



OPEN

Age-related deficits in dip-listening evident for isolated sentences but not for spoken stories

Vanessa C. Irsik¹✉, Ingrid S. Johnsrude^{1,2} & Björn Herrmann^{1,3,4}

Fluctuating background sounds facilitate speech intelligibility by providing speech ‘glimpses’ (masking release). Older adults benefit less from glimpses, but masking release is typically investigated using isolated sentences. Recent work indicates that using engaging, continuous speech materials (e.g., spoken stories) may qualitatively alter speech-in-noise listening. Moreover, neural sensitivity to different amplitude envelope profiles (ramped, damped) changes with age, but whether this affects speech listening is unknown. In three online experiments, we investigate how masking release in younger and older adults differs for masked sentences and stories, and how speech intelligibility varies with masker amplitude profile. Intelligibility was generally greater for damped than ramped maskers. Masking release was reduced in older relative to younger adults for disconnected sentences, and stories with a randomized sentence order. Critically, when listening to stories with an engaging and coherent narrative, older adults demonstrated equal or greater masking release compared to younger adults. Older adults thus appear to benefit from ‘glimpses’ as much as, or more than, younger adults when the speech they are listening to follows a coherent topical thread. Our results highlight the importance of cognitive and motivational factors for speech understanding, and suggest that previous work may have underestimated speech-listening abilities in older adults.

Speech sounds are characterized by low frequency amplitude fluctuations that are not only critical for speech intelligibility in quiet^{1–3}, but are also a useful cue for separating speech from background masking sounds^{4–6}. Aging is associated with a decline in processing temporal auditory features^{7–10}, such as the fluctuating speech envelope, which may be part of the reason older adults frequently struggle to understand speech when background masking sounds are present^{11,12}. We recently demonstrated that cortical sensitivity to a sound’s envelope fluctuations with different temporal profiles differs between younger and older adults¹³, raising the possibility that particular envelope profiles may alter how effectively a masker occludes target speech for older adults. Furthermore, we have also shown that engaging spoken narratives qualitatively alter the speech-listening experience when background noise is present¹⁴, compared to the disconnected sentence-length utterances that are typically used in speech research^{15–20}. In the current study we investigate how the temporal profile of the background masker influences intelligibility of isolated sentences compared to narrative stories in younger and older adults.

For most younger healthy individuals, amplitude variations in background masking sounds facilitate speech intelligibility compared to an energetically matched masker with a flat envelope. Fluctuating maskers are thought to enable a listener to perceive ‘glimpses’ of the target speech (c.f., “listening in the dips”)²¹. The intelligibility benefit for fluctuating over energetically matched, flat-envelope maskers is a type of “release from masking”^{4–6,22–26}. Whereas younger listeners derive a robust benefit from fluctuating maskers, older adults have been consistently demonstrated to either gain no benefit to intelligibility, or a reduced benefit (compared to younger people)^{16,19,23,24,27–31}. These findings may partially result from reduced audibility which reduces the effective amplitude-modulation depth^{20,24}, and therefore the opportunity for ‘glimpses’ of the target. However, controlling for audibility has not resulted in restoration of the release-from-masking effect in all older individuals^{18,23,28,32}. An age-related decline in temporal resolution in the auditory system may also contribute, as reduced temporal resolution can result in increased susceptibility to forward masking^{28,31}, thereby reducing the available ‘glimpses’ of the target (see also the contribution of temporal fine structure^{21,26,30,33–35}).

The amplitude envelope of speech is temporally dynamic: it varies in the rate of rise (attack) and fall (decay) over time³⁶. Sensitivity to the shape of amplitude envelopes is important for identifying and discriminating

¹Department of Psychology & The Brain and Mind Institute, The University of Western Ontario, London, ON N6A 3K7, Canada. ²School of Communication and Speech Disorders, The University of Western Ontario, London, ON N6A 5B7, Canada. ³Rotman Research Institute, Baycrest, Toronto, ON M6A 2E1, Canada. ⁴Department of Psychology, University of Toronto, Toronto, ON M5S 1A1, Canada. ✉email: irsik@uwo.ca

between different consonants (e.g., /pa/ versus /ta/)³. Previous research in younger and older human listeners¹³, and in rats³⁷, indicates that aging is associated with a relative increase in neural sensitivity to sounds with damped (sharp attack and gradual decay) compared to ramped (gradual attack and sharp decay) envelope shapes. Moreover, enhanced neural sensitivity to amplitude modulations in sounds has been linked to reduced speech intelligibility when the background sound is amplitude modulated^{38–40}. Enhanced neural sensitivity to amplitude envelopes may distort envelope cues^{41,42} or, when part of a masking stream of sound, distract an older listener and interfere with comprehension^{38,43–45}. Here, we examine whether the temporal envelope profile of the masker affects susceptibility to masking in young and older listeners. Given that older listeners exhibit greater cortical sensitivity to damped sounds^{13,37}, we anticipate that older listeners will obtain less masking release when target speech is masked by sound with a damped, compared to ramped, amplitude envelope.

Studies that investigate phenomena affecting speech understanding, such as release from masking, generally use brief, disconnected speech utterances, like isolated sentences^{15–20}. Such utterances typically lack a narrative thread and may not be very interesting to the listener. In everyday listening situations, sentences are not typically disconnected. Instead, conversational speech frequently contains inter-related narrated elements, such as stories about past events^{46–58}. Narrated descriptions of past events have been reported to occur as often as 5.4 times per hour⁵⁵. While the structure of a spoken story or narrative can vary based on the conversational circumstances⁵⁰, narrated speech generally follows a topical thread and is contextually rich. The presence of speech context and an overarching topical thread in spoken narratives may support speech understanding and ongoing attention as sentence-level context has been shown to facilitate word identification in noise for both younger and older listeners^{17,59–63}.

In addition, cognitive control research suggests that motivation is key to the investment of cognitive resources^{64–66}, such as when trying to understand speech masked by background sounds^{67–71} and an increasing body of work thus focuses on using enjoyable (i.e., motivating) stories or narratives to investigate speech listening^{72–83}. For example, although moderate speech masking decreases speech intelligibility and increases listening effort in young normally hearing listeners, story absorption and enjoyment are only minimally affected¹⁴. Critically, when a listener is motivated to understand (e.g., when listening to engaging spoken stories), they may be engaging in a different way that may promote intelligibility, compared to when they are less motivated to understand. This may particularly be the case for older adults who may not engage in tasks with low personal relevance in order to conserve resources for more personally relevant tasks^{84,85}. Engaging speech materials may thus reveal qualitative differences between age groups, particularly in the extent to which ‘speech glimpses’ or ‘dip listening’ facilitates intelligibility.

In three behavioral experiments, we use masked disconnected sentences and engaging spoken stories in order to examine how the type of speech utterance and masker temporal profile affects speech intelligibility in younger and older adults. We utilize 12-talker babble masking noise with different amplitude envelopes. The envelope could either be unmodulated (i.e., relatively flat), modulated with a damped temporal profile, or modulated with a ramped temporal profile. In Experiment 1, we examine the effect of different masker modulation types and age on intelligibility using isolated sentences. In Experiment 2, we conduct a similar investigation with an engaging story as the target speech. Given that the procedures and speech materials differ between Experiment 1 and 2, we conduct Experiment 3 using identical procedures and materials for both disconnected-sentence and story conditions.

Results

Experiment 1: release from masking is reduced in older adults for disconnected sentences. In Experiment 1, we investigate how the amplitude envelope type (modulated vs unmodulated) and envelope shape (damped vs. ramped) affect speech intelligibility using a sentence-based intelligibility paradigm. We use similar procedures to those previously used to study release from masking^{19,20,24,27,28} in order to (a) replicate previous observations that older adults benefit less from a modulated over an unmodulated masker compared to younger adults, and (b) examine whether the shape of the modulation (damped or ramped) influences the magnitude of release from masking observed.

The experiment was conducted online using Amazon Mechanical Turk (MTurk; <https://www.mturk.com/>) and Cloud Research (previously TurkPrime⁸⁶) for recruitment and Pavlovina (<https://pavlovina.org/>) to host the experiment. Younger (mean: 35.9 years; age-range: 18–49 years; 39 males, 29 females, 1 non-binary) and older adults (mean: 59.6 years; age-range: 50–71 years; 31 males, 37 females) without reported hearing or neurological issues (self-report) participated in the experiment. Based on previous work⁸⁷ and results from a separate project (see Supplemental Document), we estimate that older adults in the current study had about 7 dB HL higher audiometric pure-tone average thresholds compared to younger adults. The estimation further suggests that approximately 25% of our older adult sample may have a minor hearing impairment (see Supplemental Document), as would be expected from a group of older adults recruited from the community.

During the task, participants listened to disconnected sentences and, after each sentence, typed the words they heard into a text box. A 12-talker babble masker was added to each sentence, and the signal-to-noise ratio (SNR levels: – 10, – 8, – 6, – 4, – 2, 0, + 2 dB) and temporal profile of the amplitude envelope of the masker (unmodulated, 4-Hz amplitude-modulated: damped, ramped) were varied (Fig. 1). We calculated the proportion of correctly reported words for each envelope condition and SNR, fit a logistic function to the mean performance data (Fig. 2a), and analyzed the speech reception threshold (SNR associated with 50% correctly reported words) and slope.

To examine whether the magnitude of masking release differs between age groups, we compared the threshold and slope from the logistic function fits between modulated (unweighted average across ramped and damped shapes) and unmodulated masker conditions and between younger and older adults. Speech reception thresholds

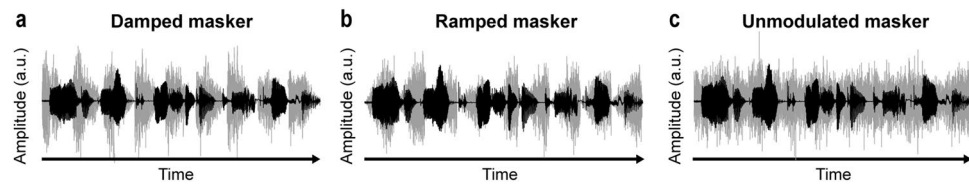


Figure 1. Stimulus design for Experiment 1. Participants listened to sentences (black) to which a 12-talker babble was added with varied SNR levels (grey). Note that for the purpose of clarity, we depict sentence and masker separately. Babble noise was amplitude modulated at a rate of 4-Hz with (a) a damped (b) a ramped envelope shape; or was (c) unmodulated.

