaging context in which they process meaningful speech, as they would in real life[97]. This will address two central questions:

1. Do older adults increase incremental processing for speech in noise? Work on the behavioral context benefit in older adults primarily comes from speech in noise, whereas the N400 literature has used speech in quiet. This leaves the possibility that older adults are able to engage incremental mechanisms in noise. We will test this hypothesis by comparing brain signatures of speech processing during speech in quiet and in noise.

2. Do older adults process contexts differently? Young adults primarily rely on a global context incorporating sentence constraints[31]. Older adults might instead rely on a different context representation, for example relying more on sublexical contexts or non-incremental processing (see measures described above in this section).

Furthermore, we can address an alternative hypothesis from the predictive coding framework: ***Are reduced N400s due to reduced representation of prediction error[98], rather than impaired incremental processing?*** Under the this framework, older adults have had more lifetime experience to fine-tune their internal language model, and may thus perform smaller updates when they mis-predict[99]. An N400 reduction then is not due to a lack of prediction, but due to a smaller brain response when predictions are disconfirmed. According to this hypothesis, we should find age-related reduction in indicators of prediction error (i.e., *surprisal*), but not in other indicators of predictive processing (*entropy*). If true, this would have very different implications than the common interpretation of the N400 results, because this would suggest that older adults have intact predictive processing.

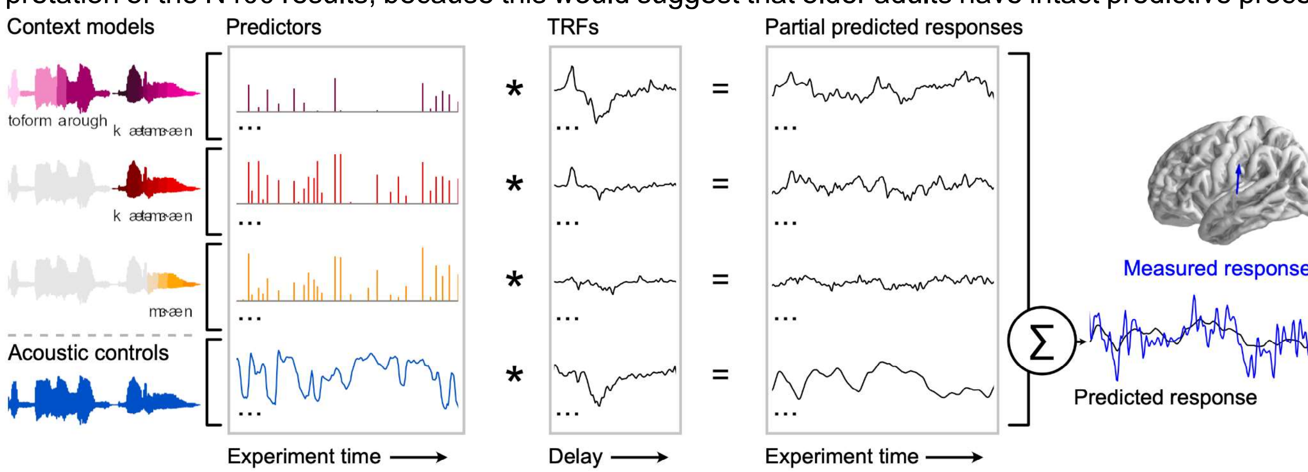

Figure 1. **Multivariate temporal response function (mTRF) method for predictive processing in continuous speech**. *Context models* are language models that use different definitions of context to predict upcoming phonemes; They are turned into time-continuous *predictors* with impulses at phoneme onset, scaled by relevant variables (e.g., phoneme surprisal); *TRFs* are estimated such that convolving the predictors with the TRFs, and summing these *partial predicted responses*, predicts the measures EEG response (see also Figure 7; graphic from Brodbeck et al., 2022).

Aim 2: Predictors of individual differences in older adults' ability to use context

Aim 2 will determine which neural mechanisms (see Aim 1) in fact aid or impair speech comprehension in older adults. For adequate power for individual difference questions we will sample 200 older adults[34]. By measuring comprehension (separately, and during EEG) as well as hearing ability (in a detailed audiological evaluation), we will determine neuro-cognitive factors related to *preserved* comprehension, i.e., comprehension ability that is better than would be expected based on hearing ability alone. This will allow us to answer the central question of Aim 2: ***Which linguistic context processing mechanisms can compensate for auditory deficits?*** EEG data allow joint estimation of the quality of subcortical[100–103] and cortical auditory representations[104–107] (see Figure 2). For example, older adults, compared to young adults, tend to have weaker subcortical auditory representations, but exaggerated early cortical responses[7,108–112]; and hearing loss is associated with delayed responses to acoustic features[107]. For speech in noise, acoustic representations reflect the quality of stream segregation[113–115,106]. We will evaluate how acoustic and linguistic representations jointly determine comprehension, relative to hearing loss. We will test the central hypothesis that relatively intact incremental processing attenuates the impact of impaired auditory processing, but we will also test whether other neural mechanisms can offset or exacerbate auditory deficits. We will also be able to address secondary questions:

Is the breakdown of incremental processing due to slowed sensory processing? While older adults have preserved offline language ability, decreased neural processing speed may make it harder for them to deploy that knowledge in real-time[116,117]. Our approach will allow detailed assessment of latencies at which information is processed by the brain[106], from acoustic to linguistic representations. E.g., while auditory responses in normal hearing older adults are not delayed compared to young adults[109], they are in listeners with hearing loss[107]. We will test whether such a delay has knock-on effects at linguistic levels, and whether increased processing latencies might thus interfere with the ability to use context incrementally. As an exploratory measure, we also collect a standardized behavioral measure of processing speed. However, we do not expect this measure to predict neural processing because auditory responses are not delayed in normal hearing older adults (behavioral processing speed is), and behavioral slowing may be due to response execution rather than sensory processing[118].

Does working memory capacity affect comprehension mechanisms? In older adults, working memory capacity may be an important cognitive predictor for speech recognition in noise[75,119,120] (although it may be primarily relevant in hearing loss[121]). Working memory could mediate comprehension in two ways: First, larger working memory capacity might allow maintaining and manipulating information to enable incremental processing[90,122].

