Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqIU5gAAAAA:7ix7iJpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQlMvhMlc by MAX PLANCK INSTITUT user on 16 May 2022

Neural Activity during Story Listening Is Synchronized across Individuals Despite Acoustic Masking

Vanessa C. Irsik[1], Ingrid S. Johnsrude[1], and Björn Herrmann[1,2,3]

Abstract

■ Older people with hearing problems often experience difficulties understanding speech in the presence of background sound. As a result, they may disengage in social situations, which has been associated with negative psychosocial health outcomes. Measuring listening (dis)engagement during challenging listening situations has received little attention thus far. We recruit young, normal-hearing human adults (both sexes) and investigate how speech intelligibility and engagement during naturalistic story listening is affected by the level of acoustic masking (12-talker babble) at different signal-to-noise ratios (SNRs). In Experiment 1, we observed that word-report scores were above 80% for all but the lowest SNR (−3 dB SNR) we tested, at which performance dropped to 54%. In Experiment 2, we calculated intersubject correlation (ISC) using EEG data to identify dynamic spatial patterns of shared neural activity evoked by the stories. ISC has been used as a neural measure of participants' engagement with naturalistic materials. Our results show that ISC was stable across all but the lowest SNRs, despite reduced speech intelligibility. Comparing ISC and intelligibility demonstrated that word-report performance declined more strongly with decreasing SNR compared to ISC. Our measure of neural engagement suggests that individuals remain engaged in story listening despite missing words because of background noise. Our work provides a potentially fruitful approach to investigate listener engagement with naturalistic, spoken stories that may be used to investigate (dis)engagement in older adults with hearing impairment. ■

INTRODUCTION

Many listening environments we encounter contain background masking sounds (Smeds, Wolters, & Rung, 2015; Olsen, 1998) that can lead to listening challenges for a significant proportion of people aged 50 years and older (Gordon-Salant, 2006; Frisina & Frisina, 1997). People who frequently experience listening difficulties may avoid environments with increased levels of background noise, resulting in social isolation (Palmer, Newsom, & Rook, 2016; Dawes et al., 2015) and negative psychosocial and physical health outcomes (Pichora-Fuller, Mick, & Reed, 2015; Wayne & Johnsrude, 2015). However, social isolation is likely subsequent to within-situation disengagement—a coping mechanism to reduce cognitive demand. A person may temporarily "zone out" in conversational situations because continuous listening is too difficult (Herrmann & Johnsrude, 2020a, 2020b; Heffernan, Coulson, Henshaw, Barry, & Ferguson, 2016). A greater understanding of the listening conditions under which a person may disengage, and development of quantitative measures of within-situation disengagement, may improve our ability to diagnose hearing problems before social isolation manifests.

Two factors are likely critical for listening engagement, which we define as "the (automatic or volitional) recruitment of executive and other cognitive resources, when speech comprehension serves a valued communication goal" (Herrmann & Johnsrude, 2020a). First, engagement depends on degree to which a listener judges that, with deliberate deployment of cognitive resources, some of the masked speech will be understandable (Herrmann & Johnsrude, 2020a; Pichora-Fuller et al., 2016; Wright, 2014; Richter, 2013; Brehm & Self, 1989). An individual may disengage if they believe speech comprehension will be impossible (Peelle, 2018; Eckert, Teubner-Rhodes, & Vaden, 2016; Richter, 2016). A second important factor is motivation. An individual who is motivated to listen, for example, because they find what they are hearing to be enjoyable and rewarding (Matthen, 2016), may engage in listening despite the presence of background sounds (Herrmann & Johnsrude, 2020a; Pichora-Fuller et al., 2016). However, typical investigations of speech comprehension involve listening to isolated sentences (Obleser, Wise, Dresner, & Scott, 2007; Duncan & Aarts, 2006; Davis & Johnsrude, 2003) that lack a topical thread, are not very interesting (e.g., He buttoned his shirt), and may not therefore foster volition to listen (i.e., conation; Reitan & Wolfson, 2000). Furthermore, listening engagement also likely develops more slowly, at time scales beyond individual sentences (Mandler & Goodman, 1982). Spoken stories, in contrast, involve event descriptions along a topical thread, are intrinsically motivating to a listener, and are common in everyday life (Dunlop & Walker, 2013). Utilizing engaging spoken stories under varying levels of acoustic masking may thus provide a

[1]The University of Western Ontario, [2]Rotman Research Institute, Toronto, Canada, [3]University of Toronto

© 2022 Massachusetts Institute of Technology

Journal of Cognitive Neuroscience 34:6, pp. 933–950
https://doi.org/10.1162/jocn_a_01842

fruitful avenue to investigate intelligibility as well as listening engagement and disengagement.

Engagement (and disengagement) can be measured behaviorally through assessment of an individual's experience (Herrmann & Johnsrude, 2020b; Dmochowski et al., 2014; Kuijpers, Hakemulder, Tan, & Doicaru, 2014; Busselle & Bilandzic, 2008, 2009). However, behavioral assessment of engagement typically requires the listener to retroactively introspect about their experiences (Herrmann & Johnsrude, 2020b). In contrast, neural activity recorded using functional imaging or EEG may provide a real-time window on engagement. Naturalistic materials such as movies or spoken stories lead to synchronized patterns of neural activity across individuals that scale with the degree of engagement (Nastase, Gazzola, Hasson, & Keysers, 2019; Nguyen, Vanderwal, & Hasson, 2019; Yeshurun et al., 2017; Hasson, Malach, & Heeger, 2010). The strength of synchrony across individuals, quantified as intersubject correlation (ISC; Dmochowski et al., 2014; Dmochowski, Sajda, Dias, & Parra, 2012; Hasson, Furman, Clark, Dudai, & Davachi, 2008; Hasson, Nir, Levy, Fuhrmann, & Malach, 2004), is stronger when stimuli are captivating or exciting (Schmälzle, Häcker, Honey, & Hasson, 2015; Hasson et al., 2010) and is predictive of behavioral measures reflecting engagement (Song, Finn, & Rosenberg, 2021; Cohen, Henin, & Parra, 2017; Dikker et al., 2017; Poulsen, Kamronn, Dmochowski, Parra, & Hansen, 2017; Dmochowski et al., 2014) and recall of the materials (Piazza, Cohen, Trach, & Lew-Williams, 2021; Song et al., 2021; Chan, Smidts, Schoots, Dietvorst, & Boksem, 2019; Davidesco et al., 2019; Cohen et al., 2018; Cohen & Parra, 2016; Stephens, Silbert, & Hasson, 2010; Hasson, Furman, et al., 2008). Conversely, ISC is reduced when individuals do not attend to naturalistic materials (Rosenkranz, Holtze, Jaeger, & Debener, 2021; Cohen et al., 2018; Ki, Kelly, & Parra, 2016; Kuhlen, Allefeld, & Haynes, 2012) or when stimuli are unstructured or temporally scrambled (Dmochowski et al., 2012; Hasson, Yang, Vallines, Heeger, & Rubin, 2008).

ISC is conceptually similar to measures that estimate the correlation between specific stimulus features (such as the amplitude envelope) and neural responses (cf. Iotzov & Parra, 2019; Ding & Simon, 2013; Peelle, Gross, & Davis, 2013). However, critically, neural synchronization to stimulus features is only one potential source that may drive shared neural patterns across subjects during naturalistic listening. Engaging narratives are designed to provide a shared conscious experience to listeners—this is driven, in part, by the recruitment of processes that enable each listener to integrate incoming information with existing schemas, refine their understanding, infer what is unsaid, and anticipate what may be coming next, while filtering out distractions (Naci, Cusack, Anello, & Owen, 2014). This manifests as ongoing engagement in the story. These cognitive processes are executive in nature because they are coordinating all mental activity as the story unfolds, enabling the listener to follow along and understand. ISC is sensitive to such experiential aspects of engaging materials (Nastase et al., 2019; Yeshurun et al., 2017; Schmälzle et al., 2015; Dmochowski et al., 2014; Naci et al., 2014; Nummenmaa et al., 2012).

In the current study, we use spoken stories to investigate how challenging listening situations affect speech intelligibility and story engagement in a group of young, normal-hearing adults. We opted to measure intelligibility and ISC in separate experiments because we suspected the inclusion of the intelligibility task may influence participants' engagement levels during story listening. Therefore, in Experiment 1, we assess intelligibility of spoken stories presented with different levels of a 12-talker babble masker. In Experiment 2, we record EEG in another group of individuals while they listen to the same stories used in Experiment 1 and investigate how synchronized activity (ISC) is affected by background masking during story listening.

EXPERIMENT 1: EXAMINING STORY INTELLIGIBILITY AND ENGAGEMENT

In Experiment 1, we investigate the extent to which participants understand spoken stories when they are masked with 12-talker babble at different signal-to-noise ratios (SNRs). We further assess overall listening engagement behaviorally using a modified narrative absorption scale (NAS; Kuijpers et al., 2014) adapted for use with spoken stories (Herrmann & Johnsrude, 2020b). Speech intelligibility ratings from Experiment 1 will be related to data from Experiment 2, in which we investigate neural signatures of story engagement, to assess the extent to which story engagement relates to speech intelligibility.

Methods

Participants

Eighty-two individuals (mean age = 28.8 years, age range = 18–36 years; 51 men, 31 women) without self-reported hearing loss, or neurological or psychiatric disorders, participated in Experiment 1. All participants were recruited from the Amazon Mechanical Turk online participant pool (www.mturk.com/) via the participant sourcing platform Cloud Research (previously TurkPrime; Litman, Robinson, & Abberbock, 2017). Participants provided informed consent, and the study protocol was approved by the University of Western Ontario's Nonmedical Research Ethics Board (REB #112574). Each individual received financial compensation of 6 USD after completion of the study (10 USD hourly rate).

Online research can be subject to increased levels of random responders as experimenters have limited control over the testing environment. However, previous work indicates that online studies generally replicate findings of in-person data collection (Buchanan & Scofield, 2018; Thomas & Clifford, 2017; Berinsky, Margolis, & Sances,

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqIU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQlMvhMlc by MAX PLANCK INSTITUT user on 16 May 2022

2014; Mason & Suri, 2012; Buhrmester, Kwang, & Gosling, 2011; Gosling, Vazire, Srivastava, & John, 2004). Twenty-one additional individuals participated in the study but were not included, because of a technical error during data recording (n = 2), reporting hearing aid usage or constant ringing in at least one ear (n = 9), not wearing headphones during the study (n = 3), or submitting random one-word answers to all questions on the intelligibility task (n = 7).

Acoustic Stimulation and Procedure

Four stories were selected from the story-telling podcast "The Moth" (themoth.org): *Reach for the Stars One Small Step at a Time* (by Richard Garriott, ~13 min), *The Bounds of Comedy* (by Colm O'Regan, ~10 min), *Nacho Challenge* (by Omar Qureshi, ~11 min), and *Discussing Family Trees in School Can Be Dangerous* (by Paul Nurse, ~10 min). Each story had 12-talker babble noise added as a masker. Babble noise was derived using the masker materials from the Revised Speech in Noise test (Bilger, 1984). Individual 5-sec babble noise snippets were randomly selected from a total set of 100 and concatenated to equal the length of each story. The SNR was pseudorandomly varied approximately every 30–33 sec throughout each story (32, 30, 33, and 30 sec, respectively, for the four stories). Five SNRs were chosen, namely, clear and +12, +7, +2, and −3 dB SNR, relative to a 12-talker babble masker. SNR was manipulated by adjusting the dB level of both the story and masker. This ensured that the overall sound level remained constant throughout a story and was similar for all stories. Three versions of SNR condition order were generated for each story to ensure that specific parts of a story were not confounded with a specific SNR. Within each version, SNR was varied pseudorandomly such that a particular SNR could not be heard twice in succession. Four 30-, 33-, or 30-sec segments per SNR were presented for stories by O'Regan, Qureshi, and Nurse, respectively, and five 32-sec segments per SNR were presented for the longest story by Garriott.