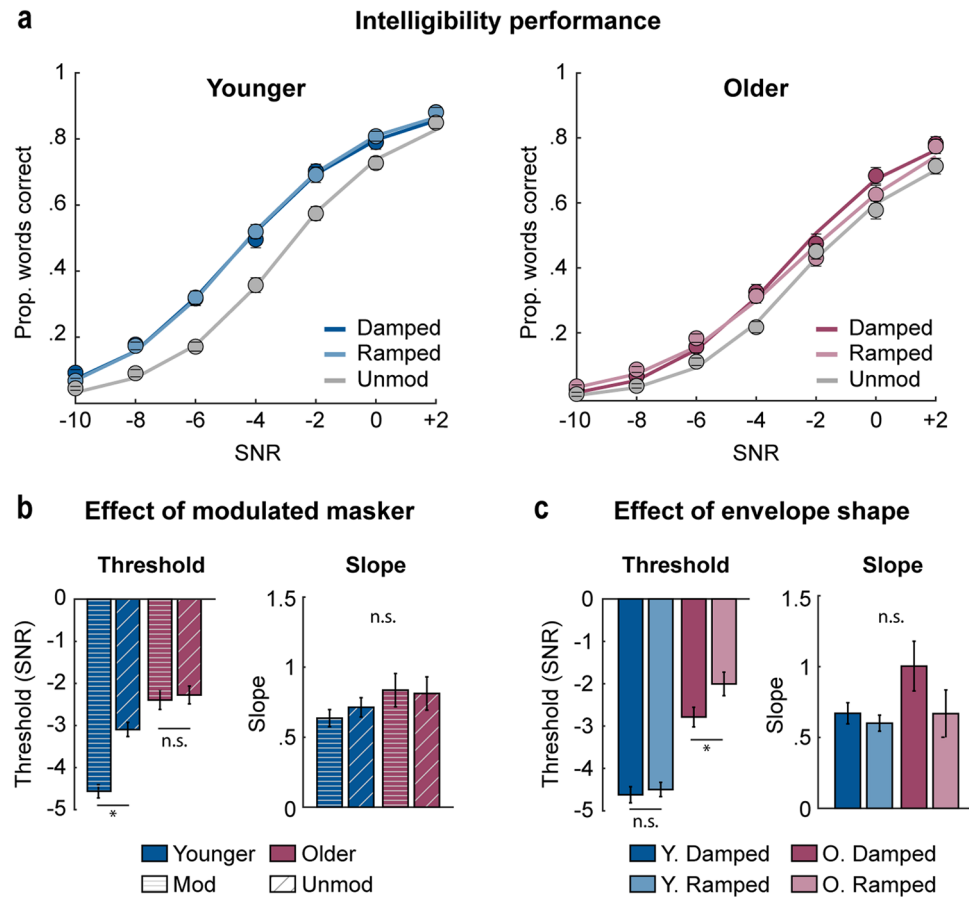


Figure 2. Intelligibility results for Experiment 1. (a) Mean proportion of correctly reported words plotted as a function of SNR ($-10, -8, -6, -4, -2, 0, +2$ dB) for younger (left) and older adults (right), and for envelope conditions (damped, ramped, unmodulated). Colored lines correspond to a logistic function fit to the proportion of correct words reported. (b) Mean threshold (left) and slope (right) coefficients from logistic function fits are plotted for the different modulation types (modulated [averaged across ramped and damped], unmodulated) and age groups (younger, older). (c) Mean threshold (left) and slope (right) coefficients are plotted for the different masker envelope shapes (damped, ramped) and age groups (younger, older). Error bars reflect the standard error of the mean. * $p < 0.05$.

for sentences with a modulated masker were lower than for unmodulated maskers [effect of modulation type: $F_{1,135} = 27.9, p = 5 \times 10^{-7}, \eta^2 p = 0.17$], consistent with previous reports of release from masking^{6,16,26–28} (we replicated this effect in a group of younger adults in a laboratory setting; Supplemental Document). Thresholds were also lower for younger compared to older adults [effect of age group: $F_{1,135} = 43.91, p = 8 \times 10^{-10}, \eta^2 p = 0.25$], consistent with older adults having more difficulty understanding speech in noise^{12,88}. We further observed a significant modulation type \times age group interaction [$F_{1,135} = 20.02, p = 2 \times 10^{-5}, \eta^2 p = 0.13$; Fig. 2b left]: Speech intelligibility was better for modulated compared to unmodulated maskers in younger individuals [$t_{68} = -8.78, p_{FDR} = 2 \times 10^{-12}, r_c = 0.73$], whereas no difference was found for older adults [$p_{FDR} = 0.63$]. This is consistent with previous research^{16,19,27,28,31} indicating that older adults experience reduced release from masking—for fluctuating

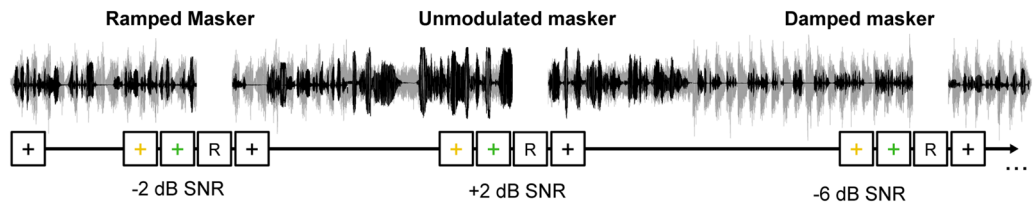


Figure 3. Stimulus design for Experiment 2. Participants listened to a spoken story (black) masked with 12-talker babble noise (grey). The amplitude envelope (damped, ramped, unmodulated) and SNR (−6, −2, +2, Clear) pseudo-randomly varied every 16-s. A fixation cross was displayed on the computer screen throughout the story and changed colors to communicate which parts of the story participants would need to report. The fixation cross turned yellow 2-s prior to the beginning of a test phrase/sentence, cueing the participant to prepare for intelligibility testing, and turned green at the start of the test phrase/sentence to indicate which phrase/sentence they should report back. The story paused with the offset of the test phrase/sentence, at which point participants would report back the phrase/sentence. The story resumed once a response was submitted.

compared to flat masker envelopes—relative to younger adults. Older adults thus do not appear to utilize glimpse listening, at least for the disconnected sentences used here. No significant differences were observed when analyzing slopes [$F < 2$, $p > 0.17$, $\eta^2 p < 0.01$, Fig. 2b right].

For the analysis of different envelope shapes (damped vs. ramped), we observed lower thresholds for damped compared to ramped envelopes [effect of envelope shape: $F_{1,135} = 8.65$, $p = 0.004$, $\eta^2 p = 0.06$], and a significant envelope shape \times age group interaction [$F_{1,135} = 4.55$, $p = 0.035$, $\eta^2 p = 0.03$; Fig. 2c left]. Speech intelligibility thresholds were better (lower) for damped compared to ramped envelope shapes for older adults [$t_{67} = -3.05$, $p_{FDR} = 0.007$, $r_e = 0.35$], but not younger adults [$p_{FDR} = 0.473$]. No significant differences were observed for slope [$F < 3$, $p > 0.11$, $\eta^2 p < 0.02$; Fig. 2c right].

The results of Experiment 1 parallel previous findings on the effect of amplitude modulations on speech intelligibility^{16,25,27,28,31,89}. We show that older individuals benefit less from a modulated over an unmodulated masker, compared to younger participants. We also observed that older, but not younger, listeners benefited when the babble background was modulated with a damped compared to a ramped envelope shape. This speech intelligibility benefit for damped temporal profiles is inconsistent with a recently proposed hypothesis based on electrophysiological work: Older adults demonstrate larger cortical responses to damped compared to ramped sounds¹³, and larger cortical responses to amplitude modulations have been linked to poorer speech intelligibility^{38–40}. Hence, we anticipated that damped babble would interfere more, not less, with the target speech. Instead, increased cortical responsivity to the damped compared to ramped masker may strengthen predictability of modulation phase, facilitating speech ‘glimpsing’.

The short, disconnected sentences used in Experiment 1 are similar to those commonly used in speech intelligibility and masking release research. However, disconnected sentences without a topical thread may be less common in everyday listening situations, where speech is commonly continuous and contains narrated elements^{46–58}. Experiment 2 was designed to investigate whether the effects obtained in Experiment 1 generalize to materials that resemble listening situations with more structured narrated elements, such as spoken stories about life events.

Experiment 2: masking release is greater for older compared to younger adults during story listening.

In Experiment 2, we investigate how the amplitude envelope type (modulated vs unmodulated) and envelope shape (damped vs. ramped) affect speech intelligibility while younger (mean: 30.1 years; age-range: 19–39 years; 37 males, 30 females) and older individuals (mean: 64.4 years; age-range: 53–80 years; 29 males, 41 females) without reported hearing or neurological issues listen to stories. Participant recruitment and testing was conducted using online platforms, as in Experiment 1. We selected a ~13-min spoken story from the story-telling podcast The Moth (<https://themoth.org>), where individuals tell stories about interesting life events. Stories are intended to be engaging and enjoyable, and are increasingly used in experimental research to study engagement with speech^{14,90–93}.

The story was masked by 12-talker babble with different amplitude envelopes (unmodulated, 4-Hz modulated: damped, ramped) and different signal-to-noise ratios (SNRs: −6, −2, +2 dB, Clear). Masker type and SNR changed approximately every 16 s (Fig. 3). The story paused pseudorandomly (approximately every 5–20 s), and participants were asked to report the last phrase/sentence that was spoken by typing into a textbox. A visual cue directed participants exactly which words they should report back (Fig. 3). We calculated the proportion of correctly reported words for each envelope condition (damped, ramped, unmodulated) and SNR (−6, −2, +2 dB, Clear) and compared the result between age groups.

Average word report significantly declined with decreasing SNR [effect of SNR: $F_{2,272} = 644.25$, $p_{GG} = 5.2 \times 10^{-80}$, $\eta^2 p = 0.83$; Fig. 4a], and older adults exhibited worse overall performance compared to younger adults [effect of age group: $F_{1,136} = 4.81$, $p = 0.03$, $\eta^2 p = 0.03$]. We also found a significant SNR \times age group interaction [$F_{2,272} = 12.46$, $p_{GG} = 5.4 \times 10^{-5}$, $\eta^2 p = 0.08$]. Follow up t-tests indicated age group differences at −6 dB SNR [$t_{136} = 2.98$, $p_{FDR} = 0.01$, $r_e = 0.25$], but not at −2 dB [$p_{FDR} = 0.212$] or +2 dB [$p_{FDR} = 0.968$]. This shows that speech intelligibility during −6 dB SNR was more challenging for older compared to younger subjects, while both groups performed equally well at −2 and +2 dB SNR.

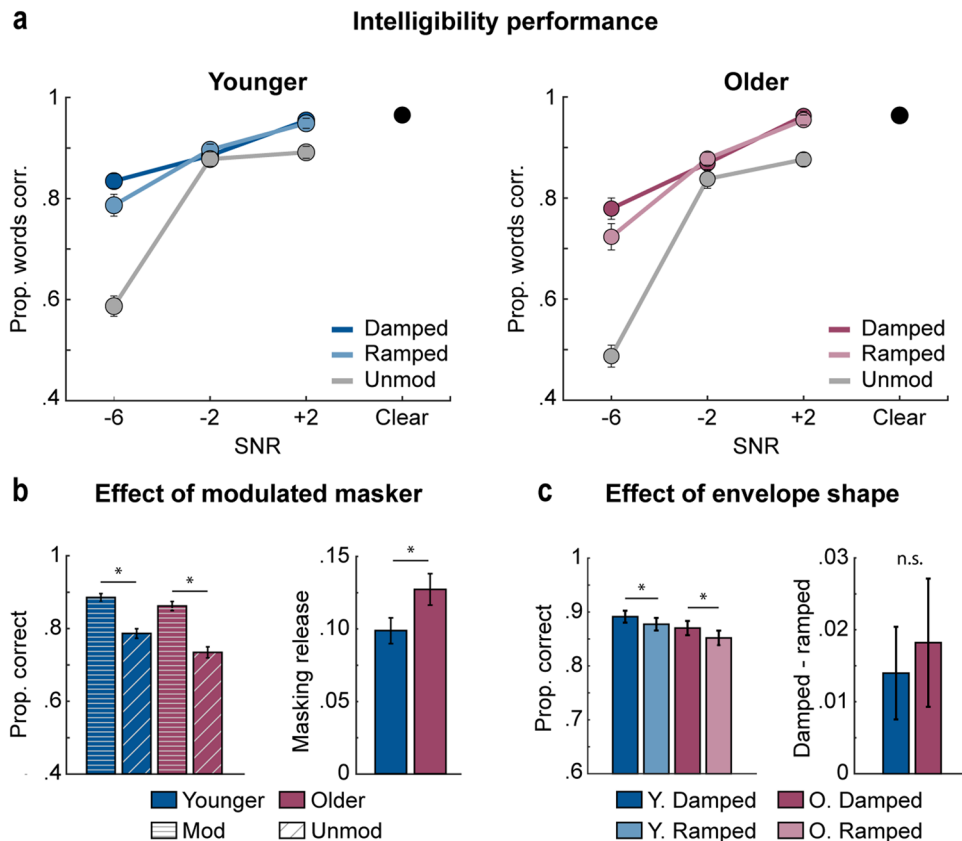


Figure 4. Intelligibility results for Experiment 2. **(a)** Mean proportion of correctly reported words are plotted as a function of SNR (-6 , -2 , $+2$ dB, Clear [black dot]) for the different envelope conditions (damped, ramped, unmodulated) and for younger (left) and older adults (right). **(b)** Mean word report (left) is plotted for the different modulation types (modulated, unmodulated) and age groups (younger, older). The difference in performance between modulated and unmodulated maskers (masking release) is plotted for both age groups (right). Plots for modulated maskers reflect the mean across damped and ramped envelope shapes. **(c)** Mean word report (left) is plotted for the different masker envelope shapes (damped, ramped) and age groups (younger, older). The difference in performance between the damped and ramped envelope shapes is plotted for both age groups (right). Error bars reflect the standard error of the mean. * $p < 0.05$.