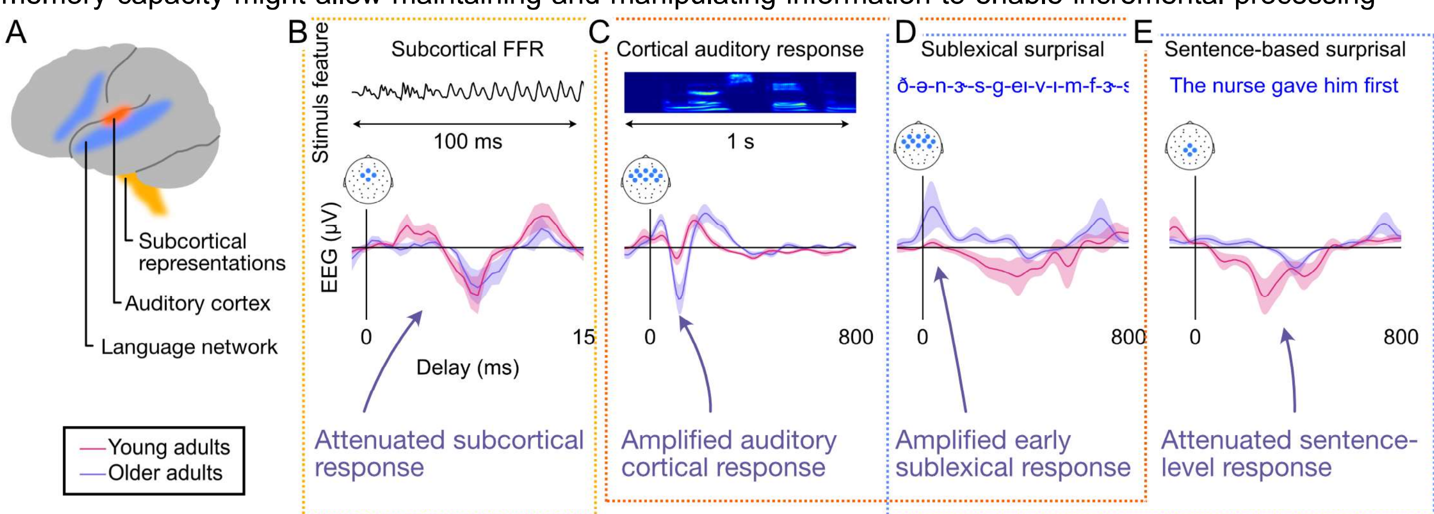

Figure 2. **Simultaneous measurement of age-related changes across different levels of speech processing in EEG (preliminary data)**. Responses from younger and older adults listening to isolated sentences replicate commonly reported age effects. (A) Using appropriate analysis parameters and predictor variables, speech processing can be measured from subcortical representations to the extended language network in the same brain responses to continuous speech. (B) Subcortical responses are seen in EEG responses that track fast (> 70 Hz) fluctuations in the sound wave with ~9 ms latency. As previously reported, older adults exhibit attenuated subcortical responses. (C) Auditory cortex responses primarily track slow (≤ 20 Hz) fluctuations in the auditory spectrogram. As previously reported, these are overrepresented in older adults. (D) Linguistic processing is measured using probabilistic language models. Our preliminary data indicate that older adults' responses track sublexical sequence probabilities more strongly than young adults'. (E) Lexical predictive processing is measured as a response that tracks how surprising speech is given the context of preceding words. Lexical predictive processing is associated with a decreased response in older adults, as predicted by the N400 literature.

Second, working memory could allow listeners to maintain information to process *after* encountering degraded input, i.e., through "post-diction"[75]. We will be able to test whether the effect of working memory capacity on comprehension is mediated through stronger responses associated with one of these neural mechanisms. In Aim 3 we will also be able to test whether working memory capacity is associated with recruitment of the domain-general multiple demand network vs the language network associated with predictive processing[123–125].

Is incremental speech processing impaired due to inability to manage lexical competition? Words differ in their neighborhood density (how many similar sounding words there are)[126]. When multiple words partially fit the acoustic input, there is a greater demand to select among competing representations[127,128]. Older adults are often impaired in tasks requiring selection among competing representation[129], including dense lexical neighborhoods[126,23,130,131]. We will test whether incremental processing is specifically impacted after high competition words. ***Is this related to inhibitory deficits?*** Inhibitory functioning has been proposed as an individual difference variable in aging[132], with better inhibitory functioning linked to better speech recognition in noise[80,133–135], and better recognition of words with larger lexical neighborhoods[23,126,136]. However, the relevance of general inhibitory deficits has been questioned[137–140], and most previous results are based on small, underpowered datasets[34]. Our sample of 200 older adults will overcome this concern and allow an adequate test of the hypothesis.

Aim 3: Brain networks for prediction and comprehension in older adults

Research using fMRI has identified a stable set of brain regions that are associated with language processing in young adults, labeled the "language network"[141–144]. This core language network is distinguished from a so-called multiple-demand network, which is engaged for difficult tasks across different domains[144–147]. While fMRI does not have the phoneme-level resolution of EEG, it can measure word-level incremental processing[123–125,148–150], which is localized primarily to the core language network in young adults[123–125], as well as phrasal (non-incremental) processing[94]. Neither of these effects has been investigated in older adults.

Compared to young adults, older adults exhibit a general shift in brain networks that are recruited during speech perception, with older adults showing less activity in auditory cortex[151–154], and/or increased activity in other frontal and parietal regions[154–156], which may involve the multiple demand network[154,155]. When considered with similar patterns in non-linguistic tasks[157,158], such brain activity may reflect compensatory processes[158] that could support preserved speech comprehension. However, it is unknown how this shift relates to specific cognitive mechanisms such as predictive processing – indeed, such age-related changes may be related to experimental tasks rather than speech comprehension per se[159], and may be confounded with neural activity related to increased listening effort due to acoustic signal degradation[160]. These considerations highlight the need for a more mechanistic understanding of age-related changes, and in a more valid task, such as continuous speech.

In Aim 3 we address this need by testing age differences in brain networks measured using fMRI during naturalistic speech comprehension, and relating them to cognitive mechanisms. We will determine whether age-related changes in brain networks can be linked to comprehension through incremental processing, or through another mediating mechanism such as working memory capacity. For instance, one possibility that would explain intact use of context despite deficient incremental processing is that older adults successfully use context, but are relying on a different mechanism, which may be implemented outside the core language network (e.g., the multiple demand network), and involve more domain-general resources such as attention and working memory.

