Brain Research 1844 (2024) 149166

Contents lists available at ScienceDirect

Brain Research

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

Functional benefits of continuous vs. categorical listening strategies on the neural encoding and perception of noise-degraded speech

Rose Rizzi [a,b], Gavin M. Bidelman [a,b,c,*]

[a] Department of Speech, Language and Hearing Sciences, Indiana University, Bloomington, IN, USA
[b] Program in Neuroscience, Indiana University, Bloomington, IN, USA
[c] Cognitive Science Program, Indiana University, Bloomington, IN, USA

ARTICLE INFO

Keywords:
Categorical perception
event-related potentials (ERPs)
speech in noise (SIN)
visual analog scale (VAS)
Listening strategy

ABSTRACT

Acoustic information in speech changes continuously, yet listeners form discrete perceptual categories to ease the demands of perception. Being a more continuous/gradient as opposed to a more discrete/categorical listener may be further advantageous for understanding speech in noise by increasing perceptual flexibility and resolving ambiguity. The degree to which a listener's responses to a continuum of speech sounds are categorical versus continuous can be quantified using visual analog scaling (VAS) during speech labeling tasks. Here, we recorded event-related brain potentials (ERPs) to vowels along an acoustic–phonetic continuum (/u/ to /a/) while listeners categorized phonemes in both clean and noise conditions. Behavior was assessed using standard two alternative forced choice (2AFC) and VAS paradigms to evaluate categorization under task structures that promote discrete vs. continuous hearing, respectively. Behaviorally, identification curves were steeper under 2AFC vs. VAS categorization but were relatively immune to noise, suggesting robust access to abstract, phonetic categories even under signal degradation. Behavioral slopes were correlated with listeners' QuickSIN scores; shallower slopes corresponded with better speech in noise performance, suggesting a perceptual advantage to noise degraded speech comprehension conferred by a more gradient listening strategy. At the neural level, P2 amplitudes and latencies of the ERPs were modulated by task and noise; VAS responses were larger and showed greater noise-related latency delays than 2AFC responses. More gradient responders had smaller shifts in ERP latency with noise, suggesting their neural encoding of speech was more resilient to noise degradation. Interestingly, source-resolved ERPs showed that more gradient listening was also correlated with stronger neural responses in left superior temporal gyrus. Our results demonstrate that listening strategy modulates the categorical organization of speech and behavioral success, with more continuous/gradient listening being advantageous to sentential speech in noise perception.

1. Introduction

Listeners are often tasked with understanding speech signals in noisy listening environments. Speech-in-noise (SIN) perception is a difficult cognitive process and a common audiologic complaint. While certain clinical populations, such as those with hearing loss (Picard et al., 1999; Plomp, 1978), cognitive deficits (Bradlow et al., 2003; Grady et al., 1989), traumatic brain injury (Hoover et al., 2017; Vander Werff & Rieger, 2019), and old age (Bergman, 1971; Humes et al., 2013), show exacerbated SIN difficulty, even normal hearing listeners can have deficits in SIN comprehension (Bharadwaj et al., 2015; Hannula et al., 2011; Ruggles et al., 2011; Tremblay et al., 2015). This large variability in SIN perception emphasizes the importance of analyzing individual differences in performance to understand how a listener's perceptual strategy might influence SIN outcomes.

Listeners have simultaneous access to both acoustic (continuous) and phonetic (categorical) cues of the speech signal (Andruski et al., 1994; Blumstein et al., 2005; McMurray et al., 2002; Miller & Volaitis, 1989; Pisoni & Tash, 1974). Different listeners might weigh information from these modes differently during speech perception tasks, such that some listeners are more categorical/discrete responders, while others are more continuous/gradient responders (Kapnoula et al., 2021; Kapnoula

Abbreviations: ERPs, Event-related potentials; SIN, speech in noise; VAS, visual analog scale; 2AFC, two alternative forced choice; STG, superior temporal gyrus.

* Corresponding author at: Department of Speech, Language and Hearing Sciences, Indiana University, Bloomington, IN, USA.
E-mail address: gbidel@iu.edu (G.M. Bidelman).

https://doi.org/10.1016/j.brainres.2024.149166
Received 19 May 2024; Received in revised form 26 July 2024; Accepted 13 August 2024
Available online 14 August 2024
0006-8993/

et al., 2017; Kong & Edwards, 2016). Theoretically, either strategy could benefit perception. First, categorical/discrete listening may be an ideal strategy for SIN perception. While acoustic information changes continuously, listeners bin speech sounds into equivalency categories to map speech acoustics to a high-level phonetic code (Liberman et al., 1967; Pisoni, 1973). This more abstract categorical code might be more resistant to degradation by noise, since it does not rely on surface features of speech that are easily washed out by noise (Bidelman et al., 2020; Bidelman et al., 2019). Indeed, more discrete listeners show less interference from informational masking in auditory streaming tasks that mimic naturalistic "cocktail party" listening scenarios (Bidelman et al., 2024). Previous work using event-related brain potentials (ERPs) has also demonstrated enhanced P2 responses to noise-degraded speech sounds with clear phonetic identities compared to ambiguous speech sounds that do not carry a clear phonetic label, suggesting the brain does not linearly code changes in acoustics but represents a categorical code (Bidelman et al., 2020; Bidelman et al., 2013; Bidelman & Walker, 2019; Bidelman & Walker, 2017). The categorical organization for speech is thought to involve a network including auditory cortex (Bidelman & Lee, 2015; Bidelman & Walker, 2019; Chang et al., 2010) and higher-order linguistic centers in the inferior frontal gyrus (IFG) (Alho et al., 2016; Myers et al., 2009) that are differentially engaged depending on task difficulty (Carter & Bidelman, 2021) and a listener's experience (e.g., language or music background; Bidelman & Lee, 2015; Bidelman & Walker, 2019).

Conversely, a gradient/continuous listening strategy might confer a perceptual advantage for SIN processing. Maintaining within-category acoustic information may allow listeners to "hedge" their bets on the speech sounds they hear and resolve ambiguity (Kapnoula et al., 2017). For instance, when speech sound identities are more uncertain having variable voice onset times (VOTs), listeners respond in a more graded fashion (Clayards et al., 2008). Kapnoula et al. (2021) found that gradient listeners were better able to recover from lexical ambiguity in garden path sentences, though SIN perception was not improved. Gradient listeners might also have more flexibility in cue-weighting when making phonetic decisions about the acoustic input (Kapnoula et al., 2017; Massaro & Cohen, 1983b; Toscano & McMurray, 2010). Neural evidence of gradient processing from ERPs has demonstrated linear changes in the N1 wave (~100 ms) with changes in VOT (Toscano et al., 2010), earlier in the time course of the response than categorical information seems to be coded (Bidelman et al., 2013). Though, graded speech representations can be maintained in the neural signal for up to 900 ms (Sarrett et al., 2020), even after category abstraction (Bidelman et al., 2013). These findings suggest the brain likely represents and maintains both within- and between-category information (Toscano et al., 2018). Imaging studies suggest that graded activation to speech cues occurs in left superior temporal gyrus (STG) (Myers et al., 2009). Thus, gradient perception may rely more heavily on sensory representations at the level of auditory cortex, while category labelling might recruit higher order linguistic resources downstream (IFG). While it is clear neural responses scale with acoustic–phonetic features of the speech signal (exogenous properties), it is unclear how they are modulated by endogenous properties of the listener. Here, we ask whether differences in listening strategy can actively modulate the neural encoding of speech and beneficially transfer to SIN processing.

One barrier to studying listening strategy is the common use of two alternative forced choice (2AFC) paradigms in speech categorization tasks. In 2AFC, listeners hear speech sounds sampled from an equidistant acoustic–phonetic continuum (e.g., /da/ to /ga/). Listeners are asked to press a button to report which of two speech sounds they heard in a forced, binary judgment. The slope of the resulting identification function is often taken as a measure of categoricity in the behavior (Bidelman, 2015; Hallé et al., 2004; Sussman, 1993; Werker & Tees, 1987; Xu et al., 2006). While identification curve slopes assessed under 2AFC can be used as a measure of listener strategy, it remains unclear whether shallower slopes reflect more/less gradiency in perception or simply noisier responses (Apfelbaum et al., 2022; Kapnoula et al., 2017; McMurray et al., 2018). Indeed, disordered populations often have shallower slopes in labeling tasks which is usually interpreted as less categorical hearing (Godfrey et al., 1981; Joanisse et al., 2000; Serniclaes et al., 2001; Serniclaes et al., 2005; Sussman, 1993; Werker & Tees, 1987). SIN deficits are also common in many of these disorders (Cunningham et al., 2001; Dole et al., 2012; Elmahallawi et al., 2021; Lagacé et al., 2010; Warrier et al., 2004; Ziegler et al., 2009), suggesting a shared mechanism between SIN abilities and categorical perception. However, it is plausible that 2AFC responses in these populations are simply less consistent, resulting in shallower identification slopes due to internal perceptual noise rather than difficulty forming categories.

More recent work has demonstrated that using a visual analog scale (VAS) yields measurements of listener strategy that are more independent of internal noise (Kapnoula et al., 2017), suggesting this method is better for quantifying listener strategy. Whereas the 2AFC task promotes categorical reporting, the VAS task allows for more open-ended responses, thereby reflecting more nuances in listeners' perception of acoustic cues (Munson et al., 2017). For example, VAS response distributions distinguish discrete from gradient listeners in that discrete listeners make more use of the endpoints of the scale and gradient listeners distribute responses across the entire scale (Kapnoula et al., 2017; Kong & Edwards, 2016; Massaro & Cohen, 1983a).

Current literature has not consistently demonstrated a perceptual or neurophysiological benefit for either categorical or continuous listening on SIN perception. It is also unclear how individuals' neural responses are modulated by task demands alone (i.e., promoting gradient responses in a VAS task and categorical responses in a 2AFC task). In the present study, we measured EEG and behavioral responses during a phoneme labelling task under 2AFC and VAS paradigms to quantify listeners' categoricity/gradiency in perception. We then assessed correspondences between listening strategy, ERPs, and standardized measures of SIN perception. Our results demonstrate that better SIN comprehension scores and stronger neural responses in left STG correspond with more gradient listening, establishing a neural-perceptual link between SIN performance and listening strategy.

2. Results

2.1. Behavioral data

We first confirmed listeners' VAS responses were subject to individual differences and thus showed evidence of different listening strategies. Fig. 1 shows responses from n = 2 representative subjects whose responses were more "gradient/continuous" and more "discrete/categorical", respectively. More gradient listeners (Fig. 1**A**) tended to respond along the entire scale to report their percept and have shallower identification curve slopes (Fig. 1**B**), while more discrete listeners' (Fig. 1**C**) responses tended to cluster around the endpoints of the continuum with steeper identification curve slopes (Fig. 1**D**). This confirms listeners do not respond uniformly, motivating us to quantify individual differences in their behavioral response patterns.

An ANOVA conducted on identification curve slopes revealed that listeners had steeper slopes (i.e., more categorical labeling) in the 2AFC compared to the VAS task [$F(1,57) = 49.55$, $p < 0.0001$, $\eta_p^2 = 0.47$] and for clean compared to noise-degraded speech [$F(1,57) = 4.91$, $p = 0.03$, $\eta_p^2 = 0.08$] (Fig. 2). The dip statistic, measured only for the VAS task blocks, decreased with noise [$F(1,209) = 18.94$, $p < 0.001$, $\eta_p^2 = 0.08$]. Behavioral VAS slopes were highly correlated with the dip statistic [$r(38) = 0.87$, $p < 0.0001$], confirming, as expected, more dichotomous responses were associated with bimodal VAS distributions.