We observed higher intelligibility for modulated compared to unmodulated maskers [effect of modulation type: $F_{1,136} = 262.2$, $p = 2 \times 10^{-33}$, $\eta^2 p = 0.66$, Fig. 4b left panel]. This release-from-masking effect (difference between modulated and unmodulated maskers) was greater for older compared to younger participants [modulation type \times age group interaction: $F_{1,136} = 4.14$, $p = 0.044$, $\eta^2 p = 0.03$; Fig. 4b right panel], although both groups showed significant release from masking [modulated vs. unmodulated: younger: $t_{67} = 11.19$, $p_{FDR} = 6 \times 10^{-17}$, $r_e = 0.81$; older: $t_{69} = 11.83$, $p_{FDR} = 6 \times 10^{-18}$, $r_e = 0.82$]. It appears that the modulated masker helped older adults to achieve a similar level of performance as younger adults [younger vs older adults for modulated masker: $p_{FDR} = 0.162$; Fig. 4b left panel], despite lower performance for the unmodulated masker [$t_{136} = 2.57$, $p_{FDR} = 0.023$, $r_e = 0.21$] (Fig. 4b right panel). This is not trivially due to a compressive effect at one or other extreme of performance: performance in the unmodulated and modulated conditions was off ceiling and floor for both age groups (Figs. 4a,b).

We also observed a modulation type \times SNR interaction [$F_{2,272} = 147.94$, $p_{GG} = 1.1 \times 10^{-36}$, $\eta^2 p = 0.52$]. The difference between modulated and unmodulated performance (masking release) was larger at -6 dB, compared both to -2 dB [$t_{137} = 15.99$, $p_{FDR} = 9.5 \times 10^{-33}$, $r_e = 0.81$], and to $+2$ dB [$t_{137} = 10.96$, $p_{FDR} = 3 \times 10^{-20}$, $r_e = 0.68$], although performance was enhanced for modulated compared to unmodulated maskers at all SNRs [-6 dB: $t_{137} = 9.19$, $p_{FDR} = 7.2 \times 10^{-16}$, $r_e = 0.62$][-2 dB: $t_{137} = 2.49$, $p_{FDR} = 0.014$, $r_e = 0.21$][$+2$ dB: $t_{137} = 18.53$, $p_{FDR} = 2 \times 10^{-38}$, $r_e = 0.85$]. The modulation type \times SNR \times age group interaction was not significant [$p = 0.695$].

Next, our analysis focused on the effects of masker envelope shape (damped, ramped) on speech intelligibility. Average word report declined with decreasing SNR [effect of SNR: $F_{2,272} = 218.43$, $p_{GG} = 4.1 \times 10^{-42}$, $\eta^2 p = 0.62$; Fig. 4a]. We additionally observed a significant SNR \times age group interaction [$F_{2,272} = 8.004$, $p_{GG} = 0.002$, $\eta^2 p = 0.06$], but did not find any significant effects during follow-up comparisons [$p_{FDRs} > 0.07$].

Consistent with Experiment 1, word report was higher when the envelope shape of the masker was damped compared to ramped [effect of envelope shape: $F_{1,136} = 8.49$, $p = 0.004$, $\eta^2 p = 0.06$; Fig. 4c left panel], and for both age groups [younger: $t_{67} = 2.17$, $p_{FDR} = 0.045$, $r_e = 0.26$; older: $t_{69} = 2.04$, $p_{FDR} = 0.045$, $r_e = 0.24$; envelope shape \times age group interaction: $p = 0.702$; Fig. 4c]. Higher speech intelligibility for damped compared to ramped maskers

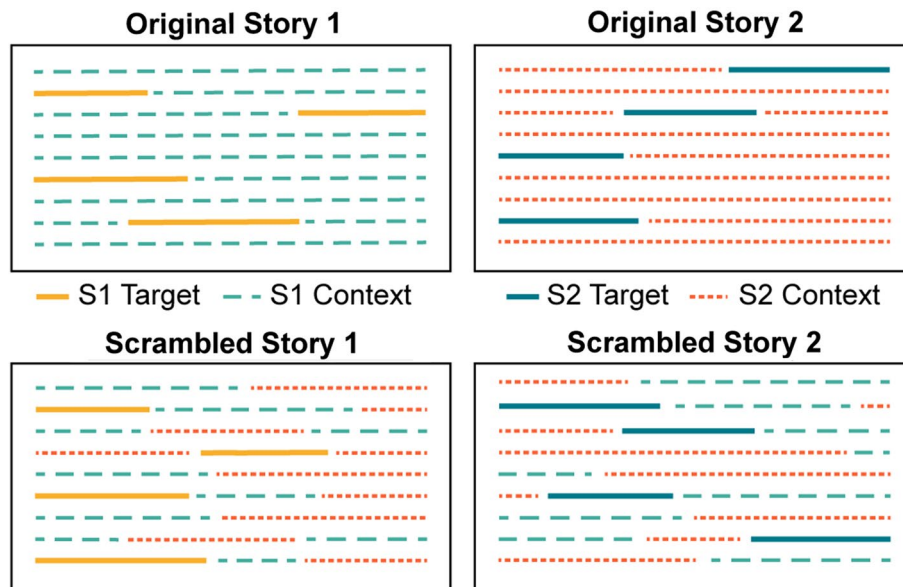


Figure 5. Stimulus schematic for Experiment 3. A schematic of the sentence structure of the original (top panel) and scrambled stories (bottom panel) is plotted. For the original stories, events were presented in their original order, whereas the scrambled stories contained the target phrases from one original story and a randomized mixture of the context sentences from both original stories. Target phrases/sentences used for intelligibility testing are plotted as solid lines (4 test sentences shown per story for simplicity), and context sentences are plotted as dashed lines. Note that the same target sentences are used in the original and scrambled stories.

was mainly driven by the most challenging SNR [6 dB: $t_{137} = 3.5$, $p_{FDR} = 0.002$, $r_c = 0.29$; envelope shape \times SNR: $F_{2,272} = 7.91$, $p_{GG} = 0.001$, $\eta^2 p = 0.06$; Fig. 4a], and not -2 dB [$p_{FDR} = 0.304$], or $+2$ dB [$p_{FDR} = 0.304$]. There were no other significant effects or interactions [$F < 2$, $p > 0.16$, $\eta^2 p < 0.01$].

Experiment 2 yielded two important findings. First, using engaging spoken stories, we show that older adults experience a larger speech intelligibility benefit from a modulated relative to an unmodulated masker compared to younger adults. This is in stark contrast to the results in Experiment 1 and the previous literature using short, disconnected sentences, which show a reduced intelligibility benefit in the presence of amplitude modulation for older compared to younger listeners^{16,18–20,23,24,28,31,32,94}. Second, both older and younger participants exhibited better intelligibility when the babble masker was modulated with a damped compared to ramped envelope, partially replicating the results of Experiment 1, in which a benefit was seen for older, but not younger, adults. The shape of the modulated masker thus does not appear to strongly interact with age or the type of speech materials used during testing.

Experiments 1 and 2 differed substantially in speech materials and task procedure. In Experiment 3, we examine the effect of stimulus material and masker envelope on speech intelligibility. To ensure that narratives and isolated sentences are closely matched, we use target phrases/sentences either embedded in coherent stories or decontextualized in “scrambled” stories for which story sentences are shuffled in time. We use identical test phrases/sentences between the coherent and scrambled stories. As a result, we can more clearly determine whether removing the narrative arc of the story systematically alters the effects of age and masker envelope on speech intelligibility.

Experiment 3: speech-intelligibility benefit for amplitude-modulated maskers depends on the speech materials in older adults. Two 10-min stories (*Wave*, by D.M. Ouellet and *Alibi*, by Kristin Butcher) were selected and recorded for use in Experiment 3. These stories were written to be highly engaging but without complex language so that readers of any level may understand and enjoy the content. Two types of each story were created: original and scrambled. Original stories presented story events in the original order. Target phrases/sentences from the original stories were identified for intelligibility testing, as in Experiment 2 (Fig. 5, top panel). A scrambled story contained the same target phrases/sentences as one of the original stories and a randomized mixture of other (context) sentences drawn from both stories (Fig. 5, bottom panel). Scrambled stories thus lack a narrative thread, but are generated such that the test phrase/sentences used for intelligibility testing are identical across both story types.

Each story was masked by 12-talker babble noise, and the signal-to-noise ratio (-6 , -2 , $+2$ dB, Clear) and amplitude envelope (unmodulated, 4-Hz modulated: damped, ramped) were manipulated. Younger (mean: 31.3 years; age-range: 21–38 years; 79 males 44 females) and older (mean: 63.2 years; age-range: 54–77 years; 44 males 77 females) adults without reported hearing or neurological issues listened to one of the four possible stories (2 original, 2 scrambled) and reported back sentences/phrases using the same online testing procedure as in Experiment 2 (Fig. 3). We calculated the proportion of correctly reported words for each story type

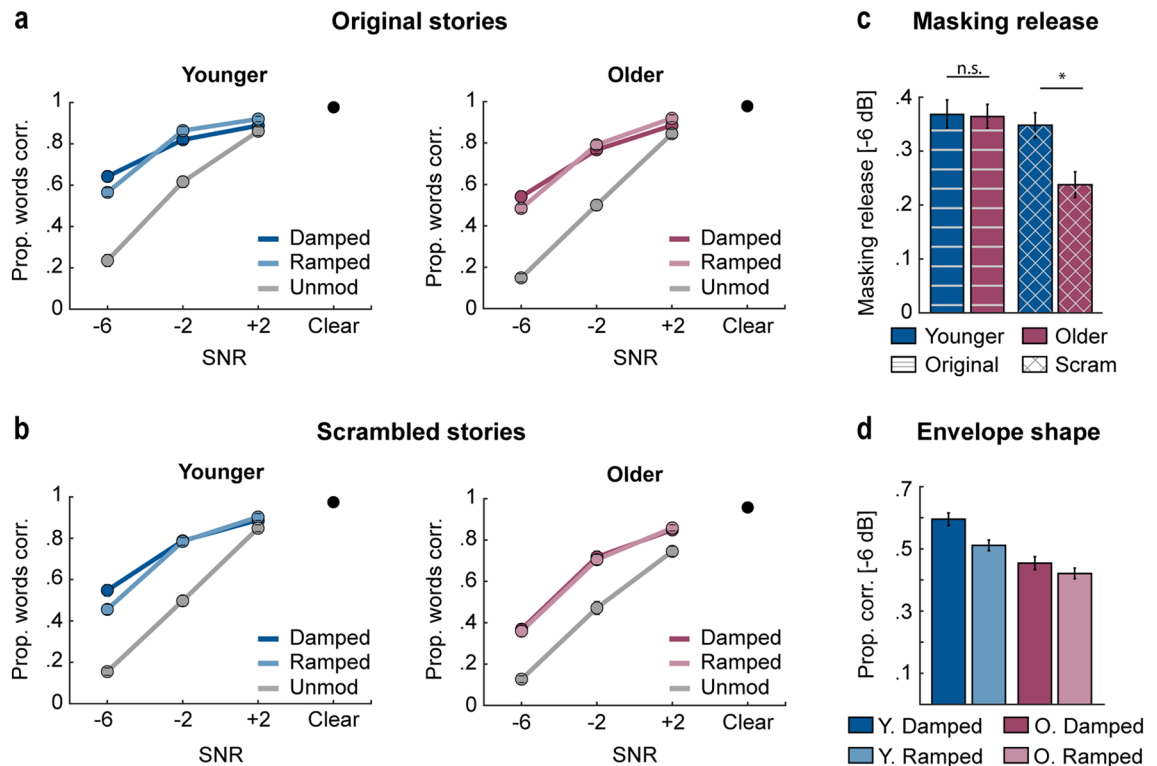


Figure 6. Intelligibility results for Experiment 3. **(a)** Mean proportion of correctly reported words for original stories is plotted as a function of SNR (-6 , -2 , $+2$ dB, Clear [black dot]) for the different envelope conditions (damped, ramped, unmodulated) and for younger (left) and older adults (right). **(b)** Mean performance for scrambled stories is plotted as a function of SNR (-6 , -2 , $+2$ dB, Clear [black dot]) for the different envelope conditions (damped, ramped, unmodulated) and for younger (left) and older adults (right). **(c)** Mean difference in intelligibility between modulated and unmodulated maskers (masking release) at -6 dB SNR is plotted for each story type (original, scrambled) and age group (younger, older). **(d)** Mean performance at -6 dB SNR is plotted for the different masker envelope shapes (damped, ramped) and age groups. Error bars reflect the standard error of the mean. * $p < 0.05$.

