

University of Groningen

## Bottom-Up Signal Quality Impacts the Role of Top-Down Cognitive-Linguistic Processing During Speech Recognition by Adults with Cochlear Implants

Moberly, Aaron C.; Lewis, Jessica H.; Vasil, Kara J.; Ray, Christin; Tamati, Terrin N.

*Published in:*

Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology

*DOI:*

[10.1097/MAO.0000000000003377](https://doi.org/10.1097/MAO.0000000000003377)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2021

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Moberly, A. C., Lewis, J. H., Vasil, K. J., Ray, C., & Tamati, T. N. (2021). Bottom-Up Signal Quality Impacts the Role of Top-Down Cognitive-Linguistic Processing During Speech Recognition by Adults with Cochlear Implants. *Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 42(10S), S33-S41. <https://doi.org/10.1097/MAO.0000000000003377>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Bottom-Up Signal Quality Impacts the Role of Top-Down Cognitive-Linguistic Processing During Speech Recognition by Adults with Cochlear Implants

\*Aaron C. Moberly, \*Jessica H. Lewis, \*Kara J. Vasil, \*Christin Ray, and \*†Terrin N. Tamati

\*Department of Otolaryngology – Head & Neck Surgery, The Ohio State University Wexner Medical Center, Columbus, Ohio, USA; and †Department of Otorhinolaryngology – Head and Neck Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

**Hypotheses:** Significant variability persists in speech recognition outcomes in adults with cochlear implants (CIs). Sensory (“bottom-up”) and cognitive-linguistic (“top-down”) processes help explain this variability. However, the interactions of these bottom-up and top-down factors remain unclear. One hypothesis was tested: top-down processes would contribute differentially to speech recognition, depending on the fidelity of bottom-up input.

**Background:** Bottom-up spectro-temporal processing, assessed using a Spectral-Temporally Modulated Ripple Test (SMRT), is associated with CI speech recognition outcomes. Similarly, top-down cognitive-linguistic skills relate to outcomes, including working memory capacity, inhibition-concentration, speed of lexical access, and nonverbal reasoning.

**Methods:** Fifty-one adult CI users were tested for word and sentence recognition, along with performance on the SMRT and a battery of cognitive-linguistic tests. The group was divided into “low-,” “intermediate-,” and “high-SMRT” groups, based on SMRT scores. Separate correlation analyses were performed for each subgroup between a composite

score of cognitive-linguistic processing and speech recognition.

**Results:** Associations of top-down composite scores with speech recognition were not significant for the low-SMRT group. In contrast, these associations were significant and of medium effect size (Spearman’s  $\rho = 0.44-0.46$ ) for two sentence types for the intermediate-SMRT group. For the high-SMRT group, top-down scores were associated with both word and sentence recognition, with medium to large effect sizes (Spearman’s  $\rho = 0.45-0.58$ ).

**Conclusions:** Top-down processes contribute differentially to speech recognition in CI users based on the quality of bottom-up input. Findings have clinical implications for individualized treatment approaches relying on bottom-up device programming or top-down rehabilitation approaches.

**Key Words:** Cochlear implant—Cognition—Sensorineural hearing loss—Spectro-temporal processing—Speech recognition.

*Otol Neurotol* 42:S33–S41, 2021.

---

Address correspondence and reprint requests to Aaron C. Moberly, M.D., The Ohio State University, Columbus, OH; E-mail: aaron.moberly@osumc.edu

Authors ACM and CR received grant support from Cochlear Americas for an unrelated investigator-initiated study of aural rehabilitation; Authors ACM, KJV, and CR serve as paid consultants for Advanced Bionics; Authors ACM and CR serve as paid consultants for Cochlear Americas; Author ACM serves on the Board of Directors and is Chief Medical Officer for Otologic Technologies.

This study was supported by the National Institutes of Health, National Institute on Deafness and Other Communication Disorders Career Development Award 5K23DC015539-02 and the American Otological Society Clinician-Scientist Award to ACM. Preparation of this manuscript was supported in part by VENI Grant No. 275-89-035 from the Netherlands Organization for Scientific Research (NWO) and funding from the President’s Postdoctoral Scholars Program (PPSP) at The Ohio State University awarded to Terrin N. Tamati.

The authors disclose no conflicts of interest.

DOI: 10.1097/MAO.0000000000003377

For postlingual adults with cochlear implants (CIs), broad variability in speech recognition outcomes persists (1). Some of this variability can be explained by clinical predictors, such as duration of deafness, preimplant speech perception scores, patient age, and electrode positioning (1,2). Difficulty in recognizing speech arises in large part because CIs provide degraded spectro-temporal “bottom-up” representations of speech (3,4). For individuals with CIs, cognitive-linguistic skills (i.e., “top-down” processes) allow the listener to compensate for this bottom-up degradation. General models of speech recognition acknowledge an interplay of bottom-up and top-down processes in speech recognition. As such, measures of speech recognition, such as conventional measures used to assess outcomes in CI patients, are tapping into a combination and interaction of bottom-up and top-down

functioning (5–7). Moreover, a listener relies more heavily on top-down processes—to a degree—when the signal is more degraded (8). However, it is unclear how bottom-up and top-down processes interact to explain outcome variability across different listeners, and in particular, adults with CIs. It also remains unclear how the experience of postlingual hearing loss impacts these bottom-up/top-down interactions in adults, while more is known regarding these effects in individuals with prelingual deafness (9). The current study explores these interactions across postlingual adult CI users.

Increased signal degradation results in greater dependence on cognitive-linguistic skills. In one explanatory framework, the Ease of Language Understanding (ELU) model (8), speech input is automatically bound in short-term memory. If this information matches speech sounds in long-term memory, automatic retrieval of the word occurs. In contrast, if there is a mismatch, controlled processing comes into play using higher-level linguistic knowledge (e.g., semantic context). Thus, the use of top-down processing primarily “kicks in” when bottom-up processing fails.