Likewise for RT speeds, there was a significant main effect of token [$F(4,361) = 12.45$, $p < 0.0001$, $\eta_p^2 = 0.12$], task [$F(1,361) = 1106.37$, $p < 0.0001$, $\eta_p^2 = 0.75$], and SNR [$F(1,361) = 7.91$, $p = 0.005$, $\eta_p^2 = 0.02$]. The task effect was partially attributed to faster responses under 2AFC

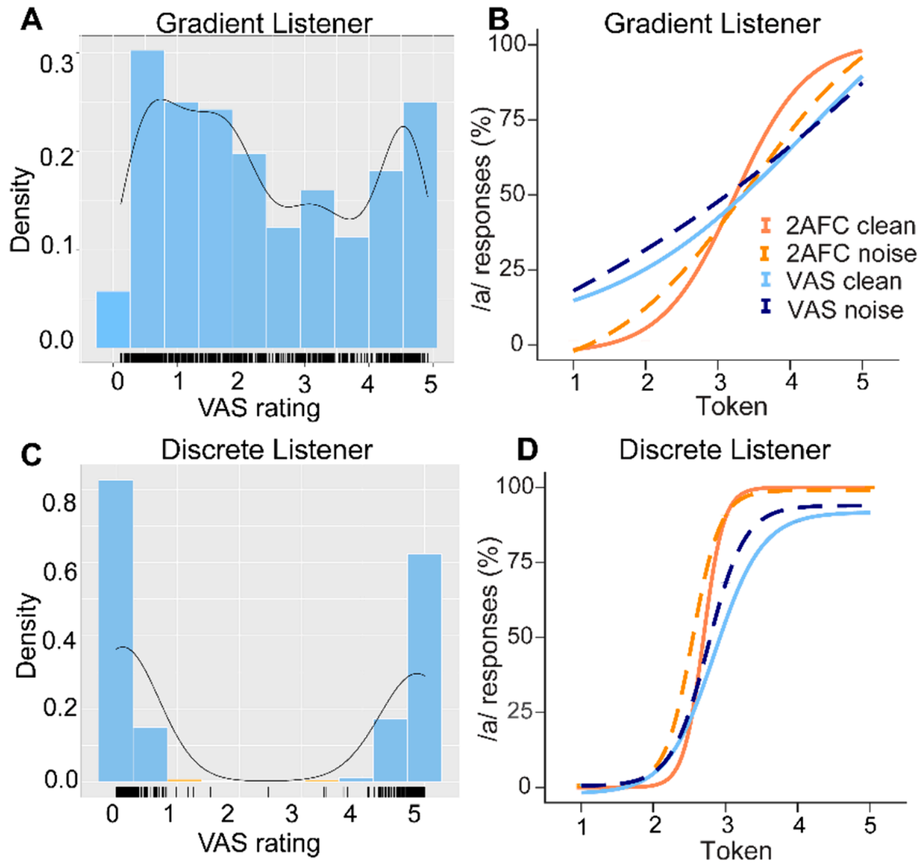

Fig. 1. Listeners vary in their response distributions during phoneme labeling. **(A,C)** Behavioral response distributions for n = 2 representative participants (clean condition) during VAS labeling. Some listeners report their percept using the entirety of the scale (**A:** more gradient listener), while others distribute their responses toward the endpoints of the scale (**C:** more discrete listener). Black lines show density plots. Tick marks on the abscissa represent individual responses along the scale. (**B, D**) Identification curves for the same n = 2 participants across all task and noise conditions. Within subjects, identification remains relatively stable across conditions but varies dramatically between listeners.

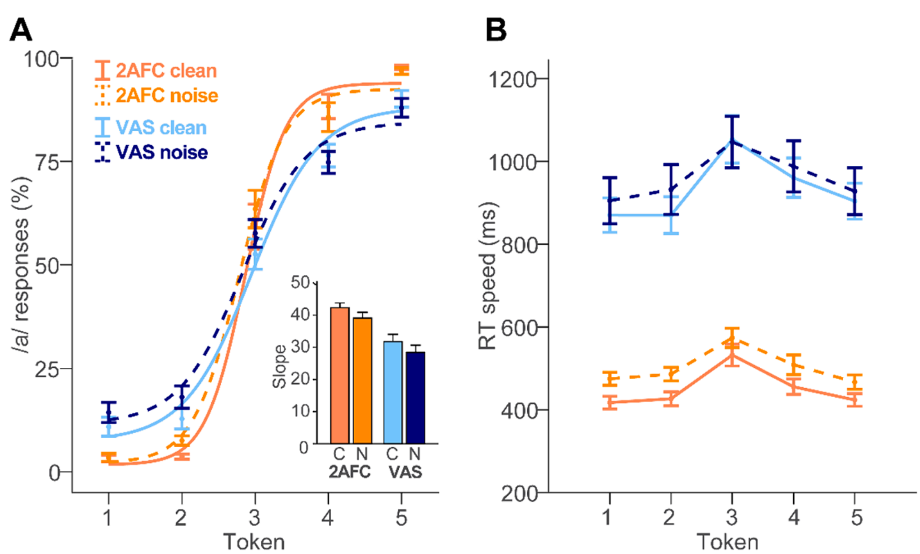

Fig. 2. Behavioral identification changes with categorization task and noise conditions. **(A)** Behavioral vowel identification follows a stair-stepped function typical for categorical perception. Steeper slopes, indicative of more categorical listening, were observed in the 2AFC clean condition. Slopes became shallower under VAS labeling and more minimally with the addition of noise. Inset, mean slope in each condition (C = clean, N = noise). **(B)** RTs are slowest for the midpoint (ambiguous) token of the continuum. Faster RTs are observed during 2AFC vs. VAS labeling and (to a lesser extent) for clean relative to noise-degraded speech. Error bars = ± 1 s.e.m.

vs. VAS labeling whereas the SNR effect was due to slightly faster responses in clean vs. noise. The token effect was attributed to the hallmark slowing in RTs for tokens near the ambiguous midpoint of the continuum (Bidelman et al., 2020; Bidelman & Walker, 2017; Pisoni & Tash, 1974). This slowing occurred regardless of task or noise SNR [contrast: mean(Tk1,2,4,5) vs. Tk3; all p-values < 0.012].

Shallower identification functions could result from weaker categorization and/or noisier responding, both of which would flatten a sigmoid function (Kapnoula et al., 2017). To test the possibility that changes in identification slopes were due to noisier responding, we calculated the standard deviation of subjects' responses to each token during the VAS conditions, which provides a proxy of sensory noise (Kapnoula et al., 2017). Response noise was not correlated with slope measures [$r(38) = 0.02$, $p = 0.92$], suggesting the steepness of listeners' identification functions was independent of sensory noise in the decision process (and therefore instead due to more/less discrete hearing).

Importantly, we found listeners' QuickSIN scores were highly correlated with categorization slopes for both tasks (2AFC: $r(38) = 0.36$, $p < 0.0001$; VAS: $r(38) = 0.30$, $p < 0.0001$) and SNR conditions (clean: $r(38) = 0.21$, $p < 0.0001$; noise: $r(38) = 0.39$, $p < 0.0001$) (Fig. 3). Listeners with lower dB SNR loss (i.e., better SIN comprehension) had more gradient (shallower) behavioral slopes, suggesting a more continuous listening strategy may be beneficial for SIN understanding.

2.2. Electrophysiological data

2.2.1. Electrode level data

Figs. 4 and 5 show the electrode-level waveforms, topographic maps, and P2 amplitudes and latencies at the central electrode cluster. Topographies are average activity across the latency window [140–320 ms] and thus, show more posterior activation than the canonical P2 topography, likely due to the active task required recruiting more neural regions during the prolonged window used here. P2 amplitudes at the central cluster differed across vowels [$F(5,437) = 3.19$, $p = 0.0078$, $\eta_p^2 = 0.04$], SNR [$F(1,437) = 54.89$, $p < 0.001$, $\eta_p^2 = 0.11$], and task [$F(1,437) = 22.5$, $p < 0.001$, $\eta_p^2 = 0.05$]. VAS amplitudes were larger than 2AFC amplitudes [$t(437) = -4.743$, $p < 0.001$] across noise conditions. There was also a main effect of noise on P2 latencies [$F(1,437) = 138.55$, $p < 0.001$, $\eta_p^2 = 0.24$]. Interestingly, P2 latencies showed a SNR x task interaction [$F(1,57) = 6.43$, $p = 0.012$, $\eta_p^2 = 0.01$], whereby VAS latencies were more prolonged by noise [$t(437) = -10.12$, $p < 0.001$] than 2AFC latencies [$t(437) = -6.53$, $p < 0.001$]. Similar effects were found at the other electrode clusters (all p-values < 0.01) with an additional main effect of vowel at all but the center cluster (all $ps < 0.0001$).

To further investigate the task x SNR interaction, we examined whether noise-related changes in the ERPs predicted the degree of listeners' behavioral categorization. To this end, we correlated P2 latency shifts produced by noise (i.e., $P2_{noise} - P2_{quiet}$) during the VAS task with listeners' behavioral slopes. We found that more gradient responders had smaller noise-related P2 latency shifts, suggesting their speech ERPs were more resilient to noise-degradation [$r(18) = 0.48$, $p = 0.032$] (Fig. 6).[1] These findings imply that neural timing to speech was less strongly impacted by noise in more gradient responders.

2.2.2. Source level data

To resolve the underlying sources that might contribute to these scalp effects, we examined source-resolved activity using CLARA distributed imaging (Carter et al., 2022; Iordanov et al., 2016). Fig. 7**A** shows correlations between the raw source activations in the P2 time window and behavioral identification across the full brain volume. Data are pooled across the 2AFC and VAS tasks to assess the overall pattern between behavioral slopes and brain activity. The cluster analysis and permutation statistics procedure revealed one significant ($p = 0.012$) spatiotemporal cluster encompassing auditory-sensory regions within left superior temporal gyrus ($_L$STG; Talairach coordinates: x = −52.5, y = −23.9, z = 2.7) (Fig. 7**B**). Both the peak amplitude ($r = -0.46$, $p < 0.0001$; Fig. 7**C**) and latency ($r = 0.26$, $p = 0.021$; Fig. 7**D**) of these $_L$STG source activations were strongly correlated with the listeners' perception.[2] That is, weaker and later STG responses were associated with steeper identification slopes and thus more categorical (discrete) hearing. Conversely, stronger and earlier STG responses were associated with shallower speech labeling and thus more continuous modes of listening.

3. Discussion

Prior work has been equivocal on whether being a more categorical or continuous listener is optimal for degraded speech perception. By measuring ERPs to clean and noise-degraded vowel sounds during behavioral tasks that require more/less continuous vs. categorical hearing, we investigated the neural mechanisms subserving the relationship between listening strategy and SIN performance. Participants' identification curve slopes served as a measure of their listening strategy independent of internal noise. We found more continuous listeners, with shallower behavioral slopes reflecting more gradient phoneme perception, performed better on the QuickSIN, suggesting a perceptual advantage for SIN comprehension afforded by a more continuous listening strategy. Our data also reveal that speech-ERPs are modulated by task structure: otherwise identical speech sounds are differentially encoded by the brain depending on post-perceptual task demands. We also found prominent noise-related changes in the ERPs that depended on listening strategy; more gradient listening corresponded with smaller shifts in latency with noise. Lastly, source analysis revealed these neural effects were attributed to changes in $_L$STG activation, whereby more gradient listening was associated with larger and faster responses in early, left lateralized auditory brain regions. Collectively, our results demonstrate both behavioral and neural benefits of a more gradient listening strategy to speech sound processing and SIN perception.