For each story, phrases/sentences ranging from four to eight words (range of durations: 0.62–3.6 sec) were selected for intelligibility (word report) testing. These test phrases/sentences did not occur during the transition period from one SNR to the next (for approximately 5 sec before, and 1 sec after, the SNR transition). Four phrases per 30- to 33-sec segment were selected, resulting in 100 phrases for the story by Garriott and 80 for each of the other three stories. Two of the four selected phrases per 30- to 33-sec segment were used as one intelligibility test set, whereas the other two of the four selected phrases were used as a second intelligibility test set (5 or 4 segments × 5 SNRs × 2 phrases/sentences × 2 intelligibility test sets = 100 or 80 phrases). Having two test sets ensured that the observed intelligibility effects were not confounded by specific phrases/sentences.

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 control; de Leeuw, 2015). The experiment code was stored at an online GitLab repository (gitlab.pavlovia .org) and hosted via Pavlovia (pavlovia.org/). During the main experiment, each participant listened to one pseudorandomly selected story (Garriott: n = 21; O'Regan: n = 22; Qureshi: n = 20; Nurse: n = 19). SNR order and intelligibility test set were randomly assigned. A black fixation cross was presented at the center of the screen throughout the story. The fixation cross turned yellow 2 sec before the beginning of a test phrase/sentence, cueing the participant to prepare for intelligibility testing (see Figure 1). 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, after which the story resumed from the beginning of the sentence most recently heard (allowing for story continuation). After the story ended, participants answered questions that assessed comprehension and rated statements about their story listening experiences (see Table 1).

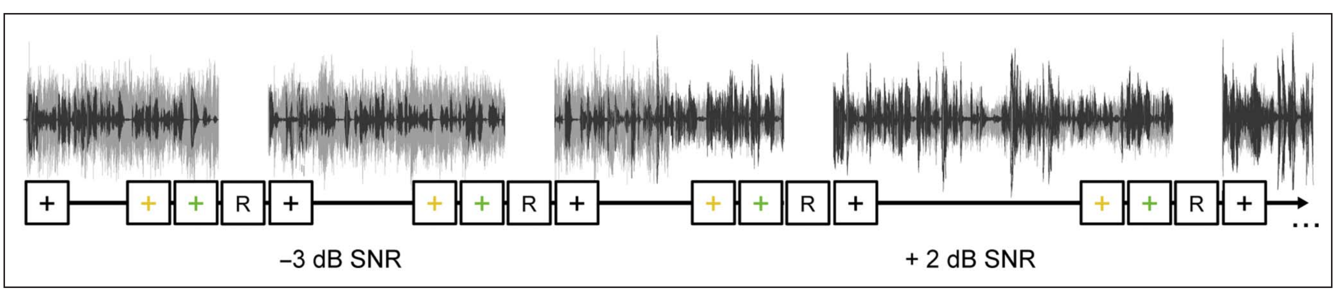

Figure 1. Schematic representation of the speech intelligibility task. Participants listened to a spoken story (black) masked with different levels of 12-talker babble noise (gray). A fixation cross was displayed on the computer screen throughout the study and changed colors to communicate which parts of the spoken story participants would need to recall. The fixation cross turned yellow 2 sec before the beginning of a test phrase/sentence, cueing the participant to prepare for intelligibility testing. The fixation cross turned green during the phrase/sentence participants would be asked to report back. The story stopped 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.

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqIU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzyIARKqQc6fOy0Tn6q3uJGHVmQIMvhMlc by MAX PLANCK INSTITUT user on 16 May 2022

Table 1. Statements of the NAS, Enjoyment, and Motivation

Dimension	*Statement*
Attention (NAS)	When I finished listening, I was surprised to see that time had gone by so fast.
Attention (NAS)	When I was listening, I was focused on what happened in the story.
Attention (NAS)	I felt absorbed in the story.
Attention (NAS)	The story gripped me in such a way that I could close myself off for things that were happening around me.
Attention (NAS)	I was listening in such a concentrated way that I had forgotten the world around me.
Emotional engagement (NAS)	When I listened to the story, I could imagine what it must be like to be in the shoes of the main character(s).
Emotional engagement (NAS)	I felt sympathy for the main character(s).
Emotional engagement (NAS)	I felt connected with the main character(s) of the story.
Emotional engagement (NAS)	I felt how the main character(s) was/were feeling.
Emotional engagement (NAS)	I felt for what happened in the story.
Mental imagery (NAS)	When I was listening to the story, I had an image of the main character(s) in mind.
Mental imagery (NAS)	When I was listening to the story, I could see the situations happening in the story being played out before my eyes.
Mental imagery (NAS)	I could imagine what the world in which the story took place looked like.
Transportation (NAS)	When I was listening to the story, it sometimes seemed as if I were in the story world too.
Transportation (NAS)	When listening to the story, there were moments in which I felt that the story world overlapped with my own world.
Transportation (NAS)	The world of the story sometimes felt closer to me than the world around me.
Transportation (NAS)	When I was finished with listening to the story, it felt like I had taken a trip to the world of the story.
Transportation (NAS)	Because all of my attention went into the story, I sometimes felt as if I could not exist separate from the story.
Enjoyment	I thought it was an exciting story.
Enjoyment	I thought it was an enthralling story.
Enjoyment	I listened to the story with great interest.
Enjoyment	I thought the story was beautiful.
Enjoyment	I thought the story was presented well.
Motivation	I was motivated to listen to the story.

Source: Herrmann and Johnsrude (2020a).

To familiarize participants with the intelligibility task, a brief practice block was presented before 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 (two trials per 30-sec segment).

Online Research Quality Assurance Measures

Participants completed two initial listening tasks at the very beginning of the online session. First, participants listened to a 15-sec stream of pink noise normalized to the same root-mean-square amplitude as the stories 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 had an opportunity to adjust their volume to a comfortable level before the intelligibility task, after which they were instructed to not make further adjustments.

Second, participants completed a headphone-check procedure to determine whether participants were wearing headphones as instructed (cf. Woods, Siegel, Traer, & McDermott, 2017). Participants performed a tone

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqlU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzyIARKqQc6fOy0Tn6q3uJGHVmQIMvhMlc by MAX PLANCK INSTITUT user on 16 May 2022

discrimination task (six 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 (antiphase tone). This task is straightforward over headphones, but difficult over loudspeakers, because the pressure waves generated from an antiphase tone interfere if heard through loudspeakers (Woods et al., 2017). If they were listening through loudspeakers, they would likely falsely select the antiphase tone as the quietest tone. No participants were excluded solely on the basis of performance on this test; however, it did provide a metric that could flag potentially noncompliant online subjects.

Assessment of Intelligibility

We calculated the proportion of correctly reported words for each SNR (clear, +12, +7, +2, −3), across the three versions of the four stories. Different or omitted words were counted as errors, but minor misspellings and incorrect grammatical number (singular vs. plural) were not. Word-report performance was assessed using a repeated-measures ANOVA (rmANOVA) with SNR (clear and +12, +7, +2, and −3 dB SNR) as the within-participant factor.

Assessment of Story Comprehension

Story comprehension was assessed using eight statements, which either correctly or incorrectly described an element from the story the participant heard. Participants were asked to categorize each statement as true or false. Comprehension performance was calculated as the proportion of statements categorized correctly. Comprehension was statistically examined using a one-sample *t* test, which tested scores against chance-level performance of 0.5.

Assessment of Story Engagement, Enjoyment, and Motivation

After the assessment of story comprehension, we also assessed how much participants were engaged with the story, how much they enjoyed the story, and how motivated they were to listen. To assess story engagement, we utilized the NAS (Kuijpers et al., 2014) adapted previously for spoken stories (Herrmann & Johnsrude, 2020b). The NAS contains statements along four dimensions (attention, emotional engagement, mental imagery, and transportation; Herrmann & Johnsrude, 2020b; Kuijpers et al., 2014), which participants rated on a 7-point scale, where 1 referred to "completely disagree" and 7 referred to "completely agree" (Table 1). Participants also rated statements about enjoyment and listening motivation on the same 7-point scale (Table 1). Rating scores for each statement were averaged separately for narrative absorption (NAS) and enjoyment. Ratings for NAS and enjoyment were statistically examined in separate one-sample *t* tests, which tested ratings against a neutral response (test value: 4). Given that motivation was assessed using a single question (Table 1), the data remain ordinal rather than continuous. Motivation was therefore assessed using a one-sample Wilcoxon signed-rank test, against a neutral response (test value: 4).

Assessment of the Relationship between Measures of Engagement, Enjoyment, and Motivation

To characterize the relationship between behavioral ratings of absorption (NAS), enjoyment, motivation, and story comprehension, we calculated correlations (Pearson) between pairs of behavioral measures (six pairs).

Experimental Design and Statistical Analysis

Statistical analyses were conducted using IBM SPSS Statistics for Windows (v24) and MATLAB (The MathWorks, Inc.). Details of the specific variables and statistical tests for each analysis can be found in the analysis subsections for each measure. In general, effects were examined using either an rmANOVA, paired samples *t* tests, or one-sample *t* tests. Behavioral ratings that were considered ordinal, not continuous, were analyzed using nonparametric tests for either one-sample (Wilcoxon signed-rank test) or independent sample (Wilcoxon rank sum) comparisons. Significant effects were followed up using *t* tests, with multiple comparisons corrected using the false discovery rate (FDR; Benjamini & Hochberg, 1995). FDR-corrected *p* values are reported as p_{FDR}. Effect sizes are reported as partial eta squared (η_p^2) for rmANOVAs and $r_{\mathrm{equivalent}}$ (r_e; Rosenthal & Rubin, 2003) 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 upon reasonable request.

Results and Discussion

Story Comprehension, Motivation, and Enjoyment Are High Despite Speech Masking

Performance on comprehension questions was significantly above chance (M = .80, SE = .02), $t(81) = 14.79$, $p = 1 \times 10^{-24}$, $r_e = .85$, which suggests participants were able to grasp details from each story despite varying the SNR and performing the speech intelligibility task while listening to each story.

We also examined participants' level of story engagement by testing mean ratings for NAS, enjoyment, and motivation against a rating of 4 (neutral response). Scores on the NAS were not significantly different from a neutral

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqIU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzyIARKqQc6fOy0Tn6q3uJGHVmQIMhMlc by MAX PLANCK INSTITUT user on 16 May 2022

response (p = .832), but enjoyment, $t(81) = 2.87$, p = .005, r_e = .3, and motivation ($V = 2103$, $p = 5 \times 10^{-11}$) ratings were higher than the neutral point (see Figure 2A). This suggests that, although participants were ambivalent regarding whether they were fully immersed in the story they heard, they appeared to enjoy listening to the story and were motivated to do so. These findings are somewhat inconsistent with previous research on story engagement (Herrmann & Johnsrude, 2020b), which found engagement to be high and largely unaffected by moderate noise levels. The presence of the intelligibility task during story listening may have altered the listening experience such that engagement was reduced to enable detail-oriented listening to ensure participants could report back specific words when prompted.

Story Intelligibility Declines as SNR Decreases

For the intelligibility task, the proportion of correctly reported words declined with decreasing SNR, $F(4, 324)$ = 176.5, $p = 2.98 \times 10^{-38}$, η_p^2 = .69 (see Figure 2B), as predicted (Duncan & Aarts, 2006; Brungart, 2001; Brungart, Simpson, Ericson, & Scott, 2001; Bronkhorst, 2000). Comparing intelligibility at successive pairs of SNRs, to identify the SNRs at which a significant drop in performance occurred, revealed that intelligibility was comparable between the clear and +12-dB SNR conditions (p_{FDR} = .077) but declined from +12 to +7 dB SNR ($t(81) = 3.33$, p_{FDR} = .002, r_e = .35), +7 to +2 dB SNR ($t(81) = 5.7$, $p_{FDR} = 3.71 \times 10^{-7}$, r_e = .54), and + 2 to −3 dB SNR ($t(81) = 14.94$, $p_{FDR} = 2.22 \times 10^{-24}$, r_e = .86).