Based on the ELU model, and supported by previous findings (10), it is clear that the quality of the bottom-up input contributes substantially to speech recognition outcomes in CI patients. The current study aimed to investigate the relative role of top-down processing in subgroups of CI users with a range of bottom-up auditory processing ability. A widely employed approach to assessing bottom-up processing is through non-linguistic behavioral measures of spectral and temporal processing, known to demonstrate variability among CI users (11–14). Pertinent to the current study, a modified spectral ripple task, the spectral-temporally modulated ripple test (SMRT), was created to control for confounds of traditional spectral ripple tasks (15). For example, potential confounds include local loudness cues, or cues at the upper or lower frequency boundaries of the stimulus (15). Multiple studies have shown significant correlations of SMRT thresholds with speech perception abilities in adult CI users (10,16,17). Thus, the SMRT has arisen as a promising tool to assess the perceptual fidelity of the signal delivered to the CI listener, with an important caveat. As discussed by multiple authors, the CI speech processor output is unpredictable for stimuli above 2.1 ripples per octave (RPO), leading to distortions related to aliasing (18–20). Thus, SMRT scores above 2.1 RPO should be interpreted cautiously or reassigned an upper limit score of 2.1 RPO. Nonetheless, taking these concerns into consideration, SMRT scores were treated in the current study as an indicator of bottom-up processing.

Previous findings have also established important contributions of individual cognitive-linguistic skills to speech recognition in CI patients. The foundation of the ELU model is that top-down processing depends on working memory (WM) (8), which is responsible for storing and manipulating information temporarily (21,22). In CI users, there is evidence that WM capacity contributes to speech recognition outcomes (23–25).

Additional cognitive-linguistic skills have been linked to speech recognition variability in adult CI users: inhibition-concentration, information-processing speed, and nonverbal reasoning. As listeners process incoming speech, lexical competitors are activated and must be inhibited (6). Faster response times during a task of inhibition-concentration predicted better sentence recognition scores in speech-shaped noise, and also relate to the ability to use semantic context (26,27). Information-processing speed for linguistic information, specifically speed of lexical access, is a likely contributor to successful speech recognition (28,29). In adult CI users, speed of lexical access contributed to recognition of multiple types of sentence materials (27,30). Lastly, nonverbal reasoning (i.e., IQ) appears to play a role in CI top-down processing. Holden and colleagues reported a correlation between a composite cognitive score (including verbal memory, vocabulary, similarities, and matrix reasoning) and word recognition outcomes in adult CI users (31). More recently, Mattingly and colleagues demonstrated a relationship between scores on the Raven’s Matrices task and recognition scores for sentences in CI users (32).

Top-down and bottom-up processes likely interact to impact speech recognition. A more recent model of degraded speech recognition, the Framework for Understanding Effortful Listening (FUEL), theorizes that the allocation of cognitive resources during degraded speech recognition is influenced by attention and motivation (33). For example, neuroimaging studies suggest that attention enhances the processing of degraded speech by engaging higher-order mechanisms that modulate auditory perceptual processing and that attention may shift towards weighting of more informative features and away from uninformative features during degraded speech perception (34,35). The FUEL model hypothesizes that cognitive resources are allocated differently when the quality of the signal is reduced. If task demand is too high or too low, then cognitive demand is lower as the listener mentally disengages from the task. Additionally, research in normal-hearing (NH) listeners suggests that reliance on top-down processing may actually decrease when the sensory degradation is too extreme (36,37). Previous studies in CI users have shown that relatively poorer performing CI users demonstrate a reduced ability to use top-down compensation and suggest a limited role for cognitive skills in CI users with poor auditory sensitivity (38,39). Similarly, Tamati et al. suggested that top-down processes may play a limited role in CI users with the poorest bottom-up auditory sensitivity (10). Thus, the ability to engage neurocognitive resources to compensate for a degraded signal in CI users likely depends on the quality of the signal.

In the current study, experienced adult CI users were tested using measures of word and sentence recognition, along with top-down cognitive-linguistic assessments of WM capacity, inhibition-concentration, speed of lexical access, and nonverbal reasoning. Bottom-up spectrotemporal processing was measured using the SMRT. The main hypothesis was that top-down processes would

contribute differentially to speech recognition outcomes, depending on the fidelity of the bottom-up processing.

## METHODS

### Participants

Fifty-one postlingually deafened experienced adult CI users were tested, with mean age of 66.8 years ( $SD = 9.8$ , range 45–87). Mean patient-reported duration of hearing loss was 38.4 years ( $SD = 19.0$ , range 4–76). All had at least 1 year of CI use (mean 6.7 years,  $SD 6.1$ , range 1–34). Twenty-seven (52.9%) were female. All were native English speakers with corrected near-vision of better than 20/40, general reading/language proficiency standard score  $\geq 70$  on the Wide Range Achievement Test (WRAT, (40)) and lack of cognitive impairment as demonstrated by a raw score  $\geq 24$  on the Mini-Mental State Examination (MMSE, (41)), delivered in a combined auditory and visual format, with instructions delivered in writing. Average screening WRAT and MMSE scores were 97.3 ( $SD 12.0$ , range 77–122) and 28.7 ( $SD 1.3$ , range 26–30), respectively. Socioeconomic status (SES) was assessed using a metric that accounted for occupational and educational levels, ranging from values of 1 (lowest) to 64 (highest) (42). Mean composite SES score was 25.9 ( $SD 14.6$ , range 6–64). Most ( $N = 46$ , 90.2%) had Cochlear CIs, with four patients (7.8%) using Advanced Bionics devices and one using a Med-EL device (2.0%). Thirteen (25.5%) had bilateral CIs, 22 (43.1%) were bimodal listeners (i.e., CI and contralateral hearing aid), and 16 (31.4%) were unilateral CI users. Unaided better-ear pure tone averages (PTA, 0.5, 1, 2, and 4 kHz) were also computed, with mean PTA 94.5 dB HL ( $SD 23.7$ ).

### General Approach and Equipment

This study was approved by the local Institutional Review Board, protocol #2015H0173, and participants underwent informed, written consent. Each participant was tested using a battery of visual cognitive-linguistic measures. Visual tasks (instead of auditory tasks) were used to avoid audibility effects on performance. Participants were also tested using sentence and isolated word recognition measures, along with the SMRT. All participants were tested in a sound-treated booth, using their standard hearing devices (one CI, two CIs, or one CI with contralateral hearing aid if typically worn). Speech materials and SMRT were presented at 68 dB SPL in quiet from a speaker 1 m away at zero degrees azimuth.