3.1. Speech categories are robust to noise

Behaviorally, we found categorization remained strong in noise with only a small reduction in behavioral slopes at more difficult SNRs. This supports behavioral findings from Bidelman et al. (2020) showing categorical representations for speech persist even in perceptually challenging levels of noise. These results suggest that binning speech sounds into categories is a robust perceptual process and may confer advantages to SIN perception (Bidelman et al., 2020; Bidelman and Carter, 2023). Higher-level category representations can be maintained even when acoustic representations are degraded by noise. RTs followed a classic pattern in labeling, slowing around the ambiguous midpoint of the continuum in both task and noise conditions; responses were fastest for speech sounds carrying a clear phonetic label (Bidelman & Walker,

[1] This correlation was still significant [$r(14) = 0.48$, $p = 0.044$] when the two participants with negative latency shifts were removed from the analysis.

[2] Apparent edge effects in source amplitudes and latencies are due to how the CLARA inverse solution prunes low amplitude neural activations. CLARA iteratively reduces the source space to minimize low amplitude activations. The method estimates the total variance of the scalp-recorded data and applies a smoothness constraint to ensure that current changes little between adjacent regions in the brain (Michel et al., 2004; Picton et al., 1999). CLARA makes source images more focal by iteratively reducing the source space during repeated estimations. On each step (x2), a spatially smoothed LORETA solution (Pascual-Marqui et al., 2002). is recomputed and voxels below a 1% max amplitude threshold are removed. Thus, some listeners showed near zero amplitude after CLARA pruning and subset cluster-based statistics to correct spurious activations across the brain volume (see *Section 5.7*).

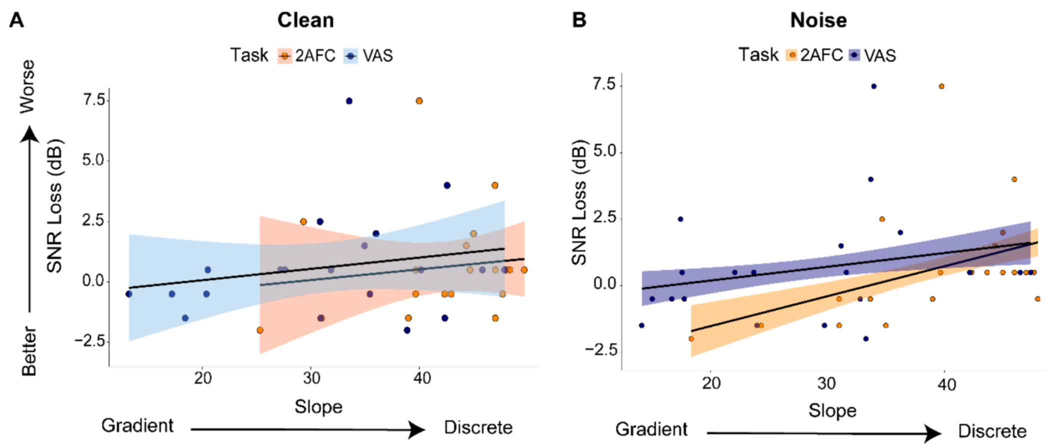

Fig. 3. Successful SIN performance (sentence-level recognition) is predicted by a gradient listening strategy (discrete phoneme identification). Pearson correlations between slope in vowel identification and QuickSIN scores (lower score = better SIN comprehension). (**A**) QuickSIN scores correlated with labeling slopes for clean and (**B**) noise-degraded phonemes under both 2AFC and VAS tasks. Shading = 95 % CI.

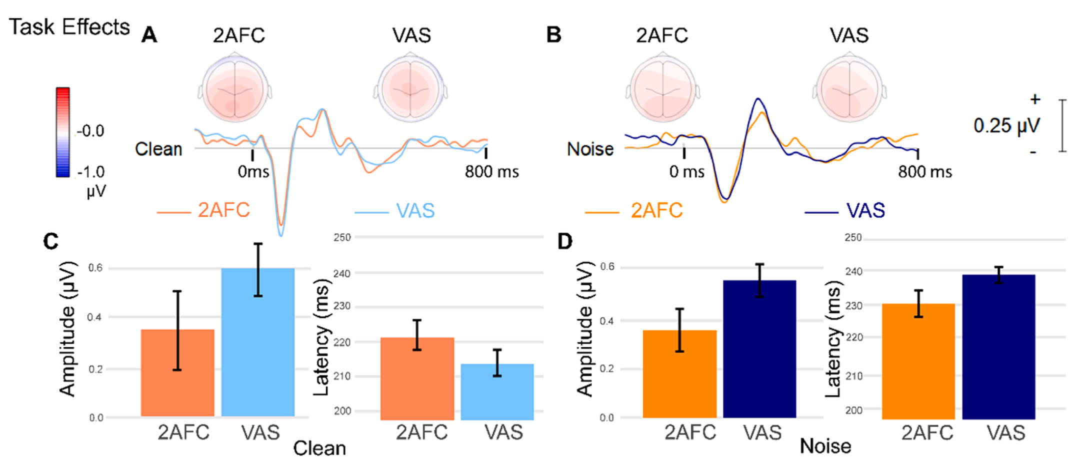

Fig. 4. Categorization task modulates cortical responses to otherwise identical speech phonemes. Grand average ERP waveforms measured at the central electrode cluster (FC1, FCz, FC2, C1, Cz, C2) and topographic maps (P2 latency window; 140–320 ms) for **(A)** Clean and **(B)** Noise-degraded speech. **(C-D)** P2 peak amplitudes and latencies (only endpoint tokens [vw1/vw5] shown). Error bars = ± 1 s.e.m.

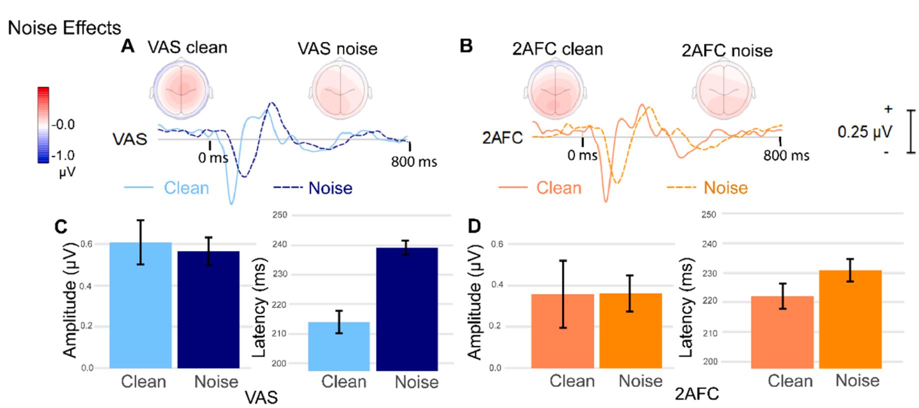

Fig. 5. Noise modulates P2 amplitudes and latencies. Note the task x noise interaction depicted by the exaggerated shift in latency in the VAS compared to the 2AFC condition (panel D). Otherwise as in Fig. 4. Error bars = ± 1 s.e.m.

2017; Pisoni & Tash, 1974). This suggests, in the broadest sense, category speech representations might provide an easy, faster readout to perceptual processing.

RT delays in the VAS relative to 2AFC task may be attributable to differences in the behavioral response requiring the use of a mouse click rather than a button press (Bidelman et al., 2024). In addition to differences in the motor response, the VAS task may have placed greater attentional demands and/or listening effort on listeners due to the use of a continuous perceptual rating scale, effects which are typically magnified in noise-degraded scenarios (Lewis & Bidelman, 2020). This

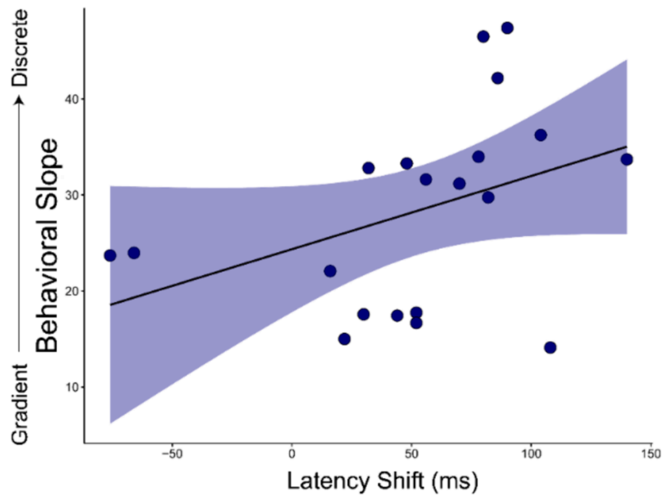

Fig. 6. Gradient listeners are more resilient to noise related changes in speech encoding. Latency shift represents the difference in P2 timing elicited in the VAS condition with and without background noise (i.e., $P2_{noise} - P2_{quiet}$). Behavioral slopes are positively correlated with noise-related shifts in P2 latency. Neural timing is less strongly impacted by noise in more gradient responders. Shading = 95 % CI.

could have resulted in the prolonged RTs we find during VAS categorization. Indeed, the use of mouse-tracking in categorization tasks has revealed subtle dynamics in perception as listeners actively weigh sensory cues to formulate their ultimate category label (Bidelman & Carter, 2023; Huette & McMurray, 2010; Viswanathan & Kelty-Stephen, 2018). The use of mouse responses in both 2AFC and VAS tasks could have teased out whether slower RTs are indeed indicative of greater attention for graded decision-making. In this vein, we found that more gradient listeners had slower RTs in general, suggesting more gradient listening may indeed require greater attentional resources. Although task demands may have influenced behavioral RT outcomes, large individual differences in listening strategy persisted in the precision of listeners' identification slopes, which are unlikely to be attributable to task motor demands. Moreover, task demands did not affect participants uniformly.

3.2. Categorization skills are related to SIN perception but show an advantage for more gradient listening

Notably, we found that more gradient listeners (as measured via phonetic categorization), achieved better SIN comprehension scores (measured via sentences in the QuickSIN). More gradient listeners may be better equipped to deal with uncertainty by maintaining within-category information to better "hedge" their bets when a speech signal is ambiguous (Kapnoula et al., 2017; McMurray et al., 2008), such as during SIN testing.

Prior work has failed to establish a consistent link between SIN performance and listening strategy. This discrepancy is likely due to differences in experimental task design. Neural and behavioral evidence from phoneme labelling tasks in Bidelman et al. (2020) suggests that stronger categorization (i.e., more discrete listening strategy) could aid in successful SIN performance. However, SIN measures in that study were based solely on a phoneme labelling in noise task (2AFC), whereas the present study used more complex SIN assays via the QuickSIN. Moreover, our use of VAS labeling to evaluate listener strategy with regard to categorization skills provides a cleaner measure than 2AFC as it is less obstructed by internal sensory noise that can confound the interpretation of identification curve slope data (Kapnoula et al., 2017). Collectively, our data support the notion that more gradient/continuous listening strategies are more beneficial to real-world SIN listening scenarios.

