Neurophysiological markers of central gain and their relationship to speech-in-noise intelligibility

Jacie R. McHaney[1], Kenneth E. Hancock[2,3], Daniel B. Polley[2,3], Aravindakshan Parthasarathy[1,4]

[1]Department of Communication Science and Disorders, University of Pittsburgh, [2]Department of Otolaryngology- Head and Neck Surgery, Harvard Medical School
[3]Eaton-Peabody Laboratories, Massachusetts Eye and Ear, [4]Department of Otolaryngology, University of Pittsburgh

Introduction

Optimal speech-in-noise (SpIN) intelligibility requires separation of a target speaker stream from multiple competing background streams[1,2]. We have previously shown that the ability to segregate these competing streams of speech depends on the fidelity of bottom-up neural representation of sensory information as well as top-down influences of effortful listening[3]. Here, we extend these results using a behaviorally relevant speech task (QuickSIN)[4] while exploring the contributions of putative central gain to bottom-up sensory coding fidelity.

Experimental methods

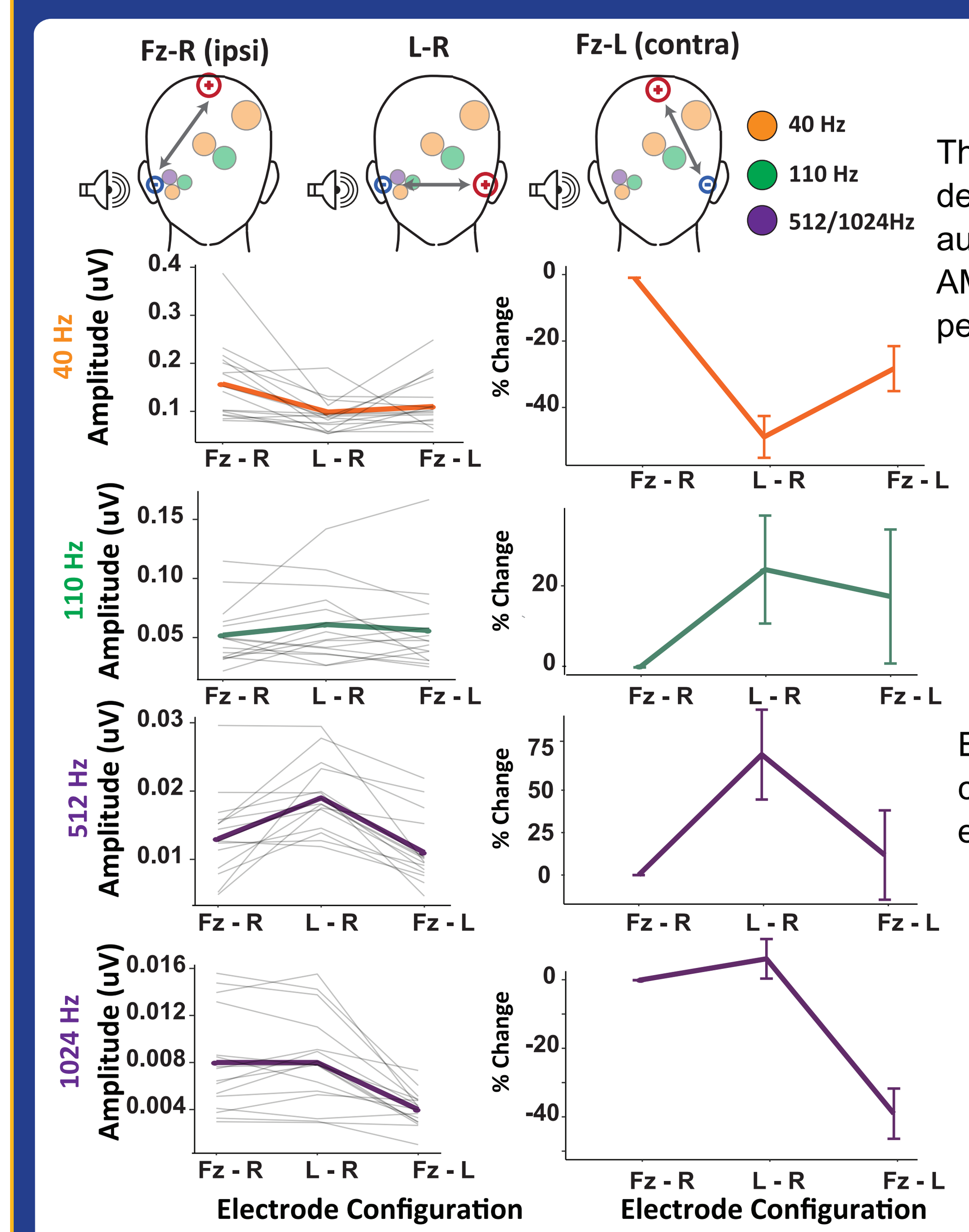
Subjects:
19 younger adults (Mean age = 28.70, **SD** = 4.15 years) particpated in the study. Normal hearing sensitivity <25 dB at octave spaced frequencies between 250-8000 Hz. All procedures were approved by the institutional review board at the Massachusetts Eye and Ear (Protocol #1006581) and Mass-General Brigham (Protocol #2019P002423).