In Aim 3, we will recruit two groups of participants according to behavioral evidence of *high* and *low* ability to benefit from linguistic context, contrasting high and low context-users to identify the neural basis of resilient comprehension[161]. *<u>fMRI paradigm</u>*: As in EEG, participants will listen to audiobook segments in quiet and in noise (*using high-end noise canceling headphones* to attenuate interference from scanner noise). We take a multi-pronged analysis approach: 1) we use a GLM block-design analysis: by contrasting speech in quiet with an acoustic baseline (speech distorted to be indecipherable and sound like poor radio reception of speech[162]) we will localize the language network (this contrast has been validated as functional localizer[162]); by contrasting speech in noise and in quiet we will identify networks that are modulated by the presence of noise. These contrasts will identify group-constrained, subject-specific functional regions of interest (fROIs)[141] for 2) event-related analysis that parallels the EEG analysis in Aims 1&2, modeling each word as an event modulated by regressors scaled by lexical-level measures (e.g. surprisal, lexical frequency, see Approach). Previous research[148,123–125], as well as our preliminary data (Figure 5) show that this approach can identify fMRI signatures of predictive language processing. 3) We ask whether behavioral outcomes (specifically, the use of context to overcome perceptual challenges) can be predicted from the fMRI measures (activation maps, and sensitivity to different context processing measures in fROIs)[163]. We predict that an individual's ability to successfully use context (i.e. to out-perform one's hearing status) will be related to increased sensitivity to incremental processing within the language network, and/or compensation from language-external regions. Together, this rich dataset will allow us to address the following questions:

Is impaired incremental processing in older adults corroborated with fMRI? Data from young adults indicates that incremental language processing is measurable with fMRI from activity throughout the language network[123,124,148,156]. Accordingly, corroborating an age-related deficit in incremental processing in fMRI will rule out the possibility that the original finding is related to an idiosyncrasy of the EEG methodology.

Do older adults compensate with language-internal or -external resources? An important question in speech comprehension in aging is the degree to which compensation for sensory declines arises from efficient use of language-internal resources, or rather due to redirection of domain-general resources to language comprehension. Evidence for language-internal compensation would be comprehension-related activity in the language network (incremental or non-incremental). Comprehension-related activity outside of the language network would be taken as evidence for language-external compensation. One possibility is that older adults recruit the domain-general multiple demand network. In young adults, the multiple demand network may be activated with controlled linguistic complexity manipulations, but it is not commonly activated in naturalistic comprehension[144,164] (although it may be recruited when speech input is acoustically challenging[127,128]). This would be consistent with the mediating effect of working memory capacity on preserved comprehension[75], since a high reading span score is associated with activity in the multiple demand network[165]. It would also explain how context is used without incremental processing, because the multiple demand network is not modulated by predictability, at least in young adults[123,124], suggesting that the increased memory for sentence information allows for more effective non-incremental context use. A second possibility is that preserved comprehension depends more on activity in the cingulo-opercular network, which has been shown to predict word recognition in noise and is thought to be related to adaptive control and the direction of attention[166].

Is activity outside the core language network in older adults during natural speech processing compensatory? The relationship between brain activity and performance in older adults is complicated. Additional brain activity could reflect compensation[156], but it could also reflect inefficient use of neural resources[167]. Our dataset will allow us to tease apart brain activity that is associated with improved performance (compensation) and brain activity that reflects inefficient strategies. Specifically, compensatory brain activity should be positively correlated with comprehension (controlling for other factors such as hearing ability, and quality of peripheral auditory representations determined from EEG), whereas maladaptive activity should be negatively correlated.

How is structural brain health related to speech comprehension and predictive processing? Aging, even healthy aging, is associated with large-scale differences in gray and white matter structure[168]. Such declines have been linked to poorer speech in noise comprehension[169] and functional localization shifts during speech processing[156]. Another age-related indicator of brain health, small white matter lesions appearing as white matter hyperintensities on structural MRI scans, often discovered incidentally, has a prevalence up to 94% at age 82 in the general population[170]. In an exploratory analysis, we ask how such individual differences in brain structure (grey and white matter densities, cortical thickness, gyrification, ventricular volume) relate to comprehension and the robustness of the prediction network. Both volumetric and cortical thickness indicators will allow assessing whether impaired predictive processing is due to structural challenges in the language network, or whether it is unrelated to those indicators of brain health, possibly being a more general age-related cognitive change.

Impact

The ability to communicate successfully in aging has wide ranging consequences, from enabling social connections, warding off loneliness and depression, to being able to seek and receive appropriate health care[171]. Preserved communication ability might thus be a protective factor against several risk factor for dementia, such as depression and poor social interactions[172,173]. Because comprehension is essential to successful communication, it is important to understand the mechanisms involved in successful speech comprehension in older adults. Current hearing aids only partially address consequences of hearing loss[174] – their failure to improve speech in noise perception might be partly responsible for low uptake[175] – and cognition is thought to play a major role in older adults' comprehenion[176,177]. This proposal targets a big gap in our understanding of these higher-level neuro-cognitive contributions to comprehension. Once neuro-cognitive factors associated with preserved comprehension are identified, it may become possible to influence them. This may take the form of cognitive training, or more advanced approaches such as neural stimulation[178–180], aimed at shifting speech related activity to a compensatory brain network. The research proposed here will establish potential targets and outcome measures for such interventions.

The measures themselves, as developed and validated here, will provide useful biomarkers. Conventional measures of speech recognition rely on behavioral responses to isolated syllables words or sentences. Such artificial tasks can lead to results that don't generalize well to more naturalistic situations[181]. There is a lack of measures that assess comprehension in engaging situations. EEG is an ideal measure because it can be taken without disrupting engagement in listening to a story, and the measures developed here allow assessing comprehension simultaneously from auditory to sentence-level linguistic processing. Furthermore, EEG is a widely available and cost-effective method.

This work will determine whether reduced incremental processing is an impairment, or whether it is an innocuous consequence of aging. If it is indeed an impairment, and possibly a biomarker for more subtle comprehension difficulties, it should be taken seriously and investigated further (for example, abnormal predictive processing may be an early warning sign of dementia[182]). The proposed work lays a foundation for this. On the other hand, the focus on incremental processing in common neural markers of speech comprehension may stem from an over-reliance on young adult subject populations, and the non-incremental measures developed here may reveal novel markers of successful speech comprehension that are more appropriate for older adults.