### Cognitive-Linguistic Measures

Each of the cognitive-linguistic measures is described below in brief, but details can be found in prior publications (10,17,27).

Working memory was assessed using a visual digit span task. Participants were scored based on their ability to recall a sequence of digits after presentation one at a time on a computer screen. Scores were percent total digits recalled in correct serial order.

Inhibition-concentration was assessed using a computerized version of the Stroop Color-Word Interference test. This test (<http://www.millisecond.com>) measures how quickly participants can identify the correct font color of a word shown. Separate trials including congruent trials (color and color word matched), incongruent trials (color and color word did not match), and control trials (single rectangle of a given color) were recorded with response times for each condition. The control trial represents a user's general processing speed; the congruent response time represents concentration ability; and the incongruent response time represents inhibitory control.

Speed of lexical access was assessed using the Test of Word Reading Efficiency (TOWRE, (43)), in which participants were asked to read as many words as accurately as possible from a list of 108 words in 45 s. The percent words correctly read was used in analyses.

Lastly, nonverbal reasoning was measured using a computer-based, timed version of Raven's Progressive Matrices, a series of progressively more difficult  $3 \times 3$  visual grid patterns presented on a touch screen with the bottom-right square unfilled. Participants were scored based on how many items they correctly completed in 10 minutes.

### Speech Recognition Measures

Speech recognition tests included the Central Institute of the Deaf (CID) W-22 isolated words (44) spoken by a single male talker; isolated words were expected to rely mostly on bottom-up input and less on cognitive-linguistic skills. Harvard Standard sentences (IEEE, (45)) are complex but meaningful sentences, also spoken by a single male talker, which were expected to rely more heavily on cognitive-linguistic skills than CID words. The third speech recognition test was the Perceptually Robust English Sentence Test Open-set (PRESTO) sentences (46), in which sentences are each spoken by a different talker (male or female) across a variety of American dialects. Performance on PRESTO was expected to depend most strongly on cognitive-linguistic skills. Details of each measure can be found in a previous report (17).

### Spectral-Temporally Modulated Ripple test

Participants' spectro-temporal processing was assessed using the Spectral-Temporally Modulated Ripple test (SMRT), developed by Aronoff and Landsberger (15). Stimuli were 202 pure-tone frequency components with amplitudes spectrally modulated by a sine wave. A three-interval, two-alternative forced-choice task was performed in which two of the intervals contained a reference signal with 20 RPO and one contained the target signal. This target, set initially at 0.5 RPO, was modified using a one-up/one-down adaptive procedure with step size of 0.2 RPO. A ripple-detection threshold was calculated based on the last six reversals of each run. Listeners selected the deviant signal. A higher score represented better spectro-temporal processing.

### Data Analyses

Data for several speech recognition (CID words and Harvard Standard sentences) and cognitive-linguistic measures (Stroop, Digit Span), as well as SMRT, were not normally distributed, based on Kolmogorov-Smirnov and Shapiro-Wilk tests of normality. Thus, nonparametric tests were used for analyses. Prior to performing main analyses, Kruskal-Wallis one-way ANOVA was performed to determine if there was any effect of side of implantation (right, left, or bilateral) on SMRT scores, which would suggest that the bottom-up input differed among those three groups. No significant effect was found ( $p = 0.359$ ). Similarly, an independent-samples Mann-Whitney  $U$  test was performed to determine whether SMRT scores differed between CI users who wore a contralateral hearing aid or did not. No significant difference was found ( $p = 0.930$ ). Thus, CI users were collapsed across these conditions for our main analyses.

The 51 CI users were divided into three groups based on SMRT performance: "low-," "intermediate-," and "high-SMRT." The "high-SMRT" group was defined as scoring  $> 2.1$  RPO ( $N = 18$ ), as discussed in the Introduction. The remaining 33 participants were divided by median split into "low-SMRT" ( $N = 16$ ) and "intermediate-SMRT" ( $N = 17$ ).

**TABLE 1.** SMRT, demographic/audiologic, and cognitive-linguistic measures for low-, intermediate-, and high-SMRT participants

Measure	Group			ANOVA <i>p</i>
	Low-SMRT (N = 16) Mean (SD)	Intermediate-SMRT (N = 17) Mean (SD)	High-SMRT (N = 18) Mean (SD)	
SMRT (RPO)	0.92 (0.23)	1.6 (0.22)	3.7 (1.1)	<0.001
Demographic/audiologic				
Age (years)	66.9 (9.4)	68.5 (10.1)	65.2 (10.2)	0.74
SES	27.3 (18.6)	22.7 (14.4)	27.6 (11.1)	0.46
Reading (standard score)	92.4 (9.4)	95.3 (11.5)	103.6 (12.4)	0.04
MMSE (raw score)	28.4 (1.4)	28.9 (1.1)	28.8 (1.3)	0.54
Residual PTA (dB HL)	98.5 (18.9)	95.8 (24.3)	89.4 (27.7)	0.67
Duration Hearing Loss (years)	51.0 (17.5)	29.6 (17.1)	35.1 (16.6)	0.004
Duration CI Use (years)	8.9 (8.7)	5.3 (3.9)	5.9 (4.4)	0.47
Cognitive-linguistic				
Working Memory (score)	40.5 (19.8)	36.9 (11.0)	47.1 (18.2)	0.18
Average inhibition-concentration (ms)	1,472 (712)	1,478 (378)	1,346 (539)	0.24
Speed of Lexical access (percent words)	69.5 (8.1)	69.5 (12.4)	77.7 (9.7)	0.04
Nonverbal Reasoning (score)	9.6 (4.1)	9.2 (3.8)	11.8 (6.1)	0.21

*P* values are results of Kruskal–Wallis one-way analyses of variance (ANOVA).

