Article:

Corps, R.E. orcid.org/0000-0001-6027-8109, Shehzad, M., McLaren, A.I. et al. (2 more authors) (2025) Effects of listening demand on turn-taking processes: Investigating turn-end prediction and verbal response planning. Acta Psychologica, 260. 105694. ISSN: 0001-6918

https://doi.org/10.1016/j.actpsy.2025.105694

Acta Psychologica 260 (2025) 105694

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Effects of listening demand on turn-taking processes: Investigating turn-end prediction and verbal response planning

Ruth E. Corps [a], Muzna Shehzad [b], Andrew I. McLaren [b], Leigh B. Fernandez [c], Lauren V. Hadley [b,*]

[a] School of Psychology, University of Sheffield, 219 Portobello, Broomhall, Sheffield, S1 4DP, UK

[b] Hearing Sciences - Scottish Section, School of Medicine, University of Nottingham, Level 3 New Lister Building, Glasgow Royal Infirmary, 16 Alexandra Parade, Glasgow, G31 2ER, UK

[c] Department of Social Sciences, Psycholinguistics Group, RPTU Kaiserslautern-Landau, Postfach 3049, 67653, Kaiserslautern, Germany

A R T I C L E I N F O

Keywords:
Turn-taking
Hearing loss
Listening demand
Prediction
Conversation

A B S T R A C T

During conversation, speakers often take turns with little gap or overlap. But older adults with hearing loss are often slower to take turns and show more variability in turn timing than older adults with normal hearing, potentially due to increased cognitive demands of processing speech with hearing loss. In this study, we investigated how hearing loss and listening demand affect turn-end prediction and verbal response planning – two processes thought to support timely turn-taking. We developed a single task that captured these two processes, instructing participants to listen to semantically predictable and unpredictable turns and press a button when they thought the turn would end before producing a verbal response. Experiment 1 validated this task, demonstrating that younger adults with normal hearing responded faster when turns were predictable rather than unpredictable. Experiment 2 investigated how listening demand and hearing loss affected these processes by comparing older adults with normal hearing (PwNH) listening at low demand (a highly intelligible level of 70 dB) to older adults with hearing loss (PwHL) listening at either low demand (i.e., 70 dB), or high demand (the lowest level above 60 dB at which the participant could report >70 % of words). Participants again responded faster for predictable than unpredictable turns. Additionally, PwHL in the low demand group showed larger predictability effects than PwNH. In contrast, PwHL in the high demand group showed no difference to PwNH in turn-end prediction but a delay in verbal response production. Together, these findings suggest that content predictability facilitates both turn-end prediction and response planning in such a 'listen-to-respond' task, and that adults with hearing loss show particular reliance on prediction when listening demand is low, which may buffer against turn-taking delays.

1. Introduction

During conversation, interlocutors are so finely coordinated that there is often little gap (typically around 200 ms) between their utterances (Stivers et al., 2009), even though language production is comparatively slow, taking around 600 ms to produce a single word (Indefrey & Levelt, 2004). Current theories agree that interlocutors achieve such coordination by predicting the semantic content of what the speaker is likely to say next (e.g., Garrod & Pickering, 2015; Levinson & Torreira, 2015).

We adopt the same approach as other studies in the literature (e.g., Corps et al., 2018; Corps et al., 2019; Magyari et al., 2014) and operationalise content predictability as semantic predictability – specifically, how strongly the preceding context constrains the semantic content of the final word of an utterance. High predictability, often quantified through measures like cloze probability (Taylor, 1953), means that the upcoming word is highly predictable in semantic content. There is much evidence that these semantic content predictions facilitate turn-taking by enabling the comprehender to determine (1) what they want to say (response planning), and (2) when to say it (turn-end prediction).

But most work has studied these processes separately, even though they typically occur together in conversation. Furthermore, this work

* Corresponding author.
E-mail addresses: r.corps@sheffield.ac.uk (R.E. Corps), muzna.shehzad@nottingham.ac.uk (M. Shehzad), leigh.fernandez@rptu.de (L.B. Fernandez), lauren.hadley1@nottingham.ac.uk (L.V. Hadley).

https://doi.org/10.1016/j.actpsy.2025.105694
Received 24 March 2025; Received in revised form 23 September 2025; Accepted 30 September 2025
Available online 14 October 2025

has tended to focus on young adults. However, research suggests that older adults with hearing loss (PwHL) are slower to take turns (Petersen, MacDonald, & Sørensen, 2022) and show more variability in their turn timing (e.g., Petersen, Walravens, & Pedersen, 2022) than older adults without hearing loss (PwNH), potentially because listening is more demanding. Greater listening demand, or lower speech intelligibility, could lead to greater listening effort (e.g., Peelle, 2018) and thus fewer cognitive resources available for prediction than PwNH (e.g., Fernandez et al., 2024). In this study, we developed a task involving both response planning and turn-end prediction to investigate the effect of hearing loss and listening demand on listeners' ability to use content predictions for response planning and turn-end prediction. In the sections that follow, we first review evidence that listeners use content predictions to plan a response, as well as to predict the speaker's turn-end. Next, we discuss how these processes may be affected by hearing loss, before describing our experiment in more detail.

1.1. Response planning

Much research has investigated the timeline of response planning. These studies have primarily focused on how quickly participants respond, with faster responses thought to reflect earlier response planning. For example, Bögels et al. (2015) measured EEG correlates during a question-answering task, in which the critical information necessary for response planning (in the following example, the critical information is *007*) was available either early (e.g., *Which character, also called 007, appears in the famous movies*?) or late (e.g., *Which character from the famous movies is also called 007?*). Participants responded faster when the critical information was available early ($M = 640$ ms) rather than late ($M = 950$ ms) and showed earlier activation in brain areas associated with speech production (e.g., Indefrey & Levelt, 2004) and motor response preparation (e.g., Babiloni et al., 1999) when the critical information was early compared to when it was late.

These results suggest that listeners plan a response early when given the opportunity to do so. Similarly, in Experiment 1, Meyer et al. (2018) had participants answer *yes/no* questions about the objects within an array displayed on-screen (e.g., *Do you have a green sweater?* while seeing an array including a sweater, a cake, a branch, and a barrel). When all objects were displayed in the same colour, and thus responses could be planned as soon as the colour was heard, participants responded earlier ($M = 215$ ms) than when objects were displayed in different colours ($M = 297$ ms), further indicating a propensity to plan early when possible.

But although these studies suggest that participants plan a response as soon as they can determine the likely answer to a question, they do not test whether the predictability of an utterance facilitates such planning. Corps et al. (2018); see also Corps et al., 2020) addressed this issue using a question-answering task in which participants responded *yes* or *no* to questions in which the linguistic context was manipulated to make the final word either predictable (e.g., *Are dogs your favourite animal*?) or unpredictable (e.g., *Would you like to go to the supermarket*?). Participants answered more quickly when the final word was predictable ($M = 379$ ms; Experiment 2b) than unpredictable ($M = 536$ ms), suggesting that they predicted the content of the speaker's question and used these predictions to plan an appropriate response.

Thus, there is evidence that content predictions facilitate response planning in PwNH. In the next section, we review evidence for turn-end prediction.

1.2. Turn-end prediction

Other studies have investigated how listeners determine the speaker's turn-end so they can respond at the appropriate moment (i.e., without overlapping with the other speaker or leaving a long gap). These studies have typically used experimental paradigms like the button-press task, in which participants listen to turns and press a button when they think the speaker is about to reach the end of their utterance. Turn-end prediction has been operationalised using two measures (e.g., Corps et al., 2018, 2019). Button-press speed (or response speed) quantifies when a participant initiates their response relative to the turn-end, reflecting the speed at which participants form an expectation about the turn-end. Button-press precision (or response precision), in contrast, reflects the accuracy with which they pinpoint the turn-end, indicating how close their responses are to the actual end.

Research using this paradigm has investigated the role of various linguistic cues in turn-end prediction. For example, De Ruiter et al. (2006) assessed turn-end prediction using a button-press paradigm, in which participants listened to turns taken from natural conversation and pressed a button when they thought the speaker was about to reach the end of their utterance. The authors either removed the lexico-syntactic information from the utterances through low-pass filtering (leaving the prosody unaltered) or removed the prosody by setting the pitch to a constant level (leaving the lexico-syntactic information unaltered). When prosody was flattened, participants responded on average 200 ms before the end of the speaker's turn, which was similar to their responses to unmodified turns and comparable to the timing of verbal responses in the original conversations. However, when lexical information was removed, participants responded too early and further away from the actual turn-end - on average 500 ms before the end of the turn. These results suggest that listeners predominantly rely on information gained from the speaker's words (i.e., semantic/syntactic information) to predict when they will finish speaking, though it is likely that some sources of prosodic information also play a role (e.g., final syllable duration; see Bögels & Torreira, 2015).

Extending from the above finding that turn-end prediction relies on semantic/syntactic processing, several studies have tested whether the content predictability within an utterance affects turn-end prediction. In one study, Magyari et al. (2014) used the same button press task to show that participants respond earlier for utterances ending in a predictable manner (e.g., *I live in a house with four women and another man; M* $= 70$ ms before the turn end) than an unpredictable manner (e.g., *She was again alone in the north; M* $= 139$ ms after the turn end). Furthermore, this work has been extended to larger timeframes (Bögels & Torreira, 2021; Corps et al., 2019), with Corps et al. (2019) demonstrating that participants respond more quickly, but not more precisely (i.e., closer to the actual turn-end) when a prior sentence provides greater semantic context for an upcoming sentence that when it does not. Specifically, people respond more quickly to *I really like Taylor Swift* after hearing *I listen to a lot of music*, than *There's one thing you need to know about me*. Hence both sentence and discourse predictability contribute to turn-end prediction.

However, not all evidence suggests that listeners use the content of the speaker's utterance to predict the turn-end. In Experiments 1a and 2a, Corps et al. (2018) had participants press a button at the moment they thought a series of predictable and unpredictable questions were about to end, as in the studies above. Unlike Magyari et al. (2014), Corps et al. (2018) found no evidence that button-press times were affected by content predictability: participants responded similarly to predictable ($M = 154$ ms before the turn end in Experiment 2a) and unpredictable utterances ($M = 99$ ms before the turn end). The discrepancy in these results may be explained by additional variables that affect turn-end prediction. Research has shown that button-press times are influenced by turn duration, such that participants respond earlier when utterances are longer in duration (e.g., De Ruiter et al., 2006). While Corps et al. (2018) accounted for effects of duration by including it as a covariate in their analyses, Magyari et al. matched the average durations of their predictable and unpredictable conditions without accounting for individual stimulus differences. Although the conditions did not significantly differ, predictable utterances were nonetheless 410 ms longer than the unpredictable utterances on average in Magyari et al.'s study, which may explain why they found earlier responses in the predictable condition.

In another study, Barthel et al. (2016); see also Barthel et al., 2017) used a task in which participants completed a confederate's pre-recorded utterances. Since participants had to name any on-screen objects that the confederate had not already named, participants could plan their response as soon as the confederate began uttering their last object name (indicated by the use of the word *and*; e.g., *I have a door and a bicycle*). Importantly, the authors also manipulated the predictability of the confederate's turn-end so that participants could or could not predict that a sentence final verb would follow the last object name. Both eye-movements and response times suggested that participants planned their response as soon as possible. However, neither of these measures were influenced by the predictability of the speaker's turn-end.

In sum, there is clear evidence that listeners use content predictions to plan a response, and some evidence that they use these predictions to determine the current speaker's turn-end. But these studies have tended to segregate response planning and turn-end prediction, investigating them separately. To fully understand the mechanisms supporting conversation, it is essential to investigate how content predictability affects turn-end prediction when listeners then act on this prediction by producing a verbal response. In our experiment, we therefore tested whether content prediction facilitates turn-end prediction by using a task that additionally required participants to act on their turn-end prediction by producing a verbal response. Participants listened to predictable and unpredictable questions, as in Corps et al. (2018), and both (1) used a button-press response to indicate when they thought the turn would end; as well as (2) verbally responded to the turn. Although this task is highly constrained, it is ecologically valid in that it captures processes that occur together in conversation, enabling us to investigate how content prediction affects turn-end prediction when these predictions are task relevant.

So far, we have focused on response planning and turn-end prediction in PwNH. In the next section, we discuss how these processes may be affected by hearing loss.