Electrophysiology:
Envelope Following Responses (EFRs) to amplitude modulated (AM) tones. Stimuli were 200 ms long, presented with a 3.1/s repetition rate at 85dB SPL. Carrier frequency was 3000Hz amplitude modulated (AM) at 40, 110, 512, and 1024 Hz. Recordings were done on a 16-channel EEG system with two foil tiptrodes positioned in the ear canals and a gold-cup electrodes placed at Fz, all referenced to ground at the nape.
Montage configurations for analysis: Fz-Right tiptrode (ipsilateral side), Left tiprode-Right tiptrode, Fz-Left tiptrode (Contralateral side). Processing: Fourth-order Butterworth filter with lowpass filter of 3000 Hz. Highpass filters of 5, 80, 200, and 300 Hz used for 40, 110, 512, and 1024 Hz AM tones, respectively. Fast Fourier transforms (FFTs) performed on averaged time domain waveforms. Maximum amplitude of FFT peak at the AM frequency rate is reported as the EFR amplitude.

SpIN Intelligibility:
QuickSIN: Six sentences masked in four-talker babble at decreasing signal-to-noise ratio (SNR) levels: 25, 20, 15, 10, 5, and 0 dB. Participants were instructed to repeat target sentence--each sentence contained 5 keywords for identification. Six lists were completed by each participant (2 practice, 4 test). Proportion of keywords correctly identified at each SNR level across the 4 test lists were averaged as a measure of SpIN intelligibillity.

Pupillometry:
Pupillary responses were sampled at 30Hz during QuickSIN. Pupillary data were processed to remove blinks. Any trial with a blink > 600 ms were removed. All other blinks were linearly interpolated from 120 ms before to 120 ms after the detected blink. Pupillary responses were averaged across all 4 test lists at each SNR. Pupillary responses at SNR 25 were not analyzed due to different physiological response because SNR 25 occurred at the beginning of each list. All other pupillary measures serve as an index of listening effort.

Changes in amplitude modulation rates emphasize varying EFR generators

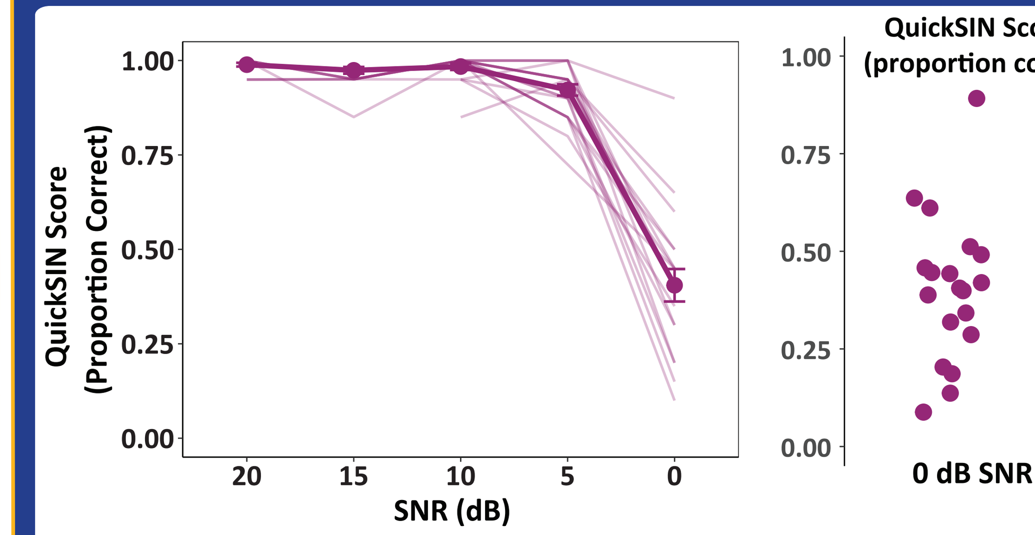


The upper limit of phase-locking decreases along the ascending auditory pathway. EFRs to faster AM rates emphasize more peripheral generators.