Our findings differ from several prior studies assessing putative relations between categorization and SIN skills. Kapnoula et al. (2021) sought to correlate success of word comprehension in noise with listening strategy measured by a VAS during a garden path sentences task. Though garden path sentences require listeners to resolve ambiguity, a theoretically similar cognitive process to listening to speech in noise, this task does not directly use degraded speech stimuli. Our study thus differs from Kapnoula et al (2021) in that we directly relate two measures of SIN perception at the phoneme and sentence levels, though we used different types of noise at each level (speech-shaped noise for phonemes and multi-talker babble for sentences). More akin to the present study, Bidelman et al. (2024) demonstrated that discrete listeners assessed by a VAS phoneme labeling task had less interference from informational masking than gradient listeners when measured in a cocktail party streaming task containing speech-on-speech maskers. While our results conflict with those of Bidelman et al. (2024), it is highly likely that different strategies could be deployed across different listening environments to optimize perceptual performance—a single listening "mode" might not be a one size fits all. While more gradient

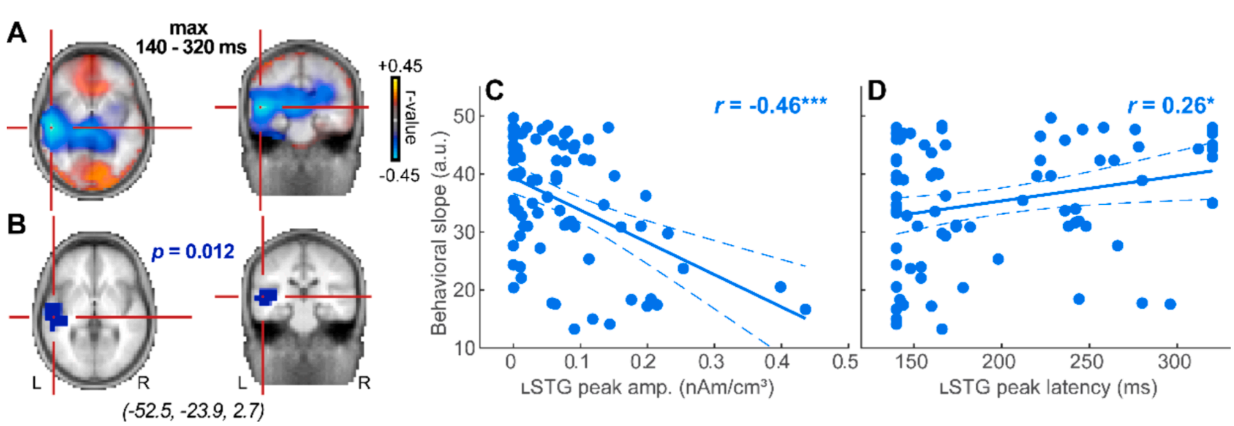

Fig. 7. Source localization of speech ERPs distinguishes discrete vs. continuous listening strategies in speech categorization. (**A**) Full-brain correlations between CLARA source activations and behavioral identification slopes (aggregating all data from the 2AFC and VAS tasks). Hot/cool colors = positive/negative associations. Statistical maps are projected onto BESA's standard brain template. (**B**) Cluster-based permutation statistics identified a single cluster in left STG driving the brain-behavior relation (cross-hair; Talairach: x = −52.5, y = –23.9, z = 2.7). Both peak amplitude (**C**) and latency (**D**) of STG source activity predict the degree of listeners' categorical hearing. Weaker/later responses are associated with more discrete phonetic labeling, and stronger/earlier responses with more gradient hearing. Dotted lines = 95 % CI, $*p < 0.05$, $***p < 0.0001$.

listeners may perform better on sentence recognition tasks in multi-talker babble, auditory streaming tasks may engage separate processes that benefit from more discrete listening strategies (Bidelman et al., 2024).

One limitation in our design is the use of isolated phoneme labeling to investigate listening strategy. In real-world listening, there are more factors that might influence speech-in-noise perception, including other sources of acoustic variability (Rogers et al., 2004; Rogers et al., 2006) and contextual information (Sheldon et al., 2008; Van Engen et al., 2014) available to a listener. Listeners may change their perceptual strategies across different listening scenarios or remain consistent. An ideal listener may be one that weighs continuous and categorical cues optimally (Kapnoula et al., 2017; Massaro & Cohen, 1983b) or even dynamically shifts their weights according to the present listening demands. Future studies using a large variety of ecological SIN assessments and informational maskers during phoneme labeling are needed to test this possibility.

We used behavioral slopes to index listener categorization strategy. Controversy exists regarding whether the slope measure of a sigmoidal identification function truly reflects listener strategy or rather represents internal noise due to perceptual uncertainty (Kapnoula et al., 2017; McMurray, 2022). That is, a shallower identification function could theoretically index either a more gradient listener or a noisier responder. Kapnoula et al. (2017) and Bidelman et al. (2024) both demonstrated that slopes calculated from VAS responses did not reflect response noise (i.e., sensory noise internal to the listener). Similarly, we found that estimates of perceptual noise via response variance were not correlated with behavioral slopes. This suggests that slopes during VAS labeling are a veridical measure of listeners' categorization strategy independent of sensory noise. If categorization slopes reflected internal noise in the decision instead of the categoricity in judgment—as suggested by Kapnoula et al. (2017) and colleagues —then less noisy responders (i.e., those with steeper slopes) should theoretically be more successful in SIN perception.[3] On the contrary, we actually find the opposite pattern: a positive correlation between QuickSIN scores and behavioral slopes. This reinforces the notion that more continuous/gradient categorizers, not noisier responders, *per se*, are more skilled SIN listeners.

If gradient listeners are more attentive in phoneme labeling as suggested by our behavioral RTs, they may also be more attentive in SIN tasks, perhaps at least partially accounting for their observed benefits in QuickSIN comprehension. However, we find this explanation unlikely since sustained attention does not necessarily predict QuickSIN and cocktail party streaming performance (Bidelman & Yoo, 2020). While attention may mediate the relationship between listening strategy and SINperception, future studies are needed to understand their independent contributions.

3.3. Speech-ERPs are modulated by task and noise

Our electrophysiological data showed that P2 evoked by otherwise identical speech stimuli was modulated depending on whether listeners were performing 2AFC or VAS labeling. This suggests that the neural encoding and early sensory representations for speech are altered by task demands. P2 amplitudes were reduced with noise, supporting prior work demonstrating reduction in P2 with background noise (Billings et al., 2009; Gustafson et al., 2019; Koerner & Zhang, 2015; Papesh et al., 2015). Interestingly, we observed larger responses in the VAS condition compared to the 2AFC condition. To our knowledge, this is the first study to demonstrate changes in the auditory-sensory ERPs with changes in post-perceptual task structure.

Listeners have access to both categorical and continuous cues simultaneously which can be used differently with varying task demands, as evidenced by RTs (Pisoni & Tash, 1974), eye tracking (Clayards et al., 2008; McMurray et al., 2018), and MRI studies (Fuhrmeister & Myers, 2021). Some tasks, including the 2AFC identification task used here, may rely only on a categorical or phonetic listening mode, while other tasks may require additional access to a more continuous or acoustic listening mode. Because the VAS task allows for a more continuous rating than the 2AFC task, it is reasonable to assume that listeners might use both categorical and continuous information simultaneously to form their responses. During VAS judgments, more neural resources might be recruited to allow access to both types of information, resulting in larger ERP amplitudes. While Toscano et al. (2018) demonstrated that gradient cues are represented earlier in the ERP time course (~N1) than categorical representations (~P2), Sarrett et al. (2020) suggested gradiency is encoded on a longer time scale, spanning that of the latency window used here. Moreover, the P2 is not solely a response to exogenous acoustics, but an early endogenous response indexing speech discrimination (Alain et al., 2010; Ben-David et al., 2011), auditory object identification (Ross et al., 2013), and category representation (Bidelman et al., 2020; Bidelman et al., 2013). Thus, it is not surprising that P2 changed with task demands, as the response reflects categorical (perceptual) and continuous (acoustic) components. Still, a novel finding is that these early auditory responses beginning at ~ 150 ms are influenced by post-perceptual mechanisms that initiate the motor response much later in time (400–800 ms).

Topographies from our P2 latency window appeared more posterior than canonical P2 topographies. Our latency window was long enough to capture noise-related shifts in latency which may have contributed to their peak activations not restricted to the vertex. Additionally, we used an active task during ERP recording, which may have resulted in other contributions to P2 aside from stimulus encoding. Prior work in speech categorization has demonstrated similar posterior activation and "post-P2" activity relating to perception rather than acoustic information (Bidelman et al., 2013; Bidelman and Alain, 2015a; Bidelman et al., 2020). This post-P2 activity may reflect matching a stimulus to a phonetic memory template (Bidelman and Alain 2015b) or attentional reorienting (perhaps a P3-like response) during an active task (Knight et al., 1989). Because these processes recruit other neural regions including medial temporal lobe and superior temporal association cortices near parietal lobe (Alain et al., 2001; Dykstra et al., 2016), the resulting topographies are more posterior than a purely sensory P2.

As expected, P2 latencies were longer in noisy than clean conditions across both tasks. Noise-related prolongation of the P2 is likely due to decreased neural synchrony due to masking noise (Billings et al., 2009; Kaplan-Neeman et al., 2006; Whiting et al., 1998). However, these latency effects were more prominent in the VAS compared to 2AFC condition. It is unlikely this reflects mere differences in speed of the motor response during VAS labeling since P2 effects were substantially earlier (600–800 ms) than listeners' RTs. Instead, stronger noise-related changes in VAS may reflect disproportionately augmented sensory effort when performing identification in noise during a graded (VAS) vs. binary (2AFC) task (Bidelman & Walker, 2017; Crowley & Colrain, 2004; Picton & Hillyard, 1974).

More critically, we found that noise-related shifts in ERP latency were behaviorally relevant; smaller latency prolongations were observed for more gradient compared to discrete listeners. This finding supports the notion that making use of continuous cues to decode ambiguous speech is advantageous, as neural timing is less disturbed by noise among more continuous listeners. It is possible that listeners who weigh continuous information more heavily when making perceptual decisions experience less change in stimulus ambiguity. Alternatively,

[3] This is not to say that "external" and "internal" noise have isomorphic effects on perception. The former is an exogenous property governed by SNR of the acoustic stimulus whereas the latter reflects endogenous, sensory noise of the observer. However, as noted by Kapnoula et al. (2017), noisier responders should have less precise internalized speech representations leading to difficulties in speech-in-noise perception. However, this is not what we find in the current data. Noisier phoneme labeling, per se, was not correlated with QuickSIN scores.

and by the logic above, more gradient listeners may also experience reduced attentional load in the presence of noise, accounting for the smaller changes we find in their ERPs.

3.4. Gradient listeners have stronger neural activation in $_L$STG

Source reconstruction revealed that the P2 effects were attributed to early activation in left-lateralized auditory brain regions. Notably, more gradient listeners had stronger neural activations in left STG. Prior work has shown that activation in the $_L$STG is greater than that in the $_R$STG during speech perception (Ramos Nuñez et al., 2020; Turkeltaub & Branch Coslett, 2010). This left > right asymmetry is largely consistent with theories of brain lateralization (Hickok & Poeppel, 2007) and hemispheric differences in categorization which suggest speech labeling is processed dominantly by the left hemisphere and music labeling by the right (Mankel et al., 2022; Zatorre et al., 1992). Additionally, larger STG responses in more graded listening could reflect increased engagement of working memory resources that help maintain and refresh the neural trace of acoustic-sensory information of the speech signal prior to labeling. Indeed, stronger sustained activity within left (but not right) auditory cortex is observed under more demanding auditory working memory loads as listeners retain verbal sounds in memory (Bidelman et al., 2021; Kumar et al., 2016). Conceivably, the leftward bias in STG source activations we find for more gradient listening could reflect heavier retention of continuous acoustic information in the auditory sensory-memory buffer prior to assigning a category label.