1.3. Conversation in people with hearing loss

Research suggests that PwHL find conversation more challenging than PwNH in terms of communication effort (e.g., Beechey et al., 2020). Spoken utterances are less well-timed, with the gaps between turns being both delayed and more variable in PwHL than in PwNH (e.g., Petersen, Walravens, & Pedersen, 2022). However, little research has directly investigated the reason that turns are less well timed in PwHL.

Hearing loss typically results in reduced audibility (e.g., Moore, 1996) and one explanation, grounded in models of listening demand and cognitive resource allocation (e.g., Peelle, 2018; Pichora-Fuller et al., 2016; Rönnberg, 2003), is that this reduced audibility increases the cognitive demands of speech processing. According to these models, cognitive resources are reallocated to perceptual processing under challenging conditions, leaving fewer resources for other higher-level processes, such as semantic prediction (see e.g., Fernandez et al., 2024, for discussion). Consistent with this hypothesis, there is evidence that challenging listening conditions delay speech processing, both in PwNH (e.g., Wagner et al., 2016) and PwHL (e.g., Wendt et al., 2015), and this delay has been attributed to increased cognitive load due to difficulty hearing (e.g., Carroll et al., 2016).

Some research suggests that PwHL show greater reliance on predictive contextual cues compared to PwNH (e.g., Benichov et al., 2012; Gordon-Salant & Fitzgibbons, 1997; Mattys et al., 2012; Pichora-Fuller et al., 1995), which frees up cognitive resources (e.g., Hunter & Humes, 2022). However, these studies are based on measures taken after a target word, and so have not directly investigated whether hearing loss impacts the time-course of semantic prediction. Fernandez et al. (2025) addressed this issue using a task in which participants listened to predictable sentences (e.g., *The waiter brings the plate*) while viewing four objects on-screen, one of which was a potential target (e.g., plate, scarf, window, carpark). Older adults could predict the sentence's final word before it was mentioned and fixated the target object more than the other three objects before it was spoken. However, these semantic predictions were delayed with hearing loss and further exacerbated when listening demand was high.

Thus, there is evidence that PwHL experience delays in prediction, which could be due to listening demands. Such delays may in turn affect the accuracy of turn-end prediction and the timing of response planning, leading to more variable turn transitions. For example, if PwHL predict more slowly or to a lesser extent than PwNH, then they may also find it harder to predict when the current speaker is likely to finish their turn, leading to responses that are produced too early or too late. Similarly, if PwHL are slower to predict what a speaker is likely to say, then they may also be slower to begin planning their own response. Previous research has not directly investigated how hearing loss affects these two processes. We addressed this issue in our experiment by investigating whether content prediction facilitates turn-end prediction and response planning within a single task, in both PwNH and PwHL.

1.4. The current study

In sum, current theories agree that interlocutors achieve coordination in conversation by predicting the content of the speaker's ongoing turn (e.g., Garrod & Pickering, 2015; Levinson & Torreira, 2015). There is clear evidence that listeners use these predictions to plan a response, and some evidence that they use these predictions to determine the current speaker's turn-end. But most of these studies have focused on the processes of response planning and turn-end prediction separately, even though they typically occur together during conversation. This paper presents a new paradigm that combines turn-end prediction and verbal response production, which we first validated on younger normal hearing listeners (Experiment 1). We then assessed how these turn-taking mechanisms differ in older adults with and without hearing loss, at different levels of listening demand, given that research suggests that hearing loss alters turn taking timing but the basis of this change is not clear (Experiment 2).

Following Fernandez et al. (2025), in Experiment 2 we included one group of older adults with normal hearing, and two groups of older adults with hearing loss. For the PwNH and one PwHL group, auditory information was presented at a highly intelligible level of 70 dB LAeq, making two low listening demand groups. For the other PwHL group, audibility was individually set to the lowest conversationally-appropriate level (i.e., >60 dB LAeq) at which the participant could repeat >70 % of words, making a high listening demand group (see Experiment 2 Methods). By including PwHL groups differing in listening demand (confirmed through differences in both stimulus presentation level and an intelligibility pre-test), we could address both the direct effect of hearing loss and the indirect effect of increased listening demand on turn-end prediction and response planning.

2. Experiment 1

2.1. Method

2.1.1. Participants

A total of 30 young adults participated in this study: 17 females and 13 males. The participants' age ranged from 18 to 25 years (age_{mean} = 20.37 age_{SD} = 2.03). All participants had normal hearing, with their better ear four frequency pure tone average (FFPTA) being <20 dB LAeq (see Fig. 1), and achieved 100 % in an intelligibility pretest (described below).

Ethical approval for the research project was obtained from the University of Nottingham Faculty of Medicine and Health Science Research Ethics Committee, with the reference number FMHS 423–1221. All participants provided written informed consent and were compensated £20 for their participation.

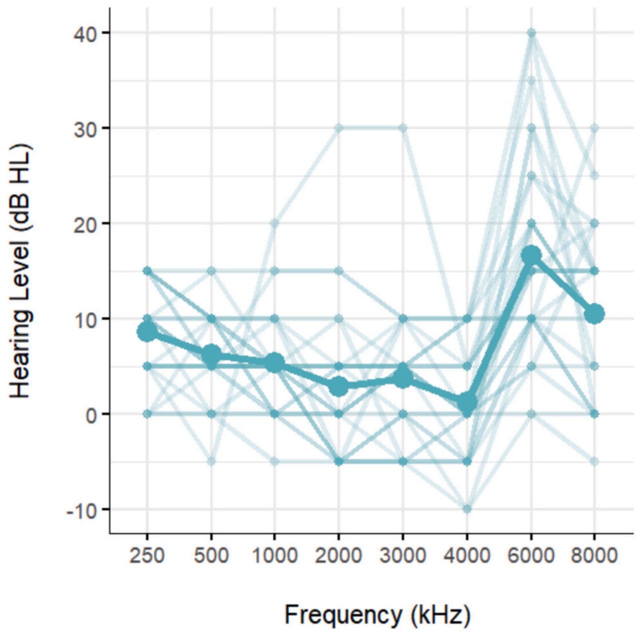

Fig. 1. Hearing thresholds averaged across the left and right ears for younger adults without hearing loss. Individual lines show individual participants, while the bold line indicates the mean hearing threshold across all participants.

2.2. Materials

We selected 52 questions (26 for each condition; see Appendix A for a full list) using a pre-test, in which 33 participants from the same population ($age_{mean} = 20.67$; 24 females, 9 males) were presented with 72 question fragments and were instructed to complete the question with a single word (i.e., we used a cloze task; Taylor, 1953).

We assessed content predictability, operationalised as semantic predictability, using two different measures that assessed the extent to which the preceding context constrained the sentence's final word. First, we computed Shannon entropy (i.e., $-\Sigma p_i \log_2(p_i)$, where p_i is the proportion of times each completion occurs for a given fragment; Shannon, 1948). Entropy is low (a minimum of 0) when completions are similar across participants and high (a maximum of 5) when responses are different. We also calculated cloze probability (Taylor, 1953), which is the percentage of participants who provided a particular completion. For the questions in the predictable condition, we selected the final word that had the highest cloze probability and was thus provided by the majority of participants (e.g., *Have you ever dyed your **hair**?*); for the unpredictable condition, we selected completions that had the lowest cloze probability (e.g., *Have you ever broken your **tooth**?,* which could also end with another body part). Stimuli in the predictable condition had higher cloze probability ($p < .001$) and lower entropy ($p < .001$) than those in the unpredictable condition (see Table 1). The conditions were matched for average word length ($p = .18$).

All questions were recorded by a native English male speaker with a Glaswegian accent. Recordings were between 692 and 3309 ms in duration (see Table 1). Duration was determined by measuring the difference between the acoustic onset of the speaker's first phoneme and the acoustic offset of the final phoneme within each question. Questions in the predictable condition were longer than those in the unpredictable condition ($p < .001$). We return to this duration difference in the Results section.

Table 1
The means (*M*) and standard deviations (*SD*) of question fragment entropy, cloze probability (%), word length, and duration (ms) for the predictable and unpredictable conditions.

Predictability		Question fragment entropy	Completion content cloze (%)	Word length	Duration (ms)
Predictable	*M*	0.25	96 %	8.65	2187
	SD	0.27	0.05 %	2.73	542
Unpredictable	*M*	2.99	4 %	7.73	1663
	SD	0.51	0.02 %	2.07	458

2.2.1. Procedure

First, participants underwent air-conduction audiometry to ensure normal hearing levels, followed by an intelligibility check. In the intelligibility check, participants were asked to repeat back a set of five pre-recorded *yes/no* questions (a mixture of predictable and unpredictable) spoken by the talker of the later main experiment. Stimuli were presented at 70 dB LAeq, the presentation level for the experiment, and intelligibility was scored by determining the number of keywords (all words beyond articles and short prepositions) that participants correctly recalled. An average intelligibility score was calculated across the five sentences, which was 100 % for all participants. Participants then took part in an unrelated visual world study, reported separately (Fernandez et al., 2025), before moving to the main task.

The main task consisted of 52 questions, presented in two blocks (stimulus order randomized), preceded by 15 practice questions to familiarize participants with the task. Participants were told: "*In this part of the experiment, you will be listening to a series of yes or no questions. Your task is to predict when the sentences will end. Press the button when you believe the question will end. Do not wait until the speaker has stopped speaking. Instead, you should press the button as soon as you expect the speaker to finish. Immediately after pressing, you should answer the question with yes or no. See this as like being on a quiz show where you press the buzzer and answer. There will be a practice section and then a main section.*" The script was run on the Matlab version R2023a (The MathWorks Inc., 2023). Participants listened to the stimuli at 70 dB LAeq using the AKG Reference Headphones K702 with the sound being transmitted using the RME Babyface Pro sound card system. Participants responded using The Black Box Toolkit USB response pad and verbal responses were spoken into a dynamic microphone with a cardioid pick-up pattern. Button-press onsets were automatically detected using the experiment script. Speech onsets (i.e., the beginning of the speaker's first phoneme) were detected using an automated algorithm dependent on a manually-selected speech amplitude threshold (which varied according to the participant's vocal level). To avoid giving participants feedback about their performance, stimuli playback was automatically stopped after a button press.

Immediately following the main experiment, participants were asked to rate their self-perceived effort using the NASA TLX questionnaire (Hart & Staveland, 1988), and the average effort score was 29.80 ($SD = 23.53$), with higher scores (a maximum of 100) indicating higher effort. At the end of the session, participants were debriefed, given the opportunity for questions and feedback, and compensated for their participation.

2.3. Data analysis

Before extracting speech onsets, all verbal response recordings were checked by plotting the response waveform in Matlab to ensure a verbal response was captured (rather than coughing or laughter, for example). Recordings that had picked up white noise were cleaned using Audacity's noise reduction feature, which reduced the white noise without altering the onset of the verbal response. Recordings were then manually checked for 20 randomly selected participants to confirm that the verbal responses were either *yes* or *no*. Two participants were discarded from the verbal response analysis because their verbal responses were not recorded due to technical issues, leaving 28 participants. All participants and trials were included in the button-press analysis.

We analysed the button-press and verbal response tasks separately because we were interested in the distinct cognitive processes of turn-end prediction and verbal response planning. The time between the button-press (the point when participants think they can answer) and the verbal response (initiation of that answer) could also reflect the time needed for planning a response. However, it is unlikely to capture the full planning process – some planning likely occurs before the predicted turn-end, particularly for predictable sentences. As a result, we analyse the two tasks separately but report an analysis of the time between the button-press and the verbal response in Appendix B.

Participants were encouraged to press the button as soon as they expected the speaker to reach the end of their turn and then immediately answer the question verbally. The task therefore encouraged both speed and accuracy, and so we analysed how quickly (response times) and precisely (response precision) participants responded in relation to the stimulus end. For the button-pressing task, we take response times to indicate how early listeners generate predictions about the turn-end, and response precision to indicate how accurately their predictions pinpoint the turn-end. Studies using verbal response tasks have primarily focused on response times, which indicate how early listeners plan a verbal response. But we also analysed response precision to indicate how close participants' verbal responses were to the turn-end.

For the analysis of response speed, response times were defined with respect to sentence offset (i.e., the end of the speaker's final phoneme). They were thus negative when participants responded before the end of the speaker's sentence and positive when they responded after the end. We evaluated the effects of content predictability on response times with linear mixed effects models (LMM; Baayen et al., 2008) using the *lmer* function of the *lme4* package (version 1.1–35.3; Bates et al., 2023) in RStudio (version 2023.06.1). Full model outputs are reported in Appendix C.