(original, scrambled), envelope condition (damped, ramped, unmodulated) and SNR (-6 , -2 , $+2$ dB, Clear) and compared the result between age groups.

Consistent with Experiment 2, average word report declined with decreasing SNR [effect of SNR: $F_{2,480} = 1665.46$, $p = 1.1 \times 10^{-216}$, $\eta^2 p = 0.87$; Fig. 6a,b]. Intelligibility was higher for original stories relative to scrambled [effect of story type: $F_{1,240} = 20.12$, $p = 1.1 \times 10^{-5}$, $\eta^2 p = 0.08$], and higher for younger than older adults [effect of age group: $F_{1,240} = 19.83$, $p = 1.3 \times 10^{-5}$, $\eta^2 p = 0.08$].

Speech intelligibility was also higher for modulated relative to unmodulated maskers [effect of modulation type: $F_{1,240} = 999.81$, $p = 2 \times 10^{-87}$, $\eta^2 p = 0.81$; release-from-masking effect]. The modulation type \times story type [$p = 0.404$], modulation type \times age group [$p = 0.698$], and the modulation type \times story type \times age group [$p = 0.051$] interactions were not significant. All remaining 2- and 3-way interactions were significant [$p_{GG} < 0.05$]. However, because the 4-way interaction was also significant [modulation type \times SNR \times story type \times age group: $F_{2,480} = 3.98$, $p_{GG} = 0.021$, $\eta^2 p = 0.02$], we analyze this 4-way interaction and do not discuss the 2- and 3-way interactions any further.

To explore the significant 4-way interaction, we first calculated difference scores between average intelligibility scores for modulated and unmodulated trials (masking release) for each participant. Using post-hoc t-tests, we examined the effect of age group and story type on masking release at each SNR level. This revealed that the 4-way interaction was driven by group differences at -6 dB SNR. At this challenging SNR, masking release for scrambled stories was larger for younger compared older adults [$t_{120} = 3.3$, $p_{FDR} = 0.008$, $r_c = 0.29$, Fig. 6c], whereas older adults benefited as much as younger adults from a modulated relative to an unmodulated masker for original stories [$p_{FDR} = 0.91$]. No differences were observed as a function of age group and story type at -2 dB SNR or $+2$ dB SNR [$p_{FDR} > 0.06$].

One potential explanation of this finding is that the reduced release from masking for older adults was simply due to the poor signal quality at -6 dB leading to fewer intelligible words, and thus, less available context specifically for the older subject group. However, this seems unlikely since performance in the unmodulated condition for scrambled stories at -6 dB SNR was not different between younger and older listeners [avg. words reported: younger: 16%; older: 13%; $p_{FDR} > 0.4$; Fig. 6b left panel vs right panel]. It is therefore unlikely that the reduced release from masking exhibited by older individuals for scrambled stories is due to less available context as a result of lower intelligibility for this condition. Furthermore, within the older group, performance in the unmodulated condition at -6 dB SNR did not differ between scrambled and original stories [$p_{FDR} > 0.4$ Fig. 6a

right panel vs 6b right panel]; therefore, the presence of context in the original stories is not solely driving the increased masking release for older adults, as such an effect should lead to better performance for both modulated and unmodulated conditions when listening to original stories. We tentatively conclude that the presence of meaningful context and perhaps the engagement that it fosters is qualitatively changing the older adults' ability to benefit from masker modulation.

Next, we investigated whether the temporal profile of the modulated masker (damped vs. ramped) affects speech intelligibility for different story types and age groups. As expected, intelligibility declined with decreasing SNR [$F_{2,480} = 1041.51$, $p_{GG} = 1.2 \times 10^{-168}$, $\eta^2p = 0.81$; Fig. 6a,b]; and was higher for original compared to scrambled stories [$F_{1,240} = 22.31$, $p = 4 \times 10^{-6}$, $\eta^2p = 0.09$]. The difference between original and scrambled stories was largest when the SNR was most challenging [-6 dB: 0.13; -2 dB: 0.06; +2 dB: 0.03] [SNR \times story type interaction: $F_{2,480} = 15.12$, $p_{GG} = 7 \times 10^{-7}$, $\eta^2p = 0.06$]. Intelligibility was also higher for younger compared to older adults [$F_{1,240} = 20.16$, $p = 1.1 \times 10^{-5}$, $\eta^2p = 0.08$]. The difference between older and younger adults was primarily observed at -6 dB [$t_{242} = 4.92$, $p_{FDR} = 5 \times 10^{-6}$, $r_e = 0.3$] and -2 dB [$t_{242} = 3.53$, $p_{FDR} = 0.0007$, $r_e = 0.22$], but not +2 dB [$p_{FDR} = 0.08$] [SNR \times age group interaction: $F_{2,480} = 13.64$, $p_{GG} = 3 \times 10^{-6}$, $\eta^2p = 0.05$].

The interaction between envelope shape \times SNR was significant [$F_{2,480} = 13.37$, $p_{GG} = 5 \times 10^{-6}$, $\eta^2p = 0.05$]. Follow-up t-tests revealed that, at -6 dB, target phrases/sentences were more intelligible when the masker was damped compared to ramped [$t_{243} = 4.46$, $p_{FDR} = 4 \times 10^{-5}$, $r_e = 0.28$]. At -2 dB there was no effect of envelope shape [$p_{FDR} = 0.276$], while at +2 dB [$t_{243} = -2.26$, $p_{FDR} = 0.04$, $r_e = 0.14$] target phrases were more intelligible if the masker envelope was ramped, compared to damped. No other effects or interactions were significant [$F < 2.5$, $p > 0.14$, $\eta^2p < 0.008$].

To summarize Experiment 3, we demonstrate that the effect of age on the benefit of a fluctuating, compared to steady-state, masker critically depends on the speech material. Older adults experience less masking release than younger adults when listening to a stream of randomized sentences (scrambled story) that are similar to those commonly used in experimental aging research (cf. Experiment 1; see also^{16,18–21,23,24,27–31,35,95}). However, for continuous speech with a topical thread, older adults benefit as much from a fluctuating masker as younger adults. These results suggest that research using disconnected sentences may systematically underestimate the speech-listening capabilities of older adults. Additionally, intelligibility is generally better when the babble masker had a damped compared to ramped temporal profile, effectively replicating Experiments 1 and 2.

General discussion

In the current study, we investigated how intelligibility of masked speech in younger and older adults is affected by the nature of speech materials (isolated sentences vs engaging stories) and by the temporal profile of the masker's amplitude envelope. We asked two specific questions: (1) Does the known age-related reduction in the speech-intelligibility benefit for modulated compared to unmodulated maskers depend on the nature of the speech materials? (2) Does speech intelligibility differ for modulated maskers with different temporal profiles (damped: sharp attack and gradual decay; ramped: gradual attack and fast decay), and does this differ between younger and older adults? We observed a reduced speech-intelligibility benefit for modulated over unmodulated background maskers in older relative to younger adults when individuals listened to disconnected sentences (Experiments 1 and 3). In marked contrast, older adults benefited more than younger adults (Experiment 2) or equally as much (Experiment 3) when they listened to engaging stories. We also generally observed better speech intelligibility for maskers with damped compared to ramped envelope shapes, suggesting that temporal profiles characterized by fast onsets and slow offsets may benefit intelligibility similarly across age groups and speech materials. Our results suggest that the well-documented deficit in 'dip listening' in older adults^{16,18–21,23,24,27–31,35,95} can be mitigated if the speech materials are engaging and contextually rich. Standard laboratory listening paradigms, utilizing disconnected sentences, elicit listening behavior that is qualitatively different from that observed when richer, continuous speech stimuli are used.

Damped maskers interfere less with speech intelligibility than ramped maskers. The current study investigated whether the envelope shape of the masker (damped vs. ramped) influences the intelligibility of target speech. This research question was motivated by recent electrophysiological work in rodents and human participants^{13,37–39}. Neural activity appears to synchronize more strongly with ramped than damped envelopes in younger people, and with damped compared to ramped envelopes in older people^{13,37}. Furthermore, increased neural synchronization to a sound with a low-frequency amplitude modulation (e.g., ~4 Hz)^{40,96,97} may specifically predict declines in speech intelligibility when masked by a modulated background sound^{38,39}. Based on these electrophysiological studies, we expected to observe reduced speech intelligibility for damped envelope shapes in older adults, and reduced intelligibility for ramped envelope shapes in younger adults. In contrast to our predictions, we generally observed better speech intelligibility for maskers with damped compared to ramped envelope shapes in both age groups, particularly when the SNR was low. Further, we did not find evidence that the effect of envelope shape differs between disconnected sentences and engaging stories (Experiment 3; Fig. 6). However, while the predictable masker rate of 4 Hz was motivated by electrophysiological work, we recognize that real-world listening situations do not typically have background maskers with predictable envelopes. Future studies could include a more ecologically valid manipulation of the amplitude envelope, such as using the temporal envelope of natural speech with salient ramped and damped envelopes by virtue of using words with those specific envelope shapes.

Release from masking is not reduced in older compared to younger adults for engaging stories. Previous research indicates that aging is associated with a decline in processing temporal sound features^{7–10,98–101}, and that temporal processing deficits may contribute to older adults experiencing difficulty

understanding speech when background noise is present^{11,12,17,100}. The persistent finding that older adults demonstrate either no benefit or a reduced speech-intelligibility benefit from a fluctuating relative to a flat envelope background masking sound^{16,19,24,27,28} has long been discussed as a prime example of temporal deficits limiting the ability of older adults to ‘glimpse’ target speech. In Experiments 1 and 3, we replicated previous findings that older adults benefit less from speech ‘glimpses’ compared to younger adults (Figs. 2b and 6c).

Critically, we also demonstrate that the ability to benefit from speech ‘glimpsing’ is only reduced in older adults when speech materials lack an overarching and engaging narrative context. When listening to engaging spoken stories, older adults demonstrated similar (Experiment 3; Fig. 6c) or even greater (Experiment 2; Fig. 4b) masking release compared to younger adults. Further, the interaction between age and speech material does not appear to have been driven by a ceiling effect in the younger participant group because we observed it at the most difficult SNR (– 6 dB) where performance was markedly lower than ceiling (Fig. 6). Our experiments demonstrate that the reduction in ‘speech glimpsing’ previously observed in older people may be specific to the speech materials commonly used in research studies, and may not generalize to listening situations with rich narrative structure.