Our results also parallel similar findings of stronger and earlier responses in (left) auditory cortex that have been associated with increased attention during speech perception (Hugdahl et al., 2003; Wong et al., 2008) and verbal working memory (Bidelman et al., 2021). The larger and faster responses we observed corresponding with more gradient listening may be at least partially attributable to increased attention among this subset of listeners. This finding along with the slower behavioral RTs among more gradient listeners suggests that gradient listening may be a more effortful process that requires greater sustained auditory attention and/or working memory.

Non-auditory regions such as inferior frontal gyrus (IFG) have been shown to predict behavioral performance in categorization tasks (Bidelman & Walker, 2019; Golestani & Zatorre, 2004; Lee et al., 2012; Meyers et al., 2008). While we therefore expected to find IFG involvement, the association between behavioral and neural measures was instead restricted to canonical auditory brain regions ($_L$STG). Behavioral tasks such as those used here inherently employ both sensory and decision-making processes. Studies suggest a functional distinction between the two operations whereby activity in auditory cortical regions (including STG) maps onto sound identification (sensory process), while inferior frontal regions map onto reaction time (decision-making process) (Binder et al., 2004; Du et al., 2014). Our findings are consistent with these functional distinctions. Activation of $_L$STG was correlated with a measure of precision (i.e., behavioral slope) rather than speed of speech identification. As such, we infer that how a signal is encoded at the level of auditory cortex may predict the degree of categorical perception a listener experiences. In support of this notion, we have shown that neural representations for speech in auditory cortex reorganize to take on more abstract, categorical organization with increased listening experience of the individual (Bidelman & Walker, 2019). Thus, it is possible that continuous feature coding in auditory temporal cortex is initially more effective in supporting speech sound identification (present study) but that over time, intensive auditory or language experience causes it to re-organize (Guenther et al., 2004) and begin supporting abstract phonetic representations for speech in and of itself (Bidelman & Lee, 2015; Bidelman & Walker, 2019; Chang et al., 2010).

4. Conclusions

We examined brain and behavioral links between two fundamental operations in speech perception: categorization and speech-in-noise listening skills. ERPs revealed task-dependent changes in early neural responses starting around 150 ms that differentiated more categorical from more gradient listeners. More gradient listeners had better SIN comprehension scores, more resilience to noise-related degradation in speech encoding, and stronger neural responses in $_L$STG than their more discrete/categorical listening peers. While a more gradient listening mode was beneficial in multiple domains here, whether our findings extend to more realistic listening environments remains to be investigated. Listeners may adapt their strategies on the fly in real world situations, switching strategies to adapt to changes in signals. Different task structures using more complex stimuli such as sentences or spatially separated streams may find different utility for the gradient strategy. Categorization and SIN deficits are common hallmarks of a variety of auditory-based disorders (Cunningham et al., 2001; Dole et al., 2012; Dole et al., 2014; Lagacé et al., 2010; Putter-Katz et al., 2008; Warrier et al., 2004). Therefore, documenting associations between these skills may provide a linking hypothesis to understand certain communication deficits (Calcus et al., 2016). Future work should examine how a gradient listening strategy might be fortified in listeners with poor SIN comprehension via auditory training and similar rehabilitation tools.

5. Experimental procedures

5.1. Participants

Our sample included N=20 English-speaking young adults (19–30 years old, 10 female/10 male) with 17.5 ± 2.4 years of education and 8.7 ± 8.0 years of self-reported formal music training. Years of musical training did not correlate with listening strategy measures or SIN performance (all $ps > 0.05$). Participants all had normal hearing (≤25 dB HL; 250–8000 Hz octave frequencies) and were mostly right-handed (73 % ± 30 %; Edinburgh Handedness Inventory; Oldfield, 1971). Participants provided written informed consent in accordance with a protocol approved by the Institutional Review Board at Indiana University and were paid $10 an hour for their time.

5.2. Stimuli and task

Prior to EEG testing, listeners completed the QuickSIN assessment (Killion et al., 2004) to measure individual SIN comprehension abilities. Sentences were presented binaurally over headphones. The average of scores from two lists of QuickSIN sentences was used to determine a listener's dB SNR loss, reflecting the SNR threshold required for 50 %-word recall.

For the EEG experiment, stimuli consisted of 5 synthetic vowel sounds along a continuum from /u/ to /a/ changing in first formant frequency (F1) (Bidelman et al., 2020; Bidelman et al., 2013). Tokens were sampled from evenly spaced points along a continuum changing F1 linearly from 430 Hz to 730 Hz (Fig. 8**A**). Tokens had identical F0 (150 Hz), F2 (1090 Hz), and F3 (2350 Hz). Tokens were 100 ms in duration gated with 10 ms ramps. Stimuli were presented using MATLAB (The MathWorks, Natick; MA, USA) coupled to a TDT RZ6 (Tucker-Davis Technologies, Alachua, FL, USA) signal processor at 75 dB SPL binaurally over shielded insert headphones (ER-2; Etymotic Research).

Vowel stimuli were presented in one of two noise conditions: clean and noise (−2.5 dB SNR). We selected this SNR based on previous findings showing speech categorization is resilient to noise down to ~ 0 dB SNR (Bidelman et al., 2020) and pilot testing, that confirmed −2.5 dB SNR hindered speech perception while still maintaining categorical identification. Based on prior work, we used a speech-shaped noise based on the long-term power spectrum (LTPS) of the vowel continuum rather than using multi-talker babble which makes the task too difficult

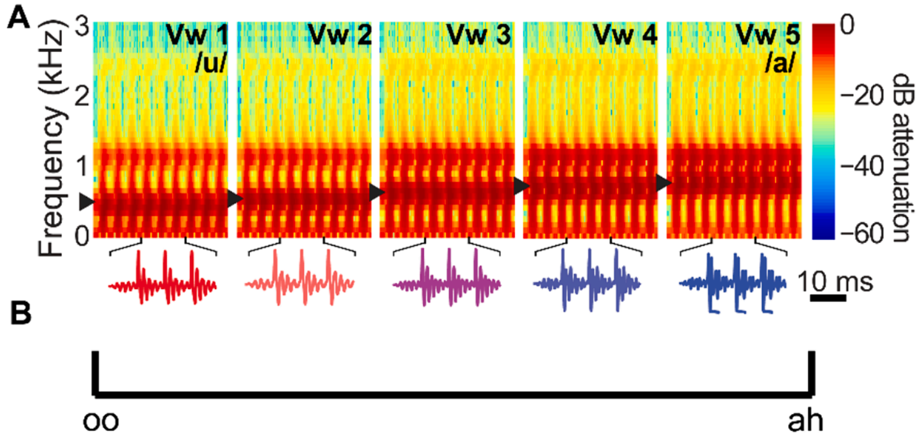

Fig. 8. Stimuli and task design. **(A)** Stimulus spectrograms. F1 was changed from 430 to 730 Hz to produce an acoustic–phonetic continuum from "oo" to "ah". Color scale represents the spectral level relative to full scale. Stimuli were presented at 75 dB SPL during the task. **(B)** Visual analog scale shown to participants during VAS task blocks. Participants were asked to click on the scale to report what they heard.

(Bidelman et al., 2020). Noise was presented continuously throughout the noise block so that it was not time-locked to the stimulus presentation (Alain et al., 2012; Bidelman & Howell, 2016).

During each block, listeners heard 150 presentations of each token and were asked to identify the vowel they heard as quickly and accurately as possible using either a (i) 2 alternative-forced choice (2AFC) binary key press or (ii) visual analog scale (VAS) response. 2AFC and VAS tasks were run in separate blocks but used otherwise identical stimuli; only the task paradigm differed. The VAS paradigm required participants to click a point along a continuous visual scale with endpoints labeled "oo" and "ah" to report their percept (Fig. 8**B**). Use of the entire analog scale was encouraged. Following listeners' behavioral response, the interstimulus interval (ISI) was jittered randomly between 800 and 1000 ms (20 ms steps, uniform distribution) to avoid rhythmic entrainment of the EEG and the anticipation of subsequent stimuli. In total, there were 4 conditions: 2AFC/VAS in clean/noise. Block order was counter-balanced between participants using a Latin square.

5.3. Behavioral data analysis

To analyze the behavioral responses, we computed the identification curve slopes for each condition, computed as the rise/run change in %-labeling between tokens straddling the midpoint category boundary (i.e., vw2, vw4). Steeper slopes are indicative of more categorical/discrete listening. To provide another quantitative measure of listener strategy, we also used the distribution of VAS responses to calculate Hartigan's dip statistic, a number that quantifies how multimodal a distribution is (Hartigan & Hartigan, 1985). Higher dip statistic values indicate a more bimodal distribution, representative of a more discrete response strategy (Bidelman et al., 2024). Behavioral speech labeling speeds (i.e., reaction times, RTs) were computed as listeners' median response latency across trials for a given condition. RTs outside 250–2500 ms were deemed outliers (e.g., fast guesses, lapses of attention) and were excluded from the analysis (Bidelman et al., 2020; Bidelman et al., 2013).

5.4. EEG recording and data processing

During each block of behavioral tasks, we recorded high density EEG using 64-channel Ag/AgCl electrodes located at 10–10 positions on the scalp (Oostenveld & Praamstra, 2001). We used Neuroscan Curry 9 software and SynAmps RT Amplifiers (Compumedics Neuroscan, Charlotte, NC) to digitize recordings at 500 Hz. Data preprocessing was then performed in BESA Research 7.1 (BESA, GmbH). During recording, the EEG was referenced to an electrode located 1 cm behind Cz. Recordings were later re-referenced to a common average reference. Single electrodes on the outer canthi of the eyes and the superior and inferior orbit recorded eye movements. We used principal component analysis to spatially correct ocular artifacts (Lins et al., 1993; Picton et al., 2000; Wallstrom et al., 2004). Additional epochs > 150 μV were rejected as artifacts. EEGs were then bandpass filtered from 2 to 30 Hz (zero-phase filters, 48 dB/octave slope). We chose to high-pass filter at 2 Hz to minimize contributions of low-frequency motor potentials since our two tasks required different movements for responses. After filtering, recordings were epoched (−200–800 ms), baselined, and ensemble averaged across trials to generate ERPs for each token per noise and task condition.

5.5. ERP analysis

To reduce the dimensionality of the electrode-level data, ERPs were quantified using 5 electrode clusters (Carter et al., 2022). We averaged activity from adjacent electrodes within each cluster area on the scalp: front left (AF3, F3, F1), front right (AF4, F2, F4), left temporal (FT7, FC5, FC3, T7, C5, C3, TP7, CP5, CP3), right temporal (FC4, FC6, FT8, C4, C6, T8, CP4, CP6, TP8), and center (FC1, FCz, FC2, C1, Cz, C2). ERPs were quantified in peak latency and amplitude in the time window of the P2 (140–320 ms). We chose to analyze the P2 peak because it occurs in the time course of the ERP when speech categories fully emerge in the brain and it is sensitive to degraded speech perception skills (Bidelman et al., 2020; Bidelman & Lee, 2015; Bidelman et al., 2013; Bidelman & Walker, 2017; Ross et al., 2013). In contrast, earlier peaks in the ERP (i.e., N1) typically reflect gradient differences in acoustics that may or may not be related to listeners' actual categorical perception of the signal (Bidelman et al., 2013; Toscano et al., 2018; Toscano et al., 2010). This window was based on visual inspection of the grand average waveform and ensured we captured both noise- and task-related related shifts in P2 latency (Bidelman et al., 2020) (see Figs. 4-5).