For the analysis of response precision, the absolute value of the response time was calculated. In this analysis, we consider how close participants respond to the sentence offset, with responses closer to zero being considered better regardless of whether they occurred before or after the actual turn-end. We initially analysed precision using LMMs, but the residuals violated assumptions of normality, likely because the distribution of response precision is truncated at zero. As a result, we instead analysed precision using generalised linear mixed effects models (GLMM) with a Gamma family and a log link function, which is appropriate for right-skewed data.

In both analyses, response times or precision were predicted by Predictability (reference level: unpredictable vs. predictable), which was contrast coded (−0.5, 0.5). Given that previous button-pressing studies have demonstrated that longer utterances tend to elicit earlier responses (e.g., De Ruiter et al., 2006) and predictable utterances were longer than unpredictable ones, we also included centered Utterance Duration and its interaction with Predictability as a fixed effect. We used the maximal random effects structure justified by our design (Barr et al., 2013), including by-participant random effects for Predictability. For all analyses, we report coefficient estimates (b), standard errors (SE), and t values for each predictor. We assume that an absolute t value of 1.96 or greater indicates significance at the 0.05 alpha level (Baayen et al., 2008). Full model outputs are reported in Appendix C. Raw data and analysis scripts are available at: https://osf.io/852tx/.

2.4. Results

2.4.1. Button press

Younger adults pressed the button on average 116 ms after the sentence end (see Fig. 2 for a breakdown by condition). They responded earlier when sentences were predictable ($M_{\text{response time}} = -14$ ms) rather than unpredictable ($M_{\text{response time}} = 246$ ms; $b = -176.72$, $SE = 46.51$, $t = -3.80$), but they were no more precise when the sentences were predictable ($M_{\text{response precision}} = 281$ ms) rather than unpredictable ($M_{\text{response precision}} = 294$ ms; $b = -0.13$, $SE = 0.15$, $t = -0.86$). Thus, younger adults responded more quickly but not more precisely when sentences were predictable. As in previous studies (e.g., De Ruiter et al., 2006), participants responded earlier ($b = -88.72$, $SE = 21.79$, $t = -4.07$) when utterances were longer in duration, although they did not respond any less precisely ($b = 0.08$, $SE = 0.07$, $t = 1.17$). There was an interaction between Utterance Duration and Predictability when analysing response precision ($b = 0.35$, $SE = 0.13$, $t = 2.67$) but not when analysing response times ($b = -3.08$, $SE = 43.57$, $t = -0.07$).

2.4.2. Verbal response

Younger adults verbally responded 207 ms after the sentence end on average (see Fig. 3 for a breakdown by condition). They responded earlier ($b = -197.33$, $SE = 53.16$, $t = -3.71$) and more precisely ($b = -0.24$, $SE = 0.12$, $t = -1.97$) when sentences were predictable ($M_{\text{response time}} = 67$ ms; $M_{\text{response precision}} = 304$ ms) rather than unpredictable ($M_{\text{response time}} = 347$ ms; $M_{\text{response precision}} = 382$ ms). Participants again responded earlier ($b = -88.44$, $SE = 24.40$, $t = -3.63$) but not more precisely ($b = 0.02$, $SE = 0.06$, $t = 0.27$) when utterances were longer in duration. There was again an interaction between Utterance Duration and Predictability when analysing response precision ($b = 0.29$, 0.12, $t = 2.36$), but not when analysing speed ($b = -0.12$, $SE = 48.79$, $t =$

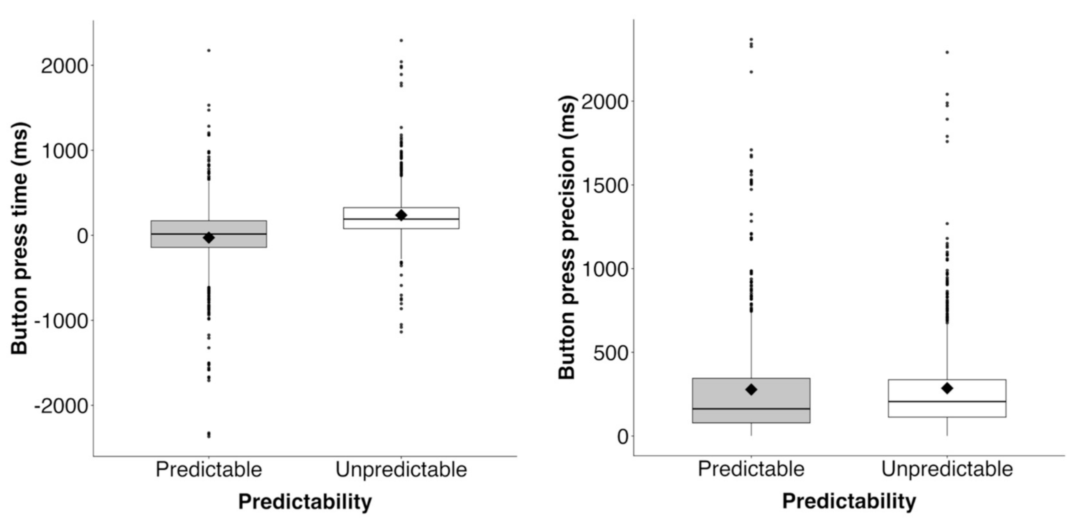

Fig. 2. Distribution of response times (left) and precision (right) for younger adults in the button-pressing task. The box plots show the median, interquartile range, and 1.5× interquartile range (whiskers), with diamond points indicating group means. Button press times were significantly earlier (but not more precise) for predictable than unpredictable sentences.

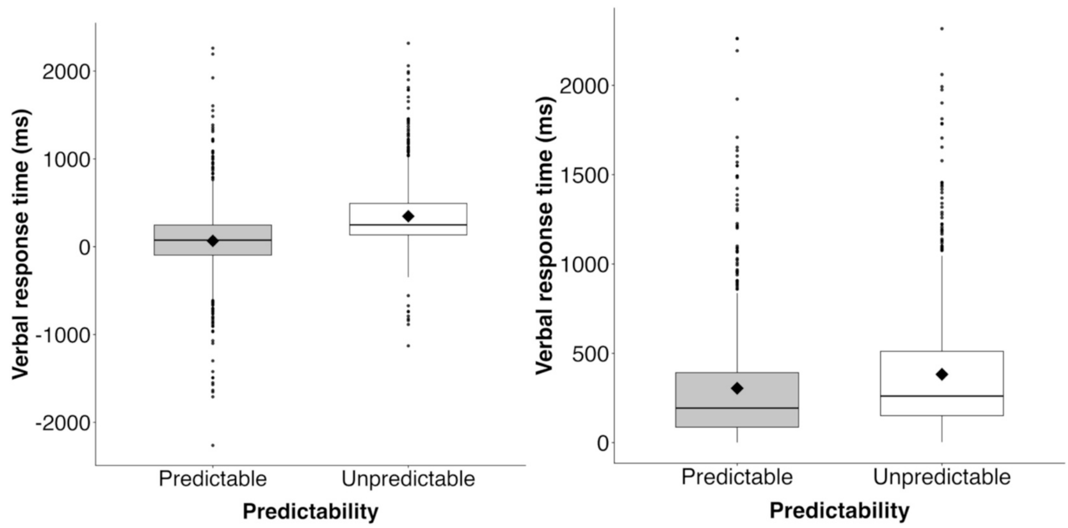

Fig. 3. Distribution of response times (left) and precision (right) for younger adults in the verbal response task. The box plots show the median, interquartile range, and 1.5× interquartile range (whiskers), with diamond points indicating group means. Verbal response times were significantly earlier and more precise for predictable than unpredictable sentences.

−0.03).

2.5. Discussion

In Experiment 1, we tested a paradigm on younger normal hearing listeners that combined turn-end prediction and verbal response production. Participants responded earlier, but not more precisely, in the button-press task when sentences were predictable rather than unpredictable. In the verbal response task, participants responded both earlier and more precisely for predictable than unpredictable sentences.

Our findings are consistent with other results reported in the literature. For example, Magyari et al. (2014) found that participants responded earlier in the button-press task when turns were predictable rather than unpredictable. Corps et al. (2019) similarly found that participants responded earlier, but not more precisely, for utterances that were predictable ($M_{\text{response time}} = 49$ ms; $M_{\text{response precision}} = 276$ ms) rather than unpredictable ($M_{\text{response time}} = 90$ ms; $M_{\text{response precision}} = 282$ ms) given the discourse context. Furthermore, numerous verbal response studies have found that participants respond earlier when utterances are predictable rather than unpredictable (e.g., Corps et al., 2018). Thus, our results are consistent with previous studies in the literature. In Experiment 2, we exploit this paradigm to assess how turn-end prediction and response planning mechanisms differ in older adults with and without hearing loss.

3. Experiment 2

In Experiment 2, we investigated how turn-end prediction and response planning are affected by listening demand and hearing loss in older adults. Given that we are investigating turn-taking processes inherent to conversation, we adopted an ecological approach, aiming to understand turn-end prediction and response planning in the sorts of listening conditions that occur in everyday conversation. As a result, we recruited one group of older adults with normal hearing (PwNH) and two groups of older adults with hearing loss (PwHL). For the PwNH and one PwHL group, auditory information was presented at a highly intelligible level of 70 dB LAeq ($\text{PwHL}_{\text{low-demand}}$). For the other PwHL group, audibility was individually set to the lowest conversationally-appropriate level at which the participant could repeat >70 % of words ($\text{PwHL}_{\text{high-demand}}$). Since conversational speech is typically around 60 dB, we chose this as our starting point and raised the presentation level to 65 dB if intelligibility was too low. This manipulation led to two hearing loss groups that were separable in terms of presentation level (see Participants section) but not intelligibility (i.e., some participants in the $\text{PwHL}_{\text{high-demand}}$ group may had similar intelligibility levels to those in the $\text{PwHL}_{\text{low-demand}}$ group). While groups were not entirely distinct, this approach allowed us to compare the two PwHL groups to the PwNH group, revealing how turn-taking processes are affected under two ecologically valid listening scenarios.

3.1. Method

3.1.1. Participants

We recruited 90 participants (37 males, 53 females), ranging from 50 to 80 years old from the Glasgow area. To ensure distinct hearing groups, we recruited 30 normal hearing participants (without a clinical diagnosis of hearing loss) who had a worse ear FFPTA of <30 dB HL, and asymmetry <20 dB HL, and hearing-impaired participants with a better ear FFPTA of 30–65 dB HL (see Fig. 4) and asymmetry <25 dB HL. Thus, the worse ear of the normal hearing participants had better hearing than the better ear of the hearing-impaired participants.

The 60 hearing-impaired participants were randomly divided into subgroups comprising 30 $\text{PwHL}_{\text{low-demand}}$ and 30 $\text{PwHL}_{\text{high-demand}}$. For the $\text{PwHL}_{\text{low-demand}}$ group, sentences were presented at a highly intelligible level of 70 dB LAeq (as in Experiment 1). Two participants were excluded because they did not achieve 100 % intelligibility even though the sentences were designed to be highly intelligible. For the $\text{PwHL}_{\text{high-demand}}$ group, audibility was individually set to the lowest conversationally-appropriate level (60 dB) at which the participant could repeat 70 % of the words. Six participants were excluded because they listened at 70 dB or higher, a presentation level that overlapped with the $\text{PwHL}_{\text{low-demand}}$ group. As a result, the final participant set included 30 participants in the PwNHgroup, 28 in the $\text{PwHL}_{\text{low-demand}}$ group, and 24 in the $\text{PwHL}_{\text{high-demand}}$ group. The demographic information for these groups is shown in Table 2.

Although the age of the PwNH group was significantly lower than the two PwHL groups ($ps < 0.001$), the difference was only six years (66 to 72). There was no significant difference in age between the two PwHL groups ($p > .5$). As expected by the demand manipulation, the $\text{PwHL}_{\text{high-demand}}$ group experienced speech at a lower presentation level and thus showed lower speech intelligibility in the pretest than both the PwNH and $\text{PwHL}_{\text{low-demand}}$ groups, ($ps < 0.01$), who did not differ from each other.