Relating Intelligibility and Behavioral Measures of Engagement and Comprehension

In an effort to quantify how intelligibility may relate to engagement and comprehension measures, we calculated correlations between intelligibility (averaged across SNRs), NAS, motivation, enjoyment, and comprehension scores (see Figure 2C). We found that intelligibility was significantly correlated with motivation (r_{80} = .33, p_{FDR} = .004) and comprehension (r_{80} = .50, $p_{FDR} = 6.7 \times 10^{-6}$) scores, such that better word-report performance was associated with higher levels of motivation to listen and better comprehension of the stories. We did not find a significant correlation between intelligibility and NAS (r_{82} = .10, p_{FDR} = .381) or intelligibility and enjoyment scores (r_{80} = .16, p_{FDR} = .223), suggesting that enjoyment/engagement with the story and being able to report the specific words spoken during the story may be independent (cf. Herrmann & Johnsrude, 2020b).

We also found a significant correlation between NAS and enjoyment (r_{80} = .77, $p_{FDR} = 3.9 \times 10^{-16}$) and between NAS and motivation (r_{80} = .55, $p_{FDR} = 3.4 \times 10^{-6}$), suggesting that higher levels of engagement were associated with greater enjoyment of the stories and motivation to listen. We further observed a significant relationship between enjoyment and motivation (r_{80} = .48, $p_{FDR} = 1.6 \times 10^{-5}$) and between enjoyment and comprehension (r_{80} = .26, p_{FDR} = .029). This indicates that enjoyment was associated with greater motivation to listen and better comprehension. No relationship was observed between NAS and comprehension (r_{80} = .11, p_{FDR} = .381) or between motivation and comprehension (r_{80} = .12, p_{FDR} = .367).

The results of Experiment 1 suggest that participants were able to maintain a relatively high level of story comprehension, were generally motivated, and enjoyed listening to the stories, despite the presence of the babble masker. Furthermore, speech intelligibility was relatively good for SNRs up to +2 dB (~80% correct words reported) but substantially declined for a higher masking level (−3 dB SNR; ~54%). The intelligibility and engagement data obtained in this experiment will be related to ISC during story listening in Experiment 2.

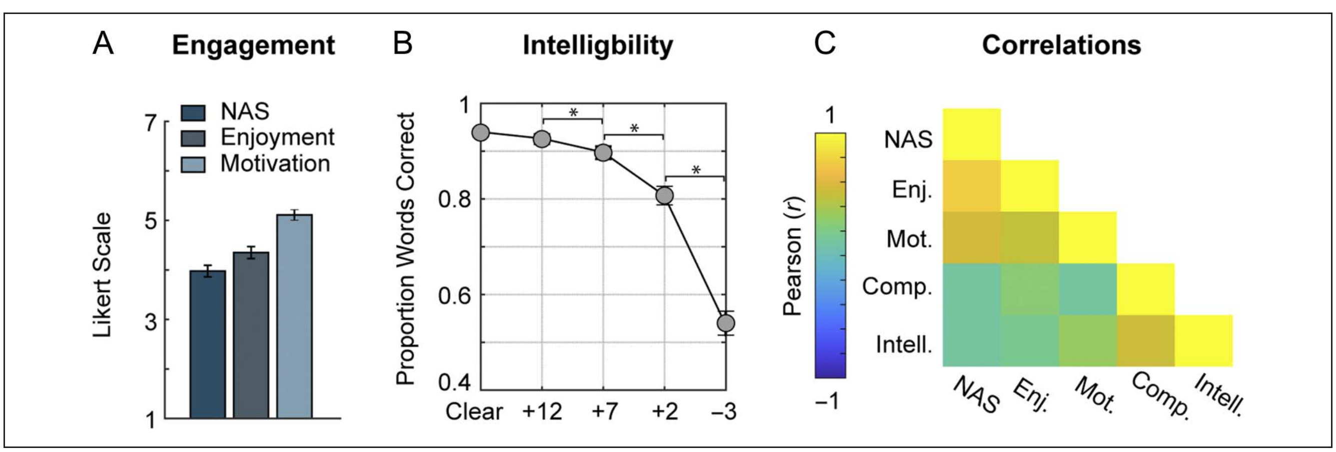

Figure 2. Behavioral measures of story comprehension, engagement, and intelligibility. (A) Mean ratings of the statements comprising the NAS, and those assessing enjoyment and motivation. Ratings were provided on a 7-point Likert scale, where 1 refers to "completely disagree" and 7 refers to "completely agree." (B) Mean proportion of correctly reported words plotted as a function of SNR (clear and +12, +7, +2, and −3 dB). (C) Correlation matrix depicting the correlation (Pearson) between behavioral measures of story comprehension, engagement, and intelligibility. Error bars reflect *SEM*. Enj. = enjoyment; Mot. = motivation; Comp. = comprehension; Intell. = intelligibility. *$p < .05$.

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqlU5gAAAAA:7ix7UpA0RSVUYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQIMmMlc by MAX PLANCK INSTITUT user on 16 May 2022

EXPERIMENT 2: EXAMINING ISC DURING STORY LISTENING

Synchronization of neural activity across different individuals—measured as ISC—has been suggested to provide a window onto engagement and shared experience with naturalistic materials (Nastase et al., 2019; Nguyen et al., 2019; Yeshurun et al., 2017; Hasson et al., 2010). In Experiment 2, we investigate how ISC during story listening is affected by 12-talker background babble noise at different SNRs.

Methods

Participants

Thirty-nine individuals (mean age = 20.3 years, age range = 18–32 years; 19 men, 20 women) without hearing loss, neurological issues, or psychiatric disorders participated in Experiment 2. All participants were recruited from Western University or the surrounding community of London, Canada. All participants provided written informed consent and were financially compensated with 10 CAD per hour. The study was conducted in accordance with the Declaration of Helsinki and the Canadian Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans (TCPS2-2014) and approved by the local health sciences research ethics board of the University of Western Ontario (REB #112015). Five additional individuals participated but were not included because of either an issue with sound delivery ($n = 2$) or a technical error during data recording ($n = 3$).

Acoustic Stimulation and Procedure

The experiment was conducted in a single-walled sound-attenuating booth (Eckel Industries). Sounds were delivered through Sennheiser (HD 25 Light) headphones at a comfortable listening level, using a Focusrite Scarlett 2i4 external soundcard controlled by a PC (Windows 10) and Psychtoolbox (Version 3) in MATLAB (R2017b).

Stimuli were the same four stories used in Experiment 1, with the same SNRs (clear, +12, +7, +2, −3) and the same three versions of SNR condition order per story. EEG was recorded while participants listened to each of the four stories consecutively (see Figure 3). Story version and story order were counterbalanced across participants. After each story, comprehension, story engagement, motivation, and enjoyment were assessed, and these data were analyzed, as in Experiment 1. After listening to all four stories, participants completed a resting block, where they sat quietly for 6 min with their eyes open while EEG was recorded. This block was used as a baseline for ISC analyses of story listening.

EEG Recording and Preprocessing

EEG was recorded from 64 active electrodes (Ag/AgCl) placed on the scalp using an electrode cap and two additional electrodes placed on both mastoids, with spacing on the scalp according to the 10/20 system (Biosemi ActiveTwo system). During data recording, all electrodes were referenced to a feedback loop of two electrodes, a Common Mode Sense active electrode and a driven passive electrode (see www.biosemi.com/faq/cms&drl.htm). EEG was recorded at 1024 Hz with an online low-pass filter of 208 Hz to focus on cortical sources.

All preprocessing was carried out offline using MATLAB software, the FieldTrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011), and custom scripts. EEG data were rereferenced to the average of the signal from both mastoids. EEG data were then notch-filtered at 60 Hz to attenuate line noise and then high-pass (0.5 Hz, 3429 points, Hann window) and low-pass (22 Hz, 211 points, Kaiser window) filtered. EEG data were also downsampled to 256 Hz. Artifacts because of eye movements (saccades, blinks) and muscle activity were removed using independent component analysis (Makeig, Bell, Jung, & Sejnowski, 1996). To further exclude additional artifacts from subsequent analyses, data segments in which the EEG signal changed by more than 80 μV within a 0.2-sec period in any channel were set to 0 μV (cf. Cohen & Parra, 2016; Dmochowski et al., 2012, 2014).

ISC

We quantified ISC using correlated component analysis (Parra, Haufe, & Dmochowski, 2019; Dmochowski et al., 2012), a signal decomposition method that identifies a set of electrode weights (i.e., spatial filters) yielding a

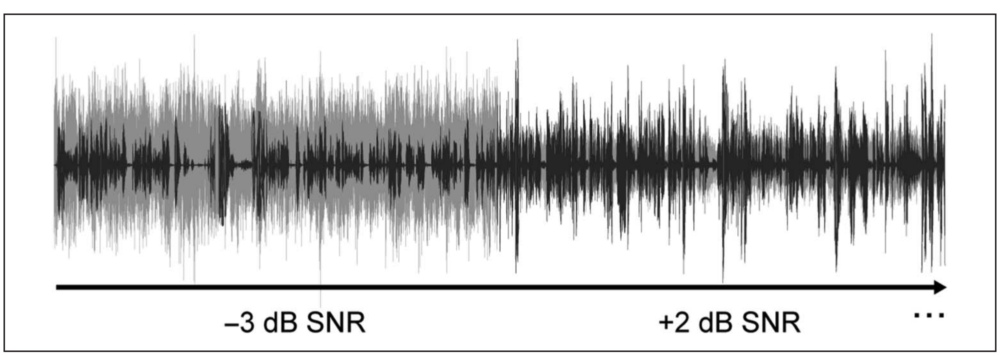

Figure 3. Stimulus design for Experiment 2. Participants listened to four consecutive spoken stories (black) masked with varied levels of 12-talker babble noise (gray); two of the SNRs are shown here. Each SNR lasts for ~30–33 sec (see text for details).

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqlU5gAAAAA:7ix7UpA0RSVUYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQlMvhMc by MAX PLANCK INSTITUT user on 16 May 2022

linear combination of electrode activation (components) maximally correlated across participants. Correlated component analysis can uncover patterns of neural activity that would not be possible with an electrode-to-electrode correlation method (Cohen & Parra, 2016; Ki et al., 2016; Dmochowski et al., 2012). For mathematical details of the method, see Parra et al. (2019); for available MATLAB scripts, see www.parralab.org/isc/; and for the adapted version of the MATLAB scripts used in the current study, see osf.io/tv7kg/. We calculated the components for each story individually to enable leave-one-out analyses (see below), such that spatial filter calculations are independent from subsequent analyses (see also Broderick, Anderson, & Lalor, 2019; Herrmann, Maess, & Johnsrude, 2018; Crosse, Di Liberto, Bednar, & Lalor, 2016). In line with previous work, we restrict our analysis to the three components with the highest overall ISC for each story (cf. Cohen & Parra, 2016; Ki et al., 2016; Dmochowski et al., 2012, 2014), because they show consistent spatial projections onto the scalp across stories (see Figure 4, left).

Separately for each of the three components (i.e., spatial filters), component weights for one story were multiplied with participants' EEG data for the other three stories. Spatially filtered data for the three stories were then concatenated, leading to one RT course that reflects the component's underlying sources (cf. Broderick et al., 2019; Crosse et al., 2016). This was repeated for all four stories, resulting in four spatially filtered EEG time courses for each component, per participant (4 concatenated datasets × 3 components = 12 spatially filtered time courses). All ISC analyses described in subsequent paragraphs were calculated after using this leave-one-out approach. ISC values were subsequently averaged across the four leave-one out iterations for each component.