RPO indicates ripples per octave; SMRT, Spectral-Temporally Modulated Ripple Test.

groups, with the cutoff occurring at a median of 1.3 RPO. Mean SMRT scores for these groups are shown in Table 1. Kruskal–Wallis one-way analyses of variance (ANOVA) confirmed that SMRT scores were highly significantly different among the three groups, with each post hoc pairwise Dunn's comparison also highly significant with Bonferroni correction (all pairwise  $p < 0.001$ ). Demographic, audiologic, and cognitive-linguistic variables were also compared among the three groups, shown in Table 1. Among demographic and audiologic variables, reading score and duration of deafness were significantly different among the three groups. On post hoc pairwise Dunn's tests with Bonferroni correction, reading scores were only significantly different between the low-SMRT and high-SMRT groups ( $p = 0.014$ ). Duration of deafness was greater for low-SMRT than both the intermediate-SMRT ( $p = 0.001$ ) and the high-SMRT ( $p = 0.022$ ) group. Generally, cognitive-linguistic scores were not significantly different among the groups, except for speed of lexical access. Pairwise comparison demonstrated that the only significant difference in speed of lexical access was between low-SMRT and high-SMRT groups ( $p = 0.016$ ).

A final set of Kruskal–Wallis ANOVA analyses was performed to compare speech recognition outcomes among the three groups, with results shown in Table 2. As expected, speech recognition scores differed among the three groups for all three

speech measures, with the low-SMRT group performing the worst, the intermediate-SMRT group scoring in the middle, and the high-SMRT group performing the best on all speech recognition tasks. Post hoc pairwise Dunn's tests with Bonferroni correction, demonstrated all pairwise comparisons to be significant ( $p < 0.05$ ), for all three speech recognition measures, except that intermediate-SMRT and high-SMRT did not differ significantly ( $p = 0.24$ ) on Harvard Standard sentence recognition.

For our main analyses shown in Results, one-tailed non-parametric Spearman's rank-order correlations were performed in each group separately between speech recognition performance measures and cognitive-linguistic processing. One-tailed tests were used based on the consistent prediction across speech materials that better cognitive-linguistic scores would be associated with better speech recognition. To account for the fact that the subgroups were small in size (each  $N < 20$ ), to reduce redundancy among variables, and to represent the variance among the measures with minimal factors, a composite score of cognitive-linguistic processing was computed. To create this composite score, the cognitive-linguistic data were subjected to a principal-components factor analysis (PCA), following Humes, Kidd, and Lentz (47). A single factor accounted for 51.39% of variance (KMO sampling adequacy statistic = 0.73; communalities = 0.65–0.86). The component

**TABLE 2.** Speech recognition measures for low-, intermediate-, and high-SMRT participants

Measure	Group			ANOVA <i>p</i>
	Low-SMRT (N = 16) Mean (SD)	Intermediate-SMRT (N = 17) Mean (SD)	High-SMRT (N = 18) Mean (SD)	
CID Words (% correct)	52.5 (15.0)	65.9 (23.1)	82.5 (11.8)	< .001
Harvard Sentences (% words)	63.0 (18.7)	77.5 (16.5)	82.7 (11.1)	< .001
PRESTO Sentences (% key words)	41.9 (16.7)	57.7 (16.6)	73.1 (18.7)	< .001

*P* values are results of Kruskal–Wallis one-way analyses of variance (ANOVA).

CID indicates Central Institute of the Deaf; SD, standard deviation; SMRT, Spectral-Temporally Modulated Ripple Test.

**TABLE 3.** Results of Spearman’s rank-order correlations among cognitive-linguistic composite score and speech recognition measures for each subgroup: Low-SMRT, Intermediate-SMRT, and High-SMRT; Spearman’s rho and p value are bolded where  $p < 0.05$

Speech Measure		Group		
		Low-SMRT (N = 16)	Intermediate-SMRT (N = 17)	High-SMRT (N = 18)
CID words (% correct)	Spearman’s rho	0.28	−0.10	−0.58
	<i>p</i>	0.15	0.35	0.007
Harvard sentences (% words)	Spearman’s rho	0.26	−0.44	−0.45
	<i>p</i>	0.17	0.04	0.03
PRESTO sentences (% key words)	Spearman’s rho	0.39	−0.46	−0.50
	<i>p</i>	0.07	0.04	0.02

CID indicates Central Institute of the Deaf; SMRT, Spectral-Temporally Modulated Ripple Test.

weights of each cognitive-linguistic measure on the single factor are as follows: Digit Span (WM) = −0.26; Stroop-control (processing speed) = 0.92; Stroop-congruent (concentration) = 0.93; Stroop-incongruent (inhibition) = 0.90; TOWRE (speed of lexical access) = −0.37; Raven’s Progressive Matrices (nonverbal reasoning) = −0.61. Note that for Stroop scores, lower values (shorter response times) represent better performance. Thus, the pattern of component weights across measures was consistent, and the resulting factor was interpreted as a single “cognitive-linguistic composite” score. Categorization of effect sizes for Spearman correlation coefficients are as follows: small effects, Spearman coefficient 0.10 to 0.29; medium effects, Spearman coefficient 0.30 to 0.49; and large effects, Spearman coefficient  $\geq 0.50$ . An  $\alpha$  of  $<0.05$  was considered significant.