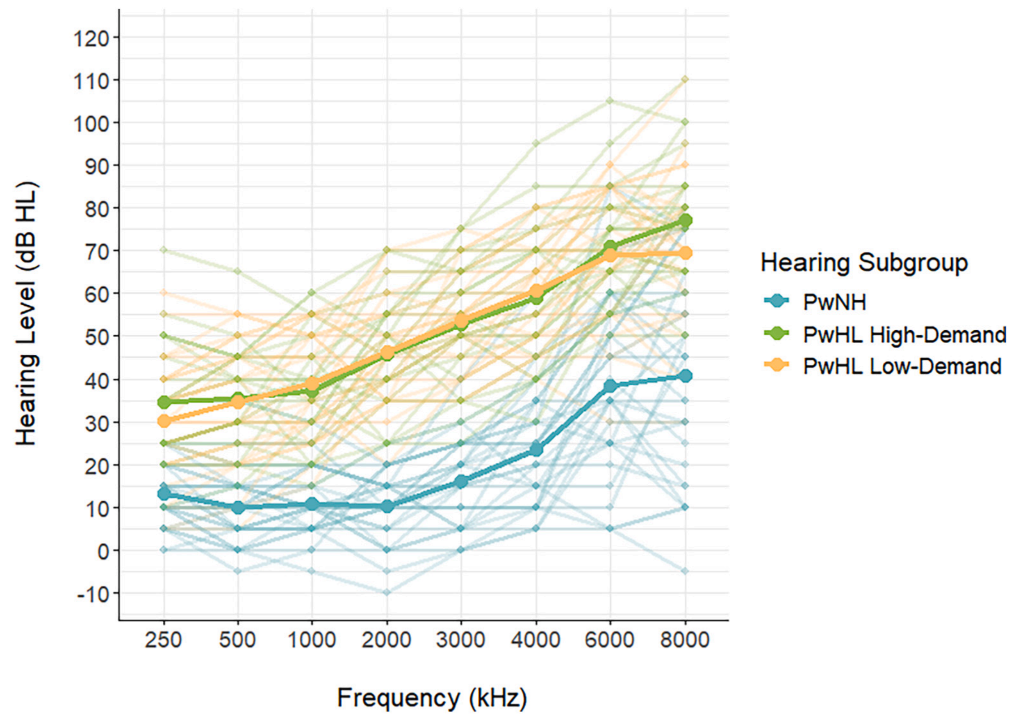

Fig. 4. Hearing thresholds averaged across the left and right ears for the PwNH, $PwHL_{low\text{-}demand}$, and $PwHL_{high\text{-}demand}$ groups. Individual lines show individual participants, while the bold lines indicate group averages.

Table 2
Demographic information for participants in Experiment 2.

	PwNH	$PwHL_{low\text{-}demand}$	$PwHL_{high\text{-}demand}$
$Age_{mean(sd)}$	$66.00_{(6.63)}$	$71.71_{(5.30)}$	$71.08_{(5.70)}$
Better ear $FFPTA_{mean(sd)}$	$13.13_{(6.74)}$	$44.46_{(8.47)}$	$43.59_{(8.67)}$
Worse ear $FFPTA_{mean(sd)}$	$16.29_{(5.59)}$	$48.66_{(8.92)}$	$48.02_{(9.62)}$
$Asymmetry_{mean(sd)}$	$3.17_{(2.82)}$	$4.38_{(3.51)}$	$4.95_{(4.42)}$
Sex	13_{males}, $17_{females}$	13_{males}, $15_{females}$	8_{males}, $16_{females}$
Intelligibility $scores_{mean (sd)}$	$100.0_{(0.00)}$	$100.00_{(0.00)}$	$96.10_{(8.23)}$
Presentation $level_{mean(sd)}$	$70_{(0.00)}$	$70_{(0.00)}$	$61.67_{(2.41)}$
Effort $scores_{mean(sd)}$	$13.60_{(20.65)}$	$20.43_{(23.70)}$	$23.46_{(32.16)}$

3.1.2. Materials

The materials were identical to Experiment 1.

3.1.3. Procedure

The procedure was identical to Experiment 1, with the exception that the participant's level of intelligibility was measured at the presentation level to be used in the main experiment, which differed according to group. For the $PwHL_{low\text{-}demand}$ group, intelligibility was measured using a set of five sentences presented at 70 dB LAeq (as in Experiment 1). All participants achieved 100 % intelligibility. For the $PwHL_{high\text{-}demand}$ group, a set of five stimuli was initially presented at 60 dB LAeq, and keyword accuracy was assessed for each set of five sentences (there were three sets in total), with the presentation level increasing progressively in 5 dB increments (with different sets of five sentences) until over 70 % of the keywords were correctly recalled. Thus, stimulus presentation for the $PwHL_{high\text{-}demand}$ group was individually set to the level where they could recall at least 70 % of the keywords. Fig. 5A shows the presentation level for each participant in the $PwHL_{high\text{-}demand}$ group, while Fig. 5B shows the intelligibility scores. Note that the majority of participants in the $PwHL_{high\text{-}demand}$ group achieved 100 % intelligibility, just like the $PwHL_{low\text{-}demand}$ group. However, we still consider these participants to belong to the high-demand group because research suggests that intelligibility does not directly equate to listening effort (e.g., Winn & Teece, 2021) and these participants listened to speech at the lowest level above 60 dB at which they could achieve at least 70 % intelligibility. They listened at a maximum level of 65 dB, lower than that presented to the low-demand group.

3.2. Data analysis

Before analysis, we removed three button-press times because they were longer than 10,000 ms and clear outliers. The verbal response times for these trials were included in the verbal response analysis, but we also excluded three verbal response times because they were also longer than 1000 ms. The data were analysed using the same procedure as Experiment 1, with Hearing subgroup as an additional predictor. Hearing subgroup had three levels (PwNH, $PwHL_{low\text{-}demand}$, and $PwHL_{high\text{-}demand}$) and we fitted one model including two contrasts: One comparing the PwNH group to the $PwHL_{high\text{-}demand}$ group (0.5, −0.5; Hearing Contrast 1) and the other comparing the PwNH group to the $PwHL_{low\text{-}demand}$ group (0.5, −0.5; Hearing Contrast 2). The full model for the older adults thus included interactions between Predictability, Hearing Contrast 1, Predictability and Hearing Contrast 2, and Predictability and Utterance Duration. Using the maximal structure resulted in a singular fit error, likely because Hearing Contrast 2 explained little variance by-item. As a result, the final models for the analysis of button-press speed included by-participant random effects for Predictability and by-item random effects for Hearing Contrast, while the models for button-press precision included by-participant random effects for Predictability. For the verbal response analysis, we included by-participant random effects for Predictability for both the response speed and precision analysis. We included by-item random effects for Hearing Contrast 1 for the response speed analysis, but only by-item intercepts for the response precision analysis because any additional random effects resulted in convergence issues.

Although not directly relevant to testing our hypotheses, we also fitted a model comparing the $PwHL_{high\text{-}demand}$ (−0.5) and $PwHL_{low\text{-}}$

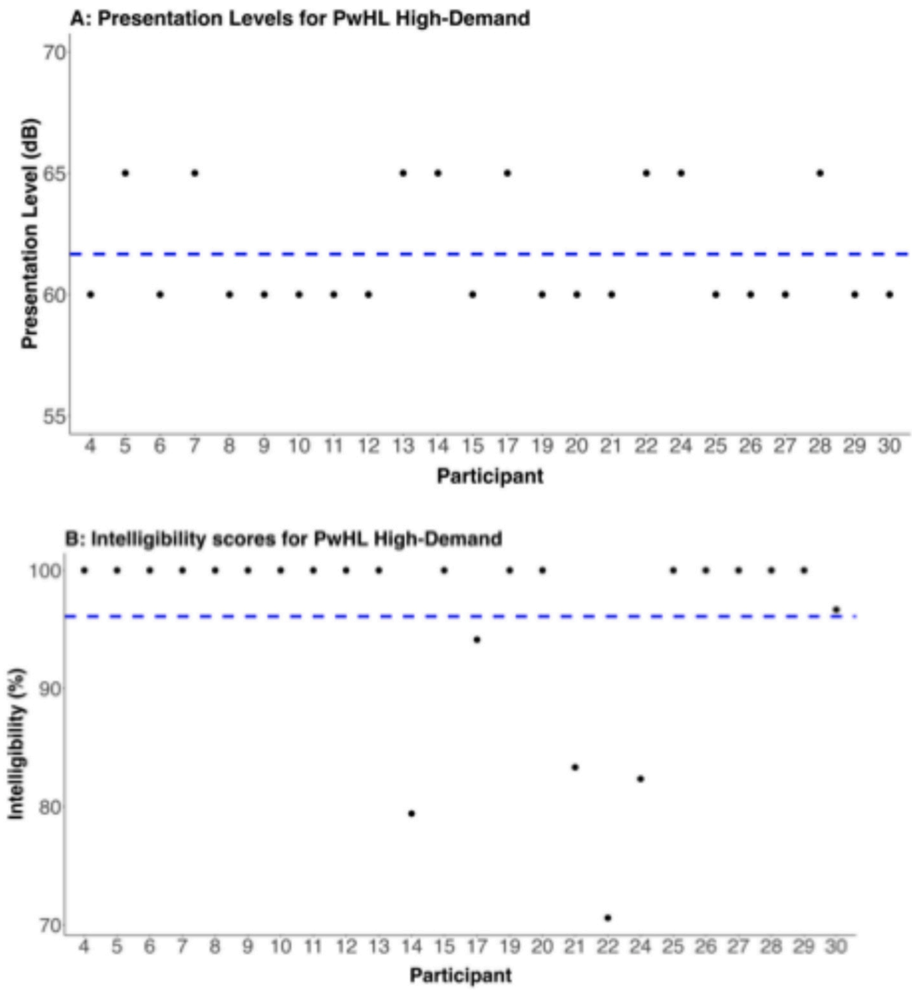

Fig. 5. Panel A shows presentation levels (dB) for each participant in the PwHL$_{\text{high-demand}}$ group and Panel B shows intelligibility scores (%). Dashed blue lines represent the mean presentation level/intelligibility across all participants. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

demand (0.5) groups (Hearing Contrast 3) for completeness. This model followed the same structure as the model including Hearing Contrasts 1 and 2, but it used the maximal random effects structure, including by-participant random effects for Predictability and by-item random effects for Hearing Contrast 3.

3.3. Results

3.3.1. Button press

Older adults (regardless of hearing status) responded on average 379 ms after the sentence end (see Fig. 6 for a breakdown by condition). As in the younger adults analysis, we found that older adults responded earlier when sentences were predictable ($M_{\text{response time}} = 244$ ms) rather than unpredictable ($M_{\text{response time}} = 477$ ms; $b = -166.88$, $SE = 50.34$, $t = -3.32$), and they were no more precise when sentences were predictable ($M_{\text{response precision}} = 466$ ms) rather than unpredictable ($M_{\text{response precision}} = 558$ ms; $b = -0.17$, $SE = 0.10$, $t = -1.73$). Participants also responded earlier ($b = -67.98$, $SE = 21.23$, $t = -3.20$), but not more precisely ($b = -0.00$, $SE = 0.05$, $t = -0.10$) when utterances were longer in duration. There was no interaction between Predictability and Utterance Duration when analysing either button-press times ($b = 24.87$, $SE = 42.45$, $t = 0.59$) or precision ($b = 0.09$, $SE = 0.09$, $t = 0.97$).

Participants in the PwNH group were no faster ($M_{\text{response time}} = 286$ ms) or more precise ($M_{\text{response precision}} = 437$ ms) than those in the PwHL$_{\text{low demand}}$ group ($M_{\text{response time}} = 351$ ms; $b = 34.13$, $SE = 115.38$, $t = 0.30$; $M_{\text{response precision}} = 532$ ms; $b = -0.06$, $SE = 0.14$, $t = -0.42$) or those in the PwHL$_{\text{high-demand}}$ group ($M_{\text{response time}} = 464$ ms; $b = -196.58$, $SE = 121.58$, $t = -1.62$; $M_{\text{response precision}} = 582$ ms; $b = -0.21$, $SE = 0.15$, $t = -1.43$). There was also no difference between the two PwHL groups in either response times ($b = -108.29$, $SE = 105.32$, $t = -1.03$) or precision ($b = -0.06$, $SE = 0.13$, $t = -0.46$).