Researchers have long concluded that the lack of benefit from speech ‘glimpses’ in older compared to younger individuals is due to increased spectrotemporal overlap between the target and masking signals in the auditory periphery (“energetic masking”), as a result of age-related hearing loss. Despite self-reports indicating the absence of hearing issues, our supplementary analysis (see Supplemental Document) indicates that our older adult sample may have slightly elevated hearing thresholds compared to younger adults (as would be expected). Elevated thresholds should be associated with reduced speech intelligibility and reduced release from masking overall, regardless of the sound type. Instead, our results suggest that the lack of benefit from speech ‘glimpses’ for isolated sentences might be related to other, perhaps more cognitive factors.

Several factors potentially contribute to the observed interaction between age and the type of materials (sentences vs stories) on release from masking. One critical difference is the higher degree of semantic context present in stories compared to disconnected sentences. Semantic context is well known to facilitate comprehension of words in disconnected sentences masked with noise for both older and younger adults^{17,59–61,102–105} and can alleviate listening effort for individuals with hearing impairment¹⁰⁶. Moreover, spoken stories, such as the ones used in the current study, have an overarching topical thread that engages listeners¹⁴, and encourages them to continuously generate, update, and integrate story events and characters into a mental model that supports ongoing attention to the story^{107–109}. This may recruit additional cognitive processes, compared to those recruited to understand isolated, unrelated sentences. The overarching narrative provides additional topic context that may enhance intelligibility, enabling participants to fill in missing information that was lost due to low SNR. Yet, context effects are unlikely to solely account for the older adults’ recovery of release from masking when listening to original stories. If this were due entirely to context effects, the added context of the original over scrambled stories should have led to better performance for the unmodulated original compared to scrambled story, and this was not observed.

Engaging spoken stories and disconnected sentences may also elicit different levels of motivation to listen. Motivation is crucial for the recruitment of cognitive resources during challenging tasks. A person will only invest cognitively if the activity is expected to be rewarding relative to the anticipated mental costs^{64–66,110}. ‘Reward’ can take many forms and can be either extrinsic; for example, monetary rewards¹¹¹ or intrinsic; through enjoyment and interest^{71,112}. Spoken stories of the kind used here have been shown to be highly enjoyable and absorbing¹⁴ and elicit synchronized brain activity across listeners⁹⁰, indicating their highly engaging nature. A recent study reported that listeners find stories as enjoyable and absorbing when they are masked by moderate background noise as when they are heard clearly, despite missing some words and finding listening more effortful in the former condition¹⁴. We speculate that older adults in the current study may have benefited from a modulated masker during story listening as much as younger adults because they enjoyed the story content, and were intrinsically motivated to invest additional cognitive resources to listen. We did not implement a measure of motivation or enjoyment following story listening, so it is not possible to relate motivation/enjoyment directly to intelligibility here. However, this interpretation is consistent with previous observations that older adults tend to engage less when tasks are less personally meaningful to them, perhaps in order to conserve mental resources^{84,85}. Our results certainly point to large qualitative differences in listening behaviors for engaging spoken stories, compared to the disconnected sentence materials that are typically used in clinical and laboratory settings. We suggest that typical research approaches with disconnected sentences may underestimate the speech-listening abilities of older adults, especially in listening situations with narrated elements.

Conclusions

Speech masked by a background sound with fluctuating amplitude is typically better understood than speech masked by sound with a relatively steady amplitude, but older adults have frequently been shown to benefit less from fluctuating maskers. This apparent deficit in the ability of older adults to ‘listen in the dips’ has been taken as a prime example of decreased temporal processing or reduced audibility in older individuals. Yet, speech intelligibility and masking release are typically investigated using short, disconnected sentences. Our results show that the release from masking depends on whether listeners are attending to disconnected sentences or to an engaging, connected, narrative. We replicated previous work showing a deficit in the speech-intelligibility benefit from amplitude fluctuations in older adults when they listened to disconnected sentences (Experiments 1 and 3). However, we further show that older adults either benefit more (Experiment 2) or similarly (Experiment 3) from modulated maskers relative to younger adults when listening to engaging spoken stories that follow a topical thread. Maskers with a damped temporal profile generally facilitated intelligibility and did not reliably interact with age or the type of speech material. Taken together, our data suggest that reduced ‘dip listening’ previously

observed in older adults does not appear to generalize to engaging spoken stories. This result highlights that at least some deficits considered to be audiological may be more related to cognitive or motivational factors, and that the nature of the listening materials qualitatively changes listening behavior. Standard laboratory listening paradigms using disconnected sentences may underestimate the speech abilities of older adults.

Materials and methods

Experiment 1. Participants. One hundred and thirty-seven individuals (mean: 47.7 years; age-range: 18–71 years; 66 males 70 females 1 non-binary) without self-reported hearing loss, neurological issues, or psychiatric disorders participated in Experiment 1. Participants below age 50 were considered part of the ‘younger’ group (mean: 35.9 years; age-range: 18–49 years; 39 males, 29 females, 1 non-binary) and the remaining participants aged 50 and older were considered part of the ‘older’ group (mean: 59.6 years; age-range: 50–71 years; 31 males, 37 females). Participants were recruited from the Amazon Mechanical Turk online participant pool (MTurk; <https://www.mturk.com/>) via the participant sourcing platform Cloud Research (previously TurkPrime⁸⁶). All participants provided informed consent prior to participation. The study was conducted in accordance with the Declaration of Helsinki, the Canadian Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans (TCPS2-2018), and approved by the local Nonmedical Research Ethics Board of the University of Western Ontario (REB #112574).

Each individual received financial compensation of \$5 USD following completion of the study (\$10 hourly rate). Twenty-seven additional individuals participated in the study but were not included either due to reporting a technical error during data recording ($N=9$), hearing aid usage or neurological issues ($N=7$), not wearing headphones ($N=2$), submitting the same one-word answers to all questions ($N=5$), or scoring at floor ($\sim 10\%$) for all levels of background noise in the intelligibility task ($N=4$). Online research can be subject to increased levels of random responders, since experimenters have less control over the testing environment compared to a laboratory setting. However, online studies have generally been shown to replicate findings of in-person data collection^{113–118} (see also Supplemental Document for the results of an in-lab pilot of Experiment 1), particularly if controls are in place to ensure compliance with study instructions.

Acoustic stimulation and procedure. All target sentences ($N=84$) were spoken by the same female talker and ranged between 8 and 10 words in length (range of durations: 1.95–3.43 s). 12-talker babble noise from the Revised Speech in Noise test (R-SPIN)¹¹⁹ was added as a masker. Babble noise was either unmodulated (flat amplitude envelope) or amplitude modulated at a rate of 4 Hz with a damped (sharp attack and gradual decay) or ramped (gradual attack and sharp decay) envelope shape (Fig. 1). The modulation frequency of 4 Hz was chosen as it falls within the range of the low-frequency speech envelope^{36,120} and for consistency with previous electrophysiology work investigating how aging affects neural synchronization to the amplitude envelope^{13,39,40,96}. Envelope shape was manipulated by varying parameters of the following equation:

$$b = t^{z-1}(1 - t) \quad (1)$$

where t is a time vector representing one cycle (0.250 s), z determines the envelope shape, and b is the resulting function used to modulate the noise. A z parameter of 2 generates a symmetrical envelope shape, while a value closer to 1 generates an envelope with a damped shape (sharp attack and gradual decay). Varying the z parameter also impacts the sharpness and half-life of each cycle. We used a z parameter of 1.15 to generate damped envelopes, each with a sharp onset and a 168.4 ms half-life^{13,37}. Ramped envelopes (gradual attack and sharp decay) were created by mirroring the vector b (Fig. 1).

The signal-to-noise ratio (SNR) between the speech signal and the background babble was manipulated by adjusting the level of the sentence (target) relative to the babble masker (SNR levels: -10 , -8 , -6 , -4 , -2 , 0 , $+2$ dB). There were 21 possible stimulus conditions (7 SNRs $\times 3$ envelope conditions = 21 stimulus conditions) that were tested in each block of trials (21 envelope conditions $\times 4$ blocks = 84 total trials). To ensure intelligibility results were not confounded by specific sentences, 21 counterbalanced versions were generated, such that each sentence was heard with every SNR and envelope combination across versions. All sentence/babble mixtures were normalized relative to the same root-mean square amplitude (RMS).

The experiment was conducted online, using custom written JavaScript/html and jsPsych code (Version 6.1.0, a high-level JavaScript library used for precise stimulus control¹²¹). The experiment code was stored at an online repository (<https://gitlab.pavlovia.org>) and hosted via Pavlovia (<https://pavlovia.org/>). A test version was randomly assigned to each participant when data files were loaded into the internet browser. Prior to the main experimental procedures, participants were instructed to use headphones and complete the tasks in a quiet room free from distractions. We did not provide specifications as to the type/brand of equipment participants should use (e.g., computer, headphone type), but took steps to ensure participants complied with the instruction to use headphones (see “Online research quality assurance measures”).

During the main task (intelligibility task), participants were instructed to listen to each sentence and, after the sentence ended, type the words that they heard into a text box. Participants had unlimited time to type each response. Once participants submitted an answer, the next sentence would begin following a brief inter-trial silent interval of 0.25 s. Participants had the opportunity to take a break after each experimental block. The total duration of the intelligibility test was therefore dependent on the typing speed and total break length for each individual, but the intelligibility test duration typically ranged between 20 to 25 min.

Online research quality assurance measures. Participants completed two initial listening tasks at the beginning of the online session. First, participants listened to a 15-s stream of pink noise normalized to the same RMS

amplitude as the sentences and were instructed to adjust their volume to a comfortable listening level. Participants had the option to replay the noise if they needed additional time to adjust their volume. This task ensured that participants could adjust their volume to a comfortable level prior to the intelligibility task, after which they were instructed to not make further adjustments.

Wearing headphones during the main experiment (intelligibility task) is an important condition of participation, since it can limit the influence of nearby distractions and help preserve stimulus characteristics, such as signal-to-noise ratio. In addition to explicitly asking participants whether they complied with instructions to wear headphones, participants also completed a headphone-check procedure (second listening task) to determine whether they were wearing headphones¹²². During the headphone-check procedure, participants performed a tone discrimination task (6 trials; ~ 2 min total duration), in which they determined which of three consecutive 200-Hz sine tones was the quietest. The three tones differed such that one was presented at the comfortable listening level, one at - 6 dB relative to the other two tones, and one at the comfortable listening level with a 180° phase difference between the left and right headphone channels (anti-phase tone). This task is straightforward over headphones, but difficult over loudspeakers, because the pressure waves generated from an anti-phase tone interfere¹²². If they were listening through loudspeakers, they would likely erroneously select the anti-phase tone as the quietest tone. This task provides another metric (in addition to self-report of headphone use), that could be used to flag participants who may not have been complying with instructions. No participants were excluded solely on the basis of performance on this test. Participants were excluded if they explicitly reported not wearing headphones during the task (N = 2).

Statistical analysis. Statistical analyses were conducted using IBM SPSS Statistics (version 27) for Windows and MATLAB (version 2018a). Details of the specific variables and statistical tests can be found in analysis subsections for each measure. The false-positive rate for multiple comparisons was controlled using false discovery rate (FDR)¹²³. FDR corrected p-values are reported as p_{FDR} . Effect sizes are reported as partial eta squared (η^2p) for rmANOVAs and $r_{\text{equivalent}}$ (r_e)¹²⁴, for t-tests. Greenhouse–Geisser corrected p-values are reported when sphericity assumptions have not been met (reported as p_{GG}). This experiment was not preregistered. Data are available at the project website on the Open Science Framework (OSF: <https://osf.io/swy57/>). All figures were generated by the authors using MATLAB and Adobe Illustrator (version 2019).