**RESULTS**

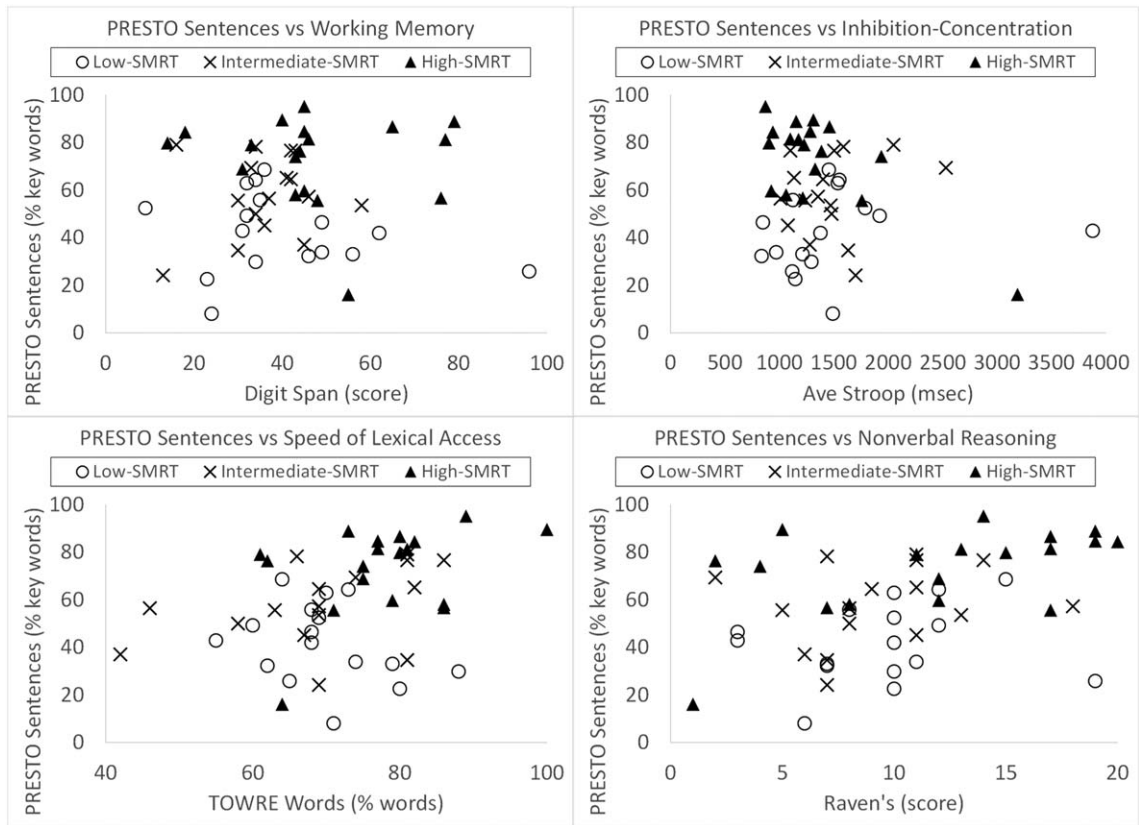
To address our main hypothesis, correlation analyses were performed among cognitive-linguistic composite scores and speech recognition for each group separately. Results are shown in Table 3. Note that, as a result of incorporating Stroop response times (with shorter response time representing better performance) into the creation of the composite cognitive-linguistic score through PCA, a *negative* correlation coefficient suggests that better cognitive-linguistic processing is associated with better speech recognition. Differential relations between cognitive-linguistic composite scores and speech recognition scores were found among the three groups. (1) For the low-SMRT group, no significant correlations were found among the composite cognitive-linguistic score and any speech recognition measure. (2) For the intermediate-SMRT group, significant correlations were found between the composite cognitive-linguistic score and both sentence recognition types, Harvard Standard and PRESTO, and these were of medium effect size. (3) For the high-SMRT group, significant correlations were found between the composite cognitive-linguistic score and all three speech recognition measures. For CID words and PRESTO sentences, these correlations were of large effect size; for Harvard Standard sentences, the correlation was of medium effect size.

As discussed above, a composite cognitive-linguistic score was used for our main analyses because the subgroups were small in size. However, scatterplots of participant

data are shown in Figure 1 for the three SMRT subgroups, representing PRESTO sentence recognition scores versus each of the cognitive-linguistic tasks of working memory, inhibition-concentration, speed of lexical access, and nonverbal reasoning. PRESTO scores were chosen for Figure 1 because PRESTO was the speech recognition measure for which we expected the strongest effect of cognitive-linguistic skills. Exploratory one-tailed Spearman correlation analyses were performed among each speech recognition measure and each cognitive-linguistic skill, with results shown in Table 4; these were treated as exploratory because they led to a large number of correlation analyses for such small subgroups. Nonetheless, these exploratory analyses provide some evidence for which cognitive-linguistic measures might be contributing most strongly to speech recognition performance for the three subgroups and for the three speech recognition outcomes. Overall, only a few weak to moderate-sized correlations were identified among speech recognition measures and cognitive-linguistic measures in the low-SMRT group (i.e., CID word recognition and nonverbal reasoning). More moderate-sized correlations were identified broadly among speech recognition and cognitive-linguistic measures in the intermediate-SMRT and high-SMRT groups, with the patterns of relations appearing to differ between the two groups. For example, inhibition-concentration related more to speech recognition in the high-SMRT group, while working memory related more to speech recognition in the intermediate-SMRT group. However, it should be noted that these results should be interpreted with caution in these small subgroups of participants.

**DISCUSSION**

Cognitive-linguistic functions contribute to the ability to recognize degraded speech (e.g., (48,49)), and specifically so for adult CI users (10,27). Because the input provided by a CI is spectro-temporally limited, CI users must tap into top-down cognitive-linguistic functions to understand speech. However, most prior studies on this topic in adult CI listeners have been limited by small sample sizes, and/or have failed to capture measures of both top-down functions as well as the quality of bottom-up input. One study used discriminant analysis approaches



**FIG. 1.** Scatterplots of PRESTO sentence recognition scores versus the individual cognitive-linguistic measures for the three subgroups based on bottom-up SMRT performance. SMRT indicates Spectral-Temporally Modulated Ripple Test.

**TABLE 4.** Results of Spearman's rank-order correlations among cognitive-linguistic measures, divided by SMRT subgroup; Spearman's rho and p value are bolded where  $p < 0.05$