There was an interaction in terms of response time between Hearing Contrast 2 (PwNH vs. PwHL$_{\text{low-demand}}$) and Predictability ($b = 118.28$, $SE = 46.755$, $t = 2.54$). We followed up this interaction using the emmeans package to compute simple pairwise comparisons testing for a Predictability effect in each Subgroup with Bonferroni corrections. The model returned positive estimates because it subtracted the unpredictable effect from the predictable effect in each subgroup. We found that both PwNH ($b = 111.00$, $SE = 54.30$, $p = .04$) and PwHL$_{\text{low-demand}}$ ($b = 229.00$, $SE = 54.80$, $p < .001$) were faster to respond to predictable (PwNH $M_{\text{response time}} = 177$ ms; PwHL$_{\text{low-demand}}$ $M_{\text{response time}} = 206$ ms) than unpredictable questions (PwNH $M_{\text{response time}} = 395$ ms; PwHL$_{\text{low-demand}}$ $M_{\text{response time}} = 496$ ms), but the predictability effect was larger for participants in the PwHL$_{\text{low-demand}}$ group than those in the PwNH group.

There was also an interaction between Hearing Contrast 3 (PwHL$_{\text{low-demand}}$ vs. PwHL$_{\text{high-demand}}$) and Predictability when analysing response times ($b = -106.26$, $SE = 51.12$, $t = -2.08$). We followed up this interaction using the same procedure as above, finding a predictability effect for the PwHL$_{\text{low-demand}}$ ($b = 215.00$, $SE = 56.90$, $p < .001$) but not the PwNH$_{\text{high-demand}}$ group ($M_{\text{response time}}$ predictable $= 372$ ms, M_{response}

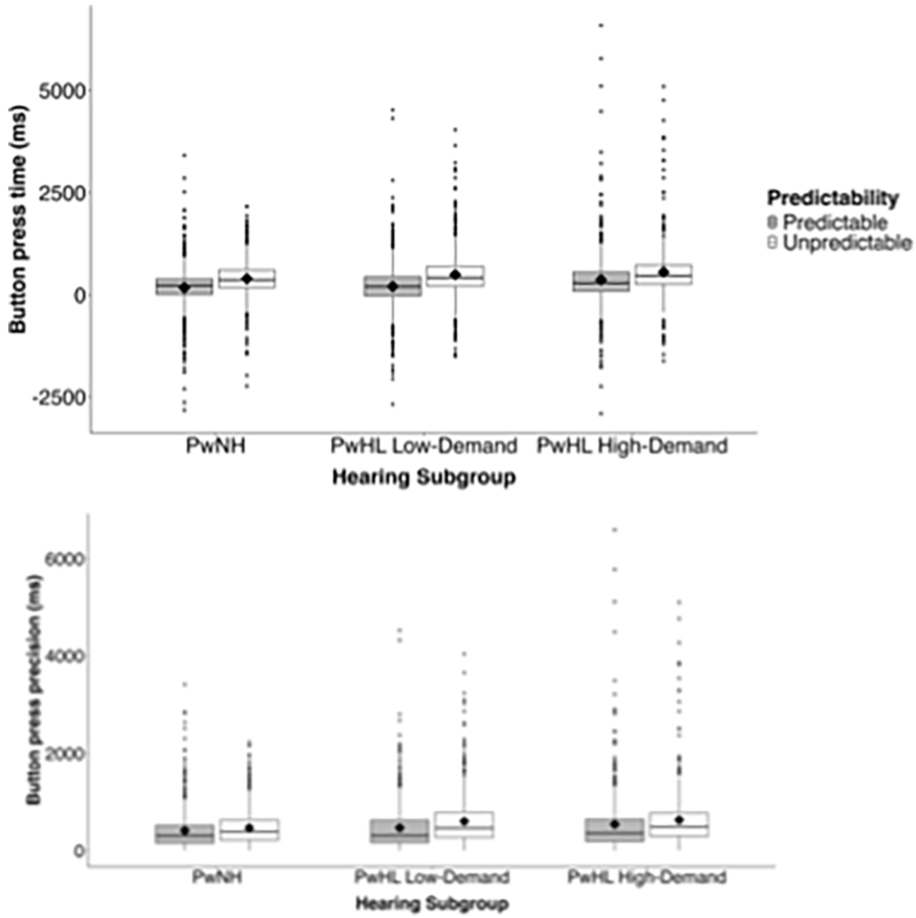

Fig. 6. Distribution of button-press times (top) and precision (bottom) for older adults in the button-pressing task. The box plots show the median, interquartile range, and 1.5× interquartile range (whiskers), with diamond points indicating group means. Button press times were significantly earlier (but not more precise) for predictable than unpredictable sentences. While there were no overall group differences, PwHL$_{\text{low-demand}}$ showed a greater predictability benefit than PwNH and PwHL$_{\text{high-demand}}$ in button-press speed.

$_{\text{time}}$ unpredictable = 558 ms; $b = 109.00$, $SE = 57.70$, $p = .06$). Thus, the PwHL$_{\text{low-demand}}$ group showed a larger predictability effect than both the PwNH and PwHL$_{\text{high-demand}}$ groups.

There was no interaction between Hearing Contrast 1 and Predictability (PwNH vs. PwHL$_{\text{high-demand}}$; $b = -92.97$, $SE = 62.03$, $t = -1.50$) when analysing response times, and none of the Hearing Contrasts interacted with Predictability when analysing response precision (all ts < 1.48). For readability, a summary of the main findings are presented in Table 3.

3.3.2. *Verbal response*

On average, older adults (regardless of hearing status) verbally responded 889 ms after the sentence end (see Fig. 7 for a breakdown by condition). They responded earlier and more precisely when questions were predictable ($M_{\text{response time}} = 742$ ms; $M_{\text{response precision}} = 826$ ms) rather than unpredictable ($M_{\text{response time}} = 1001$ ms; $M_{\text{response precision}} = 1023$ ms; $b = -208.22$, $SE = 63.26$, $t = -3.29$; $b = -0.20$, $SE = 0.08$, $t = -2.55$). Participants did not respond earlier ($b = -51.07$, $SE = 31.67$, $t = -1.61$) or less precisely ($b = -0.03$, $SE = 0.04$, $t = -0.76$) when utterances were longer in duration. There was no interaction between Utterance Duration and Predictability when analysing response times ($b = 4.33$, $SE = 63.35$, $t = 0.07$) or precision ($b = 0.04$, $SE = 0.07$, $t = 0.48$).

Unlike the button-pressing task, participants in the PwNH group responded faster ($M_{\text{response time}} = 664$ ms) and more precisely ($M = 745$ ms) than those in the PwHL$_{\text{high-demand}}$ group ($M_{\text{response time}} = 1039$ ms, $b = -309.99$, $SE = 148.34$, $t = -2.09$; $M_{\text{response precision}} = 1078$ ms, $b = -0.36$, $SE = 0.15$, $t = -2.41$), but there was again no overall difference between the PwNH and PwHL$_{\text{low-demand}}$ groups ($M_{\text{response time}} = 987$ ms, $b = -130.92$, $SE = 141.20$, $t = -0.93$; $M_{\text{response precision}} = 986$ ms, $b = -0.11$, $SE = 0.14$, $t = -0.75$), nor between the two PwHL groups

Table 3
Summary of the main findings for button-press times and precision in Experiment 2.

Effect	Button-press times	Button-press precision
Predictability	Earlier responses for predictable than unpredictable sentences	No more precise for predictable than unpredictable sentences
Hearing Contrast 1: PwNH vs. PwHL$_{\text{high-demand}}$	No difference	
Hearing Contrast 2: PwNH vs. PwHL$_{\text{low-demand}}$	No difference	
Hearing Contrast 3: PwHL$_{\text{high-demand}}$ vs. PwHL$_{\text{low-demand}}$	No difference	
Hearing Contrast 1 * Predictability	No interaction	
Hearing Contrast 2 * Predictability	Interaction: Larger predictability effect for PwHL$_{\text{low-demand}}$ than PwNH	No interaction
Hearing Contrast 3 * Predictability	Interaction: Larger predictability effect for PwHL$_{\text{low-demand}}$ than PwHL$_{\text{high-demand}}$	No interaction

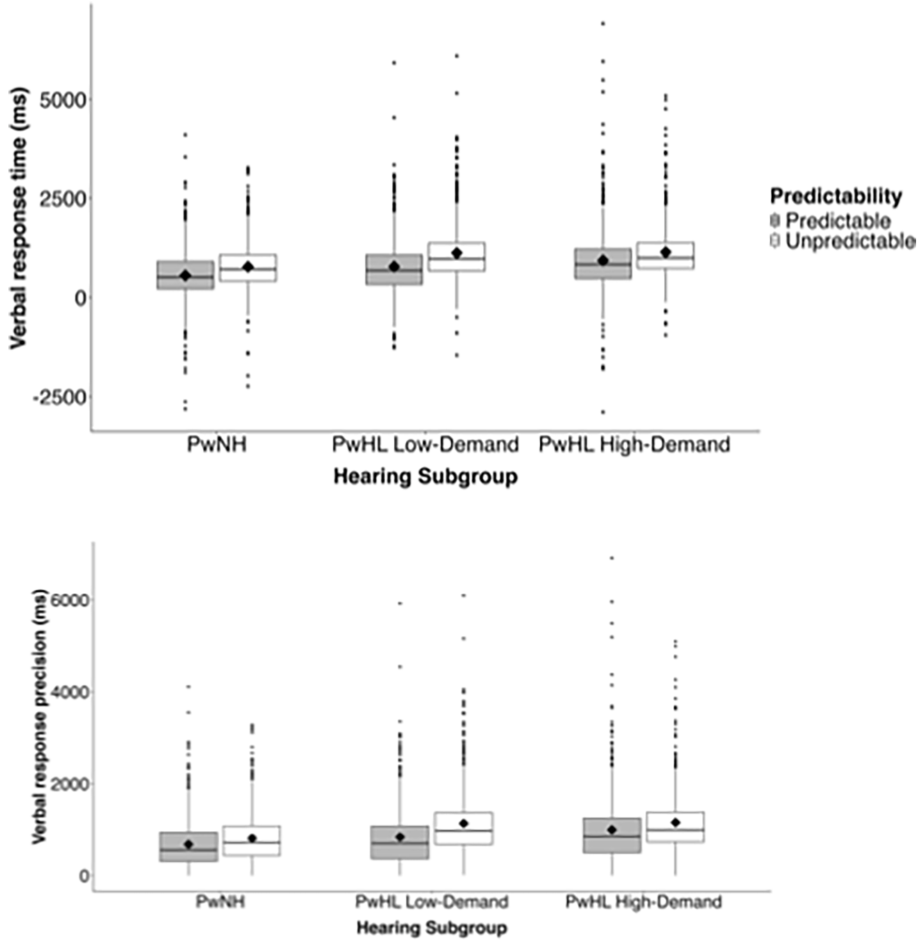

Fig. 7. Distribution of verbal response times (top) and precision (bottom) for older adults in the question-answering task. The box plots show the median, interquartile range, and 1.5× interquartile range (whiskers), with diamond points indicating group means. Verbal response times were significantly earlier and more precise for predictable than unpredictable sentences and PwNH responded faster and more precisely than the $\text{PwHL}_{\text{high-demand}}$ group. Again $\text{PwHL}_{\text{low-demand}}$ showed a greater predictability benefit than PwNH in verbal response speed and precision and a greater predictability benefit than $\text{PwHL}_{\text{high-demand}}$ in verbal response speed.

(response time: $b = -70.69$, $SE = 134.13$, $t = -0.53$, response precision: $b = -0.12$, $SE = 0.13$, $t = -0.84$).