5.6. Statistical analysis

For behavioral data and ERPs, we used linear mixed model ANOVAs (R; lme4 package; version 1.1–31) to test differences in outcome variables (slope, dip statistic, RT, P2 amplitude, P2 latency). Multiple comparisons were corrected using Tukey-adjusted contrasts with an overall $\alpha = 0.05$. Vowel, task, and noise conditions were fixed effects and subjects served as a random effect. We used Pearson's correlations to characterize behavior-behavior and brain-behavior relationships. Effect sizes are reported as partial eta squared (η_p^2) and degrees of freedom were calculated using Satterthwaite's method.

5.7. Source analysis

We used Classical Low Resolution Electromagnetic Tomography Analysis Recursively Applied (CLARA) [BESA® Research (v7.1)] (Iordanov et al., 2016; Iordanov et al., 2014; Scherg et al., 2019) with a 4-shell ellipsoidal head model (conductivities of 0.33 [brain], 0.33 [scalp], 0.0042 [bone], and 1.00 [cerebrospinal fluid] (Berg & Scherg, 1994) to determine the intracerebral sources that account for continuous vs. discrete listening strategies in speech categorization (e.g., Alain et al., 2023; Bidelman et al., 2018; Carter et al., 2022). Source images were computed for endpoint (vw1/vw5) tokens within the 140–320 ms (~P2 wave) analysis window, where task and noise effects were maximal in the scalp ERPs (see Figs. 4-5). CLARA models the inverse solution as a large collection of elementary dipoles distributed over nodes on a mesh of the cortical volume. The algorithm estimates the total variance of the scalp data and applies a smoothness constraint to ensure current changes minimally between adjacent brain regions (Michel et al., 2004; Picton et al., 1999). CLARA renders more focal source images by iteratively reducing the source space during repeated estimations. On each iteration (x2), a spatially smoothed LORETA solution (Pascual-Marqui et al., 2002) was recomputed and voxels below a 10 % max amplitude threshold were removed. This provided a spatial weighting term for each voxel on the subsequent step. Two iterations were used with a voxel size of 7 mm in Talairach space and regularization (parameter accounting for noise) set to 0.01 % singular value decomposition. Source activations were visualized on BESA's adult brain template (Richards et al., 2016), providing a distributed image describing the P2 activation across the entire brain volume.

We used cluster-based permutation tests (Maris & Oostenveld, 2007) implemented in BESA Statistics (2.1) to examine correlations between the neural source and behavior measures and identify anatomical locations within the full-brain volume that predicted listeners' degree of categorical hearing (see Fig. 7**A**). For each voxel, a Pearson's correlation was computed between neural (CLARA source activations) and behavioral (identification slopes) responses. Statistical maps were corrected for multiple comparisons across space by building voxel clusters that control the familywise error rate via a Monte-Carlo resampling technique (Maris & Oostenveld, 2007). We used an alpha level of $\alpha = 0.001$ and N=1000 permutations for cluster building. This more stringent alpha level allowed for separation from nearby sources. To better visualize the brain-behavior relations, we then extracted peak CLARA activations and latencies from each significant cluster in the brain volume and regressed these values against listeners' behavioral identification slopes (see Fig. 7**C, D**).

Author contributions

R. R. and G.M.B. designed the experiment, R.R. collected the data, R. R. and G.M.B. analyzed the data and wrote the paper.

CRediT authorship contribution statement

Rose Rizzi: Writing – original draft, Visualization, Formal analysis, Data curation, Conceptualization. **Gavin M. Bidelman:** Writing – original draft, Visualization, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the National Institute on Deafness and Other Communication Disorders, United States (R01DC016267 to G.M. B.).

Ethical statement

Participants all provided written informed consent to participate and all procedures were performed in compliance with the Indiana University Institutional Review Board (protocol number 14860, approved April 7, 2022).

References

Alain, C., Arnott, S.R., Picton, T.W., 2001. Bottom-up and top-down influences on auditory scene analysis: evidence from event-related brain potentials. J. Exp. Psychol. Hum. Percept. Perform. 27 (5), 1072–1089. https://doi.org/10.1037//0096-1523.27.5.1072.
Alain, C., Campeanu, S., Tremblay, K., 2010. Changes in sensory evoked responses coincide with rapid improvement in speech identification performance. J. Cogn. Neurosci. 22 (2), 392–403. https://doi.org/10.1162/jocn.2009.21279.
Alain, C., McDonald, K., Van Roon, P., 2012. Effects of age and background noise on processing a mistuned harmonic in an otherwise periodic complex sound. Hear. Res. 283 (1–2), 126–135. https://doi.org/10.1016/j.heares.2011.10.007.
Alain, C., Göke, K., Shen, D., Bidelman, G.M., Bernstein, L.J., Snyder, J.S., 2023. Neural alpha oscillations index context-driven perception of ambiguous vowel sequences. iScience 26 (12), 108457. https://doi.org/10.1016/j.isci.2023.108457.
Alho, J., Green, B.M., May, P.J.C., Sams, M., Tiitinen, H., Rauschecker, J.P., Jääskeläinen, I.P., 2016. Early-latency categorical speech sound representations in the left inferior frontal gyrus. Neuroimage 129, 214–223. https://doi.org/10.1016/j.neuroimage.2016.01.016.
Andruski, J.E., Blumstein, S.E., Burton, M., 1994. The effect of subphonetic differences on lexical access. Cognition 52 (3), 163–187. https://doi.org/10.1016/0010-0277(94)90042-6.
Apfelbaum, K.S., Kutlu, E., McMurray, B., Kapnoula, E.C., 2022. Don't force it! Gradient speech categorization calls for continuous categorization tasksa). J. Acoust. Soc. Am. 152 (6), 3728–3745. https://doi.org/10.1121/10.0015201.
Ben-David, B.M., Campeanu, S., Tremblay, K.L., Alain, C., 2011. Auditory evoked potentials dissociate rapid perceptual learning from task repetition without learning. Psychophysiology 48 (6), 797–807. https://doi.org/10.1111/j.1469-8986.2010.01139.x.
Berg, P., Scherg, M., 1994. A fast method for forward computation of multiple-shell spherical head models. Electroencephalogr. Clin. Neurophysiol. 90 (1), 58–64.
Bergman, M., 1971. Hearing and aging: implications of recent research findings. Audiology 10 (3), 164–171. https://doi.org/10.3109/00206097109072554.
Bharadwaj, H.M., Masud, S., Mehraei, G., Verhulst, S., Shinn-Cunningham, B.G., 2015. Individual differences reveal correlates of hidden hearing deficits. J. Neurosci. 35 (5), 2161–2172. https://doi.org/10.1523/jneurosci.3915-14.2015.
Bidelman, G.M., 2015. Induced neural beta oscillations predict categorical speech perception abilities. Brain Lang. 141, 62–69. https://doi.org/10.1016/j.bandl.2014.11.003.
Bidelman, G.M., Alain, C., 2015a. Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. J. Neurosci. 35 (3), 1240–1249. https://doi.org/10.1523/jneurosci.3292-14.2015.
Bidelman, G.M., Alain, C., 2015b. Hierarchical neurocomputations underlying concurrent sound segregation: Connecting periphery to percept. Neuropsychologia 68, 38–50. https://doi.org/10.1016/j.neuropsychologia.2014.12.020.
Bidelman, G.M., Bush, L.C., Boudreaux, A.M., 2020. Effects of noise on the behavioral and neural categorization of speech [Brief Research Report]. Front. Neurosci. 14 https://doi.org/10.3389/fnins.2020.00153.
Bidelman, G.M., Brown, J.A., Bashivan, P., 2021. Auditory cortex supports verbal working memory capacity. Neuroreport 32 (2), 163–168.
Bidelman, G.M., Carter, J.A., 2023. Continuous dynamics in behavior reveal interactions between perceptual warping in categorization and speech-in-noise perception. Front. Neurosci. 17 (1032369), 1–13. https://doi.org/10.3389/fnins.2023.1032369.
Bidelman, G.M., Walker, B., 2019. Plasticity in auditory categorization is supported by differential engagement of the auditory-linguistic network. Neuroimage 201, 116022. https://doi.org/10.1016/j.neuroimage.2019.116022.
Bidelman, G.M., Yoo, J., 2020. Musicians show improved speech segregation in competitive, multi-talker cocktail party scenarios [Brief Research Report]. Front. Psychol. 11 (1927), 1–11. https://doi.org/10.3389/fpsyg.2020.01927.
Bidelman, G.M., Moreno, S., Alain, C., 2013. Tracing the emergence of categorical speech perception in the human auditory system. Neuroimage 79, 201–212. https://doi.org/10.1016/j.neuroimage.2013.04.093.
Bidelman, G.M., Davis, M.K., Pridgen, M.H., 2018. Brainstem-cortical functional connectivity for speech is differentially challenged by noise and reverberation. Hear. Res. 367, 149–160. https://doi.org/10.1016/j.heares.2018.05.018.
Bidelman, G. M., Bernard, F., & Skubic, K. (2024). Hearing in categories aids speech streaming at the "cocktail party". *bioRxiv*, 2024.2004.2003.587795. https://doi.org/10.1101/2024.04.03.587795.