Speech Measure		Top-down Cognitive-Linguistic Measure				
		Working Memory (score)	Average Inhibition-Concentration (ms)	Speed of Lexical Access (% Correct)	Nonverbal Reasoning (Score)	
Low-smrt group	CID words (% correct)	Spearman's rho	-0.18	0.31	-0.08	0.52
		<i>p</i>	0.25	0.12	0.39	0.02
	Harvard sentences (% words)	Spearman's rho	0.11	0.27	-0.23	0.30
		<i>p</i>	0.34	0.16	0.20	0.13
	PRESTO sentences (% key words)	Spearman's rho	-0.07	0.40	-0.31	0.26
		<i>p</i>	0.40	0.06	0.12	0.17
Intermediate-SMRT group	CID words (% correct)	Spearman's rho	0.19	0.09	0.48	0.33
		<i>p</i>	0.23	0.36	0.03	0.10
	Harvard sentences (% words)	Spearman's rho	0.44	-0.35	0.35	0.44
		<i>p</i>	0.05	0.09	0.09	0.05
	PRESTO sentences (% key words)	Spearman's rho	0.49	-0.40	0.30	0.35
		<i>p</i>	0.03	0.07	0.13	0.09
High-SMRT group	CID words (% correct)	Spearman's rho	-0.01	-0.49	0.42	0.42
		<i>p</i>	0.49	0.02	0.05	0.05
	Harvard sentences (% words)	Spearman's rho	-0.05	-0.47	0.48	0.22
		<i>p</i>	0.42	0.03	0.02	0.19
	PRESTO sentences (% key words)	Spearman's rho	-0.03	-0.35	0.37	0.52
		<i>p</i>	0.46	0.08	0.07	0.01

CID indicates Central Institute of the Deaf; SMRT, Spectral-Temporally Modulated Ripple Test.

to determine that both spectro-temporal resolution (using the SMRT) and nonverbal reasoning (using the Raven's Progressive Matrices) contributed to differentiating high- vs low-performers of sentence recognition in 21 adult CI users (of note, that sample overlapped with the current sample) (10). Moreover, that study suggested only a limited role of cognitive functioning in those CI users who had relatively poor bottom-up input. A similar conclusion was made by Bhargava et al. who examined top-down processing using phonemic restoration in CI users, and found that CI users could take advantage of top-down processing only when provided with better bottom-up input (39).

The current study sought to investigate whether variability in spectro-temporal processing experienced across adult CI users would lead to differences in contributions of top-down cognitive-linguistic skills to individual variability in outcomes. Findings were generally supportive of our main hypothesis. For the low-SMRT group, no significant associations were found. For the intermediate-SMRT group, the composite cognitive-linguistic score was associated with better recognition performance on both sentence materials. Interestingly, for the high-SMRT group, the composite cognitive-linguistic score was associated with better recognition performance on both sentence types but also on isolated word recognition. This was somewhat surprising, as we assumed that isolated word recognition would predominantly rely on bottom-up processing. Our findings, however, suggest that top-down processes contribute significantly to both sentence and word recognition, at least when the bottom-up input is of relatively high quality. Overall, our findings suggest that listeners with at least "intermediate" bottom-up input can capitalize on top-down functions for speech recognition.

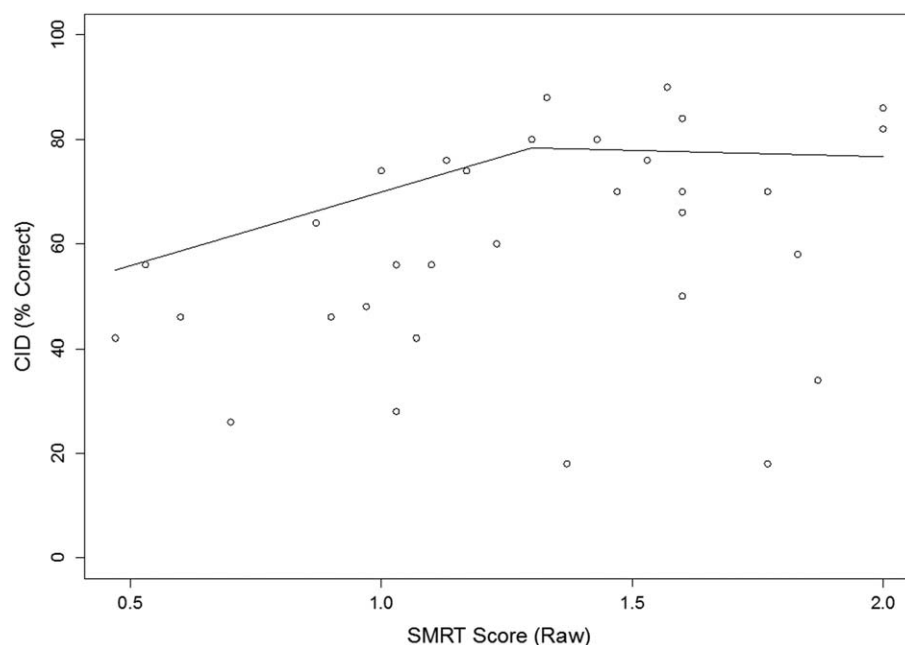
A question that follows, and one that will require more data to answer, is whether there might be a "threshold" of bottom-up processing ability above which top-down functions are able to contribute to speech recognition and, conversely, below which cognitive-linguistic skills cannot compensate. If this were the case, it would have ramifications for patient care. For example, for individuals receiving a signal better than the proposed "threshold", auditory training approaches that encourage the listener to capitalize on top-down processes (e.g., combined "auditory-cognitive" training, or auditory training with passages that vary in use of semantic context) could be effective (50,51). In contrast, for individuals with bottom-up abilities that are worse than threshold, it might be more important to seek device re-programming or use of accessories that improve the sensory input prior to implementing training approaches.

To explore this threshold concept as a supplemental analysis, we performed a segmented linear regression with two segments, carried out on the relationship between CID word recognition and SMRT score across the low-SMRT and intermediate-SMRT groups ( $N = 33$ ). This analysis served to find a breakpoint of the two segments beyond which the relation between CID word recognition (chosen because it relies heavily on access to

acoustic-phonetic information) and SMRT score changes. The high-SMRT individuals were excluded, because SMRT scores  $> 2.1$  RPO are not reliably related to spectro-temporal resolution in CI users, and using the reassigned SMRT score of 2.1 RPO for all high-SMRT individuals would not be very meaningful. For completeness, the segmented regression analysis was repeated using data from all 51 participants, and the breakpoint remained at 1.3 RPO. This analysis suggested a breakpoint of SMRT score around 1.3 RPO where the relationship between SMRT score and CID word recognition changed from a strong positive correlation (below 1.3 RPO) to no longer positive (above 1.3 RPO), as displayed in Figure 2. This analysis provides evidence for a stronger relation between spectro-temporal processing and CID word recognition below 1.3 RPO, with potentially top-down factors contributing more strongly above 1.3 RPO. This breakpoint threshold should be interpreted with caution since the source of this differing relation is unclear, and confirmation will require a larger study.