As in the button-press task, there was an interaction between Hearing Contrast 2 (PwNH vs. $\text{PwHL}_{\text{low-demand}}$) and Predictability ($b = 174.13$, $SE = 51.89$, $t = 3.36$) when analysing response times, which we followed up using the same procedure as the button-press analysis. $\text{PwHL}_{\text{low-demand}}$ were faster to respond to predictable ($M_{\text{response time}} = 779$ ms) than unpredictable questions ($M_{\text{response time}} = 1122$ ms; $b = 299.00$, $SE = 68.60$, $p < .001$), while PwNH were not significantly faster to respond to predictable ($M_{\text{response time}} = 555$ ms) than unpredictable questions ($M_{\text{response time}} = 774$ ms; $b = 125.00$, $SE = 68.10$, $p = .07$). Response precision showed a similar pattern, with an interaction between Hearing Contrast 2 and Predictability ($b = 0.26$, $SE = 0.09$, $t = 2.71$) that was driven by a significant predictability effect for $\text{PwHL}_{\text{low-demand}}$ (predictable $M_{\text{response precision}} = 838$ ms, unpredictable $M_{\text{response precision}} = 1135$ ms; $b = 0.34$, $SE = 0.09$, $p < .001$) but not for PwNH (predictable $M_{\text{response precision}} = 678$ ms, unpredictable $M_{\text{response precision}} = 812$ ms; $b = 0.08$, $SE = 0.09$, $p = .39$).

There was also an interaction between Hearing Contrast 3 and Predictability when analysing response times ($\text{PwHL}_{\text{high-demand}}$ vs. $\text{PwHL}_{\text{low-demand}}$; $b = -140.00$, $SE = 52.89$, $t = -2.65$), which was driven by a larger predictability effect for $\text{PwHL}_{\text{low-demand}}$ (predictable $M_{\text{response time}} = 779$ ms, unpredictable $M_{\text{response time}} = 1122$ ms; $b = 277.00$, $SE = 67.70$, $p < .001$) than $\text{PwHL}_{\text{high-demand}}$ (predictable $M_{\text{response time}} = 935$ ms, unpredictable $M_{\text{response time}} = 1143$ ms; $b = 137.00$, $SE = 68.50$, $p = .05$).There was no such interaction for response precision ($SE = -0.18$, $SE = 0.10$, $t = -1.88$). Thus, the $\text{PwHL}_{\text{low-demand}}$ group again showed a larger predictability effect than both the PwNH and $\text{PwHL}_{\text{high-demand}}$ groups.

There was no interaction between Hearing Contrast 1 and Predictability when analysing response times (PwNH vs. $\text{PwHL}_{\text{high-demand}}$; $b = -100.33$, $SE = 68.19$, $t = -1.47$), nor did Predictability interact with Hearing Contrast 1 ($b = -0.10$, $SE = 0.10$, $t = -1.04$) when analysing response precision. A summary of the main findings are presented in Table 4.

3.4. Discussion

In Experiment 2, we used the paradigm validated with younger adults in Experiment 1 to investigate the basis of poorer timed conversational turn taking in older adults with hearing loss. As in Experiment 1, older adults (regardless of hearing status) responded earlier when utterances were predictable rather than unpredictable in both the button-press and verbal response tasks. In the verbal response task, we also found that older adults responded more precisely when utterances were predictable than unpredictable.

In the button-press task, there were no overall group differences between PwNH, $\text{PwHL}_{\text{low-demand}}$, and $\text{PwHL}_{\text{high-demand}}$, but $\text{PwHL}_{\text{low-demand}}$ showed larger effects of utterance predictability than either PwNH or $\text{PwHL}_{\text{high-demand}}$ when analysing response times. We replicated this interaction in the verbal response task, but also found that $\text{PwHL}_{\text{low-demand}}$ showed larger effects of predictability than PwNH when we

Table 4
Summary of the main findings for verbal response times and precision in Experiment 2.

Effect	Verbal response times	Verbal response precision
Predictability	Earlier and more precise responses for predictable than unpredictable sentences.	
Hearing Contrast 1: PwNH vs. $PwHL_{high\text{-}demand}$	PwNH were faster and more precise than $PwHL_{high\text{-}demand}$	
Hearing Contrast 2: PwNH vs. $PwHL_{low\text{-}demand}$	No difference	
Hearing Contrast 3: $PwHL_{high\text{-}demand}$ vs. $PwHL_{low\text{-}demand}$	No difference	
Hearing Contrast 1 * Predictability	No interaction	
Hearing Contrast 2 * Predictability	Interaction: Larger predictability effect for $PwHL_{low\text{-}demand}$ than PwNH	
Hearing Contrast 3 * Predictability	Interaction: Larger predictability effect for $PwHL_{low\text{-}demand}$ than $PwHL_{high\text{-}demand}$	No interaction

analysed response precision. There was also an overall difference between PwNH and $PwHL_{high\text{-}demand}$ in the verbal response task, with the latter showing slower and less precise responses regardless of predictability. These findings suggest that when listening is minimally demanding, people with hearing loss may proactively engage beneficial context-based listening strategies (i.e., relying more on prediction). On the other hand, when listening is more demanding, people with hearing loss may have fewer resources available for such compensation. We discuss these findings in more detail in the General Discussion.

4. General discussion

In two experiments, we investigated how content predictability supports turn-end prediction and response planning in PwNH and PwHL who either listened at lower or higher levels of listening demand. We used a novel paradigm in which participants listened to predictable and unpredictable sentences and pressed a button when they thought the turn would end, before subsequently verbally responding. This task was developed to capture the simultaneous processes of turn-end prediction and response planning present in natural conversation.

In Experiment 1, younger adults without hearing loss responded earlier, but not more precisely, in the button-press task when turns were predictable rather than unpredictable, suggesting that predictability facilitated the timing of generating a prediction about the turn-end but not the accuracy of pinpointing its occurrence. In contrast, participants responded earlier and more precisely for predictable than unpredictable turns in the verbal response task, suggesting that predictability facilitated both the speed of planning a verbal response and the precision of pinpointing when this response should be produced. We replicated these results in older adults with and without hearing loss in Experiment 2, but we also found that PwHL who listened at higher levels of demand were overall slower to respond verbally than PwNH. Importantly, we found that PwHL who listened at lower levels of demand showed stronger effects of content predictability than PwNH or PwHL listening at higher levels of demand when analysing button-press times and verbal response times and precision. This finding suggests that content predictability facilitates turn-end prediction and response planning in adults with hearing loss, particularly when listening demand is low.

It is notable that predictability influenced the precision of verbal responses, but not button-presses, even though we expected both measures to be affected by the accuracy of turn-end prediction. This difference likely occurred because the tasks differ in their cognitive demands and sensitivity to the full scope of turn-end prediction processes. The button-press task simply involved indicating the turn-end, and although predictability might affect the speed of forming this expectation the ultimate precision of determining this moment may be more affected by general motor activation than semantic predictability. In contrast, producing a verbal response does not only involve timing a response – participants must also retrieve, formulate, and articulate the linguistic content of an appropriate answer. Predictability likely affects not only the timing of turn-end prediction (similar to the button-press task) but also contributes to the efficiency of subsequent response planning, especially since theories suggest predictability facilitates lexical access (e.g., Kuperberg & Jaeger, 2016). As a result, the verbal response task provided a broader scope for predictability to exert its influence.

The slower and less precise verbal responses for the $PwHL_{high\text{-}demand}$ group compared to the PwNH group is consistent with studies that have shown that PwHL are slower (e.g., Petersen, MacDonald, & Sørensen, 2022) and more variable at taking turns than PwNH, an effect that is reduced when listening demand is reduced via use of hearing aids (Petersen & Parker, 2024). Thus the paradigm taps into several of the same processes to those that occur in real conversation. The simultaneous delay in timing and reduced precision of verbal responses for the $PwHL_{high\text{-}demand}$ group, despite no difference in button press timing or precision could arise because listening demand draws on the same resources as those required for response planning but not temporal prediction. In this case, PwHL may be able to predict when a turn will end (as reflected in their button-press timing) but show a delay in formulating (measured using verbal response times) and initiating (measured using verbal response precision) their verbal responses due to the increased cognitive demand from listening. Recent work showing that PwHL are slower to predict spoken content when listening demand is high (Fernandez et al., 2025) supports the idea that listening demand will affect the cognitive resources available for either temporal prediction or verbal response planning.

Importantly, PwHL who listened at lower levels showed stronger effects of content predictability in both the button-press (as evidenced by faster button-press times) and verbal response tasks than PwNH and $PwHL_{high\text{-}demand}$ (as evidenced by faster and more precise verbal responses), suggesting that content predictability facilitates both turn-end prediction and verbal response planning in adults with hearing loss when turns are highly intelligible and listening demand is low. This finding is consistent with models of listening demand and cognitive resource allocation (e.g., Peelle, 2018; Pichora-Fuller et al., 2016; Rönnberg, 2003), which claim that adverse listening conditions can increase the cognitive demands of speech processing, even when intelligibility is high (e.g., Zekveld et al., 2011). Our findings suggest that although PwHL may have fewer cognitive resources available due to increased listening effort (as evidenced by the overall group difference between $PwHL_{high\text{-}demand}$ and PwNH), they proactively use beneficial context-based listening strategies (i.e., relying more on semantic prediction) to buffer against potential difficulties, consistent with several previous studies demonstrating that PwHL show greater reliance on conextual cues than PwNH (e.g., Benichov et al., 2012). These contextual cues lead to a larger facilitative effect of content predictability in PwHL when listening demand is low. When listening demand is high, however, we did not observe a comparable facilitative effect of content predictability, potentially because the higher cognitive demands reduced the resources available to fully engage in predictive processing and benefit from content predictability (e.g., Fernandez et al., 2025). Alternatively, cognitive demand may be so high that any facilitative effects of content predictability may be insufficient to provide a relative advantage.

It is worth noting that some participants in the $PwHL_{high\text{-}demand}$ group had similar intelligibility levels as the $PwHL_{low\text{-}demand}$ group. Research suggests that even though intelligibility and listening effort are related, it is entirely possible to have differences in effort even if intelligibility is similar across groups (Winn & Teece, 2021). As a result, we aimed for high intelligibility across all our participant groups because had participants in our $PwHL_{high\text{-}demand}$ group consistently missed a substantial proportion of the words, any observed delays in responses could have simply been attributed to poor audibility – if participants could not hear the sentence, then they could not determine its end or plan a response. Even though intelligibility was high in both our PwHL groups, our results suggest our manipulation was still strong enough to elicit group differences, which we can confidently attribute to effects of listening demand and hearing loss rather than a mere byproduct of missing words due to insufficient audibility.

Although the goal of this work was not to explicitly examine effects of age on turn-end prediction and verbal response planning, we did conduct an exploratory analysis to address whether the older adult groups showed particularly prominent differences in button-press and verbal response performance in relation to the younger adults (the full analysis is available in Appendix C and at https://osf.io/852tx/). There was no difference between the groups when analysing button-press times, but younger adults without hearing loss were more precise in this task than the three groups of older adults. They were also faster and more precise in the verbal response task.

In sum, we found that content predictability facilitated both turn-end prediction and verbal response planning in younger adults (Experiment 1) and older adults with and without hearing loss (Experiment 2). Importantly, our exploration of hearing loss revealed interesting differences in use of predictable content dependent on listening demand. PwHL who listened at lower levels of demand showed stronger effects of content predictability than PwNH in both turn-end prediction and verbal response, suggesting predictability benefits for people with hearing loss when they had adequate capacity to use such information. PwHL listening at higher levels of demand, on the other hand, showed similar effects of content predictability to PwNH in relation to turn-end prediction, but overall delays in preparing their verbal responses. Together these findings indicate that use of predictable context for people with hearing loss may vary depending on listening demands, and thus that hearing loss may not only impact speech intelligibility, but also the reliance on, and speed of, specific conversationally-relevant cognitive processes.

CRediT authorship contribution statement

Ruth E. Corps: Writing – review & editing, Writing – original draft, Formal analysis. **Muzna Shehzad:** Methodology, Data curation. **Andrew I. McLaren:** Methodology, Data curation, Conceptualization. **Leigh B. Fernandez:** Writing – review & editing, Methodology, Conceptualization. **Lauren V. Hadley:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition.

Financial

This work was supported by a UKRI Future Leaders Fellowship [grant number MR/T041471/1]; and the Medical Research Council [grant number MR/X003620/1].

Declaration of competing interest

There are no conflicts of interest, financial, or otherwise.

Appendix A. Experimental stimuli

Table A1
The predictable and unpredictable sentences used in both experiments.