Finally, the data collected for this project will be shared in a public repository, responding to the need for large-scale neural data on speech comprehension in aging. Shared data will include raw EEG and deidentified MRI data, together with the careful characterization of individual differences in hearing, cognition, and language function. For such a repository, continuous narrative speech is a particularly versatile stimulus with potential for testing many different hypotheses through appropriate language models. We expect that this will enable more insights into the relationship between functional and structural age-related changes in speech processing.

General implications for speech processing

In addition to the questions related to speech processing in older adults, the results from this project will have implications for our understanding of human speech processing in general.

The role of incremental processing and prediction in human language processing: Widely cited theories of brain function argue that these are fundamental to information processing in perception and cognition[3,89,88]. Predictive processing has been instrumental in developing computational language models[183,184], although its importance in human language processing has also been questioned[90]. If older adults can achieve superior language comprehension (relative to their sensory acuity) with decreased incremental or predictive processing, then that would provide some evidence against the central mechanistic importance (predictive processing might still be important for language acquisition, but might not be essential for perception per se).

The role of inhibition in lexical processing is subject to debate. TRACE, a classical model of human speech recognition, relies on competition between representations at each level through lateral inhibition[185], and this competition may be linked to the *entropy* (uncertainty) variable used here[186]. However, theoretical considerations suggest competition may not be necessary for speech recognition[82]. These views thus make opposing predictions as to whether inhibitory functioning should be related to speech recognition ability or not. Our test of the relationship between inhibitory functioning and speech processing will advance this debate.

INNOVATION

The innovation of this project arises from the methodology (specifically the use of a TRF approach in continuous speech, and convergent data from EEG and fMRI) and the theoretical innovation of the question itself.

Methodological Innovation: In contrast to most prior research on this topic, we will study speech processing using continuous, meaningful speech, ensuring that measurements are generalizable to situations in which older adults face challenges in their everyday lives. Recent research on speech perception in older adults has highlighted the importance of this, by showing that the factors affecting speech comprehension in older adults in more artificial experimental paradigms do not generalize well to the more realistic situation of engaging, continuous narrative speech[181].

While the TRF approach has become established to measure acoustic and word-level features, using it to measure incremental speech processing at the phoneme level is a recent innovation, spear-headed by work by PI Brodbeck[29–31,187]. The proposed work will be the first application of this to individual differences and lifetime changes in predictive processing. Furthermore, this work will develop novel measures for non-incremental speech processing in continuous speech.

Though not a novel idea, something that sets this proposal apart from other research is the detailed characterization of a large sample for a high-powered dataset. The inclusion of different modalities (behavioral, EEG, fMRI, structural MRI) will enable more convergent evidence. For example, by acquiring EEG and fMRI data in the same participants, with similar tasks, we will be able to verify that measured effects are not idiosyncrasies of imaging modality (while EEG and fMRI operate at very different time-scales, our preliminary data demonstrate that our innovative language models, which connect the time-scales of word and phoneme recognition, allow reliably assessing incremental speech processing in both modalities).

Theoretical Innovation. While it has long been recognized that the use of linguistic context aids comprehension, in both younger and older adults, the neural mechanisms supporting individual differences in the use of context in aging are not well understood. The outcome of this project will be a rich neurobiological dataset which affords an answer to key questions about effective language processing strategies across the lifespan.

APPROACH

Preliminary data

Aims 1&2 (EEG): To demonstrate the feasibility of the proposed measures, we re-analyzed a dataset from a recently published N400 EEG study[77] with the mTRF method. In the original study, 15 younger and 15 older adults (age, mean ± standard deviation: younger 21.2 ± 2.6; older: 64.4 ± 4.9) listened to 100 sentences each, spoken with an American English accent. Older adults showed a robust context benefit, recognizing more sentence-final words after sentences with high compared to low constraint. For each subject, a separate mTRF model was estimated to predict EEG responses to the entire sentence (not only the final word) with 5-fold cross-validation, using the complete set of predictors described below (*Methods*, *EEG experiment (Aim 1)*). The total audio duration per participant was only 3 min 42 sec – drastically less than the proposed experiments.

The analysis demonstrates the feasibility of revealing age differences across the speech processing hierarchy (Figure 2). To assess sentence-level predictive processing, we trained an n-gram model on a subtitle corpus (sources dated 1900-2007)[188]. Surprisal, word-uncertainty and phoneme-uncertainty each independently improved cross-validated EEG predictions, demonstrating the feasibility of distinguishing their contributions. The response functions show a pattern similar to the N400, with a peak at centro-posterior sites at a delay of 300 ms (Figure 3 left; the peak is slightly earlier because it is modeled as delay relative to phonemes, instead of word-onset). The peak is larger for young adults than for older adults, replicating results from the N400 literature, although the difference is less pronounced for word entropy than surprisal, suggesting that these may dissociate (Figure 3). On the other hand, the preliminary data also demonstrate the potential for finding compensatory activity in older adults, with stronger responses to sublexical phoneme predictability (Figure 2-D).

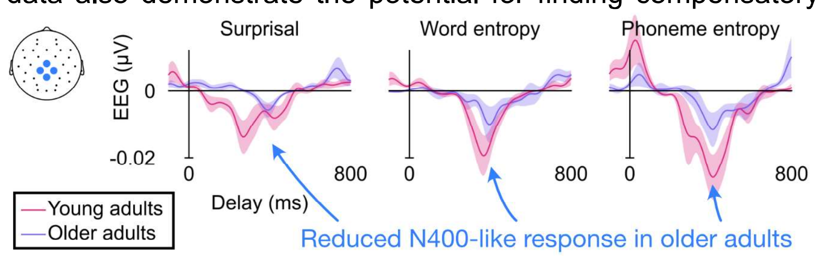

Figure 3. **Preliminary data: older adults show significantly attenuated responses in indicators of incremental and predictive processing**. Y-axis shows µV *per unit surprisal/entropy* (uncertainty).

Preliminary data also show potential for measuring non-incremental speech processing (Figure 4). A model of phrasal processing that assumes that listeners accumulate information (and corresponding maintenance demands) until they reach phrase boundaries[95] significantly improved EEG predictions (when added to all other predictors). Corresponding brain responses exhibit a late component that is specifically enhanced in older adults, suggesting that this may be a biomarker for enhanced context processing that does not involve incremental processing.