There are several limitations of the current study. First, the study sample presented here overlaps with smaller samples in our previous reports (10,17,27), and findings may not generalize to other CI samples. Second, although our current sample is of relatively large size in the adult CI literature, breaking the group into three subgroups based on SMRT scores resulted in relatively small subgroups ( $N < 20$  for each). To account for these small subgroups, we correlated speech recognition outcomes with a single composite measure of cognition, which was relatively strongly weighted towards performance on the Stroop task. Ideally, we would be able to examine the relationships among each independent cognitive-linguistic measure and speech recognition for each subgroup; to do so reasonably, a larger sample size will be required. Similarly, in this small sample, we could not account for all the factors that might contribute to group differences and outcome variability; for example, duration of deafness and reading scores differed among the three groups, and the impacts of these factors deserves additional exploration. Additionally, there may be cognitive effects that contribute to performance on the SMRT test, suggested to some degree by our finding that the high-SMRT group generally (but not necessarily significantly) performed better than the other two subgroups on most cognitive-linguistic measures. Kirby and colleagues also found that better general cognitive abilities were associated with better SMRT thresholds in children with hearing aids (52). This potential impact of cognitive functioning on SMRT scores may be even more relevant when using this test in older adults, since there is a clear link between advancing age and declines in cognitive processing (53). Finally, it should be noted that hearing loss, itself, may impact top-down functions as well as top-down/bottom-up interactions, which has been demonstrated clearly as a result of congenital deafness (54,55). Although this study was not designed to directly investigate those effects of hearing loss, future studies are warranted to determine those effects in individuals with post-lingual hearing loss.





**FIG. 2.** Scatterplot of CID word recognition versus SMRT performance for the low- and intermediate-SMRT groups ( $N = 33$ ). Lines show results of the segmented linear regression with two segments ( $N = 33$ ). A breakpoint occurs at SMRT score around 1.3 ripples per octave (RPO). CID indicates Central Institute of the Deaf; RPO, ripples per octave; SMRT, Spectral-Temporally Modulated Ripple Test.

## CONCLUSIONS

Top-down cognitive-linguistic functions contribute differentially to speech recognition outcomes among adult CI users depending on the quality of the bottom-up input. With poor spectro-temporal processing abilities, CI users cannot capitalize effectively on top-down processing. In contrast, at intermediate and higher fidelities of spectro-temporal processing, listeners are able to take advantage of cognitive-linguistic processes to assist them in recognizing speech. This evidence that the complex top-down/bottom-up interactions in speech recognition for CI users have ramifications for individualized approaches to auditory rehabilitation in this clinical population.

## REFERENCES

- Blamey P, Artieres F, Başkent D, et al. Factors affecting auditory performance of postlingually deaf adults using cochlear implants: an update with 2251 patients. *Audiol Neurootol* 2013;18:36–47.
- Hoppe U, Hocke T, Hast A, Iro H. Cochlear implantation in candidates with moderate-to-severe hearing loss and poor speech perception. *Laryngoscope* 2021;131:E940–5.
- Henry BA, Turner CW, Behrens A. Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired and cochlear implant listeners. *J Acoust Soc Am* 2005;118:1111–21.
- Zeng FG. Trends in cochlear implants. *Trends Amplif* 2004;8:1–34.
- Tuenerhoff J, Noppeney U. When sentences live up to your expectations. *NeuroImage* 2016;124:641–53.
- Luce PA, Pisoni DB. Recognizing spoken words: the neighborhood activation model. *Ear Hear* 1998;19:1–36.
- Norris D, McQueen JM, Cutler A. Prediction, Bayesian inference and feedback in speech recognition. *Lang Cogn Neurosci* 2016;31:4–18.
- Rönnerberg J, Lunner T, Zekveld A, et al. The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front Syst Neurosci* 2013;7:31.
- Kral A, Kronenberger WG, Pisoni DB, O'Donoghue GM. Neurocognitive factors in sensory restoration of early deafness: a connectome model. *Lancet Neurol* 2016;15:610–21.
- Tamati TN, Ray C, Vasil KJ, Pisoni DB, Moberly AC. High- and low-performing adult cochlear implant users on high-variability sentence recognition: differences in auditory spectral resolution and neurocognitive functioning. *J Am Acad Audiol* 2020;31:324–35.
- Henry BA, Turner CW. The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *J Acoust Soc Am* 2003;113:2861–73.
- Won JH, Drennan WR, Rubinstein JT. Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *J Assoc Res Otolaryngol* 2007;8:384–92.
- Drennan WR, Anderson ES, Won JH, Rubinstein JT. Validation of a clinical assessment of spectral ripple resolution for cochlear-implant users. *Ear Hear* 2014;35:e92.
- Davies-Venn E, Nelson P, Souza P. Comparing auditory filter bandwidths, spectral ripple modulation detection, spectral ripple discrimination, and speech recognition: normal and impaired hearing. *J Acoust Soc Am* 2015;138:492–503.
- Aronoff JM, Landsberger DM. The development of a modified spectral ripple test. *J Acoust Soc Am* 2013;134:EL217–22.
- Lawler M, Yu J, Aronoff J. Comparison of the spectral-temporally modulated ripple test with the Arizona Biomedical Institute Sentence Test in cochlear implant users. *Ear Hear* 2017;38:760.
- Moberly AC, Vasil KJ, Wucinich TL, et al. How does aging affect recognition of spectrally degraded speech? *Laryngoscope* 2018;128 (Suppl 5). doi: 10.1002/lary.27457.
- O'Brien E, Winn MB. Aliasing of spectral ripples through CI processors: a challenge to the interpretation of correlation with speech recognition scores. InPoster session presented at the conference on implantable auditory prostheses, Tahoe, CA; 2017.
- DiNino M, Arenberg JG. Age-related performance on vowel identification and the spectral-temporally modulated ripple test in children with normal hearing and with cochlear implants. *Trends Hear* 2018;22. doi: 10.1177/2331216518770959.