Predictability	Sentence
Predictable	Do you celebrate Christmas on the twenty-fifth of December?
	If I wear sunglasses, will they keep the sun out of my eyes?
	Are pandas the colours black and white?
	Have you ever seen a spider with less than eight legs?
	Do you regularly borrow books from the library?
	Is a piano a musical instrument?
	Do you think most students will pass their exams?
	Did you wake up at 9 o'clock this morning?
	When meeting someone new, do you shake their hand?
	Have you ever dyed your hair?
	Can most fish breathe under water?
	Is red your favourite colour?
	Should I go to the zoo if I want to see a lot of different animals?
	In your life, have you ever failed an exam?
	Is Spring your favourite season of the year?
	Do genies grant wishes?
	Does the Queen live at Buckingham Palace?
	Should I buy my friend a present for her birthday?
	Do you wash your hair every day?
	Have you passed your driving test?
	To cook a cake, will I need to put it in the oven?
	Do you think surfers are scared of being bitten by a shark?
	Can you type without looking at the keyboard?
	While eating, have you ever accidentally bitten your tongue?
	Do you live in a house with other people
	Is a unicorn a horse with a horn?
Unpredictable	Would you like to go for a walk in the countryside?
	Are you very scared of thunder?
	Is an orange the same colour as a mango?
	Today, do you think I should wear a tie?
	Do most people have two names?
	Do you live far away from the pub?
	Are you really looking forward to dinner?
	Have you ever broken your tooth?

(continued on next page)

Table A1 (*continued*)

Predictability	Sentence
	Would you like to take an evening class?
	Do you need to go to the supermarket to buy some cereal?
	Have you ever seen a wild elephant?
	Do you have a big sister?
	Are you in your third year of retirement?
	Have you ever visited the city of Sheffield?
	Are you in a mood?
	Are you doing anything strenuous?
	Do you think you are good at dancing?
	In your opinion, do you think you are a good driver?
	In the past, have you had a lot of different hairstyles?
	Do you have two feet?
	Are you allergic to cats?
	Have you ever injured your eye?
	Would you like to see a picture of my holiday?
	Have you ever played a game of bridge?
	Do you like to eat a lot of cheese?
	Have you ever been on a donkey?

Appendix B. Analysis of the time between button-press and verbal response onset

As a proxy of verbal response time we conducted an additional analysis of the time between button-press (i.e., the point when participants thought they could answer) and the verbal response (initiation of that answer).

Experiment 1: Younger adults

As in our main analyses, we used linear mixed effects models (LMM) such that the time between button-press and verbal response was predicted by Predictability (reference level: unpredictable vs. predictable; −0.5, 0.5). We initially fitted a model using the maximal random effects structure (Barr et al., 2013), but it returned a singular fit error likely because Predictability explained little by-participant variance. As a result, the final model included by-participant and by-item random intercepts only. On average, younger adults produced a verbal response 103 ms after their button-press. There was no difference in the time between button-press and verbal response for predictable ($M = 95$ ms) and unpredictable sentences ($M = 110$ ms; $b = -15.75$, $SE = 11.18$, $t = -1.41$).

Experiment 2: Older adults

LMMs were fitted using the same procedure as Experiment 1, but additionally included Hearing subgroup and its interaction with Predictability. As in our main analysis, Hearing subgroup had three levels (PwNH, $\text{PwHL}_{\text{low-demand}}$, and $\text{PwHL}_{\text{high-demand}}$) and we fitted one model including two contrasts: One comparing the PwNH group to the $\text{PwHL}_{\text{high-demand}}$ group (0.5, −0.5; Hearing Contrast 1) and the other comparing the PwNH group to the $\text{PwHL}_{\text{low-demand}}$ group (0.5, −0.5; Hearing Contrast 2). The maximal structure resulted in a singular fit error, likely because Hearing Contrast 1 explained little by-item variance and Predictability explained little by-participant variance. As a result, the final model included by-item random effects for Hearing Contrast 2 only. We also fitted a similar model comparing the $\text{PwHL}_{\text{high-demand}}$ (−0.5) and $\text{PwHL}_{\text{low-demand}}$ groups (Hearing Contrast 3), using the maximal random effects structure.

On average, older adults produced a verbal response 510 ms after their button-press. As in the analysis of the younger adults, there was again no difference for predictable ($M = 497$ ms) and unpredictable sentences ($M = 523$ ms; $b = -26.34$, $SE = 19.29$, $t = -1.32$). The PwNH group ($M = 378$ ms) did not differ from $\text{PwHL}_{\text{high-demand}}$ ($M = 573$ ms; $b = -112.81$, $SE = 90.27$, $T = -1.25$) or $\text{PwHL}_{\text{low-demand}}$ ($M = 599$ ms; $b = -164.01$, $SE = 86.93$, $t = -1.89$), and there was no difference between the two PwHL groups ($b = 37.33$, $SE = 82.34$, $t = 0.45$). There was also no interaction between Predictability and Hearing Contrast 1 (PwNH vs. $\text{PwHL}_{\text{high-demand}}$; $b = -9.04$, $SE = 28.16$, $t = -0.32$) or Hearing Contrast 3 ($\text{PwHL}_{\text{high-demand}}$ vs. $\text{PwHL}_{\text{low-demand}}$; $b = -36.33$, $SE = 25.12$, $t = -1.45$), but there was an interaction with Hearing Contrast 2 (PwNH vs. $\text{PwHL}_{\text{low-demand}}$ $b = 59.36$, $SE = 28.87$, $t = 2.06$). We followed up this interaction using the emmeans package to compute simple pairwise comparisons testing for a Predictability effect in each Subgroup with Bonferroni corrections. The model returned positive estimates because it subtracted the unpredictable effect from the predictable effect in each subgroup. We found that the time between button-press and verbal response in the $\text{PwHL}_{\text{low-demand}}$ group was smaller for predictable ($M = 571$ ms) than unpredictable sentences ($M = 627$ ms; $b = 56.35$, $SE = 24.80$, $p = .02$), but there was no such difference for the PwNH group (M predictable = 378 ms, M unpredictable = 378 ms; $b = -3.01$, $SE = 24.40$, $p = .90$).

Appendix C. Full model outputs

Table C1
Younger adults – Full model outputs for the analysis of button-press times, button-press precision, verbal response times, and verbal response precision. RE var. = Random effects variance; (p) stands for random effects by participants; (i) stands for random effects by items. All predictors are defined in the Data Analysis section for each experiment.

Fixed effect	Estimate	SE	*t* value	RE variance
Button-press times				
Intercept	116.83	47.09	2.48	(p) 52647; (i) 16,200
Duration	−88.72	21.79	−4.07	–
Predictability (predictable vs. unpredictable)	−176.72	46.51	−3.80	(p) 9399
Duration x Predictability	−3.08	43.57	−0.07	–
Button-press precision				
Intercept	5.36	0.11	49.65	(p) 0.15; (i) 0.10
Duration	0.08	0.07	1.17	–
Predictability (predictable vs. unpredictable)	−0.13	0.15	−0.86	(p) 0.10
Duration x Predictability	0.35	0.13	2.67	–
Verbal response times				
Intercept	207.12	51.27	4.04	(p) 57374; (i) 20,039
Duration	−88.44	24.40	−3.63	–
Predictability (predictable vs. unpredictable)	−197.33	53.16	−3.71	(p) 14160
Duration x Predictability	−0.12	48.79	−0.00	–
Verbal response precision				
Intercept	5.58	0.11	52.11	(p) 0.15; (i) 0.09
Duration	0.02	0.06	0.27	–
Predictability (predictable vs. unpredictable)	−0.24	0.12	−1.97	(p) 0.68
Duration x Predictability	0.29	0.12	2.36	–

Table C2
Older adults – Full model outputs for the analysis of button-press times, button-press precision, verbal response times, and verbal response precision. RE var. = Random effects variance; (p) stands for random effects by participants; (i) stands for random effects by items. All predictors are defined in the Data Analysis section for each experiment.

Fixed effect	Estimate	SE	*t*/*p* value	RE variance
Button-press times				
Intercept	361.70	47.50	7.62	(p) 133570; (i) 24,190
Duration	−67.98	21.23	−3.20	–
Hearing subgroup 1 (PwNH vs. $PwHL_{high\text{-}demand}$)	−196.58	121.58	−1.62	(i) 19,515
Hearing subgroup 2 (PwNH vs. $PwHL_{low\text{-}demand}$)	34.13	115.38	0.30	–
Hearing subgroup 3 ($PwHL_{low\text{-}demand}$ vs. $PwHL_{high\text{-}demand}$)	−108.29	105.32	−1.03	(i) 11,352
Predictability (predictable vs. unpredictable)	−166.88	50.34	−3.32	(p) 6649
Predictability x Hearing subgroup 1 (PwNH vs. $PwHL_{high\text{-}demand}$)	−92.97	62.03	−1.50	–
Predictability x Hearing subgroup 2 (PwNH vs. $PwHL_{low\text{-}demand}$)	118.28	46.55	2.54	–
Predictability x Hearing subgroup 3 ($PwHL_{low\text{-}demand}$ vs. $PwHL_{high\text{-}demand}$)	−106.26	51.12	−2.08	–
Duration x Predictability	24.87	42.45	0.59	–
Predictability x Hearing subgroup 2: PwNH – Predictability	111.00	54.30	0.04	–
Predictability x Hearing subgroup 2: $PwHL_{low\text{-}demand}$ – Predictability	229.00	54.80	<0.001	–
Predictability x Hearing Subgroup 3: $PwHL_{low\text{-}demand}$ – Predictability	215.00	56.90	<0.001	–
Predictability x Hearing Subgroup 3: $PwHL_{high\text{-}demand}$ – Predictability	109.00	57.70	0.06	–
Button-press precision				
Intercept	6.07	0.07	90.91	(p) 0.12; (i) 0.05
Duration	0.00	0.05	0.92	–
Hearing subgroup 1 (PwNH vs. $PwHL_{high\text{-}demand}$)	−0.21	0.15	−1.43	–
Hearing subgroup 2 (PwNH vs. $PwHL_{low\text{-}demand}$)	−0.06	0.14	−0.42	–
Hearing subgroup 3 ($PwHL_{low\text{-}demand}$ vs. $PwHL_{high\text{-}demand}$)	−0.06	0.13	−0.46	(i) 0.03
Predictability (predictable vs. unpredictable)	−0.17	0.10	−1.73	(p) 0.09
Predictability x Hearing subgroup 1 (PwNH vs. $PwHL_{high\text{-}demand}$)	−0.05	0.13	−0.34	–
Predictability x Hearing subgroup 2 (PwNH vs. $PwHL_{low\text{-}demand}$)	0.19	0.13	1.48	–
Predictability x Hearing subgroup 3 ($PwHL_{low\text{-}demand}$ vs. $PwHL_{high\text{-}demand}$)	−0.12	0.13	−0.90	–
Duration x Predictability	0.09	0.09	0.97	–
Verbal response times				
Intercept	883.82	58.72	15.05	(p) 200753; (i) 22,551
Duration	−51.07	31.67	−1.61	–
Hearing subgroup 1 (PwNH vs. $PwHL_{high\text{-}demand}$)	−309.99	148.34	−2.09	(i) 22,551

(*continued on next page*)

Table C2 (*continued*)

Fixed effect	Estimate	SE	*t/p* value	RE variance
Hearing subgroup 2 (PwNH vs. $PwHL_{low\text{-}demand}$)	−130.92	141.20	−0.93	–
Hearing subgroup 3 ($PwHL_{low\text{-}demand}$ vs. $PwHL_{high\text{-}demand}$)	−70.69	134.13	−0.53	(i) 7760
Predictability (predictable vs. unpredictable)	−208.22	63.26	−3.29	(p) 7231
Predictability x Hearing subgroup 1 (PwNH vs. $PwHL_{high\text{-}demand}$)	−100.33	68.19	−1.47	–
Predictability x Hearing subgroup 2 (PwNH vs. $PwHL_{low\text{-}demand}$)	174.13	51.89	3.36	–
Predictability x Hearing subgroup 3 ($PwHL_{low\text{-}demand}$ vs. $PwHL_{high\text{-}demand}$)	−139.99	52.89	−2.65	–
Duration x Predictability	4.33	63.35	0.07	–
Predictability x Hearing subgroup 2: PwNH - Predictability	125.00	68.10	0.06	–
Predictability x Hearing subgroup 2: $PwHL_{low\text{-}demand}$ - Predictability	299.00	68.60	< 0.001	–
Predictability x Hearing subgroup 3: $PwHL_{low\text{-}demand}$ – Predictability	277.00	67.70	<0.001	–
Predictability x Hearing subgroup 3: $PwHL_{high\text{-}demand}$ – Predictability	137.00	68.50	0.05	–
Verbal response precision				
Intercept	6.70	0.06	106.12	(p) 0.07; (i) 0.02
Duration	−0.03	0.04	−0.76	–
Hearing subgroup 1 (PwNH vs. $PwHL_{high\text{-}demand}$)	−0.36	0.15	−2.41	–
Hearing subgroup 2 (PwNH vs. $PwHL_{low\text{-}demand}$)	−0.11	0.14	−0.75	–
Hearing subgroup 3 ($PwHL_{low\text{-}demand}$ vs. $PwHL_{high\text{-}demand}$)	−0.10	0.14	−0.75	(i) 0.01
Predictability (predictable vs. unpredictable)	−0.20	0.08	−2.55	(p) 0.03
Predictability x Hearing subgroup 1 (PwNH vs. $PwHL_{high\text{-}demand}$)	−0.10	0.10	−1.04	–
Predictability x Hearing subgroup 2 (PwNH vs. $PwHL_{low\text{-}demand}$)	0.26	0.09	2.71	–
Predictability x Hearing subgroup 3 ($PwHL_{low\text{-}demand}$ vs. $PwHL_{high\text{-}demand}$)	−0.18	0.10	−1.88	–
Duration x Predictability	0.04	0.07	0.48	–
Predictability x Hearing subgroup 2: PwNH - Predictability	0.08	0.09	0.39	–
Predictability x Hearing subgroup 2: $PwHL_{low\text{-}demand}$ - Predictability	0.34	0.09	<0.001	–

Table C3

Full model outputs for the combined analysis of button-press times, button-press precision, verbal response times, and verbal response precision. RE var. = Random effects variance; (p) stands for random effects by participants; (i) stands for random effects by items. All predictors are defined in the Data Analysis section for each experiment.