For an overall ISC analysis independent of SNR conditions, each participant's time course was correlated with the time course of each of the other participants, resulting in $N - 1$ correlations per participant. ISC for a unique participant was calculated as the mean across these $N - 1$ correlation values. To test whether story listening led to higher ISC than resting state activity in these "story-listening" networks, the EEG data from the resting block were projected through each story's spatial filters using the leave-one-out approach described above. Resting ISC was calculated similarly to story-listening ISC. Mean story-listening ISC was compared to resting ISC using a paired sample *t* test, separately for each of the three components. Note that results from using surrogate-data testing (Lancaster, Iatsenko, Pidde, Ticcinelli, & Stefanovska, 2018), by circularly shifting RT courses of participants, mirrored results from the resting-state contrasts.

To examine ISC as a function of SNR, we concatenated EEG segments with the same SNR and spatially filtered using the above leave-one-out approach. This resulted in a spatially filtered EEG time series for each participant, SNR condition, and spatial-filter component. Separately for each SNR, each participant's time course was correlated with the time course of each of the other participants who heard the same story version (i.e., same order of SNR conditions) and subsequently averaged across the resulting $N - 1$ correlations for each participant. We analyzed ISC using rmANOVAs with SNR (clear and +12, +7, +2, and −3 dB) as one within-participant factor, separately for each of the three components.

Assessment of the Relationship between ISC and Behavioral Measures

To characterize the relationship between ISC and behavioral ratings of absorption (NAS), enjoyment, motivation, and story comprehension, we calculated correlations

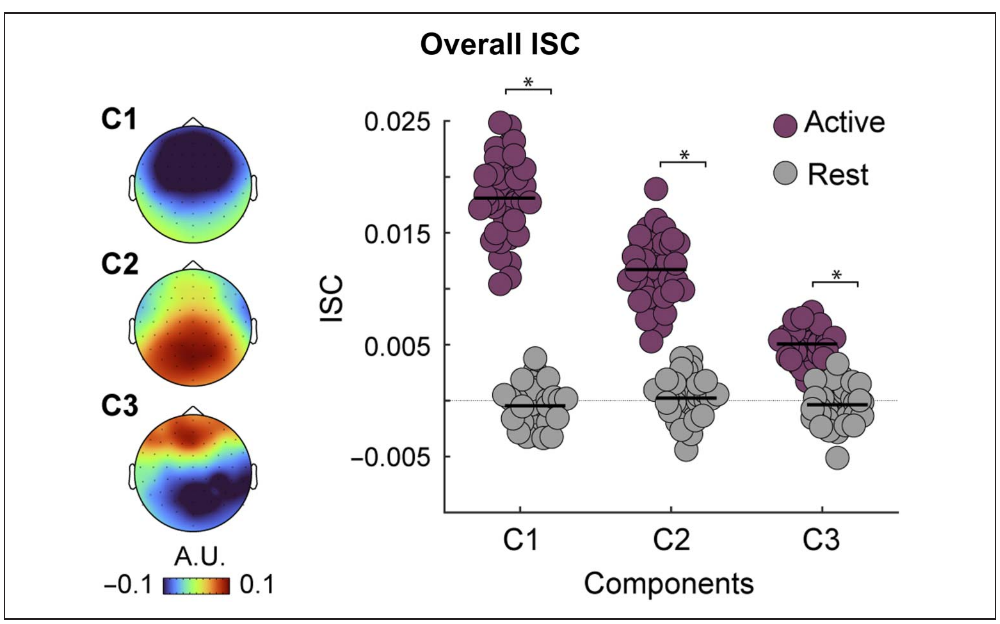

Figure 4. Overall ISC. (Left) Scalp projections of the three most correlated components. (Right) Overall ISC is plotted as a function of component number (Component 1, Component 2, Component 3) and listening condition (active: listening to stories; rest: relaxing with eyes open). Dots represent data points for individual participants. Black solid lines indicate means. The thin, dashed line marks ISC of zero. $^*p < .05$.

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqlU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQIMnMlc by MAX PLANCK INSTITUT user on 16 May 2022

(Pearson) between overall ISC for each of the three components and these behavioral measures.

Comparison of ISC and Speech Intelligibility

To directly compare word-report performance in Experiment 1 with ISC during story listening in Experiment 2, we transformed both word-report scores and ISC values to z scores. We restricted our analysis to the component with the highest ISC (Component 1; see Figure 4, right). To assess the extent to which the effect of SNR differentially impacts intelligibility and ISC, we fit an exponential function to the z-scored word-report and ISC values using the following equation:

$$y = a * e^{(-x*\tau)} + s \quad (1)$$

where y is the exponential function, a is the parameter for amplitude, x is a vector of linearly spaced values (1–5) corresponding to each SNR (clear and +12, +7, +2, and −3 dB SNR), τ (tau) controls the decay strength, and the s parameter allows the function to shift along the y axis. Note that the inclusion of s was essential to fit the exponential function to z-scored data. To detect differences in how strongly intelligibility and ISC declined with SNR, we analyzed the τ (tau) parameter using an independent t test, with experiment measure (intelligibility [Experiment 1], ISC [Experiment 2]) as the grouping variable. To make sure our results are robust to the type of function fit, we also fit a quadratic function to the z-scored word-report and ISC values and compared quadratic and linear coefficients using two separate independent t tests, with experiment measure (intelligibility [Experiment 1], ISC [Experiment 2]) as the grouping variable.

Finally, to capture how each nonclear SNR condition (+12, +7, +2, −3 dB) changed relative to the clear condition, we calculated difference scores between each SNR level and the clear condition, separately for word-report scores and ISC values. Resulting values were submitted to a mixed design ANOVA with measure (intelligibility, ISC) as the between-participant variable and SNR (+12, +7, +2, −3 dB) as the within-participant variable.

Results and Discussion

Comprehension and Engagement Are High Despite Story Masking

For story comprehension, similar to Experiment 1, participants' performance was significantly higher than chance level (M = .86, SE = .01), $t(38) = 28.77, p = 2.1 \times 10^{-27}$, $r_e = .98$, indicating that they attended and grasped details from the stories despite the varying babble masker.

We examined participants' level of story engagement by testing mean ratings for each engagement category (NAS, enjoyment, motivation) against a rating of 4 (neutral response). In contrast to Experiment 1, scores on the NAS, $t(38) = 3.62$, $p = .001$, $r_e = .51$, enjoyment, $t(38) = 7.28, p = 1.02 \times 10^{-8}$, $r_e = .76$, and motivation, $t(38) = 8.9, p = 7.6 \times 10^{-11}, r_e = .82$, were all significantly higher than a neutral response (see Table 2). This indicates that participants were engaged, enjoyed the stories, and felt motivated to listen. As a follow-up, we directly compared behavioral ratings between Experiments 1 and 2 using either separate independent t tests (NAS, enjoyment) or a Wilcoxon rank sum test (motivation), with Experiment as the grouping factor (Experiment 1, Experiment 2). We found that narrative absorption, $t(119) = 2.7, p_{FDR} = .008, r_e = .24$, enjoyment, $t(119) = 4.14, p_{FDR} = 2 \times 10^{-4}, r_e = .35$, and motivation ($z = 2.77$, $p_{FDR} = .008$) were higher in Experiment 2 compared to Experiment 1 (Table 2), perhaps because Experiment 2 was in person, whereas Experiment 1 was online and Experiment 2 did not require participants to provide word reports during the story.

Table 2. Mean Engagement (NAS), Enjoyment, and Motivation Ratings for Experiments 1 and 2

	NAS	*Enjoyment*	*Motivation*
	M ± *SE*	*M* ± *SE*	*M* ± *SE*
Experiment 1	3.98 ± 0.12	4.35 ± 0.12	5.11 ± 0.11
Experiment 2	4.5 ± 0.14	5.23 ± 0.17	5.59 ± 0.18

ISC Is Higher during Active Listening Compared to Rest

Figure 4 (left) shows the scalp projections of the first three components (Parra, Spence, Gerson, & Sajda, 2005): Topographical distributions are consistent with previous work (Cohen & Parra, 2016; Ki et al., 2016; Dmochowski et al., 2012, 2014). The first component has a frontocentral scalp distribution, which is consistent with a source in auditory cortex (Picton, John, Dimitrijevic, & Purcell, 2003; Näätänen & Picton, 1987), but may also reflect contributions from frontal cortex. The second component has a parietal scalp distribution, suggesting a source in parietal cortex or posterior auditory cortex. The third component shows a parieto-occipital peak with frontal polarity reversal.

For each component, we observed ISC values that, although appearing small, were within an expected range reported previously (Ki et al., 2016; Dmochowski et al., 2012, 2014). ISC was stronger during story listening compared to rest for all three components (Component 1: $t(38) = 31.6, p = 7.3 \times 10^{-29}, r_e = .98$; Component 2: $t(38) = 19.6, p = 1.7 \times 10^{-21}, r_e = .95$; Component 3: $t(38) = 16.7, p = 4.6 \times 10^{-19}, r_e = .94$; see right panel). ISC for surrogate data (generated by circularly shifting participants' time courses during the stories) also yielded significantly lower ISC values compared to unshifted (story listening) data for each component. That the ISC was stronger during story listening compared to rest for the three strongest

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqIU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQIMhMlc by MAX PLANCK INSTITUT user on 16 May 2022

spatial components suggests that it is driven by patterns of shared neural activity evoked by the stories.

ISC Declines More Strongly with Challenging SNRs

To examine whether ISC is affected by SNR, we conducted an rmANOVA. Consistent with behavioral performance from Experiment 1, we observed that ISC declined with decreasing SNR for all three components (Component 1: $F(4, 152) = 20.64, p = 1.3 \times 10^{-13}, \eta_p^2 = .35$; Component 2: $F(4, 152) = 18.66, p = 1.7 \times 10^{-12}, \eta_p^2 = .33$; Component 3: $F(4, 152) = 3.22, p = .014, \eta_p^2 = .08$; Figure 5).

We then compared ISC at successive pairs of SNRs to identify the point at which a significant drop in ISC occurred. For Component 1, ISC strength did not differ between clear and + 12 dB SNR ($p_{FDR} = .859$), or between +12 and +7 dB SNR ($p_{FDR} = .703$), but decreased significantly from +7 to +2 dB SNR, $t(38) = 2.7, p_{FDR} = .02, r_e = .4$, and from +2 to −3 dB SNR, $t(38) = 5.7, p_{FDR} = 5.2 \times 10^{-6}, r_e = .68$. An identical pattern was observed for Component 2, such that ISC strength did not differ between clear and + 12 dB SNR ($p_{FDR} = .452$), or between +12 and +7 dB SNR ($p_{FDR} = .498$), but decreased significantly from +7 to +2 dB SNR, $t(38) = 2.54, p_{FDR} = .03, r_e = .38$, and from +2 to −3 dB SNR, $t(38) = 4.72, p_{FDR} = .0001, r_e = .61$. No differences were observed between successive SNRs for Component 3 after correcting for multiple comparisons ($p_{FDR} > .14$). Together, this suggests that ISC for Components 1 and 2 remained fairly stable at moderate SNRs but decreased most substantially at the lowest SNRs. Neural ISC correlates with the degree of engagement with naturalistic materials (Nastase et al., 2019; Nguyen et al., 2019; Yeshurun et al., 2017; Hasson et al., 2010), and our data may thus suggest that listeners remained engaged during naturalistic listening even when about 10% of words were missed (+7 dB SNR; see Figure 2 for intelligibility data) but started disengaging when approximately 20% of words were missed (+2 dB SNR).