20. Resnick JM, Horn DL, Noble AR, Rubinstein JT. Spectral aliasing in an acoustic spectral ripple discrimination task. *J Acoust Soc Am* 2020;147:1054–8.
21. Baddeley A. Working memory. *Science* 1992;255:556–9.
22. Daneman M, Carpenter PA. Individual differences in working memory and reading. *J Mem Lang* 1980;19:450.
23. Tao D, Deng R, Jiang Y, Galvin JJ III, Fu QJ, Chen B. Contribution of auditory working memory to speech understanding in mandarin-speaking cochlear implant users. *PLoS One* 2014;9:e99096.
24. Moberly AC, Pisoni DB, Harris MS. Visual working memory span in adults with cochlear implants: some preliminary findings. *World J Otorhinolaryngol Head Neck Surg* 2017;3:224–30.
25. O'Neill ER, Krefit HA, Oxenham AJ. Cognitive factors contribute to speech perception in cochlear-implant users and age-matched normal-hearing listeners under vocoded conditions. *J Acoust Soc Am* 2019;146:195–210.
26. Moberly AC, Houston DM, Castellanos I. Non-auditory neuro-cognitive skills contribute to speech recognition in adults with cochlear implants. *Laryngosc Invest Otolaryngol* 2016;1:154–62.
27. Moberly AC, Reed J. Making sense of sentences: top-down processing of speech by adult cochlear implant users. *Journal of Speech. Lang Hear Res* 2019;62:2895–905.
28. Marslen-Wilson W. Issues of process and representation in lexical access. In *Cognitive models of speech processing: The second Sperlonga meeting* 1993, pp. 187–210.
29. McClelland JL, Elman JL. The TRACE model of speech perception. *Cogn Psychol* 1986;18:1–86.
30. Tamati TN, Vasil KJ, Kronenberger W, Pisoni DB, Moberly AC, Ray C. Word and nonword reading efficiency in post-lingually deafened adult cochlear implant users. *Otol Neurotol* 2021;42:e272–8.
31. Holden LK, Finley CC, Firszt JB, et al. Factors affecting open-set word recognition in adults with cochlear implants. *Ear Hear* 2013;34:342.
32. Mattingly JK, Castellanos I, Moberly AC. Nonverbal reasoning as a contributor to sentence recognition outcomes in adults with cochlear implants. *Otol Neurotol* 2018;39:e956.
33. Pichora-Fuller MK, Kramer SE, Eckert MA, et al. Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). *Ear Hear* 2016;37:5S–27S.
34. Wild CJ, Yusuf A, Wilson DE, Peelle JE, Davis MH, Johnsrude IS. Effortful listening: the processing of degraded speech depends critically on attention. *J Neurosci* 2012;32:14010–21.
35. Sohoglu E, Davis MH. Perceptual learning of degraded speech by minimizing prediction error. *Proc Natl Acad Sci* 2016;113:E1747–56.
36. Mattys SL, Brooks J, Cooke M. Recognizing speech under a processing load: dissociating energetic from informational factors. *Cogn Psychol* 2009;59:203–43.
37. Clopper CG. Effects of dialect variation on the semantic predictability benefit. *Lang Cogn Processes* 2012;27:1002–20.
38. Chatterjee M, Peredo F, Nelson D, Başkent D. Recognition of interrupted sentences under conditions of spectral degradation. *J Acoust Soc Am* 2010;127:EL37–41.
39. Bhargava P, Gaudrain E, Başkent D. Top-down restoration of speech in cochlear-implant users. *Hear Res* 2014;309:113–23.
40. Robertson GJ. Wide-Range Achievement Test. *Corsini Encyclopedia Psychol* 2010;1–2.
41. Folstein MF, Folstein SE, McHugh PR. Mini-mental state. A practical method for grading the cognitive state of patients. *J Psychiatric Res* 1975;86:189–98.
42. Nitttrouer S, Burton LT. The role of early language experience in the development of speech perception and phonological processing abilities: evidence from 5-year-olds with histories of otitis media with effusion and low socioeconomic status. *J Commun Disord* 2005;38:29–63.
43. Torgesen JK, Rashotte CA, Wagner RK. *TOWRE: Test of word reading efficiency*. Toronto, Ontario: Psychological Corporation; 1999.
44. Hirsh IJ, Davis H, Silverman SR, Reynolds EG, Eldert E, Benson RW. Development of materials for speech audiometry. *J Speech Hear Disord* 1952;17:321–37.
45. IEEE. *IEEE Recommended Practice for Speech Quality Measurements*. New York: Institute for Electrical and Electronic Engineers; 1969.
46. Gilbert JL, Tamati TN, Pisoni DB. Development, reliability, and validity of PRESTO: a new high-variability sentence recognition test. *J Am Acad Audiol* 2013;24:26–36.
47. Humes LE, Kidd GR, Lentz JJ. Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Front Syst Neurosci* 2013;7:55.
48. Pichora-Fuller MK, Schneider BA, Daneman M. How young and old adults listen to and remember speech in noise. *J Acoust Soc Am* 1995;97:593–608.
49. Janse E, Ernestus M. The roles of bottom-up and top-down information in the recognition of reduced speech: evidence from listeners with normal and impaired hearing. *J Phonetics* 2011;39:330–43.
50. Ingvalson EM, Young NM, Wong PC. Auditory-cognitive training improves language performance in prelingually deafened cochlear implant recipients. *Int J Pediatric Otorhinolaryngol* 2014;78:1624–31.
51. Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N. Reversal of age-related neural timing delays with training. *Proc Natl Acad Sci* 2013;110:4357–62.
52. Kirby BJ, Spratford M, Klein KE, McCreery RW. Cognitive abilities contribute to spectro-temporal discrimination in children who are hard of hearing. *Ear Hear* 2019;40:645.
53. Salthouse TA. Selective review of cognitive aging. *J Int Neuro-psychol Soc* 2010;16:754.
54. Yusuf PA, Hubka P, Tillein J, Kral A. Induced cortical responses require developmental sensory experience. *Brain* 2017;140:3153–65.
55. Yusuf PA, Hubka P, Tillein J, Vinck M, Kral A. Deafness weakens interareal couplings in the auditory cortex. *Front Neurosci* 2021;14:1476.