Assessment of intelligibility. We calculated the proportion of correctly reported words for each SNR (- 10, - 8, - 6, - 4, - 2, 0, + 2 dB) and envelope condition (unmodulated, damped, ramped). Different or omitted words were counted as errors, but minor misspellings and incorrect grammatical number (singular vs. plural) were not. A logistic function was fit to the proportion of correctly reported words using the following equation:

$$y = \frac{K}{(1 + e^{-r(x-x_0)})} \quad (2)$$

where K sets the curves maximum value, r is the slope, x_0 is the inflection point or the speech reception threshold associated with 50% proportion of correct words, and x refers to the SNR values (- 10, - 8, - 6, - 4, - 2, 0, + 2 dB). We analyzed two parameters from each fit, the threshold and slope.

To examine differences in masking release as a function of age, we calculated the threshold and slope from the logistic function fit, separately for modulated (averaged across damped and ramped) and unmodulated trials. Threshold and slope were analyzed in separate mixed design repeated-measures analyses of variance (rmANOVAs) with modulation type (modulated, unmodulated) as a within-subject factor and age group (younger, older) as a between-subjects factor.

To analyze differences in speech intelligibility due to envelope shape (damped, ramped), thresholds and slopes from the logistic function fits were analyzed in separate rmANOVAs with envelope shape (damped, ramped) as a within-subjects factor and age group (younger, older) as a between-subjects factor.

Experiment 2. Participants. One hundred and thirty-eight younger (mean: 30.1 years; age-range: 19–39 years; 37 males, 30 females) and older individuals (mean: 64.4 years; age-range: 53–80 years; 29 males, 41 females) without self-reported hearing loss, neurological issues, or psychiatric disorders participated in Experiment 2. All participants were recruited using identical procedures to Experiment 1, except that individuals who participated in Experiment 1 were precluded from participating in Experiment 2. Each individual received financial compensation of \$6 USD following completion of the study (\$10 hourly rate). Twenty-three additional individuals participated in the study but were not included either due to reporting a technical error during data recording (N = 6), hearing aid usage or neurological issues (N = 5), not wearing headphones (N = 4), identifying as a non-native English speaker (N = 2), or scoring ~ 50% or below on the intelligibility task when there was no masker (i.e., during clear speech; N = 6), suggesting participants were not attending during the task.

Acoustic stimulation and procedure. Acoustic stimulation and task procedures were adapted from a task developed previously⁹⁰. One story (male talker) from “The Moth” story-telling podcast (<https://themoth.org>) was used as the target speech (*Reach for the Stars One Small Step at a Time*; by Richard Garriott, ~ 13 min). The target story had 12-talker babble noise added as a masker (R-SPIN)¹¹⁹. Babble noise could either be unmodulated (flat amplitude envelope) or amplitude modulated at a rate of 4 Hz with a ramped (gradual rise and sharp fall) or damped (sharp rise and gradual fall) envelope shape. Envelope shape was altered using identical parameters to Experiment 1 [cf. Equation (1)]. The signal-to-noise ratio (SNR) was manipulated by adjusting the dB level of both the story and masker. There were 3 possible envelope conditions (unmodulated, damped, ramped) and 3

different SNRs (-6 , -2 , $+2$ dB SNR) along with a condition in which no masker was heard (clear), resulting in 10 total possible stimulus conditions (3 envelopes \times 3 SNRs + clear = 10 conditions). Stimulus condition was pseudo-randomly varied approximately every 16 s (see Fig. 3) throughout each story. The length of the 16-s time window was determined by dividing the total duration of the story (in seconds) by the total number of trials. Each of the 10 stimulus conditions (3 envelopes \times 3 SNRs + clear) were heard a total of 5 times over the course of the story (16 s \times 10 conditions \times 5 repetitions = \sim 13 min). Three versions of condition order were generated to ensure that specific parts of the story were not confounded with a particular SNR and envelope combination. Within each version, SNR and envelope shape was varied pseudo-randomly such that a particular combination of SNR and envelope shape could not be heard twice in succession.

Phrases/sentences ranging from 4 to 7 words (range of durations: 0.85–2.6 s) were selected from the target story for intelligibility testing. These test phrases/sentences did not occur during the transition period from one SNR to the next (for approximately 1-s before and after the SNR transition). Two phrases/sentences per 16-s segment were selected, resulting in 100 possible test phrases for the target story (10 conditions \times 5 repetitions \times 2 phrases/sentences). One of the two selected phrases/sentences per 16-s segment was assigned to one intelligibility test set, whereas the other selected phrase/sentence was assigned to a second intelligibility test set (50 phrases/sentences per set). Having two test sets ensured that any observed intelligibility effects were not confounded by item (specific phrases/sentences) effects.

The experiment was conducted online using custom written JavaScript/html and jsPsych code hosted via Pavlovia (<https://pavlovia.org/>). During the main experiment, each participant listened to the target story and completed the intelligibility task. The condition order and intelligibility test set were randomly assigned to participants at the beginning of the experiment. Participants were instructed to use headphones and complete the tasks in a quiet room free from distractions. During story listening, a black fixation cross was presented at the center of the screen throughout the story. The fixation cross turned yellow two seconds prior to the beginning of a test phrase/sentence, cueing the participant to prepare for intelligibility testing (see Fig. 3). The fixation cross then turned green for the duration of the test phrase in the story, indicating to the participant the phrase they would be asked to report back. The story stopped with the offset of the test phrase, and an input text box appeared on the screen. Participants were asked to type their answer into the text box (no time limit), after which the story resumed from the beginning of the sentence most recently heard (allowing for story continuation). The total duration of the intelligibility task ranged between 25 to 30 min.

In order to familiarize participants with the intelligibility task, a brief practice block was presented prior to the main experiment. Participants heard a \sim 3-min story (a shortened version of *A Shoulder Bag to Cry On* by Laura Zimmerman), without added babble noise, and performed 12 trials of the intelligibility task (2 trials per 30-s segment, practice duration: \sim 5 min).

Online research quality assurance measures. Participants completed two initial listening tasks at the very beginning of the online session, as in Experiment 1. These preliminary tasks were meant to give the participant an opportunity to adjust their volume to a comfortable listening level and to provide a metric, aside from self-report, which could flag participants who may not be complying with instructions to wear headphones (headphone check). No participants were excluded solely on the basis of performance on this test, but were automatically excluded if they explicitly reported not wearing headphones during the task ($N=4$). These tasks are described in Experiment 1.

Assessment of intelligibility. We calculated the proportion of correctly reported words for each envelope condition (damped, ramped, unmodulated) and SNR (-6 , -2 , $+2$ dB, Clear) across the three versions of the target story. Different or omitted words were counted as errors, but minor misspellings, and incorrect grammatical number (singular vs. plural) were not. Contractions were also accepted as correct when the target contained the written out form of the contraction.

To analyze differences in masking release between age groups, mean performance for modulated (averaged across damped and ramped) and unmodulated trials were calculated and submitted to an rmANOVA with modulation type (modulated, unmodulated) and SNR (-6 , -2 , $+2$ dB) as within-subject factors and age group (younger, older) as the between-subjects factor.

To examine the effect of envelope shape (damped, ramped) mean performance for damped and ramped trials were calculated and submitted to an rmANOVA with envelope shape (damped, ramped) and SNR (-6 , -2 , $+2$ dB) as within-subject factors and age group (younger, older) as a between-subjects factor.

Experiment 3. Participants. Two hundred and forty-four younger (mean: 31.3 years; age-range: 21–38 years; 79 males 44 females) and older individuals (mean: 63.2 years; age-range: 54–77 years; 44 males 77 females) without self-reported hearing loss, neurological issues, or psychiatric disorders participated in Experiment 3. Note that a higher number of participants were recruited for Experiment 3 than Experiments 1 and 2, because of the additional experimental factor: speech material type. All participants were recruited using identical procedures to Experiment 1 and 2, except that individuals who participated in Experiment 1 or 2 were precluded from participating in Experiment 3. Each individual received financial compensation of \$5 USD following completion of the study (\$10 hourly rate). Thirty-seven additional individuals participated in the study but were not included either due to reporting a technical error during data recording ($N=15$), neurological issues ($N=7$), not wearing headphones ($N=9$), submitting random one-word answers to all questions ($N=3$), or scoring \sim 50% or below on the intelligibility task when there was no masker (i.e., for clear speech; $N=3$), suggesting participants were not attending during the task.

Acoustic stimulation and procedure. Stories were adapted from the content of two books (Story 1: *Wave*, by D.M. Ouellet; Story 2: *Alibi*, by Kristin Butcher) that were written to be engaging while avoiding complex language so that readers of any level may understand and enjoy the content. Shortened versions of the original stories were created and recorded by a female talker (duration of each story: ~ 10 min). Target phrases for the word-report task were identified in each of the two stories, as in Experiment 2 (see Fig. 5, top panel: solid lines). These phrases/sentences ranged from 4 to 7 words in length (range of durations: 0.66–2.05 s). Two phrases in each 15-s segment of the story were selected, resulting in 80 possible test phrases for story 1 and 80 possible test phrases for story 2. One of the two selected phrases per 15-s segment were assigned to one intelligibility test set, whereas the other selected phrases/sentences were assigned to a second intelligibility test set. This resulted in 4 total intelligibility test sets (2 per story), each comprising 40 test phrases/sentences. Having two intelligibility test sets for each story ensured that any observed effects were not confounded by the effects of specific word report items.

Half of the listeners performed the intelligibility task with the test phrases/sentences naturally embedded in the stories in their original, coherent form. The other half performed the intelligibility task with the test phrases/sentences embedded in “scrambled stories.” Four scrambled stories (one for each story and intelligibility test set: 2 stories × 2 intelligibility test sets) were created by embedding target phrases in a randomized mixture of other sentences drawn from both stories (see Fig. 5, bottom panel), such that an equal proportion of materials from each of the two stories entered each scrambled story version. The scrambled story condition therefore serves as an approximation of listening to disconnected sentences (cf. Experiment 1), since shuffling and intermixing the sentences limits any contextual relation between the embedded target phrases and the filler/contextual materials. In this design, each listener heard and reported sentences from only one of eight possible story conditions (2 stories × 2 intelligibility test sets × 2 story type [original, scrambled]), and we measure word-report performance on exactly the same material when it is presented in an engaging story versus decontextualized as disjointed sentences.

Each original and scrambled story was masked by 12-talker babble noise (R-SPIN)¹¹⁹. The SNR (– 6, – 2, + 2 dB, clear), and envelope condition (ramped, damped, unmodulated) varied pseudo-randomly as in Experiment 2, with the exception that the stimulus condition changed every 15 s (instead of the 16 s period used in Experiment 2), since the stories used here were shorter in duration. Each of the 10 stimulus conditions (3 envelopes × 3 SNRs + clear) were heard four times over the course of the story (15 s × 10 conditions × 4 repetitions = ~ 10 min). Three different stimulus condition orders were generated for each story to ensure that specific parts of a story were not confounded with a particular SNR and envelope combination. Within each version, SNR and envelope shape were varied pseudo-randomly such that a particular combination of SNR and envelope shape could not be heard twice in succession.

The experiment was conducted online using custom written JavaScript/html and jsPsych code hosted via Pavlovia (<https://pavlovia.org/>). Each participant was pseudo-randomly assigned to one of the 8 story conditions described (2 stories × 2 intelligibility test sets × story type [original, scrambled]) and to one of the three stimulus condition orders. Participants were instructed to use headphones and complete the tasks in a quiet room free from distractions. In the main experiment, the participant listened to a story and completed the same intelligibility task used in Experiment 2 (see Fig. 3). Participants had unlimited time to submit each response. The total duration of the intelligibility test ranged between 15 to 20 min. In order to familiarize participants with the intelligibility task, a brief practice block was presented prior to the main experiment. Participants heard a ~ 3-min story (a shortened version of *A Shoulder Bag to Cry On* by Laura Zimmerman), without added babble noise, and performed 12 trials of the intelligibility task (2 trials per 30-s segment, practice duration: ~ 5 min).