Bidelman, G.M., Howell, M., 2016. Functional changes in inter- and intra-hemispheric cortical processing underlying degraded speech perception. Neuroimage 124, 581–590. https://doi.org/10.1016/j.neuroimage.2015.09.020.
Bidelman, G.M., Lee, C.-C., 2015. Effects of language experience and stimulus context on the neural organization and categorical perception of speech. Neuroimage 120, 191–200.
Bidelman, G.M., Sigley, L., Lewis, G.A., 2019. Acoustic noise and vision differentially warp the auditory categorization of speech. J. Acoust. Soc. Am. 146 (1), 60. https://doi.org/10.1121/1.5114822.
Bidelman, G.M., Walker, B.S., 2017. Attentional modulation and domain-specificity underlying the neural organization of auditory categorical perception. Eur. J. Neurosci. 45 (5), 690–699. https://doi.org/10.1111/ejn.13526.
Billings, C.J., Tremblay, K.L., Stecker, G.C., Tolin, W.M., 2009. Human evoked cortical activity to signal-to-noise ratio and absolute signal level. Hear. Res. 254 (1–2), 15–24. https://doi.org/10.1016/j.heares.2009.04.002.
Binder, J.R., Liebenthal, E., Possing, E.T., Medler, D.A., Ward, B.D., 2004. Neural correlates of sensory and decision processes in auditory object identification. Nat. Neurosci. 7 (3), 295–301. https://doi.org/10.1038/nn1198.
Blumstein, S.E., Myers, E.B., Rissman, J., 2005. The perception of voice onset time: an fMRI investigation of phonetic category structure. J. Cogn. Neurosci. 17 (9), 1353–1366. https://doi.org/10.1162/0898929054985473.
Bradlow, A.R., Kraus, N., Hayes, E., 2003. Speaking clearly for children with learning disabilities: sentence perception in noise. J. Speech Lang. Hear. Res. 46 (1), 80–97. https://doi.org/10.1044/1092-4388(2003/007).
Calcus, A., Lorenzi, C., Collet, G., Colin, C., Kolinsky, R., 2016. Is there a relationship between speech identification in noise and categorical perception in children with dyslexia? J. Speech Lang. Hear. Res. 59 (4), 835–852. https://doi.org/10.1044/2016_JSLHR-H-15-0076.
Carter, J.A., Bidelman, G.M., 2021. Auditory cortex is susceptible to lexical influence as revealed by informational vs. energetic masking of speech categorization. Brain Res. 1759, 147385. https://doi.org/10.1101/2020.10.20.347724.
Carter, J.A., Buder, E.H., Bidelman, G.M., 2022. Nonlinear dynamics in auditory cortical activity reveal the neural basis of perceptual warping in speech categorization. JASA Express Lett. 2 (4), 045201 https://doi.org/10.1121/10.0009896.
Chang, E.F., Rieger, J.W., Johnson, K., Berger, M.S., Barbaro, N.M., Knight, R.T., 2010. Categorical speech representation in human superior temporal gyrus. Nat. Neurosci. 13 (11), 1428–1432. https://doi.org/10.1038/nn.2641.
Clayards, M., Tanenhaus, M.K., Aslin, R.N., Jacobs, R.A., 2008. Perception of speech reflects optimal use of probabilistic speech cues. Cognition 108 (3), 804–809. https://doi.org/10.1016/j.cognition.2008.04.004.
Crowley, K.E., Colrain, I.M., 2004. A review of the evidence for P2 being an independent component process: age, sleep and modality. Clin. Neurophysiol. 115 (4), 732–744. https://doi.org/10.1016/j.clinph.2003.11.021.
Cunningham, J., Nicol, T., Zecker, S.G., Bradlow, A., Kraus, N., 2001. Neurobiologic responses to speech in noise in children with learning problems: Deficits and strategies for improvement. Clin. Neurophysiol. 112 (5), 758–767 https://doi.org/S1388-2457(01)00465-5 [pii].
Dole, M., Hoen, M., Meunier, F., 2012. Speech-in-noise perception deficit in adults with dyslexia: Effects of background type and listening configuration. Neuropsychologia 50 (7), 1543–1552. https://doi.org/10.1016/j.neuropsychologia.2012.03.007.
Dole, M., Meunier, F., Hoen, M., 2014. Functional correlates o fthes peech-in-noise perception impairment in dyslexia: An MRI study. Neuropsychologia 60, 103–114.
Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2014. Noise differentially impacts phoneme representations in the auditory and speech motor systems. PNAS 111 (19), 7126–7131. https://doi.org/10.1073/pnas.131873811.
Dykstra, A.R., Halgren, E., Gutschalk, A., Eskandar, E.N., Cash, S.S., 2016. Neural correlates of auditory perceptual awareness and release from informational masking recorded directly from human cortex: A case study [Original Research]. Front. Neurosci. 10 https://doi.org/10.3389/fnins.2016.00472.
Elmahallawi, T.H., Gabr, T.A., Darwish, M.E., Seleem, F.M., 2021. Specific language impairment: speech perception in relation to noise. Egypt. J. Otolaryngol. 37 (1), 52. https://doi.org/10.1186/s43163-021-00115-8.
Fuhrmeister, P., Myers, E.B., 2021. Structural neural correlates of individual differences in categorical perception. Brain Lang. 215, 104919 https://doi.org/10.1016/j.bandl.2021.104919.
Godfrey, J.J., Syrdal-Lasky, A.K., Millay, K.K., Knox, C.M., 1981. Performance of dyslexic children on speech perception tests. J. Exp. Child Psychol. 32 (3), 401–424. https://doi.org/10.1016/0022-0965(81)90105-3.
Golestani, N., Zatorre, R.J., 2004. Learning new sounds of speech: reallocation of neural substrates. Neuroimage 21 (2), 494–506. https://doi.org/10.1016/j.neuroimage.2003.09.071.
Grady, C.L., Grimes, A.M., Patronas, N., Sunderland, T., Foster, N.L., Rapoport, S.I., 1989. Divided attention, as measured by dichotic speech performance, in dementia of the Alzheimer type. Arch. Neurol. 46 (3), 317–320. https://doi.org/10.1001/archneur.1989.00520390083021.
Guenther, F.H., Nieto-Castanon, A., Ghosh, S.S., Tourville, J.A., 2004. Representation of sound categories in auditory cortical maps. J. Speech Lang. Hear. Res. 47 (1), 46–57. https://doi.org/10.1044/1092-4388(2004/005).
Gustafson, S.J., Billings, C.J., Hornsby, B.W.Y., Key, A.P., 2019. Effect of competing noise on cortical auditory evoked potentials elicited by speech sounds in 7- to 25-year-old listeners. Hear. Res. 373, 103–112. https://doi.org/10.1016/j.heares.2019.01.004.
Hallé, P.A., Chang, Y.-C., Best, C.T., 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. J. Phon. 32 (3), 395–421. https://doi.org/10.1016/S0095-4470(03)00016-0.
Hannula, S., Bloigu, R., Majamaa, K., Sorri, M., Mäki-Torkko, E., 2011. Self-reported hearing problems among older adults: prevalence and comparison to measured hearing impairment. J. Am. Acad. Audiol. 22 (8), 550–559. https://doi.org/10.3766/jaaa.22.8.7.
Hartigan, J.A., Hartigan, P.M., 1985. The dip test of unimodality. Ann. Stat. 13 (1), 70–84. https://doi.org/10.1214/aos/1176346577, 15.
Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8 (5), 393–402 https://doi.org/nrn2113 [pii]10.1038/nrn2113.
Hoover, E.C., Souza, P.E., Gallun, F.J., 2017. Auditory and cognitive factors associated with speech-in-noise complaints following mild traumatic brain injury. J. Am. Acad. Audiol. 28 (4), 325–339. https://doi.org/10.3766/jaaa.16051.
Huette, S., McMurray, B., 2010. Continuous dynamics of color categorization. Psychon. Bull. Rev. 17 (3), 348–354. https://doi.org/10.3758/PBR.17.3.348.
Hugdahl, K., Thomsen, T., Ersland, L., Morten Rimol, L., Niemi, J., 2003. The effects of attention on speech perception: An fMRI study. Brain Lang. 85 (1), 37–48. https://doi.org/10.1016/S0093-934X(02)00500-X.
Humes, L., Kidd, G., Lentz, J., 2013. Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults [Original Research]. Front. Syst. Neurosci. 7 https://doi.org/10.3389/fnsys.2013.00055.
Iordanov, T., Hoechstetter, K., Berg, P., Paul-Jordanov, I., & Scherg, M. (2014). *CLARA: classical LORETA analysis recursively applied* OHBM 2014.
Iordanov, T., Bornfleth, H., Hoechstetter, K., Lanfer, B., & Scherg, M. (2016). *Performance of Cortical LORETA and Cortical CLARA Applied to MEG Data* Biomag 2016.
Joanisse, M.F., Manis, F.R., Keating, P., Seidenberg, M.S., 2000. Language deficits in dyslexic children: speech perception, phonology, and morphology. J. Exp. Child Psychol. 77 (1), 30–60. https://doi.org/10.1006/jecp.1999.2553.
Kaplan-Neeman, R., Kishon-Rabin, L., Henkin, Y., Muchnik, C., 2006. Identification of syllables in noise: electrophysiological and behavioral correlates. J. Acoust. Soc. Am. 120 (2), 926–933. https://doi.org/10.1121/1.2217567.
Kapnoula, E.C., Winn, M.B., Kong, E.J., Edwards, J., McMurray, B., 2017. Evaluating the sources and functions of gradiency in phoneme categorization: an individual differences approach. Human Percept. Perform. 43 (9), 1594–1611. https://doi.org/10.1037/xhp0000410.
Kapnoula, E.C., Edwards, J., McMurray, B., 2021. Gradient activation of speech categories facilitates listeners' recovery from lexical graden paths, but not perception of speech-in-noise. Human Percept. Perform. 47 (4), 578–595. https://doi.org/10.1037/xhp0000900.
Killion, M.C., Niquette, P.A., Gudmundsen, G.I., Revit, L.J., Banerjee, S., 2004. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 116 (4 Pt 1), 2395–2405. https://doi.org/10.1121/1.1784440.
Knight, R.T., Scabini, D., Woods, D.L., Clayworth, C.C., 1989. Contributions of temporal-parietal junction to the human auditory P3. Brain Res. 502 (1), 109–116. https://doi.org/10.1016/0006-8993(89)90466-6.
Koerner, T.K., Zhang, Y., 2015. Effects of background noise on inter-trial phase coherence and auditory N1–P2 responses to speech stimuli. Hear. Res. 328, 113–119. https://doi.org/10.1016/j.heares.2015.08.002.
Kong, E.J., Edwards, J., 2016. Individual differences in categorical perception of speech: Cue weighting and executive function. J. Phon. 59, 40–57. https://doi.org/10.1016/j.wocn.2016.08.006.
Kumar, S., Joseph, S., Gander, P.E., Barascud, N., Halpern, A.R., Griffiths, T.D., 2016. A brain system for auditory working memory. J. Neurosci. 36 (16), 4492–4505. https://doi.org/10.1523/jneurosci.4341-14.2016.
Lagacé, J., Jutras, B., Gagné, J.P., 2010. Auditory processing disorder and speech perception problems in noise: finding the underlying origin. Am. J. Audiol. 19 (1), 17–25. https://doi.org/10.1044/1059-0889(2010/09-0022).
Lee, Y.-S., Turkeltaub, P., Granger, R., Raizada, R.D.S., 2012. Categorical speech processing in Broca's area: An fMRI study using multivariate pattern-based Analysis. J. Neurosci. 32 (11), 3942–3948. https://doi.org/10.1523/jneurosci.3814-11.2012.
Lewis, G., Bidelman, G.M., 2020. Autonomic nervous system correlates of speech categorization revealed through pupillometry. Front. Neurosci. 13 (1418), 1–10. https://doi.org/10.3389/fnins.2019.01418.
Liberman, A.M., Cooper, F.S., Shankweiler, D.P., Studdert-Kennedy, M., 1967. Perception of the speech code. Psychol. Rev. 74 (6), 431–461. https://doi.org/10.1037/h0020279.
Lins, O.G., Picton, T.W., Berg, P., Scherg, M., 1993. Ocular artifacts in recording EEGs and event-related potentials II: Source dipoles and source components. Brain Topogr. 6 (1), 65–78. https://doi.org/10.1007/BF01234128.
Mankel, K., Shrestha, U., Tipirneni-Sajja, A., Bidelman, G.M., 2022. Functional plasticity coupled with structural predispositions in auditory cortex shape successful music category learning [Original Research]. Front. Neurosci. 16 https://doi.org/10.3389/fnins.2022.897239.
Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164 (1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.
Massaro, D.W., Cohen, M.M., 1983a. Categorical or continuous speech perception: A new test. Speech Comm. 2 (1), 15–35. https://doi.org/10.1016/0167-6393(83)90061-4.
Massaro, D.W., Cohen, M.M., 1983b. Phonological context in speech perception. Percept. Psychophys. 34 (4), 338–348. https://doi.org/10.3758/BF03203046.
McMurray, B., 2022. The myth of categorical perceptiona. J. Acoust. Soc. Am. 152 (6), 3819–3842. https://doi.org/10.1121/10.0016614.
McMurray, B., Tanenhaus, M.K., Aslin, R.N., 2002. Gradient effects of within-category phonetic variation on lexical access. Cognition 86 (2), B33–B42. https://doi.org/10.1016/S0010-0277(02)00157-9.
McMurray, B., Aslin, R.N., Tanenhaus, M.K., Spivey, M.J., Subik, D., 2008. Gradient sensitivity to within-category variation in words and syllables. J. Exp. Psychol. Hum. Percept. Perform. 34 (6), 1609–1631. https://doi.org/10.1037/a0011747.