Finally, we compared the SNR condition with the lowest ISC (−3 dB SNR) to ISC during rest to determine whether the reduced synchronization observed for the most challenging listening condition was still greater than average synchronization during rest. We observed stronger ISC during the −3 dB SNR condition for Component 1, $t(38) = -21.44, p_{FDR} = 2.4 \times 10^{-22}, r_e = .96$, Component 2, $t(38) = -17.74, p_{FDR} = 8.4 \times 10^{-20}, r_e = .94$, and Component 3, $t(38) = -11.49, p_{FDR} = 6.2 \times 10^{-14}, r_e = .88$. Therefore, although ISC was significantly reduced at −3 dB relative to the other less challenging SNRs, observing stronger ISC at −3 dB compared to rest may suggest participants were still somewhat engaged. These results are consistent with previous behavioral work suggesting that listeners continue to engage with spoken stories despite the presence of masking and reduced speech intelligibility (Herrmann & Johnsrude, 2020b).

Relating ISC and Behavioral Measures of Engagement and Comprehension

Correlations between overall ISC for each component (Component 1, Component 2, Component 3) and behavioral ratings of absorption (NAS), enjoyment, motivation, and story comprehension were not significant after correcting for multiple comparisons ($r_{37} < .3, p_{FDR} > .22$).

Consistent with Experiment 1, we found a significant correlation between NAS and enjoyment ($r_{37} = .81, p_{FDR} = 6.6 \times 10^{-9}$) and between NAS and motivation ($r_{37} = .70, p_{FDR} = 4 \times 10^{-6}$). Higher levels of absorption were associated with greater enjoyment of the stories and motivation to listen. We also found a significant relationship between enjoyment and motivation ($r_{37} = .86, p_{FDR} = 5.8 \times 10^{-11}$): Enjoyment was associated with greater motivation to listen. NAS did not correlate with story comprehension ($r_{37} = .21, p_{FDR} = .192$), and no other significant correlations were found ($r_{37} < .23, p_{FDR} > .36$).

Assessing the Relative Effect of SNR on Intelligibility and ISC

We directly compared the effect of SNR on intelligibility and ISC by fitting an exponential function *z*-scored word-report values and *z*-scored ISC values for the component with the strongest ISC (Component 1) and then compared the resulting decay coefficients between word report and

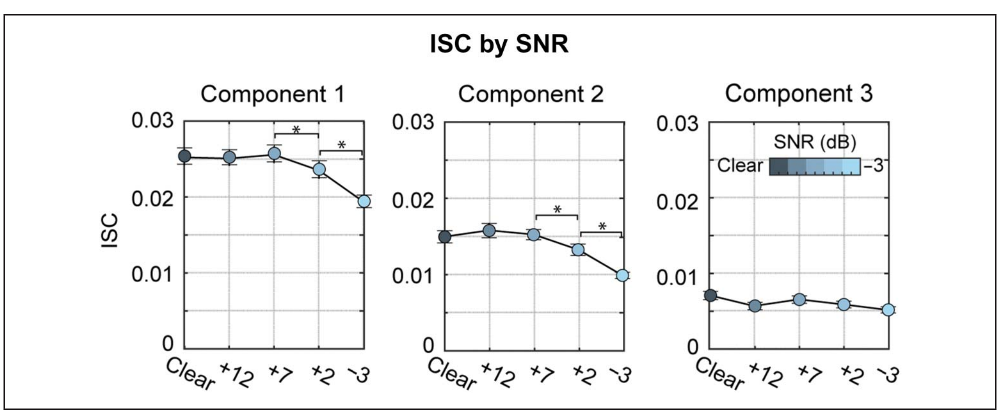

Figure 5. ISC as a function of SNR. (A) Mean ISC is plotted as a function of SNR (clear and +12, +7, +2, and −3 dB) for each component (Component 1, Component 2, Component 3). Error bars reflect *SEM*. *$p < .05$.

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqIU5gAAAAA:7ix7UpA0RSVUYwDQUbkxig2ckok0OB-FZbmU6NaSHzyIARKqQc6fOy0Tn6q3uJGHVmQlMvhMIc by MAX PLANCK INSTITUT user on 16 May 2022

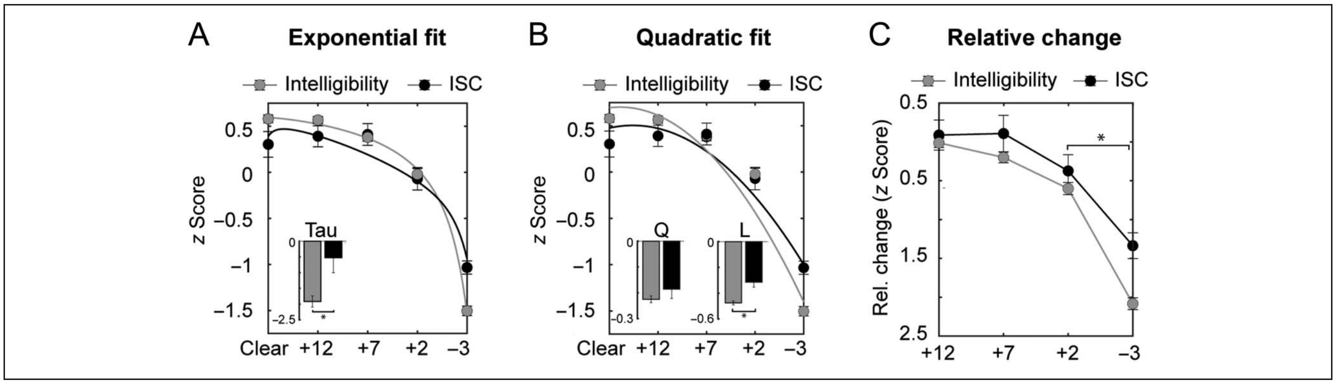

Figure 6. Normalized measures of intelligibility and ISC. (A) Mean normalized word report (Experiment 1) and ISC (Experiment 2: Component 1) are plotted as a function of SNR (clear and +12, +7, +2, and −3 dB). Solid lines correspond to an exponential function fit to normalized means. Mean tau coefficients (decay strength) are plotted as a function of experiment measure (intelligibility, ISC). (B) Mean normalized word report (Experiment 1) and ISC (Experiment 2: Component 1) are replotted as a function of SNR (clear and +12, +7, +2, and −3 dB). Solid lines correspond to a quadratic function fit to normalized means. Mean quadratic (left plot) and linear (right plot) coefficients are plotted as a function of experiment measure (intelligibility, ISC). (C) The change in normalized word report (z score) and ISC for each SNR relative to clear. Error bars reflect *SEM*. Rel. = relative. $^*p < 0.05$.

ISC. We observed a larger τ coefficient for the intelligibility task relative to ISC, $t(119) = -21.44$, $p = .001$, $r_e = .29$, suggesting that speech intelligibility declined more strongly with SNR than ISC (Figure 6A). To confirm that our results are robust to the type of function used to describe the data, we also fit a quadratic function and compared the resulting quadratic and linear coefficients between word report and ISC (Figure 6B). The quadratic coefficients did not differ between word report and ISC ($p = .21$). However, we observed more negative linear coefficients for the intelligibility task relative to ISC, $t(119) = 4.71$, $p = 7 \times 10^{-6}$, $r_e = .4$, further confirming that speech intelligibility declined more strongly with SNR than ISC.

To capture at which SNR intelligibility and ISC diverged, we calculated difference scores, separately for mean-normalized word report and ISC, between each SNR and the clear condition and analyzed the resulting z scores using an ANOVA. The relative change from clear for both word report and ISC decreased significantly as SNR decreased (effect of SNR: $F(3, 357) = 197.5$, $p = 1.8 \times 10^{-75}$), but the decline from 2 to −3 dB SNR was larger for word report than for ISC (Task × SNR interaction: $F(3, 357) = 5.99$, $p = .001$, $\eta_p^2 = .05$; within-subject contrast: $F(1, 119) = 7.76$, $p_{FDR} = .01$, $\eta_p^2 = .06$; see Figure 6C).

GENERAL DISCUSSION

We investigated how acoustic masking affects speech intelligibility, engagement, and neural activity during story listening. In Experiment 1, we observed that speech intelligibility declined with decreasing SNR, such that intelligibility was relatively high (80% or higher) at all but the lowest SNR (−3 dB), at which word-report accuracy declined to 54%. In Experiment 2, we recorded EEG while individuals listened to spoken stories and investigated ISC as a possible window into engagement. We observed that ISC remained stable across all but the lowest SNRs, whereas speech intelligibility measured in Experiment 1 declined more rapidly with decreasing SNR. Our analyses suggest that speech intelligibility is less robust to changes in SNR than synchronized neural activity (ISC). Speech intelligibility may thus be an insufficient predictor of naturalistic speech listening. Our study suggests that individuals may continue to engage with naturalistic, spoken stories even under adverse listening conditions that reduce intelligibility by 10%.

Story Comprehension and Overall Engagement Remain High despite Background Noise

In both experiments, participants scored well above chance on the comprehension questions, indicating that they were able to encode details from the stories despite regular changes in the level of background noise. Participants' self-rated story absorption in Experiment 1 (our behavioral measure of engagement) was neutral, but they reported being motivated and enjoyed listening to the stories. In Experiment 2, participants reported increased levels of absorption, enjoyment, and motivation relative to neutral, which is consistent with previous research suggesting engagement and motivation for interesting narratives are largely resistant to moderate levels of noise (Herrmann & Johnsrude, 2020b). In addition, participants in Experiment 2 reported higher levels of absorption, enjoyment, and motivation relative to Experiment 1 (see Table 2).

The differences in behavioral engagement metrics between experiments may have several explanations. Engagement with narrative materials requires individuals to generate and maintain mental models of events, situations, and characters (Oatley, 2016; Zwaan, 2016; Mar &

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqIU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqOc6fOy0Tn6q3uJGHVmQlMvhMc by MAX PLANCK INSTITUT user on 16 May 2022

Oatley, 2008; Zwaan, Langston, & Graesser, 1995). Disruptions to mental model generation and maintenance may reduce engagement with a narrative (Busselle & Bilandzic, 2009). In Experiment 1, stories were briefly discontinued about every 10–20 sec and participants performed the intelligibility task, whereas stories were played without interruption in Experiment 2. Moreover, Experiment 1 was conducted online, whereas Experiment 2 was conducted in-person in the laboratory, with possibly fewer distractions for participants. Generation and maintenance of a mental model may thus have been more disrupted in Experiments 1 than 2, leading to lower narrative absorption scores in Experiment 1 (Table 2). Alternatively, the different levels of reported engagement between the two experiments may reflect differences in listening strategy. The word-report task in Experiment 1 may have led listeners to adopt a detail-oriented strategy to ensure they could report back specific words when prompted. In contrast, the absence of a dual task in Experiment 2 may have motivated a "gist" or "global" strategy (Harding, Cooke, & König, 2007), in which individuals focus on understanding the message but are not remembering words verbatim. The conditions of Experiment 2, in which people listened without a secondary intelligibility task, are more similar to real life and suggest that, consistent with previous work (Herrmann & Johnsrude, 2020b), engagement with spoken stories—which likely relies on gist perception—is unaffected even if a listener misses 10% of words.

Speech Intelligibility Declines with Decreasing SNR

We utilized spoken stories masked with varied levels of background noise to approximate naturalistic listening conditions and found that speech intelligibility decreased as the SNR decreased. This is consistent with previous studies investigating how speech intelligibility changes with increasing speech degradation or with masking by short speech utterances, such as disconnected sentences (Wild, Davis, & Johnsrude, 2012; Wild, Yusuf, et al., 2012; Akeroyd, 2008; Obleser et al., 2007; Duncan & Aarts, 2006; Davis & Johnsrude, 2003). This is also consistent with other studies using indirect measures of speech intelligibility during naturalistic listening, such as using comprehension or fill-in-the-blank questions (Xia, Kalluri, Micheyl, & Hafter, 2017; Best, Keidser, Buchholz, & Freeston, 2016; MacPherson & Akeroyd, 2013; Power, Foxe, Forde, Reilly, & Lalor, 2012; Humes & Dubno, 2010). One critical difference is that our paradigm directly measures intelligibility by asking participants to report back phrases just after being heard, which allows for a more direct comparison with previous psychoacoustic work using isolated speech utterances.