Fixed effect	Estimate	SE	*t* value	RE variance
Button-press times				
Intercept	186.96	62.62	2.99	(p) 125047; (i) 12,285
Duration	−81.40	21.45	−3.79	–
Participant Group 1 (younger adults with normal hearing vs. PwNH)	−193.85	179.89	−1.09	–
Participant Group 2 (younger adults with normal hearing vs. $PwHL_{low\text{-}demand}$)	−193.61	142.05	−1.36	–
Participant Group 3 (younger adults with normal hearing vs. $PwHL_{high\text{-}demand}$	−673.83	214.93	−3.14	(i) 485,150
Predictability (predictable vs. unpredictable)	−183.59	52.01	−3.53	–
Predictability x Participant Group 1 (younger adults with normal hearing vs. PwNH)	−83.97	104.05	−0.81	–
Predictability x Participant Group 2 (younger adults with normal hearing vs. $PwHL_{low\text{-}demand}$)	59.62	105.89	0.56	–
Predictability x Participant Group 3 (younger adults with normal hearing vs. $PwHL_{high\text{-}demand}$)	−158.64	222.48	−0.71	–
Duration x Predictability	10.51	42.91	0.25	–
Button-press precision				
Intercept	5.42	0.09	61.12	(p) 0.14; (i) 0.06
Duration	0.05	0.05	0.98	–
Participant Group 1 (younger adults with normal hearing vs. PwNH)	−0.95	0.23	−4.07	(i) 0.10
Participant Group 2 (younger adults with normal hearing vs. $PwHL_{low\text{-}demand}$)	−1.25	0.18	−7.03	(i) 0.19
Participant Group 3 (younger adults with normal hearing vs. $PwHL_{high\text{-}demand}$	−1.48	0.26	−5.62	(i) 0.40
Predictability (predictable vs. unpredictable)	−0.10	0.12	−0.82	(p) 0.10
Predictability x Participant Group 1 (younger adults with normal hearing vs. PwNH)	0.11	0.25	0.45	–
Predictability x Participant Group 2 (younger adults with normal hearing vs. $PwHL_{low\text{-}demand}$)	0.43	0.25	1.74	–
Predictability x Participant Group 3 (younger adults with normal hearing vs. $PwHL_{high\text{-}demand}$)	0.28	0.31	0.09	–
Duration x Predictability	0.25	0.10	2.43	–
Verbal response times				
Intercept	280.36	73.32	3.82	(p) 180688; (i) 19,886
Duration	−64.64	25.77	−2.51	–
Participant Group 1 (younger adults with normal hearing vs. PwNH)	−769.14	211.95	−3.63	–
Participant Group 2 (younger adults with normal hearing vs. $PwHL_{low\text{-}demand}$)	−1228.75	156.83	−7.84	–
Participant Group 3 (younger adults with normal hearing vs. $PwHL_{high\text{-}demand}$)	−1638.76	247.31	−6.63	(i) 512,527
Predictability (predictable vs. unpredictable)	−214.08	59.69	−3.63	–
Predictability x Participant Group 1 (younger adults with normal hearing vs. PwNH)	−111.07	107.54	−1.03	–
Predictability x Participant Group 2 (younger adults with normal hearing vs. $PwHL_{low\text{-}demand}$)	136.86	109.45	1.25	–
Predictability x Participant Group 3 (younger adults with normal hearing vs. $PwHL_{high\text{-}demand}$)	−144.45	228.99	−0.63	–
Duration x Predictability	−2.19	51.54	−0.04	–
Verbal response precision				
Intercept	5.69	0.08	69.07	(p) 0.10; (i) 0.03

(*continued on next page*)

Table C3 (*continued*)

Fixed effect	Estimate	SE	*t* value	RE variance
Duration	0.01	0.04	0.18	–
Participant Group 1 (younger adults with normal hearing vs. PwNH)	−1.53	0.22	−6.93	–
Participant Group 2 (younger adults with normal hearing vs. $PwHL_{low\text{-}demand}$)	−2.00	0.16	−12.79	(i) 0.13
Participant Group 3 (younger adults with normal hearing vs. $PwHL_{high\text{-}demand}$)	−2.36	0.25	−9.35	(i) 0.19
Predictability (predictable vs. unpredictable)	−0.22	0.09	−2.52	–
Predictability x Participant Group 1 (younger adults with normal hearing vs. PwNH)	−0.11	0.10	−1.09	–
Predictability x Participant Group 2 (younger adults with normal hearing vs. $PwHL_{low\text{-}demand}$)	0.29	0.16	1.79	–
Predictability x Participant Group 3 (younger adults with normal hearing vs. $PwHL_{high\text{-}demand}$)	−0.01	0.20	−0.73	–
Duration x Predictability	0.14	0.08	1.71	–

Data availability

A link to the data/code is provided in the manuscript.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Babiloni, C., Carducci, F., Cincotti, F., Rossini, P. M., Neuper, C., Pfurtscheler, G., & Babiloni, F. (1999). Human movement-related potentials vs desynchronization of EEG alpha rhythm: A high-resolution EEG study. *Neuroimage, 10*, 658–665.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Barthel, M., Meyer, A. S., & Levinson, S. C. (2017). Next speakers plan their turn early and speak after turn-final "go-signals". *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.00393

Barthel, M., Sauppe, S., Levinson, S. C., & Meyer, A. S. (2016). The timing of utterance planning in task-oriented dialogue: Evidence from a novel list-completion paradigm. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01858

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2023). Lme4: Linear mixed-effects models using "Eigein" and S4 (R package version 1.1-35.3). Retrieved from http://CRAN.R-project.org/package=lme4.

Beechey, T., Buchholz, J. M., & Keidser, G. (2020). Hearing impairment increases communication effort during conversations in noise. *Journal of Speech, Language, and Hearing Research, 63*, 305–320.

Benichov, J., Cox, L. C., Tun, P. A., & Wingfield, A. (2012). Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability, and cognitive function. *Ear and Hearing, 33*, 250–256.

Bögels, S., Magyari, L., & Levinson. (2015). Neural signatures of response planning occur midway through an incoming question in conversation. *Scientific Reports, 5*. https://doi.org/10.1038/srep12881

Bögels, S., & Torreira, F. (2015). Listeners use intonational phrase boundaries to project turn ends. *Journal of Phonetics, 52*, 46–57.

Bögels, S., & Torreira, F. (2021). Turn-end estimation in conversational turn-taking: The roles of context and prosody. *Discourse Processes, 58*, 903–924.

Carroll, R., Uslar, V., Brand, T., & Ruigendijk, E. (2016). Processing mechanisms in hearing-impaired listeners: Evidence from reaction times and sentence interpretation. *Ear and Hearing, 37*, e391–e401.

Corps, R. E., Crossley, A., Gambi, C., & Pickering, M. J. (2018). Early preparation during turn-taking: Listeners use content predictions to determine what to say but not when to say it. *Cognition, 175*, 77–95.

Corps, R. E., Gambi, C., & Pickering, M. J. (2020). How do listeners time response articulation when answering questions? The role of speech rate. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*, 781–802.

Corps, R. E., Pickering, M. J., & Gambi, C. (2019). Predicting turn-ends in discourse context. *Language, Cognition and Neuroscience, 34*, 615–627.

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language, 82*, 515–535.

Fernandez, L. B., Pickering, M. J., Naylor, G., & Hadley, L. V. (2024). Uses of linguistic context in speech listening: Does acquired hearing loss lead to reduced engagement of prediction? *Ear and Hearing, 45*, 1107–1114.

Fernandez, L. B., Shehzad, M., & Hadley, L. V. (2025). Effects of hearing loss on semantic prediction: Delayed prediction for intelligible speech when listening is demanding. *Ear and Hearing. Advance online publication*. https://doi.org/10.1098/AUD.0000000000001679

Garrod, S., & Pickering, M. J. (2015). The use of content and timing to predict turn transitions. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00751

Gordon-Salant, S., & Fitzgibbons, P. J. (1997). Selected cognitive factors and speech recognition performance among young and elderly listeners. *Journal of Speech, Language, and Hearing Research, 40*, 423–431.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in Psychology, 52*, 139–183.

Hunter, C. R., & Humes, L. E. (2022). Predictive sentence context reduces listening effort in older adults with and without hearing loss and with high and low working memory capacity. *Ear and Hearing, 43*, 1164–1177.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*, 101–144.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31*, 32–59.

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00731

Magyari, L., Bastiaansen, M. C. M., De Ruiter, J. P., & Levinson, S. C. (2014). Early anticipation lies behind speed of response in conversation. *Journal of Cognitive Neuroscience, 26*, 2530–2539.

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes, 27*, 953–978.

Meyer, A. S., Alday, P. M., Decuyper, C., & Knudsen, B. (2018). Working together: Contributions of corpus analyses and experimental psycholinguistics to understanding conversation. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.00525

Moore, B. C. (1996). Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. *Ear and Hearing, 17*, 133–161.

Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear and Hearing, 39*, 204–214.

Petersen, E. B., MacDonald, E. N., & Sørensen, J. M. A. (2022). The effects of hearing-aid amplification and noise on conversational dynamics between normal-hearing and hearing-impaired talkers. *Trends in Hearing, 26*, 1–18.

Petersen, E. B., & Parker, D. (2024). Speak up: how hearing loss and the lack of hearing aids affect conversations in quiet. *Journal of Speech, Language, and Hearing Research, 67*, 2782–2793.

Petersen, E. B., Walravens, E., & Pedersen, A. K. (2022). *Real-life listening in the lab: Does wearing hearing aids affect the dynamics of group conversation? Proceedings of the 26th workshop on the semantics and pragmatics of dialogue* (pp. 221–224).

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and Hearing, 37*, 5S–27S.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America, 97*, 593–608.

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *International Journal of Audiology, 42*, S68–S76.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 76*, 379–423.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences, 106*, 10587–10592.

Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism Bulletin, 30*, 415–433.

Wagner, A., Pals, C., de Blecourt, C. M., Sarampalis, A., & Basknet, D. (2016). *Does signal degradation affect top-down processing of speech? In physiology, psychoacoustics and cognition in Normal and impaired hearing*. Switzerland (pp. 297–306).

Wendt, D., Kollmeier, B., & Brand, T. (2015). How hearing impairment affects sentence comprehension: Using eye fixations to investigate the duration of speech processing. *Trends in Hearing, 19*. https://doi.org/10.1177/2331216515594149

Winn, M. B., & Teece, K. H. (2021). Listening effort is not the same as speech intelligibility score. *Trends in Hearing, 25*. https://doi.org/10.1177/23312165211027688

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear and Hearing, 32*, 498–510.