McMurray, B., Danelz, A., Rigler, H., Seedorff, M., 2018. Speech categorization develops slowly through adolescence. Dev. Psychol. 54 (8), 1472–1491. https://doi.org/10.1037/dev0000542.
Meyers, E.M., Freedman, D.J., Kreiman, G., Miller, E.K., Poggio, T., 2008. Dynamic population coding of category information in inferior temporal and prefrontal cortex. J. Neurophysiol. 100 (3), 1407–1419. https://doi.org/10.1152/jn.90248.2008.
Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG source imaging. Clin. Neurophysiol. 115 (10), 2195–2222. https://doi.org/10.1016/j.clinph.2004.06.001.
Miller, J.L., Volaitis, L.E., 1989. Effect of speaking rate on the perceptual structure of a phonetic category. Percept. Psychophys. 46 (6), 505–512. https://doi.org/10.3758/BF03208147.
Munson, B., Schellinger, S.K., Edwards, J., 2017. Bias in the perception of phonetic detail in children's speech: A comparison of categorical and continuous rating scales. Clin. Linguist. Phon. 31 (1), 56–79. https://doi.org/10.1080/02699206.2016.1233292.
Myers, E.B., Blumstein, S.E., Walsh, E., Eliassen, J., 2009. Inferior frontal regions underlie the perception of phonetic category invariance. Psychol. Sci. 20 (7), 895–903. https://doi.org/10.1111/j.1467-9280.2009.02380.x.
Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9 (1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4.
Oostenveld, R., Praamstra, P., 2001. The five percent electrode system for high-resolution EEG and ERP measurements. Clin. Neurophysiol. 112 (4), 713–719. https://doi.org/10.1016/s1388-2457(00)00527-7.
Papesh, M.A., Billings, C.J., Baltzell, L.S., 2015. Background noise can enhance cortical auditory evoked potentials under certain conditions. Clin. Neurophysiol. 126 (7), 1319–1330. https://doi.org/10.1016/j.clinph.2014.10.017.
Pascual-Marqui, R.D., Esslen, M., Kochi, K., Lehmann, D., 2002. Functional imaging with low-resolution brain electromagnetic tomography (LORETA): a review. Methods Find. Exp. Clin. Pharmacol. 24 Suppl C, 91–95.
Picard, M., Banville, R., Barbarosie, T., Manolache, M., 1999. Speech audiometry in noise-exposed workers: The SRT-PTA relationship revisited. Audiology 38 (1), 30–43. https://doi.org/10.3109/00206099909073000.
Picton, T.W., Hillyard, S.A., 1974. Human auditory evoked potentials. II: Effects of attention. Electroencephalogr. Clin. Neurophysiol. 36, 191–200. https://doi.org/10.1016/0013-4694(74)90156-4.
Picton, T.W., Alain, C., Woods, D.L., John, M.S., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., Trujillo, N.J., 1999. Intracerebral sources of human auditory-evoked potentials. Audiol. Neuro Otol. 4 (2), 64–79 https://doi.org/13823 [pii].
Picton, T.W., van Roon, P., Armilio, M.L., Berg, P., Ille, N., Scherg, M., 2000. The correction of ocular artifacts: a topographic perspective. Clin. Neurophysiol. 111 (1), 53–65. https://doi.org/10.1016/S1388-2457(99)00227-8.
Pisoni, D.B., 1973. Auditory and phonetic memory codes in the discrimination of consonants and vowels. Percept. Psychophys. 13 (2), 253–260. https://doi.org/10.3758/BF03214136.
Pisoni, D.B., Tash, J., 1974. Reaction times to comparisons within and across phonetic categories. Percept. Psychophys. 15 (2), 285–290. https://doi.org/10.3758/BF03213946.
Plomp, R., 1978. Auditory handicap of hearing impairment and the limited benefit of hearing aids. J. Acoust. Soc. Am. 63 (2), 533–549. https://doi.org/10.1121/1.381753.
Putter-Katz, H., Adi-Bensaid, L., Feldman, I., Hildesheimer, M., 2008. Effects of speech in noise and dichotic listening intervention programs on central auditory processing disorders. J. Basic Clin. Physiol. Pharmacol. 19 (3–4), 301–316. https://doi.org/10.1515/jbcpp.2008.19.3-4.301.
Ramos Nuñez, A.I., Yue, Q., Pasalar, S., Martin, R.C., 2020. The role of left vs. right superior temporal gyrus in speech perception: An fMRI-guided TMS study. Brain Lang. 209, 104838 https://doi.org/10.1016/j.bandl.2020.104838.
Richards, J.E., Sanchez, C., Phillips-Meek, M., Xie, W., 2016. A database of age-appropriate average MRI templates. Neuroimage 124 (Pt B), 1254–1259. https://doi.org/10.1016/j.neuroimage.2015.04.055.
Rogers, C.L., Dalby, J., Nishi, K., 2004. Effects of noise and proficiency on intelligibility of Chinese-Accented English. Lang. Speech 47 (2), 139–154. https://doi.org/10.1177/00238309040470020201.
Rogers, C.L., Lister, J.J., Febo, D.M., Besing, J.M., Abrams, H.B., 2006. Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. Appl. Psycholinguist. 27 (3), 465–485. https://doi.org/10.1017/S014271640606036X.
Ross, B., Jamali, S., Tremblay, K.L., 2013. Plasticity in neuromagnetic cortical responses suggests enhanced auditory object representation. BMC Neurosci. 14 (1), 151. https://doi.org/10.1186/1471-2202-14-151.
Ruggles, D., Bharadwaj, H., Shinn-Cunningham, B.G., 2011. Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. PNAS 108 (37), 15516–15521. https://doi.org/10.1073/pnas.1108912108.
Sarrett, M.E., McMurray, B., Kapnoula, E.C., 2020. Dynamic EEG analysis during language comprehension reveals interactive cascades between perceptual processing and sentential expectations. Brain Lang. 211, 104875 https://doi.org/10.1016/j.bandl.2020.104875.
Scherg, M., Berg, P., Nakasato, N., Beniczky, S., 2019. Taking the EEG back into the brain: The power of multiple discrete sources [Original Research]. Front. Neurol. 10 (855) https://doi.org/10.3389/fneur.2019.00855.
Serniclaes, W., Sprenger-Charolles, L., Carré, R., Demonet, J.F., 2001. Perceptual discrimination of speech sounds in developmental dyslexia. J. Speech Lang. Hear. Res. 44 (2), 384–399. https://doi.org/10.1044/1092-4388(2001/032).
Serniclaes, W., Ventura, P., Morais, J., Kolinsky, R., 2005. Categorical perception of speech sounds in illiterate adults. Cognition 98 (2), B35–B44. https://doi.org/10.1016/j.cognition.2005.03.002.
Sheldon, S., Pichora-Fuller, M.K., Schneider, B.A., 2008. Priming and sentence context support listening to noise-vocoded speech by younger and older adults. J. Acoust. Soc. Am. 123 (1), 489–499. https://doi.org/10.1121/1.2783762.
Sussman, J.E., 1993. Perception of formant transition cues to place of articulation in children with language impairments. J. Speech Lang. Hear. Res. 36 (6), 1286–1299. https://doi.org/10.1044/jshr.3606.1286.
Toscano, J.C., Anderson, N.D., Fabiani, M., Gratton, G., Garnsey, S.M., 2018. The time-course of cortical responses to speech revealed by fast optical imaging. Brain Lang. 184, 32–42. https://doi.org/10.1016/j.bandl.2018.06.006.
Toscano, J.C., McMurray, B., 2010. Cue integration with categories: weighting acoustic cues in speech using unsupervised learning and distributional statistics. Cognit. Sci. 34 (3), 434–464. https://doi.org/10.1111/j.1551-6709.2009.01077.x.
Toscano, J.C., McMurray, B., Dennhardt, J., Luck, S.J., 2010. Continuous perception and graded categorization: electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. Psychol. Sci. 21 (10), 1532–1540. https://doi.org/10.1177/0956797610384142.
Tremblay, K.L., Pinto, A., Fischer, M.E., Klein, B.E., Klein, R., Levy, S., Tweed, T.S., Cruickshanks, K.J., 2015. Self-reported hearing difficulties among adults with normal audiograms: the beaver dam offspring study. Ear Hear. 36 (6), e290–e299. https://doi.org/10.1097/aud.0000000000000195.
Turkeltaub, P.E., Branch Coslett, H., 2010. Localization of sublexical speech perception components. Brain Lang. 114 (1), 1–15. https://doi.org/10.1016/j.bandl.2010.03.008.
Van Engen, K.J., Phelps, J.E.B., Smiljanic, R., Chandrasekaran, B., 2014. Enhancing speech intelligibility: interactions among context, modality, speech style, and masker. J. Speech Lang. Hear. Res. 57 (5), 1908–1918. https://doi.org/10.1044/JSLHR-H-13-0076.
Vander Werff, K.R., Rieger, B., 2019. Impaired auditory processing and neural representation of speech in noise among symptomatic post-concussion adults. Brain Inj. 33 (10), 1320–1331. https://doi.org/10.1080/02699052.2019.1641624.
Viswanathan, N., Kelty-Stephen, D.G., 2018. Comparing speech and nonspeech context effects across timescales in coarticulatory contexts. Atten. Percept. Psychophys. 80 (2), 316–324. https://doi.org/10.3758/s13414-017-1449-8.
Wallstrom, G.L., Kass, R.E., Miller, A., Cohn, J.F., Fox, N.A., 2004. Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods. Int. J. Psychophysiol. 53 (2), 105–119. https://doi.org/10.1016/j.ijpsycho.2004.03.007.
Warrier, C.M., Johnson, K.L., Hayes, E.A., Nicol, T., Kraus, N., 2004. Learning impaired children exhibit timing deficits and training-related improvements in auditory cortical responses to speech in noise. Exp. Brain Res. 157 (4), 431–441. https://doi.org/10.1007/s00221-004-1857-6.
Werker, J.F., Tees, R.C., 1987. Speech perception in severely disabled and average reading children. Canad. J. Psychol./Revue Canadienne De Psychologie 41 (1), 48–61. https://doi.org/10.1037/h0084150.
Whiting, K.A., Martin, B.A., Stapells, D.R., 1998. The effects of broadband noise masking on cortical event-related potentials to speech sounds /ba/ and /da. Ear Hear. 19 (3), 218–231. https://doi.org/10.1097/00003446-199806000-00005.
Wong, P.C.M., Uppunda, A.K., Parrish, T.B., Dhar, S., 2008. Cortical mechanisms of speech perception in noise. J. Speech Lang. Hear. Res. 51 (4), 1026–1041. https://doi.org/10.1044/1092-4388(2008/075.
Xu, Y., Gandour, J.T., Francis, A.L., 2006. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. J. Acoust. Soc. Am. 120 (2), 1063–1074. https://doi.org/10.1121/1.2213572.
Zatorre, R.J., Evans, A.C., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch discrimination in speech processing. Science 256 (5058), 846–849. https://doi.org/10.1126/science.1589767.
Ziegler, J.C., Pech-Georgel, C., George, F., Lorenzi, C., 2009. Speech-perception-in-noise deficits in dyslexia. Dev. Sci. 12 (5), 732–745. https://doi.org/10.1111/j.1467-7687.2009.00817.x.