We further demonstrate that intelligibility was relatively high at moderate SNRs (>80% words reported for +2 dB SNR and higher) but was particularly poor when the SNR was low (~55% words reported for −3 dB SNR). Performance at moderate SNRs may have benefited from the influence of contextual information, as semantic or lexical context can facilitate intelligibility when the speech signal is degraded or an individual has impaired hearing (Holmes, Folkeard, Johnsrude, & Scollie, 2018; Desjardins & Doherty, 2014; Obleser et al., 2007; Dubno, Ahlstrom, & Horwitz, 2000; Pichora-Fuller, Schneider, & Daneman, 1995). This strategy was perhaps not possible for the lowest SNR, at which only about half of the words could be understood—the context gleaned from this proportion of sentences may not be sufficient to enhance comprehension.

ISC Is Stronger When Listening to Stories Compared to Rest

ISC captures the aspects of neural activity that are synchronized across participants, typically while they watch a movie or listen to a spoken story (Dmochowski et al., 2012, 2014; Hasson, Furman, et al., 2008; Hasson et al., 2004). We observed that ISC was stronger during story listening compared to rest for all three spatial components. This is consistent with previous studies showing that ISC is stronger when attending to a coherent narrative compared to when attention is directed elsewhere (Rosenkranz et al., 2021; Cohen et al., 2018; Ki et al., 2016; Kuhlen et al., 2012) or compared to rest (Wilson, Molnar-Szakacs, & Iacoboni, 2008; Hasson et al., 2004). ISC is often thought of as a neural signature of engagement, because ISC is stronger for captivating or exciting stimuli (Schmälzle et al., 2015; Hasson et al., 2010) and predicts behavioral engagement measures (Song et al., 2021; Cohen et al., 2017; Dikker et al., 2017; Poulsen et al., 2017; Dmochowski et al., 2014). Observing significant ISC despite the presence of background noise in the current study may suggest that participants were able to remain engaged despite missing bits of the story during periods when the SNR was low.

Critically, the term "engagement" is not always used in the same manner. In literature and media studies, engagement is typically defined as a multidimensional construct that includes attention, emotional engagement, mental imagery, and "transportation," the experience of being absorbed in the world of the story (Herrmann & Johnsrude, 2020b; Kuijpers et al., 2014; Busselle & Bilandzic, 2008, 2009; Green, Brock, & Kaufman, 2004; Cohen, 2001; see Table 1). Engagement measured as synchronized neural activity during movie watching or story listening is thought to be driven by a shared experience across individuals (Nastase et al., 2019; Nguyen et al., 2019; Yeshurun et al., 2017; Dmochowski et al., 2012; Hasson et al., 2010) and thus may converge with more experiential, cognitive engagement dimensions. Our finding that activity is less synchronized across participants when they rest and their minds are not focused on a story (Figure 4) may suggest that EEG-recorded ISC is perhaps sensitive to one or more of the aforementioned

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqlU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQlMvhMlc by MAX PLANCK INSTITUT user on 16 May 2022

experiential dimensions of engagement. However, this remains to be more directly addressed in future research.

ISC Is More Robust to Changes in SNR than Speech Intelligibility

We observed that ISC was reduced for speech listening during challenging SNRs, but not for moderate SNRs. Specifically, ISC did not differ between clear, +12 dB, and +7 dB SNR, despite our separate behavioral experiment indicating participants missed approximately 10% of words for +7 dB SNR (Figure 2). However, ISC significantly declined from +7 to +2 dB SNR and from +2 to −3 dB SNR. Despite similar overall patterns, comparing ISC and intelligibility directly demonstrated that word-report performance declined more strongly with decreasing SNR compared to ISC (Figure 6A–C). This is consistent with related work that investigated the relation between speech intelligibility and neural synchronization with the amplitude envelope of speech. Synchronization with the speech envelope appears robust to moderate changes in SNR or the presence of a competing speech stream, while intelligibility declines (Ding & Simon, 2012, 2013). Additionally, whereas ISC was lowest at −3 dB SNR, ISC for all three components was still stronger than during rest, which may suggest participants were still somewhat engaged when speech intelligibility was ~55%. This is consistent with a recent behavioral study demonstrating that listeners were as absorbed during a story that was continuously masked by babble at +4 dB SNR as they were during a clear story, despite much higher rated listening effort for the former (Herrmann & Johnsrude, 2020b). The current data provide neural evidence that across-participant synchronization during story listening is unaffected by masking that reduces intelligibility to about 90% but begins to decrease when intelligibility reduces to 80%. We speculate that observing synchronized activity despite acoustic masking may have resulted from participants accurately perceiving the story gist, and enjoying it, which perpetuated their intrinsic motivation to continue listening and stay engaged (Herrmann & Johnsrude, 2020a; Eckert et al., 2016; Matthen, 2016; Richter, 2016).

Listening and trying to understand masked speech is supported by processes such as sustained and selective attention (Wild, Yusuf, et al., 2012; Davis & Johnsrude, 2007; Obleser et al., 2007), which, when recruited, result in effortful listening (Pichora-Fuller et al., 2016). When the SNR is so poor that the gist is lost, or attentional resources become depleted (Hornsby, Naylor, & Bess, 2016; Ivarsson & Arlinger, 1994), this may lead to temporary disengagement or attentional lapses (Heffernan et al., 2016; Hallberg & Carlsson, 1991). This is consistent with our observation of lowest, but still significant, ISC at the most difficult SNR (−3 dB SNR) for which participants missed about 45% of words on average. Our study suggests that a listener can engage and follow a story's thread as long as intelligibility is over 80%. However, one limitation of the current study is that intelligibility and EEG-measured ISC were assessed in separate groups of participants. Future studies may consider using a within-participant design to more directly assess how intelligibility relates to synchronized activity across participants.

Relating ISC with Behavioral Measures of Engagement

To assess the relationship between the behavioral and neural measures of engagement, we calculated correlations between the overall ISC for each component (Component 1, Component 2, Component 3) and behavioral ratings of absorption (NAS), enjoyment, and motivation but did not observe any significant relationships after correcting for multiple statistical comparisons ($r_{37} < .3$, $p_{FDR} > .22$). This is inconsistent with previous literature that demonstrates that ISC strength is predictive of engagement (Song et al., 2021; Cohen et al., 2017; Dikker et al., 2017; Poulsen et al., 2017; Dmochowski et al., 2014) as well as other measures dependent on sustained attention, such as subsequent recall of the materials (Piazza et al., 2021; Song et al., 2021; Chan et al., 2019; Davidesco et al., 2019; Cohen et al., 2018; Cohen & Parra, 2016; Stephens et al., 2010; Hasson, Furman, et al., 2008). We speculate that the behavioral measure of engagement did not correlate with ISC because we did not continuously assess changes in behavioral engagement throughout each story, as has been done previously (Song et al., 2021; Cohen et al., 2017; Dmochowski et al., 2014). Post hoc measures of behavioral engagement (as we used here) have also been used on occasion: for example, relating ISC with engagement in the classroom (Dikker et al., 2017; Poulsen et al., 2017), but the listening conditions did not vary systematically, as in the current study. We changed SNR every 30 sec—this may have led to dynamic changes in engagement (especially at high masking levels), which would be reflected in ISC but which would be difficult to capture behaviorally with a single post hoc measure. Furthermore, in contrast to previous work, we used a story absorption scale (Herrmann & Johnsrude, 2020b; Kuijpers et al., 2014) as a behavioral engagement measure, but whether this accurately reflects neural engagement is a matter for further research. It would be helpful to combine our story paradigm with a continuous measure of behavioral engagement to further explore this topic (cf. Song et al., 2021).

Using Narratives to Approximate Realistic Listening Scenarios

Everyday listening situations involve speech material composed of sentences that relate to each other and that are interesting to the listener. Such listening situations are commonly subject to increased levels of background noise but are also rich with many positive aspects of listening that may drive motivation to listen (Herrmann &

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqIU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQlMvhMlc by MAX PLANCK INSTITUT user on 16 May 2022

Johnsrude, 2020a; Matthen, 2016). Similarly, the spoken stories used here are intrinsically motivating to a listener, reflecting such everyday listening situations (Dunlop & Walker, 2013). Our work shows, in line with recent studies (Polonenko & Maddox, 2021; Broderick, Di Liberto, Anderson, Rofes, & Lalor, 2020; Brodbeck, Jiao, Hong, & Simon, 2020; Erb, Schmitt, & Obleser, 2020; Broderick et al., 2019; Broderick, Anderson, Di Liberto, Crosse, & Lalor, 2018), that utilizing naturalistic, spoken stories to investigate speech listening provides a useful avenue to investigate listening in ecologically valid conditions.

Conclusions

We utilized naturalistic, spoken stories to investigate how challenging listening affects intelligibility and listener engagement. We used ISC of EEG activity as a measure of listener engagement. ISC was unaffected by moderately challenging SNRs, despite speech intelligibility scores suggesting listeners missed approximately 10% of words. Our data further show that speech intelligibility declines faster than electrophysiologically measured listening engagement under masked speech conditions. Our work provides a potentially fruitful approach with naturalistic, spoken stories to investigate intelligibility and engagement during masked speech listening. Our approach may importantly open future avenues to investigate (dis) engagement in older adults with hearing impairment.

Acknowledgments

This research was supported by the Canadian Institutes of Health Research (MOP133450 to I. S. Johnsrude). B. H. was supported by a BrainsCAN Tier I postdoctoral fellowship (Canada First Research Excellence Fund) and the Canada Research Chair program.

Reprint requests should be sent to Vanessa C. Irsik, The Brain and Mind Institute, The University of Western Ontario, London, ON N6A 5B7, Canada, or via e-mail: virsik@uwo.ca.

Author Contributions

Vanessa C. Irsik: Conceptualization; Data curation; Formal analysis; Methodology; Visualization; Writing—Original draft; Writing—Review & editing. Ingrid S. Johnsrude: Conceptualization; Writing—Review & editing. Björn Herrmann: Conceptualization; Methodology; Visualization; Writing—Review & editing.

Funding Information

Ingrid S. Johnsrude, Canadian Institutes of Health Research (https://dx.doi.org/10.13039/501100000024), grant number: MOP133450. Björn Herrmann: Canada Research Chairs (https://dx.doi.org/10.13039/501100001804).

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance.

REFERENCES

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, *47(Suppl. 2)*, S53–S71. https://doi.org/10.1080/14992020802301142, PubMed: 19012113

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B: Methodological*, *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berinsky, A. J., Margolis, M. F., & Sances, M. W. (2014). Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *American Journal of Political Science*, *58*, 739–753. https://doi.org/10.1111/ajps.12081

Best, V., Keidser, G., Buchholz, J. M., & Freeston, K. (2016). Development and preliminary evaluation of a new test of ongoing speech comprehension. *International Journal of Audiology*, *55*, 45–52. https://doi.org/10.3109/14992027.2015.1055835, PubMed: 26158403

Bilger, R. C. (1984). *Manual for the clinical use of the Revised SPIN test*. University of Illinois Press.