Online research quality assurance measures. As in Experiment 1 and 2, participants completed two initial listening tasks at the very beginning of the online session. These preliminary tasks were meant to give the participant an opportunity to adjust their volume to a comfortable listening level and to provide a metric, aside from self-report, which could flag participants who may not be complying with instructions to wear headphones (headphone check). No participants were excluded solely on the basis of performance on this test, but were automatically excluded if they explicitly reported not wearing headphones during the task (N = 9). Specific methods are described in Experiment 1.

Assessment of intelligibility. We calculated the proportion of correctly reported words for each envelope type (damped, ramped, unmodulated) and SNR condition (– 6, – 2, + 2 dB, Clear), separately for original and scrambled stories, and separately for each version of the word-report task for each story. Different or omitted words were counted as errors, but minor misspellings, and incorrect grammatical number (singular vs. plural) were not. Contractions were also accepted as correct when the target contained the written-out form of the contraction.

Effects of modulation type were tested using an ANOVA (within-subjects factors modulation type (modulated [averaged across ramped and damped], unmodulated) and SNR (– 6, – 2, + 2) and the between-subjects factors story type (original, scrambled) and age group (younger, older).

Effects of envelope shape were analyzed using an rmANOVA with the within-subjects factors envelope shape (damped, ramped) and SNR (– 6, – 2, + 2) and the between-subjects factors story type (unaltered, scrambled) and age group (younger, older).

Received: 8 October 2021; Accepted: 23 March 2022

Published online: 07 April 2022

References

- Drullman, R., Festen, J. M. & Plomp, R. Effect of temporal envelope smearing on speech reception. *J. Acoust. Soc. Am.* **95**, 1053–1064 (1994).
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J. & Ekelid, M. Speech recognition with primarily temporal cues. *Am. Assoc. Adv. Sci.* **270**, 303–304 (1995).
- van der Horst, R., Leeuw, A. R. & Dreschler, W. A. Importance of temporal-envelope cues in consonant recognition. *J. Acoust. Soc. Am.* **105**, 1801–1809 (1999).
- Festen, J. M. & Plomp, R. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* **88**, 1725–1736 (1990).
- Miller, G. A. & Licklider, J. C. R. The intelligibility of interrupted speech. *J. Acoust. Soc. Am.* **22**, 167–173 (1950).
- Cooke, M. A glimpsing model of speech perception in noise. *J. Acoust. Soc. Am.* **119**, 1562–1573 (2006).
- Gordon-Salant, S. & Fitzgibbons, P. J. Profile of auditory temporal processing in older listeners. *J. Speech Lang. Hear. Res.* **42**, 300–311 (1999).
- Ruggles, D., Bharadwaj, H. & Shinn-Cunningham, B. G. Why middle-aged listeners have trouble hearing in everyday settings. *Curr. Biol.* **22**, 1417–1422 (2012).
- Grose, J. H., Mamo, S. K., Buss, E. & Hall, J. W. Temporal processing deficits in middle age. *Am. J. Audiol.* **24**, 91–93 (2015).
- Bharadwaj, H. M., Verhulst, S., Shaheen, L., Charles Liberman, M. & Shinn-Cunningham, B. G. Cochlear neuropathy and the coding of supra-threshold sound. *Front. Syst. Neurosci.* **8**, 1–18 (2014).
- Frisina, D. R. & Frisina, R. D. Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hear. Res.* **106**, 95–104 (1997).
- Gordon-Salant, S. Speech perception and auditory temporal processing performance by older listeners: Implications for real-world communication. *Semin. Hear.* **27**, 264–268 (2006).
- Irsik, V. C., Almanaseer, A., Johnsrude, I. S. & Herrmann, B. Cortical responses to the amplitude envelopes of sounds change with age. *J. Neurosci.* **41**, 5045–5055 (2021).
- Herrmann, B. & Johnsrude, I. S. Absorption and enjoyment during listening to acoustically masked stories. *Trends Hear.* **24**, 1–18 (2020).
- Davis, M. H. & Johnsrude, I. S. Hierarchical processing in spoken language comprehension. *J. Neurosci.* **23**, 3423–3431 (2003).
- Dubno, J. R., Horwitz, A. R. & Ahlstrom, J. B. Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing. *J. Acoust. Soc. Am.* **111**, 2897–2907 (2002).
- Pichora-Fuller, M. K., Schneider, B. A. & Daneman, M. How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* **97**, 593–608 (1995).
- Stuart, A. & Phillips, D. P. Deficits in auditory temporal resolution revealed by a comparison of word recognition under interrupted and continuous noise masking. *Semin. Speech Lang.* **19**, 333–343 (1998).
- Summers, V. & Molis, M. R. Speech recognition in fluctuating and continuous maskers: Effects of hearing loss and presentation level. *J. Speech Lang. Hear. Res.* **47**, 245–256 (2004).
- Turner, C. W., Souza, P. E. & Forget, L. N. Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners. *J. Acoust. Soc. Am.* **97**, 2568–2576 (1995).
- Moore, B. C. J. The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J. Assoc. Res. Otolaryngol.* **9**, 399–406 (2008).
- Vestergaard, M. D., Fyson, N. R. C. & Patterson, R. D. The mutual roles of temporal glimpsing and vocal characteristics in cocktail-party listening. *J. Acoust. Soc. Am.* **130**, 429–439 (2011).
- Gustafsson, H. A. & Arlinger, S. D. Masking of speech by amplitude-modulated noise. *J. Acoust. Soc. Am.* **95**, 518–529 (1994).
- Stuart, A., Phillips, D. P. & Green, W. B. Word recognition performance in continuous and interrupted broad-band noise by normal hearing and simulated hearing-impaired listeners. *Am. J. Otol.* **16**, 658–663 (1995).
- Füllgrabe, C., Berthommier, F. & Lorenzi, C. Masking release for consonant features in temporally fluctuating background noise. *Hear. Res.* **211**, 74–84 (2006).
- Gnansia, D., Jourdes, V. & Lorenzi, C. Effect of masker modulation depth on speech masking release. *Hear. Res.* **239**, 60–68 (2008).
- Bacon, S. P., Opie, J. M. & Montoya, D. Y. The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *J. Speech Lang. Hear. Res.* **41**, 549–563 (1998).
- George, E. L. J., Festen, J. M. & Houtgast, T. Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* **120**, 2295–2311 (2006).
- Lorenzi, C., Husson, M., Ardoint, M. & Debrulle, X. Speech masking release in listeners with flat hearing loss: Effects of masker fluctuation rate on identification scores and phonetic feature reception. *Int. J. Audiol.* **45**, 487–495 (2006).
- Lorenzi, C. & Moore, B. C. J. Role of temporal envelope and fine structure cues in speech perception: A review. *Proc. Int. Symp. Audit. Audiol. Res.* **1**, 263–272 (2008).
- Dubno, J. R., Horwitz, A. R. & Ahlstrom, J. B. Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing. *J. Acoust. Soc. Am.* **113**, 2084–2094 (2003).
- Eisenberg, L. S., Dirks, D. D. & Bell, T. S. Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. *J. Speech Hear. Res.* **38**, 222–233 (1995).
- Gilbert, G., Bergeras, I., Voillery, D. & Lorenzi, C. Effects of periodic interruptions on the intelligibility of speech based on temporal fine-structure or envelope cues. *J. Acoust. Soc. Am.* **122**, 1336–1339 (2007).
- Gnansia, D., Péan, V., Meyer, B. & Lorenzi, C. Effects of spectral smearing and temporal fine structure degradation on speech masking release. *J. Acoust. Soc. Am.* **125**, 4023–4033 (2009).
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S. & Moore, B. C. J. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc. Natl. Acad. Sci. USA.* **103**, 18866–18869 (2006).
- Rosen, S. Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. Lond. B.* **336**, 367–373 (1992).
- Herrmann, B., Parthasarathy, A. & Bartlett, E. L. Ageing affects dual encoding of periodicity and envelope shape in rat inferior colliculus neurons. *Eur. J. Neurosci.* **45**, 299–311 (2017).
- Millman, R. E., Mattys, S. L., Gouws, A. D. & Prendergast, G. Magnified neural envelope coding predicts deficits in speech perception in noise. *J. Neurosci.* **37**, 7727–7736 (2017).
- Goossens, T., Vercammen, C., Wouters, J. & van Wieringen, A. Neural envelope encoding predicts speech perception performance for normal-hearing and hearing-impaired adults. *Hear. Res.* **370**, 189–200 (2018).
- Goossens, T., Vercammen, C., Wouters, J. & van Wieringen, A. Aging affects neural synchronization to speech-related acoustic modulations. *Front. Aging Neurosci.* **8**, 1–16 (2016).
- Moore, B. C. J. & Glasberg, B. R. Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech. *J. Acoust. Soc. Am.* **94**, 2050–2062 (1993).
- Schlittenlacher, J. & Moore, B. C. J. Discrimination of amplitude-modulation depth by subjects with normal and impaired hearing. *J. Acoust. Soc. Am.* **140**, 3487–3495 (2016).