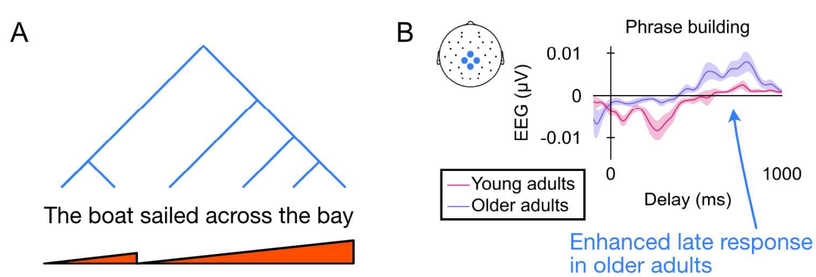

Figure 4. **Non-incremental processing may be enhanced in older adults**. (A) Model of brain responses increasing with maintenance demands as syntactic phrases grow. (B) This predictor is associated with an early peak in younger adults, but exhibits an increased extended response in older adults, which may reflect enhanced non-incremental language processing in older adults.

Aim 3 (fMRI): We demonstrate the feasibility of measuring incremental word-by-word perception of natural speech with a dataset collected by our group (79 subjects who listened to a 10 minute long podcast; same scanner and acquisition parameters as described below). Using GPT-2[183], a state-of-the-art computational language model, we determined context-based surprisal for each word. We then created two regressors of interest: *speech rate*, i.e., each word corresponds to an impulse of size 1[148], and surprisal, where each word corresponds to an impulse whose size is its surprisal. Predictors were generated at 200 Hz, convolved with the canonical hemodynamic response function, and then downsampled to the rate of fMRI (1 Hz). This revealed a reliable distinction between 1) a core-auditory network, centered on auditory cortex, with activity primarily linked to speech rate regardless of predictability, and 2) a network of predictive language processing, activating temporal and frontal regions classically associated with the language network. <u>This demonstrates the feasibility of identifying brain regions related to incremental speech processing in continuous, natural speech in event-related fMRI analysis.</u>

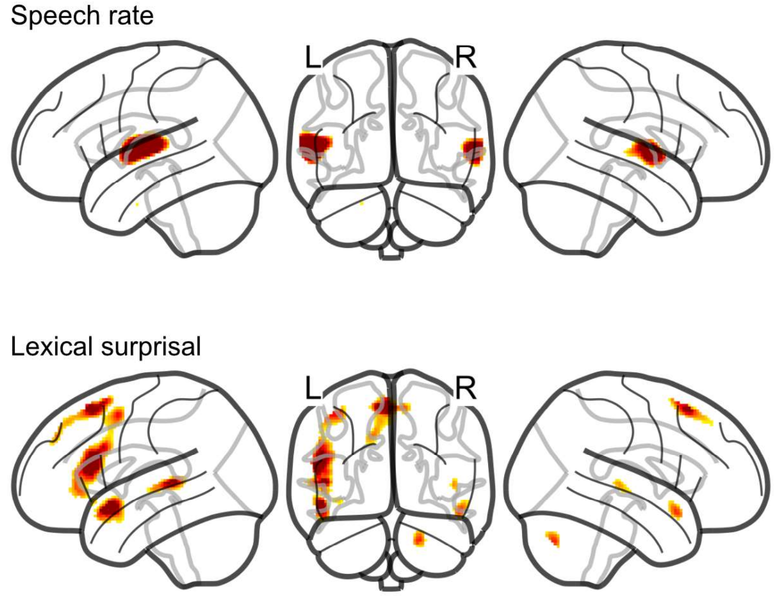

Figure 5. **Preliminary fMRI data demonstrate feasibility of localizing incremental processing in continuous speech**. Activation centered on auditory cortex tracks speech rate regardless of predictability; a fronto-temporal network of regions with strong overlap with the classical language network tracks incremental processing, formalized as surprisal ($p \leq .05$ family-wise error).

Methods

Overall recruitment strategy is shown in Figure 6. All participants will undergo an initial behavioral assessment. Those that pass screening for hearing loss and cognitive impairment will be asked to participate in the EEG study. Recruitment will continue until we reach the targeted sample size. We will then selectively recruit for fMRI.

<u>*Young adults*</u>: We will initially recruit a sample of 50 young adults (ages 18-40) from the UConn community to serve as control group (native speakers of American English with no known neurological or language disorders).

<u>*Older adults*</u>: We will aim to recruit a community sample of 200 older adults, aged 60+ for EEG. Exclusion criteria based on cognitive and hearing deficits are intended to eliminate participants whose deficits might prevent them from demonstrating reliable behavioral data, but allow for sufficient variability to investigate the effects of individual differences in hearing on predictive processing. <u>*Hearing loss*</u>: For participants with mild hearing loss, stimuli will be amplified in a frequency-specific manner based on their audiogram using the National Acoustics Laboratory-Revised Profound algorithm (NAL-RP)[189,190]. Participants with hearing loss that would require more amplification than physically possible (approximately 40-50 dB at any frequency) will be excluded from the EEG study. Participants with asymmetric hearing loss (≥10 dB difference at two adjacent frequencies) will also be excluded. <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, in consultation with co-I Thomas Hinchey, a Clinical Professor of Audiology. Hearing aid characteristics will be determined by measuring the output of the participants' hearing aids in response to a sample stimulus (same speaker as in the EEG experiment) using a probe microphone procedure. If the required level of amplification below 4 kHz would be technically impossible, they will be excluded.

The rationale for stimulus amplification in participants with hearing loss is to control for differences in acoustic perception as much as possible. This is because we are interested in the ability to use context and not in the recognition of acoustic cues *per se*: If hearing-impaired listeners fail to recognize acoustic cues, then they cannot build the necessary context representations. On the other hand, prolonged experience with hearing loss may lead to a change in context processing, which will be assessed (at least in an exploratory way) by comparison with hearing-aid aid users, whose daily life experience has been with amplified sounds instead.