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology*, *40*, 109–131. https://doi.org/10.1146/annurev.ps.40.020189.000545, PubMed: 2648973

Brodbeck, C., Jiao, A., Hong, L. E., & Simon, J. Z. (2020). Neural speech restoration at the cocktail party: Auditory cortex recovers masked speech of both attended and ignored speakers. *PLoS Biology*, *18*, 1–22. https://doi.org/10.1371/journal.pbio.3000883, PubMed: 33091003

Broderick, M. P., Anderson, A. J., Di Liberto, G. M., Crosse, M. J., & Lalor, E. C. (2018). Electrophysiological correlates of semantic dissimilarity reflect the comprehension of natural, narrative speech. *Current Biology*, *28*, 803–809. https://doi.org/10.1016/j.cub.2018.01.080, PubMed: 29478856

Broderick, M. P., Anderson, A. J., & Lalor, E. C. (2019). Semantic context enhances the early auditory encoding of natural speech. *Journal of Neuroscience*, *39*, 7564–7575. https://doi.org/10.1523/JNEUROSCI.0584-19.2019, PubMed: 31371424

Broderick, M. P., Di Liberto, G., Anderson, A., Rofes, A., & Lalor, E. (2020). Dissociable electrophysiological measures of natural language processing reveal differences in speech comprehension strategy in healthy ageing. *Scientific Reports*, *11*, 1–12. https://doi.org/10.1101/2020.04.17.046201, PubMed: 33091003

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqlU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzyIARKqQc6fOy0Tn6q3uJGHVmQlMvhMlc by MAX PLANCK INSTITUT user on 16 May 2022

Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica united with Acustica*, *86*, 117–128.
Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, *109*, 1101–1109. https://doi.org/10.1121/1.1345696, PubMed: 11303924
Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *Journal of the Acoustical Society of America*, *110*, 2527–2538. https://doi.org/10.1121/1.1408946, PubMed: 11757942
Buchanan, E. M., & Scofield, J. E. (2018). Methods to detect low quality data and its implication for psychological research. *Behavior Research Methods*, *50*, 2586–2596. https://doi.org/10.3758/s13428-018-1035-6, PubMed: 29542063
Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*, 3–5. https://doi.org/10.1177/1745691610393980, PubMed: 26162106
Busselle, R., & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory*, *18*, 255–280. https://doi.org/10.1111/j.1468-2885.2008.00322.x
Busselle, R., & Bilandzic, H. (2009). Measuring narrative engagement. *Media Psychology*, *12*, 321–347. https://doi.org/10.1080/15213260903287259
Chan, H. Y., Smidts, A., Schoots, V. C., Dietvorst, R. C., & Boksem, M. A. S. (2019). Neural similarity at temporal lobe and cerebellum predicts out-of-sample preference and recall for video stimuli. *Neuroimage*, *197*, 391–401. https://doi.org/10.1016/j.neuroimage.2019.04.076, PubMed: 31051296
Cohen, J. (2001). Defining identification: A theoretical look at the identification of audiences with media characters. *Mass Communication and Society*, *4*, 245–264. https://doi.org/10.1207/S15327825MCS0403_01
Cohen, S. S., Henin, S., & Parra, L. C. (2017). Engaging narratives evoke similar neural activity and lead to similar time perception. *Scientific Reports*, *7*, 1–10. https://doi.org/10.1038/s41598-017-04402-4, PubMed: 28676688
Cohen, S. S., Madsen, J., Touchan, G., Robles, D., Lima, S. F. A., Henin, S., et al. (2018). Neural engagement with online educational videos predicts learning performance for individual students. *Neurobiology of Learning and Memory*, *155*, 60–64. https://doi.org/10.1016/j.nlm.2018.06.011, PubMed: 29953947
Cohen, S. S., & Parra, L. C. (2016). Memorable audiovisual narratives synchronize sensory and supramodal neural responses. *eNeuro*, *3*, 1–11. https://doi.org/10.1523/ENEURO.0203-16.2016, PubMed: 27844062
Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The multivariate temporal response function (mTRF) toolbox: A MATLAB toolbox for relating neural signals to continuous stimuli. *Frontiers in Human Neuroscience*, *10*, 1–14. https://doi.org/10.3389/fnhum.2016.00604, PubMed: 27965557
Davidesco, I., Laurent, E., Valk, H., West, T., Dikker, S., Milne, C., et al. (2019). Brain-to-brain synchrony between students and teachers predicts learning outcomes. *BioRxiv*. https://doi.org/10.1101/644047
Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *Journal of Neuroscience*, *23*, 3423–3431. https://doi.org/10.1523/JNEUROSCI.23-08-03423.2003, PubMed: 12716950
Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: Top–down influences on the interface between audition and speech perception. *Hearing Research*, *229*, 132–147. https://doi.org/10.1016/j.heares.2007.01.014, PubMed: 17317056
Dawes, P., Emsley, R., Cruickshanks, K. J., Moore, D. R., Fortnum, H., Edmondson-Jones, M., et al. (2015). Hearing loss and cognition: The role of hearing aids, social isolation and depression. *PLoS One*, *10*, 1–9. https://doi.org/10.1371/journal.pone.0119616, PubMed: 25760329
de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*, 1–12. https://doi.org/10.3758/s13428-014-0458-y, PubMed: 24683129
Desjardins, J. L., & Doherty, K. A. (2014). The effect of hearing aid noise reduction on listening effort in hearing-impaired adults. *Ear and Hearing*, *35*, 600–610. https://doi.org/10.1097/AUD.0000000000000028, PubMed: 24622352
Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., et al. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*, *27*, 1375–1380. https://doi.org/10.1016/j.cub.2017.04.002, PubMed: 28457867
Ding, N., & Simon, J. Z. (2012). Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *Journal of Neurophysiology*, *107*, 78–89. https://doi.org/10.1152/jn.00297.2011, PubMed: 21975452
Ding, N., & Simon, J. Z. (2013). Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *Journal of Neuroscience*, *33*, 5728–5735. https://doi.org/10.1523/JNEUROSCI.5297-12.2013, PubMed: 23536086
Dmochowski, J. P., Bezdek, M. A., Abelson, B. P., Johnson, J. S., Schumacher, E. H., & Parra, L. C. (2014). Audience preferences are predicted by temporal reliability of neural processing. *Nature Communications*, *5*, 1–9. https://doi.org/10.1038/ncomms5567, PubMed: 25072833
Dmochowski, J. P., Sajda, P., Dias, J., & Parra, L. C. (2012). Correlated components of ongoing EEG point to emotionally laden attention—A possible marker of engagement? *Frontiers in Human Neuroscience*, *6*, 1–9. https://doi.org/10.3389/fnhum.2012.00112, PubMed: 22623915
Dubno, J. R., Ahlstrom, J. B., & Horwitz, A. R. (2000). Use of context by young and aged adults with normal hearing. *Journal of the Acoustical Society of America*, *107*, 538–546. https://doi.org/10.1121/1.428322, PubMed: 10641662
Duncan, K. R., & Aarts, N. L. (2006). A comparison of the HINT and Quick SIN tests. *Journal of Speech-Language Pathology and Audiology*, *30*, 86–94.
Dunlop, W. L., & Walker, L. J. (2013). The life story: Its development and relation to narration and personal identity. *International Journal of Behavioral Development*, *37*, 235–247. https://doi.org/10.1177/0165025413479475
Eckert, M. A., Teubner-Rhodes, S., & Vaden, K. I. (2016). Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. *Ear and Hearing*, *37*, 101S–110S. https://doi.org/10.1097/AUD.0000000000000300, PubMed: 27355759
Erb, J., Schmitt, L. M., & Obleser, J. (2020). Temporal selectivity declines in the aging human auditory cortex. *eLife*, *9*, e55300. https://doi.org/10.7554/eLife.55300, PubMed: 32618270
Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hearing Research*, *106*, 95–104. https://doi.org/10.1016/S0378-5955(97)00006-3, PubMed: 9112109
Gordon-Salant, S. (2006). Speech perception and auditory temporal processing performance by older listeners: Implications for real-world communication. *Seminars in Hearing*, *27*, 264–268. https://doi.org/10.1055/s-2006-954852
Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American*

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqlU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQIMvhMlc by MAX PLANCK INSTITUT user on 16 May 2022

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqlU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQlMnMlc by MAX PLANCK INSTITUT user on 16 May 2022

Psychologist, *59*, 93–104. https://doi.org/10.1037/0003-066X.59.2.93, PubMed: 14992636

Green, M. C., Brock, T. C., & Kaufman, G. F. (2004). Understanding media enjoyment: The role of transportation into narrative worlds. *Communication Theory*, *14*, 311–327. https://doi.org/10.1111/j.1468-2885.2004.tb00317.x

Hallberg, L. R., & Carlsson, S. G. (1991). A qualitative study of strategies for managing a hearing impairment. *British Society of Audiology*, *25*, 201–211. https://doi.org/10.3109/03005369109079853, PubMed: 1873586

Harding, S., Cooke, M., & König, P. (2007). Auditory gist perception: An alternative to attentional selection of auditory streams? In L. Paletta & E. Rome (Eds.), *Attention in cognitive systems. Theories and systems from an interdisciplinary viewpoint* (Vol. 4840, pp. 399–416). Springer. https://doi.org/10.1007/978-3-540-77343-6_26

Hasson, U., Furman, O., Clark, D., Dudai, Y., & Davachi, L. (2008). Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron*, *57*, 452–462. https://doi.org/10.1016/j.neuron.2007.12.009, PubMed: 18255037

Hasson, U., Malach, R., & Heeger, D. J. (2010). Reliability of cortical activity during natural stimulation. *Trends in Cognitive Sciences*, *14*, 40–48. https://doi.org/10.1016/j.tics.2009.10.011, PubMed: 20004608

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, *303*, 1634–1640. https://doi.org/10.1126/science.1089506, PubMed: 15016991

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *Journal of Neuroscience*, *28*, 2539–2550. https://doi.org/10.1523/JNEUROSCI.5487-07.2008, PubMed: 18322098

Heffernan, E., Coulson, N. S., Henshaw, H., Barry, J. G., & Ferguson, M. A. (2016). Understanding the psychosocial experiences of adults with mild–moderate hearing loss: An application of Leventhal's self-regulatory model. *International Journal of Audiology*, *55*, S3–S12. https://doi.org/10.3109/14992027.2015.1117663, PubMed: 26754550

Herrmann, B., & Johnsrude, I. S. (2020a). A model of listening engagement (MoLE). *Hearing Research*, *397*, 108016. https://doi.org/10.1016/j.heares.2020.108016, PubMed: 32680706

Herrmann, B., & Johnsrude, I. S. (2020b). Absorption and enjoyment during listening to acoustically masked stories. *Trends in Hearing*, *24*, 1–18. https://doi.org/10.1177/2331216520967850, PubMed: 33143565

Herrmann, B., Maess, B., & Johnsrude, I. S. (2018). Aging affects adaptation to sound-level statistics in human auditory cortex. *Journal of Neuroscience*, *38*, 1989–1999. https://doi.org/10.1523/JNEUROSCI.1489-17.2018, PubMed: 29358362

Holmes, E., Folkeard, P., Johnsrude, I. S., & Scollie, S. (2018). Semantic context improves speech intelligibility and reduces listening effort for listeners with hearing impairment. *International Journal of Audiology*, *57*, 483–492. https://doi.org/10.1080/14992027.2018.1432901, PubMed: 29415585

Hornsby, B. W. Y., Naylor, G., & Bess, F. H. (2016). A taxonomy of fatigue concepts and their relation to hearing loss. *Ear and Hearing*, *37*, 136S–144S. https://doi.org/10.1097/AUD.0000000000000289, PubMed: 27355763

Humes, L. E., & Dubno, J. R. (2010). Factors affecting speech understanding in older adults. In S. Gordon-Salant, R. D. Frisina, A. N. Popper, & R. R. Fay (Eds.), *The aging auditory system* (pp. 211–257). Springer. https://doi.org/10.1007/978-1-4419-0993-0_8

Iotzov, I., & Parra, L. C. (2019). EEG can predict speech intelligibility. *Journal of Neural Engineering*, *16*, 036008. https://doi.org/10.1088/1741-2552/ab07fe, PubMed: 30776785

Ivarsson, U. S., & Arlinger, S. D. (1994). Speech recognition in noise before and after a work-day's noise exposure. *Scandinavian Audiology*, *23*, 159–163. https://doi.org/10.3109/01050399409047502, PubMed: 7997832

Ki, J. J., Kelly, S. P., & Parra, L. C. (2016). Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. *Journal of Neuroscience*, *36*, 3092–3101. https://doi.org/10.1523/JNEUROSCI.2942-15.2016, PubMed: 26961961

Kuhlen, A. K., Allefeld, C., & Haynes, J. D. (2012). Content-specific coordination of listeners' to speakers' EEG during communication. *Frontiers in Human Neuroscience*, *6*, 1–15. https://doi.org/10.3389/fnhum.2012.00266, PubMed: 23060770

Kuijpers, M. M., Hakemulder, F., Tan, E. S., & Doicaru, M. M. (2014). Exploring absorbing reading experiences: Developing and validating a self-report scale to measure story world absorption. *Scientific Study of Literature*, *4*, 89–122. https://doi.org/10.1075/ssol.4.1.05kui

Lancaster, G., Iatsenko, D., Pidde, A., Ticcinelli, V., & Stefanovska, A. (2018). Surrogate data for hypothesis testing of physical systems. *Physics Reports*, *748*, 1–60. https://doi.org/10.1016/j.physrep.2018.06.001

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*, 433–442. https://doi.org/10.3758/s13428-016-0727-z, PubMed: 27071389

MacPherson, A., & Akeroyd, M. A. (2013). The Glasgow Monitoring of Uninterrupted Speech Task (GMUST): A naturalistic measure of speech intelligibility in noise. *Proceedings of Meetings on Acoustics*, *19*, 1–6. https://doi.org/10.1121/1.4799865

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic signals. In D. Touretzky, M. Mouzer, & M. Hasselmo (Eds.), *Advances in neural information processing systems* (Vol. 8, pp. 145–151). Cambridge, MA: MIT Press.