43. Henry, K. S., Kale, S. & Heinz, M. G. Noise-induced hearing loss increases the temporal precision of complex envelope coding by auditory-nerve fibers. *Front. Syst. Neurosci.* **8**, 20 (2014).
44. Kale, S. & Heinz, M. G. Envelope coding in auditory nerve fibers following noise-induced hearing loss. *J. Assoc. Res. Otolaryngol.* **11**, 657–673 (2010).
45. Zhong, Z., Henry, K. S. & Heinz, M. G. Sensorineural hearing loss amplifies neural coding of envelope information in the central auditory system of chinchillas. *Hear. Res.* **309**, 55–62 (2014).
46. Schiffrin, D. How a story says what it means and does. *Text Interdiscip. J. Study Discourse* **4**, 313–346 (1984).
47. Jefferson, G. Sequential aspects of storytelling in conversation. *Stud. Org. Convers. Interact.* **1**, 219–248 (1978).
48. Ochs, E. & Taylor, C. Family narrative as political activity. *Discourse Soc.* **3**, 301–340 (1992).
49. Pasupathi, M., Lucas, S. & Coombs, A. Conversational functions of autobiographical remembering: Long-married couples talk about conflicts and pleasant topics. *Discourse Process.* **34**, 163–192 (2002).
50. Ervin-Tripp, S. M. & Küntay, A. C. The occasioning and structure of conversational stories. in *Conversation: Cognitive, communicative and social perspectives* 133–166 (John Benjamins, 1997). doi:<https://doi.org/10.1075/tsl.34.06erv>.
51. Bohanek, J. G. et al. Narrative interaction in family dinnertime conversations. *Merrill. Palmer. Q.* **55**, 488–515 (2009).
52. Eisenberg, A. R. Learning to describe past experiences in conversation. *Discourse Process.* **8**, 177–204 (1985).
53. Fivush, R., Bohanek, J. G. & Zaman, W. Personal and intergenerational narratives in relation to adolescents' well-being. *New Dir. Child Adolesc. Dev.* **2011**, 45–57 (2011).
54. McLean, K. C., Pasupathi, M. & Pals, J. L. Selves creating stories creating selves: A process model of self-development. *Personal. Soc. Psychol. Rev.* **11**, 262–278 (2007).
55. Mullen, M. K. & Yi, S. The cultural context of talk about the past: Implications for the development of autobiographical memory. *Cogn. Dev.* **10**, 407–419 (1995).
56. Ochs, E. & Capps, L. Narrating the self. *Annu. Rev. Anthropol.* **25**, 19–43 (1996).
57. Ochs, E., Smith, R. & Taylor, C. Detective stories at dinnertime: Problem-solving through co-narration. *Cult. Dyn.* **2**, 238–257 (1989).
58. Ochs, E., Taylor, C., Rudolph, D. & Smith, R. Storytelling as a theory-building activity. *Discourse Process.* **15**, 37–72 (1992).
59. Kalikow, D. N., Stevens, K. N. & Elliott, L. L. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J. Acoust. Soc. Am.* **61**, 1337–1351 (1977).
60. Dubno, J. R., Ahlstrom, J. B. & Horwitz, A. R. Use of context by young and aged adults with normal hearing. *J. Acoust. Soc. Am.* **107**, 538–546 (2000).
61. Nitttrouer, S. & Boothroyd, A. Context effects in phoneme and word recognition by young children and older adults. *J. Acoust. Soc. Am.* **87**, 2705–2715 (1990).
62. Sheldon, S., Pichora-Fuller, M. K. & Schneider, B. A. Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *J. Acoust. Soc. Am.* **123**, 489–499 (2008).
63. Cohen, G. & Faulkner, D. Word recognition: Age differences in contextual facilitation effects. *Br. J. Psychol.* **74**, 239–251 (1983).
64. Botvinick, M. & Braver, T. Motivation and cognitive control: From behavior to neural mechanism. *Annu. Rev. Psychol.* **66**, 83–113 (2015).
65. Kool, W., McGuire, J. T., Rosen, Z. B. & Botvinick, M. Decision making and the avoidance of cognitive demand. *J. Exp. Psychol. Gen.* **139**, 665–682 (2010).
66. Yee, D. M. & Braver, T. S. Interactions of motivation and cognitive control. *Curr. Opin. Behav. Sci.* **19**, 83–90 (2018).
67. Herrmann, B. & Johnsrude, I. S. A model of listening engagement (MoLE). *Hear. Res.* **397**, 108016 (2020).
68. Eckert, M. A., Teubner-Rhodes, S. & Vaden, K. I. Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. *Ear Hear.* **37**, 101S–110S (2016).
69. Peelle, J. E. Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear Hear.* **39**, 204–214 (2018).
70. Pichora-Fuller, M. K. et al. Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear Hear.* **37**, 5S–27S (2016).
71. Matthen, M. Effort and displeasure in people who are hard of hearing. *Ear Hear.* **37**, 28S–34S (2016).
72. Lalor, E. C. & Foxe, J. J. Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *Eur. J. Neurosci.* **31**, 189–193 (2010).
73. Ki, J. J., Kelly, S. P. & Parra, L. C. Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. *J. Neurosci.* **36**, 3092–3101 (2016).
74. Polonenko, M. J. & Maddox, R. K. Exposing distinct subcortical components of the auditory brainstem response evoked by continuous naturalistic speech. *Elife* **10**, e62329 (2021).
75. Schmäzle, R., Häcker, F. E. K., Honey, C. J. & Hasson, U. Engaged listeners: Shared neural processing of powerful political speeches. *Soc. Cogn. Affect. Neurosci.* **10**, 1137–1143 (2015).
76. Broderick, M. P., Anderson, A. J., Di Liberto, G. M., Crosse, M. J. & Lalor, E. C. Electrophysiological correlates of semantic dissimilarity reflect the comprehension of natural, narrative speech. *Curr. Biol.* **28**, 803–809 (2018).
77. Fiedler, L., Wöstmann, M., Herbst, S. K. & Obleser, J. Late cortical tracking of ignored speech facilitates neural selectivity in acoustically challenging conditions. *Neuroimage* **186**, 33–42 (2019).
78. Keitel, A., Ince, R. A. A., Gross, J. & Kayser, C. Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks. *Neuroimage* **147**, 32–42 (2017).
79. Puvvada, K. C. & Simon, J. Z. Cortical representations of speech in a multitalker auditory scene. *J. Neurosci.* **37**, 9189–9196 (2017).
80. Brodbeck, C., Jiao, A., Hong, L. E. & Simon, J. Z. Neural speech restoration at the cocktail party: Auditory cortex recovers masked speech of both attended and ignored speakers. *PLoS Biol.* **18**, 1–22 (2020).
81. Broderick, M. P., Anderson, A. J. & Lalor, E. C. Semantic context enhances the early auditory encoding of natural speech. *J. Neurosci.* **39**, 7564–7575 (2019).
82. Broderick, M. P., Di Liberto, G., Anderson, A., Rofes, A. & Lalor, E. Dissociable electrophysiological measures of natural language processing reveal differences in speech comprehension strategy in healthy ageing. *Sci. Rep.* **11**, 1–12 (2020).
83. Erb, J., Schmitt, L. M. & Obleser, J. Temporal selectivity declines in the aging human auditory cortex. *Elife* **9**, e55300 (2020).
84. Hess, T. M. & Ennis, G. E. Assessment of adult age differences in task engagement: The utility of systolic blood pressure. *Motiv. Emot.* **38**, 844–854 (2014).
85. Hess, T. M. Selective engagement of cognitive resources. *Perspect. Psychol. Sci.* **9**, 388–407 (2014).
86. Litman, L., Robinson, J. & Abberbock, T. TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behav. Res. Methods* **49**, 433–442 (2017).
87. Smits, C., Kapteyn, T. S. & Houtgast, T. Development and validation of an automatic speech-in-noise screening test by telephone. *Int. J. Audiol.* **43**, 15–28 (2004).
88. Schneider, B. A., Daneman, M. & Murphy, D. R. Speech comprehension difficulties in older adults: Cognitive slowing or age-related changes in hearing?. *Psychol. Aging* **20**, 261–271 (2005).
89. Bernstein, J. G. W. & Brungart, D. S. Effects of spectral smearing and temporal fine-structure distortion on the fluctuating-masker benefit for speech at a fixed signal-to-noise ratio. *J. Acoust. Soc. Am.* **130**, 473–488 (2011).

90. Irsik, V. C., Johnsrude, I. S. & Herrmann, B. Synchronized neural activity indexes engagement with spoken stories under acoustic masking. *bioRxiv* (2021).
91. Dmochowski, J. P., Sajda, P., Dias, J. & Parra, L. C. Correlated components of ongoing EEG point to emotionally laden attention: A possible marker of engagement?. *Front. Hum. Neurosci.* **6**, 1–9 (2012).
92. Hasson, U., Malach, R. & Heeger, D. J. Reliability of cortical activity during natural stimulation. *Trends Cogn. Sci.* **14**, 40–48 (2010).
93. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).
94. Léger, A. C., Moore, B. C. J. & Lorenzi, C. Temporal and spectral masking release in low- and mid-frequency regions for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* **131**, 1502–1514 (2012).
95. Peters, R. W., Moore, B. C. J. & Baer, T. Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J. Acoust. Soc. Am.* **103**, 577–587 (1998).
96. Goossens, T., Vercammen, C., Wouters, J. & van Wieringen, A. The association between hearing impairment and neural envelope encoding at different ages. *Neurobiol. Aging* **74**, 202–212 (2019).
97. Herrmann, B., Buckland, C. & Johnsrude, I. S. Neural signatures of temporal regularity processing in sounds differ between younger and older adults. *Neurobiol. Aging* **83**, 73–85 (2019).
98. Fitzgibbons, P. J. & Gordon-Salant, S. Age effects on duration discrimination with simple and complex stimuli. *J. Acoust. Soc. Am.* **98**, 3140–3145 (1995).
99. Fitzgibbons, P. J. & Gordon-Salant, S. Auditory temporal order perception in younger and older adults. *J. Speech, Lang. Hear. Res.* **41**, 1052–1060 (1998).
100. Pichora-Fuller, M. K., Schneider, B. A., MacDonald, E., Pass, H. E. & Brown, S. Temporal jitter disrupts speech intelligibility: A simulation of auditory aging. *Hear. Res.* **223**, 114–121 (2007).
101. Schneider, B. A. & Pichora-Fuller, M. K. Age-related changes in temporal processing: Implications for speech perception. *Semin. Hear.* **22**, 227–238 (2001).
102. Obleser, J., Wise, R. J. S., Dresner, M. A. & Scott, S. K. Functional integration across brain regions improves speech perception under adverse listening conditions. *J. Neurosci.* **27**, 2283–2289 (2007).
103. Davis, M. H., Ford, M. A., Kherif, F. & Johnsrude, I. S. Does semantic context benefit speech understanding through ‘top-down’ processes? Evidence from time-resolved sparse fMRI. *J. Cogn. Neurosci.* **23**, 3914–3932 (2011).
104. Miller, G. A., Heise, G. A. & Lichten, W. The intelligibility of speech as a function of the context of the test materials. *J. Exp. Psychol.* **41**, 329–335 (1951).
105. Bashford, J. A., Riener, K. R. & Warren, R. M. Increasing the intelligibility of speech through multiple phonemic restorations. *Percept. Psychophys.* **51**, 211–217 (1992).
106. Holmes, E., Folkeard, P., Johnsrude, I. S. & Scollie, S. Semantic context improves speech intelligibility and reduces listening effort for listeners with hearing impairment. *Int. J. Audiol.* **57**, 483–492 (2018).
107. Mar, R. A. & Oatley, K. The function of fiction is the abstraction and simulation of social experience. *Perspect. Psychol. Sci.* **3**, 173–192 (2008).
108. Zwaan, R. A. Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychon. Bull. Rev.* **23**, 1028–1034 (2016).
109. Zwaan, R. A., Langston, M. C. & Graesser, A. C. The construction of situation models in narrative comprehension: An event-indexing model. *Psychol. Sci.* **6**, 292–297 (1995).
110. Botvinick, M., Huffstetler, S. & McGuire, J. T. Effort discounting in human nucleus accumbens. *Cogn. Affect. Behav. Neurosci.* **9**, 16–27 (2009).
111. Bénabou, R. & Tirole, J. Intrinsic and extrinsic motivation. *Rev. Econ. Stud.* **70**, 489–520 (2003).
112. Ryan, R. M. & Deci, E. L. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemp. Educ. Psychol.* **25**, 54–67 (2000).
113. Gosling, S. D., Vazire, S., Srivastava, S. & John, O. P. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *Am. Psychol.* **59**, 93–104 (2004).
114. Buhrmester, M., Kwang, T. & Gosling, S. D. Amazon’s Mechanical Turk: A new source of inexpensive, yet high-quality, data?. *Perspect. Psychol. Sci.* **6**, 3–5 (2011).
115. Thomas, K. A. & Clifford, S. Validity and Mechanical Turk: An assessment of exclusion methods and interactive experiments. *Comput. Human Behav.* **77**, 184–197 (2017).
116. Berinsky, A. J., Margolis, M. F. & Sances, M. W. Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *Am. J. Pol. Sci.* **58**, 739–753 (2014).
117. Buchanan, E. M. & Scofield, J. E. Methods to detect low quality data and its implication for psychological research. *Behav. Res. Methods* **50**, 2586–2596 (2018).
118. Mason, W. & Suri, S. Conducting behavioral research on Amazon’s Mechanical Turk. *Behav. Res. Methods* **44**, 1–23 (2012).
119. Bilger, R. C. *Manual for the clinical use of the Revised SPIN test*. (University of Illinois Press, London, 1984).
120. Edwards, E. & Chang, E. F. Syllabic (~2-5Hz) and fluctuation (~1-10Hz) ranges in speech and auditory processing. *Hear. Res.* **305**, 113–134 (2013).
121. de Leeuw, J. R. jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behav. Res. Methods* **47**, 1–12 (2015).
122. Woods, K. J. P., Siegel, M. H., Traer, J. & McDermott, J. H. Headphone screening to facilitate web-based auditory experiments. *Attent. Percept. Psychophys.* **79**, 2064–2072 (2017).
123. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
124. Rosenthal, R. & Rubin, D. B. *r* equivalent: A simple effect size indicator. *Psychol. Methods* **8**, 492–496 (2003).

Acknowledgements

This research was supported by the Canadian Institutes of Health Research (MOP133450 to I.S. Johnsrude). B.H. was supported by the Canada Research Chair program.

Author contributions

V.C.I., I.S.J., and B.H. designed the experiments. V.C.I. conducted the experiments, analyzed the data, and wrote the paper. V.C.I., I.S.J., and B.H. edited the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-09805-6>.

Correspondence and requests for materials should be addressed to V.C.I.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022