<u>*Background noise*</u> will be 20-talker babble noise, selected for the following properties: 1) it is relatively uniform acoustically; 2) it does not contain individually audible words, which would be distracting and induce more variable neural responses; 3) yet, when it is used as speech masker, comprehension has good correspondence to subjective reports of daily life speech in noise hearing difficulties[191]. Noise type and signal to noise

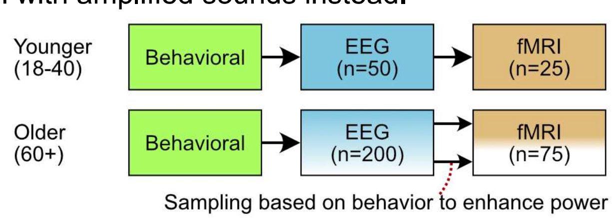

Figure 6. **Recruitment strategy**.

ratio (SNR) will be kept constant across experiments to enable more direct comparisons. *SNR conditions*: The purpose of the background noise is to introduce an acoustic challenge to encourage context use, but it should not impair comprehension to a degree where context use breaks down because too little of the context is understood. For setting subject-specific SNRs we will use the speech in noise tests conducted as part of the initial evaluation (which use the same masker noise and allow estimating % correct recognition as a function of SNR). We will include two SNR conditions: (1) SNR equal to 90% sentence in noise recognition, to control comprehension across subjects (i.e., an EEG effect in better context users cannot be attributed to comprehending more of the stimulus). However, in this condition, better context users will generally be assigned a lower SNR and thus receive more distorted acoustic-phonetic information. To control for this, we will also use (2) an SNR corresponding to 90% syllable in noise recognition, to control the availability of acoustic-phonetic information across subjects. Together, the two conditions will allow us to tease apart effects associated with the ability to use context from effects of comprehension and acoustic-phonetic perception. While there is a need to understand more realistic acoustic backgrounds, and more severe challenge from lower SNRs, the influence from those on neural responses to speech is too poorly understood at present to assess interactions with predictive processing – but we hope to address them in future research.

Sound delivery: In EEG, Etymotic ER-30 insert earphones deliver sound through plastic tubes, providing magnetic insulation required for reliable recording of subcortical responses (these earphones are used successfully in Dr. Anderson's lab). A downside is relative attenuation of frequencies above 4000 Hz. To derive consistent measures across different paradigms (behavioral, EEG, fMRI) we will filter stimuli in behavioral and fMRI paradigms to approximate the earphones' transfer function.

Behavioral Measures

In a first visit, the following audiological and behavioral measures will be collected for all participants.

- **Audiological assessment**:
  - Pure tone thresholds at 125, 250, 500 Hz and 1, 2, 3, 4, 6, 8, 9, 10, 11.2, 12.5, 14, 16, 18, 20 kHz.
  - Distortion Product Otoacoustic Emissions (DPOAE) testing to assess outer hair cell function.
  - Hearing case history to determine history of noise exposure, hearing aid status, tinnitus, etc.
  - Perceived daily-life communication difficulty (self-assessment questionnaire, SSQ)[192].
  - Click auditory brainstem response, as a broad measure of synchrony in the auditory system, to detect synaptopathy and rule out retrocochlear pathologies.
- **Cognitive screening**: MOCA[193] to screen for cognitive impairment (cutoff score <26).
- **Language screening**: A battery of tasks that has been found to identify history of developmental language disorder in adults with good sensitivity and specificity[194].
- **Speech recognition in noise**: Clinically-validated tests of speech recognition in noise will be used to measure the effect of background noise on different components of speech recognition:
  - **Acoustic speech recognition**: The California Syllables Test (CaST)[195] measures recognition of meaningless CVC syllables in the presence of background noise (here, 20-talker babble). It provides a measure of acoustic processing in the absence of a linguistic context.
  - **Speech recognition with context**: The influence of linguistic context will be measured with sentence recognition. Sentences contain linguistic context which, on average, improves recognition over acoustics alone[22]. Comparison with the CaST will be used to estimate individual-specific linguistic context benefit.
  - **Speech recognition with context and informational masking**: Sentences from the same pool, but presented in the 4-talker babble noise (QuickSIN[196]). This will be used to estimate the addition cost of recognizable voices in the background, which constitutes an additional challenge[197].
  - **Context-benefit**: to directly assess the benefit listeners derive from linguistic context, we also contrast sentence-final word recognition in sentences with high vs. low contextual constraint[77] in 20-talker babble.
- **Working memory**: We administer the reading span test[198,199], the letter-monitoring task[200] (both frequently correlated with speech in noise comprehension[119,201]) and the NIH Toolbox List Sorting Test[202]. Including multiple tests allows extracting a more reliable, common working memory factor.
- **Inhibitory functioning**:
  - **Speech-specific**: Two linguistic tasks which have previously been found to be correlated with word recognition in high-density lexical neighborhoods[23].
  - **Non-speech**: The Stroop color naming assesses inhibitory functioning in a different modality (reading/production) – it may also predict the influence of background noise on speech perception[133].
- **Processing speed** will be measured with the NIH Toolbox Pattern Comparison Processing Speed Test[203].

- **Handedness** will be recorded with the Edinburgh Inventory[204].

EEG experiment (Aim 1)

Sound will be presented through Etymotic ER30 insert earphones (same waveform in both ears). *Stimuli* will be taken from an audiobook, selected for low probability of being familiar to either group. Samples will be played in chronological order to maximize engagement, and interrupted approximately every 60 seconds for comprehension questions (pre-tested as valid measure of comprehension). Segments will alternate between speech in quiet and in noise. After the experiment, subjects will also be asked for a subjective judgement of difficulty and listening effort[160] and familiarity with the book, author and genre (familiarity did not affect results in a previous study[33]).

EEG analysis will use MNE-Python[205] and the Eelbrain toolkit[206] developed by PI Brodbeck for the multivariate temporal response function (mTRF) approach (see Figure 7 and Figure 1). The mTRF approach allows modeling EEG data as a linear function of an arbitrary number of time-continuous predictor variables. A model fit results in two main outcomes (for each subject): *predictive power* (proportion of the variability in the EEG that the model can explain), and the TRFs themselves (describing the EEG voltage as a function of delay relative to the stimulus). The former is used to ascertain, using cross-validation, that each predictor actually explains something about the EEG signal that cannot be explained from other predictors; the latter is used to evaluate the temporal characteristics of the brain response, similar to event-related potentials.