Mandler, J. M., & Goodman, M. S. (1982). On the psychological validity of story structure. *Journal of Verbal Learning and Verbal Behavior*, *21*, 507–523. https://doi.org/10.1016/S0022-5371(82)90746-0

Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, *3*, 173–192. https://doi.org/10.1111/j.1745-6924.2008.00073.x

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44*, 1–23. https://doi.org/10.3758/s13428-011-0124-6, PubMed: 21717266

Matthen, M. (2016). Effort and displeasure in people who are hard of hearing. *Ear and Hearing*, *37(Suppl. 1)*, 28S–34S. https://doi.org/10.1097/AUD.0000000000000292, PubMed: 27355767

Näätänen, R., & Picton, T. W. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, *24*, 375–425. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x, PubMed: 3615753

Naci, L., Cusack, R., Anello, M., & Owen, A. M. (2014). A common neural code for similar conscious experiences in different individuals. *Proceedings of the National Academy of Sciences, U.S.A.*, *111*, 14277–14282. https://doi.org/10.1073/pnas.1407007111, PubMed: 25225384

Nastase, S. A., Gazzola, V., Hasson, U., & Keysers, C. (2019). Measuring shared responses across subjects using intersubject correlation. *Social Cognitive and Affective Neuroscience*, *14*, 669–687. https://doi.org/10.1093/scan/nsz037, PubMed: 31099394

Nguyen, M., Vanderwal, T., & Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. *Neuroimage*, *184*, 161–170. https://doi.org/10.1016/j.neuroimage.2018.09.010, PubMed: 30217543

Nummenmaa, L., Glerean, E., Viinikainen, M., Jääskeläinen, I. P., Hari, R., & Sams, M. (2012). Emotions promote social interaction by synchronizing brain activity across individuals. *Proceedings of the National Academy of Sciences, U.S.A.*, *109*, 9599–9604. https://doi.org/10.1073/pnas.1206095109, PubMed: 22623534

Oatley, K. (2016). Fiction: Simulation of social worlds. *Trends in Cognitive Sciences*, *20*, 618–628. https://doi.org/10.1016/j.tics.2016.06.002, PubMed: 27449184

Obleser, J., Wise, R. J. S., Dresner, M. A., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *Journal of Neuroscience*, *27*, 2283–2289. https://doi.org/10.1523/JNEUROSCI.4663-06.2007, PubMed: 17329425

Olsen, W. O. (1998). Average speech levels and spectra in various speaking/listening conditions: A summary of the Pearson, Bennett, & Fidell (1977) report. *American Journal of Audiology*, *7*, 21–25. https://doi.org/10.1044/1059-0889(1998/012), PubMed: 26649514

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 156869. https://doi.org/10.1155/2011/156869, PubMed: 21253357

Palmer, A. D., Newsom, J. T., & Rook, K. S. (2016). How does difficulty communicating affect the social relationships of older adults? An exploration using data from a national survey. *Journal of Communication Disorders*, *62*, 131–146. https://doi.org/10.1016/j.jcomdis.2016.06.002, PubMed: 27420152

Parra, L. C., Haufe, S., & Dmochowski, J. P. (2019). Correlated components analysis—Extracting reliable dimensions in multivariate data. *ArXiv*. https://arxiv.org/abs/1801.08881

Parra, L. C., Spence, C. D., Gerson, A. D., & Sajda, P. (2005). Recipes for the linear analysis of EEG. *Neuroimage*, *28*, 326–341. https://doi.org/10.1016/j.neuroimage.2005.05.032, PubMed: 16084117

Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear and Hearing*, *39*, 204–214. https://doi.org/10.1097/AUD.0000000000000494, PubMed: 28938250

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex*, *23*, 1378–1387. https://doi.org/10.1093/cercor/bhs118, PubMed: 22610394

Piazza, E. A., Cohen, A., Trach, J., & Lew-Williams, C. (2021). Neural synchrony predicts children's learning of novel words. *Cognition*, *214*, 104752. https://doi.org/10.1016/j.cognition.2021.104752, PubMed: 33965782

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., et al. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and Hearing*, *37(Suppl. 1)*, 5S–27S. https://doi.org/10.1097/AUD.0000000000000312, PubMed: 27355771

Pichora-Fuller, M. K., Mick, P., & Reed, M. (2015). Hearing, cognition, and healthy aging: Social and public health implications of the links between age-related declines in hearing and cognition. *Seminars in Hearing*, *36*, 122–139. https://doi.org/10.1055/s-0035-1555116, PubMed: 27516713

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America*, *97*, 593–608. https://doi.org/10.1121/1.412282, PubMed: 7860836

Picton, T. W., John, M. S., Dimitrijevic, A., & Purcell, D. (2003). Human auditory steady-state responses. *International Journal of Audiology*, *42*, 177–219. https://doi.org/10.3109/14992020309101316, PubMed: 12790346

Polonenko, M. J., & Maddox, R. K. (2021). Exposing distinct subcortical components of the auditory brainstem response evoked by continuous naturalistic speech. *eLife*, *10*, e62329. https://doi.org/10.7554/eLife.62329, PubMed: 33594974

Poulsen, A. T., Kamronn, S., Dmochowski, J., Parra, L. C., & Hansen, L. K. (2017). EEG in the classroom: Synchronised neural recordings during video presentation. *Scientific Reports*, *7*, 1–9. https://doi.org/10.1038/srep43916, PubMed: 28266588

Power, A. J., Foxe, J. J., Forde, E. J., Reilly, R. B., & Lalor, E. C. (2012). At what time is the cocktail party? A late locus of selective attention to natural speech. *European Journal of Neuroscience*, *35*, 1497–1503. https://doi.org/10.1111/j.1460-9568.2012.08060.x, PubMed: 22462504

Reitan, R. M., & Wolfson, D. (2000). Conation: A neglected aspect of neuropsychological functioning. *Archives of Clinical Neuropsychology*, *15*, 443–453. https://doi.org/10.1016/S0887-6177(99)00043-8, PubMed: 14590220

Richter, M. (2013). A closer look into the multi-layer structure of motivational intensity theory. *Social and Personality Psychology Compass*, *7*, 1–12. https://doi.org/10.1111/spc3.12007

Richter, M. (2016). The moderating effect of success importance on the relationship between listening demand and listening effort. *Ear and Hearing*, *37*, 111S–117S. https://doi.org/10.1097/AUD.0000000000000295, PubMed: 27355760

Rosenkranz, M., Holtze, B., Jaeger, M., & Debener, S. (2021). EEG-based intersubject correlations reflect selective attention in a competing speaker scenario. *Frontiers in Neuroscience*, *15*, 1–12. https://doi.org/10.3389/fnins.2021.685774, PubMed: 34194296

Rosenthal, R., & Rubin, D. B. (2003). r equivalent: A simple effect size indicator. *Psychological Methods*, *8*, 492–496. https://doi.org/10.1037/1082-989X.8.4.492, PubMed: 14664684

Schmälzle, R., Häcker, F. E. K., Honey, C. J., & Hasson, U. (2015). Engaged listeners: Shared neural processing of powerful political speeches. *Social Cognitive and Affective Neuroscience*, *10*, 1137–1143. https://doi.org/10.1093/scan/nsu168, PubMed: 25653012

Smeds, K., Wolters, F., & Rung, M. (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *Journal of the American Academy of Audiology*, *26*, 183–196. https://doi.org/10.3766/jaaa.26.2.7, PubMed: 25690777

Song, H., Finn, E. S., & Rosenberg, M. D. (2021). Neural signatures of attentional engagement during narratives and its consequences for event memory. *Proceedings of the National Academy of Sciences, U.S.A.*, *118*, e2021905118. https://doi.org/10.1073/pnas.2021905118, PubMed: 34385312

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker–listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences, U.S.A.*, *107*, 14425–14430. https://doi.org/10.1073/pnas.1008662107, PubMed: 20660768

Thomas, K. A., & Clifford, S. (2017). Validity and Mechanical Turk: An assessment of exclusion methods and interactive experiments. *Computers in Human Behavior*, *77*, 184–197. https://doi.org/10.1016/j.chb.2017.08.038

Wayne, R. V., & Johnsrude, I. S. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing Research Reviews*, *23*, 154–166. https://doi.org/10.1016/j.arr.2015.06.002, PubMed: 26123097

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqIU5gAAAAA:7ix7tJpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQlMhMlc by MAX PLANCK INSTITUT user on 16 May 2022

Wild, C. J., Davis, M. H., & Johnsrude, I. S. (2012). Human auditory cortex is sensitive to the perceived clarity of speech. *Neuroimage*, *60*, 1490–1502. https://doi.org/10.1016/j.neuroimage.2012.01.035, PubMed: 22248574

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: The processing of degraded speech depends critically on attention. *Journal of Neuroscience*, *32*, 14010–14021. https://doi.org/10.1523/JNEUROSCI.1528-12.2012, PubMed: 23035108

Wilson, S. M., Molnar-Szakacs, I., & Iacoboni, M. (2008). Beyond superior temporal cortex: Intersubject correlations in narrative speech comprehension. *Cerebral Cortex*, *18*, 230–242. https://doi.org/10.1093/cercor/bhm049, PubMed: 17504783

Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, & Psychophysics*, *79*, 2064–2072. https://doi.org/10.3758/s13414-017-1361-2, PubMed: 28695541

Wright, R. A. (2014). Presidential address 2013: Fatigue influence on effort-considering implications for self-regulatory restraint. *Motivation and Emotion*, *38*, 183–195. https://doi.org/10.1007/s11031-014-9406-5

Xia, J., Kalluri, S., Micheyl, C., & Hafter, E. (2017). Continued search for better prediction of aided speech understanding in multi-talker environments. *Journal of the Acoustical Society of America*, *142*, 2386–2399. https://doi.org/10.1121/1.5008498, PubMed: 29092591

Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C. J., et al. (2017). Same story, different story: The neural representation of interpretive frameworks. *Psychological Science*, *28*, 307–319. https://doi.org/10.1177/0956797616682029, PubMed: 28099068

Zwaan, R. A. (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin and Review*, *23*, 1028–1034. https://doi.org/10.3758/s13423-015-0864-x, PubMed: 26088667

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*, 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/6/933/2019254/jocn_a_01842.pdf?casa_token=Yu5sBrqIU5gAAAAA:7ix7UpA0RSVJYwDQUbkxig2ckok0OB-FZbmU6NaSHzylARKqQc6fOy0Tn6q3uJGHVmQIMmMIc by MAX PLANCK INSTITUT user on 16 May 2022