Predictor variables (for more detailed descriptions see Brodbeck's prior work[31], which has shown significant predictive power for each individual variable): an *acoustic model* (gammatone spectrogram and acoustic onset spectrogram[106]) is included to measure and control for acoustic processing[207]. *Linguistic predictors* are created through time-series with impulses at the onset of each phoneme, scaled to reflect a specific property of that phoneme. The *sublexical model* estimates tracking of sublexical language statistics. It is implemented with a 5-gram model, trained on phoneme sequences in a natural speech corpus, unaware of word boundaries. For each phoneme, the sublexical model includes two quantities: how much new *information* this phoneme provides (formally: *surprisal*[208]) and the degree of *uncertainty* about what phoneme will come next (*entropy*[208]). The *word-form model* estimates neural processes related to word recognition that do not use the preceding context. It is based on an implementation of the cohort model of word-recognition[209]: At word onset, the "cohort" is initialized as a probability distribution over the whole lexicon, with the probability of each word corresponding to its lexical frequency. For each subsequent phoneme, the probability distribution is updated by retaining only words that are consistent with that phoneme. For each phoneme, this model defines three quantities: the amount of information contributed by the phoneme for determining the current word (*surprisal*), and the *entropies* regarding the identity of the next phoneme and the current word. *Context-dependent lexical models* estimate neural processes related to word recognition that make use of the preceding sentence context. They are implemented like the word-

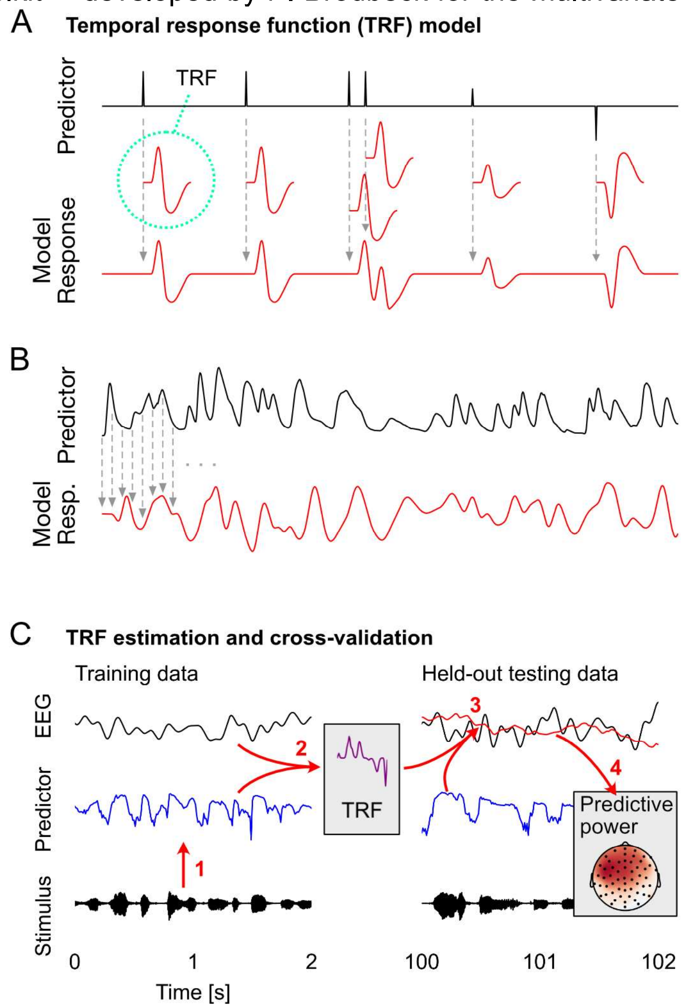

Figure 7. **The multivariate temporal response function (mTRF) approach for analyzing EEG responses to continuous speech**. (A) The temporal response function model assumes that the response is a continuous linear function of one (or several) predictor variables. Convolving a temporal response function (TRF) with impulses allows modeling overlapping responses to categorical events such as phonemes or word. (B) Continuously varying predictor for modeling features such as acoustic envelope. (C) Analysis proceeds as follows: 1) generate time-continuous predictor variables based on the stimulus; 2) Estimate a TRF to predict the aligned EEG response (one channel is shown); 3) Use that TRF to predict held-out data; 4) Evaluate the predictive power (i.e., proportion of the held-out data that can be explained by the TRF) of different models to determine which representations (predictors) are affecting neural responses.

form model, except that the probability distribution over the lexicon is initialized with context-dependent probability, estimated from a language model (see next paragraph). These measures have been validated with MEG[29,31,187], in a published EEG study[30], the preliminary data included here, and two unpublished EEG datasets. In each case we found significant brain activity related to these measures of incremental speech processing at all levels.

Finally, models of continuous speech processing are continually improved, and such refinements will be taken into account. We will specifically pursue directions described in the following three paragraphs, as they may be particularly relevant for older adults. While these are exploratory, they do not add a multiple comparison problem: The selection of the most appropriate model will be based on which set of predictors is best at predicting EEG data across participants, which is orthogonal to the main goal of this proposal, the analysis of individual differences.

Influence of language experience We will determine whether the optimal language model depends on listeners' age. For this we will train different language models using large corpora with timestamps (ranging 1930-2019)[210–212]. For example, a language model trained on texts from 1960-1990 may predict 60+ year old adults' brain responses better than a language model trained on texts from 2000 onwards, or one trained on web-based texts[183]. We currently plan to use n-gram models because we have successfully used them for M/EEG responses[31,33], but will also consider more advanced models such as LSTMs[213,214], and comparison with state-of-the art language models trained on web texts like GPT-2[183].

Non-incremental measures: parsing Following a few published reports[94–96], we will test non-incremental models based on linguistic tree structure obtained using the Stanford parser[215]. These models assume that, rather than integrating each word immediately, words are maintained and chunked into phrases. In preliminary data we tested a model based on increasing maintenance demands with each successive word in top-level phrases[95], finding evidence for enhanced responses in older adults compared to younger adults (Figure 4).

Non-incremental measures: ambiguity processing According to one model[75], when the acoustic signal is distorted to a degree such that bottom-up word recognition fails, this triggers an ambiguity resolution processes that recruits additional cognitive resources to draw on context. This is called "post-diction" because it assumes use of context only after problematic input is encountered, as oppposed prediction which uses context proactively. According to this model, we should see increased neural activity in response to acoustic distortion and ambiguity. We will test this with two approaches: *Speech intelligibility*: For speech in noise, we will estimate time-dependent speech intelligibility using the envelope regression method[216] (with other methods as fallback, e.g.[217,218]). Intelligibility can be calculated as value between 0 (speech is unintelligible due to masking) and 1 (speech is unmasked), using moving windows as narrow as 300 ms[216], and can take into account frequency-specific hearing loss. We will test for increased neural responses as a function of low intelligibility. We hypothesize that individuals who are able to recruit more brain activity in response to ambiguities are more successful at resolving these ambiguities. We will also explore combining the intelligibility estimates with phoneme transcripts to derive probabilistic phoneme sequences (i.e., each phoneme position is characterized by a probability distribution over phonemes indtead of a single phoneme). This allows estimating ambiguity in the linguistic interpretation based on the Bayesian Shortlist-B model[37]. *Vowel ambiguity*: previous work of Co-I Myers has identified a measure of acoustic ambiguity for vowels in continuous speech and shown that this measure significant modulates fMRI responses in speech in quiet[128]; we will generate a corresponding predictor for both EEG and fMRI, assuming increased brain activity in response to more ambiguous vowels as above.

Subcortical responses measurable with scalp EEG synchronize to fast oscillations in the speech waveform[219,220], and are measured with a TRF analysis by first high-pass filtering the signals to remove low frequency noise while retaining the speaker's fundamental frequency (e.g., 80 Hz), and then estimating a TRF for predicting the EEG signal from the speech waveform[100–103].

Power: 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[221]. Because the size of meaningful effect differs substantially across fields, we rely on an estimate based on effects commonly reported in gerontology[34]. 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 the initial comparison of older and young adults we estimate 50 subjects per group. The effect-size of the age difference in the sentence-based TRF peak amplitudes is large ($d \geq 0.99$ in preliminary data, >95% power with group size of 28) but we expect more heterogeneity in our sample with less stringent criteria for e.g. hearing loss.

fMRI experiment (Aim 2)

Once 50 older adults have completed behavioral assessment, we will estimate population distributions for preserved comprehension relative to hearing ability (see Behavioral Measures above). We will then start recruiting participants from the endpoints of our 200-strong EEG sample to maximize power in the fMRI design[161], with a goal of at least 35 older adults in each group (>80% power to detect an effect with $r = 0.2$[161]; younger adults are expected to perform generally well, and we will attempt to match them to resilient comprehenders).

Data acquisition will be at the BIRC facility using a 3T Prisma Scanner with a 64-channel head coil. We will collect two high-resolution structural scans (MPRAGE, FLAIR) and high-resolution functional scans (2x2x2.5mm) with an iPAT acceleration factor that allows for full-brain coverage in 1 sec per scan. Stimuli are delivered with active noise-cancelling headphones to minimize interference from scanner noise (OptoAcoustics OptoActive).

Stimuli will be audiobook segments, interrupted infrequently for comprehension questions as in EEG. Audiobook segment length will be shortened to ~40 seconds, and alternating speech in quiet and noise, to optimize for block design analysis[222] (approximately 25 segments per condition for >16 minutes of speech per condition). For the *language localizer*, acoustically degraded passages will be interspersed with the audiobook stimuli (~8 blocks of 20 sec each)[162].

Blocked analysis: Brain activity related to speech processing in quiet and in noise will be identified using a GLM approach with box-car regressors for the different conditions, convolved with the canonical hemodynamic response function (HRF). Contrasts will be used to define group-based, subject-specific functional regions of interest (fROIs)[141] in relevant brain regions (speech in quiet > acoustic baseline; speech in noise > speech in quiet). *Contingency plan*: We will explore (and account for) individual and age-related differences in the HRF[163] by modeling several orthogonal basis functions. If subject-specific fROIs cannot be established, we will resort to whole-brain analysis and anatomical fROIs based on group analysis and published research.

Event-related analysis: Regressors will be generated from (oversampled) time-series with impulses at word-onsets, convolved with the HRF. Impulses will be scaled uniformly to model speech rate, and with language model based regressors to model predictive processing (e.g., lexical surprisal and entropy). Besides a whole brain analysis, we will assess evidence for predictive processing in the fROIs.

Predicting behavior: We will train support vector machines to predict which individuals exhibit resilient comprehension, from functional (predictive and non-predictive processing) and structural (e.g., grey matter thickness) maps, and test models using *k*-fold cross-validation. We also train models to predict resilient comprehension from measures in fROIs to determine anatomical regions whose activation is relevant. *Contingency plans*: we will compare fROI results with whole-brain voxel-wise models; if individual maps are too heterogeneous, we will explore multimodal surface matching to map subjects into a more comparable anatomical space[223]. If training of predictive models fails entirely, we will resort to group comparison based on high vs. low resilient comprehension.

Structural analysis: We collect both MPRAGE and FLAIR sequences for automatic segmentation, estimation of grey matter volume and detection of white matter hyperintensities in FreeSurfer[224] (quality will be assessed with visual inspection and comparison with other algorithms[225,226], which also provide fallback options).

Timeline

The planned timeline of activities is shown in Figure 8. Based on this plan, we expect that this project will allow us to write the following manuscripts: *Year 2*: Once 50 older adult EEG datasets are available, a manuscript comparing mechanisms of speech processing in older vs. younger adults. *Year 3-4*: Once EEG data collection wraps up, individual differences in older adults based on EEG. Year 4-5: age effects in fMRI; correspondence between fMRI and EEG; predicting resilient comprehension from MRI (functional and structural).

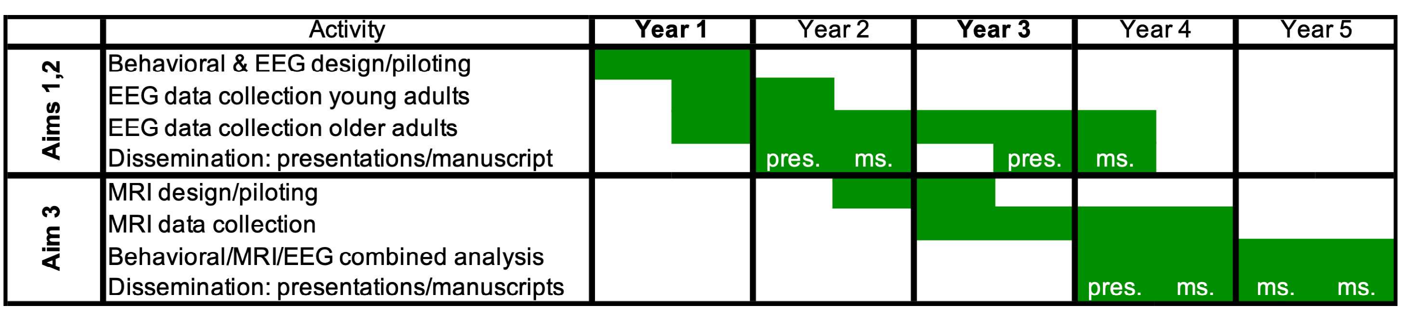

Figure 8. **Timeline for activities**. Abbreviations: pres. conference presentation; ms. manuscript preparation.