Evidence for this is seen when comparing EFRs from different electrode configurations.

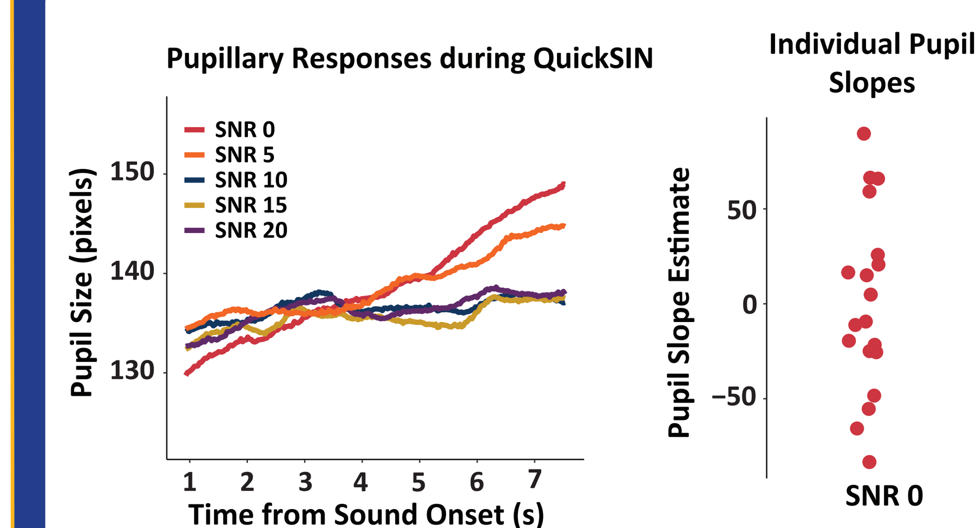
Hence, measuring EFRs at multiple AM rates from 40-1024Hz give us a measure of information transfer along the ascending auditory pathway. The Fz-R montage was used for further analysis.

"Normal" hearing listeners exhibited significant variability in SpIN performance



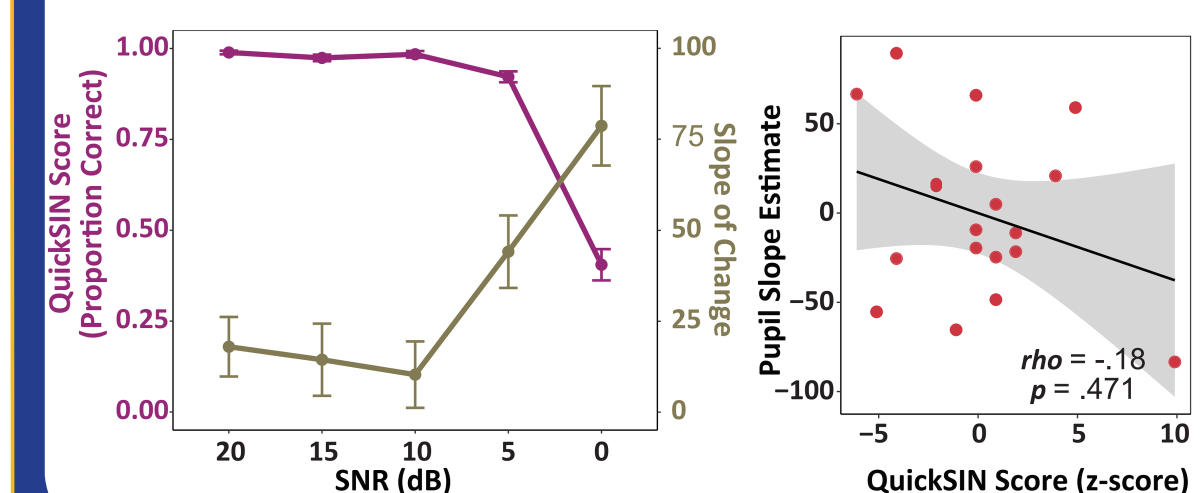
SpIN intelligibility measured using QuickSIN were near-ceiling for easier SNR levels (20, 15, 10, and 5 dB) in listeners with normal hearing thresholds. However, performance dropped for the most difficult SNR of 0 dB, with significant individual differences

Decrease in SpIN performance accompanied by increase in listening effort



Pupillary responses from the time window of 930-7530 ms after onset of the speech stimulus were analyzed using growth curve analysis (GCA)[5]. GCA uses orthogonal polynomial time terms to capture functional forms of the pupil response. A GCA was fit using a first-order orthogonal polynomial (ot1) to capture the slope of the pupillary response for each SNR, representing listening effort:

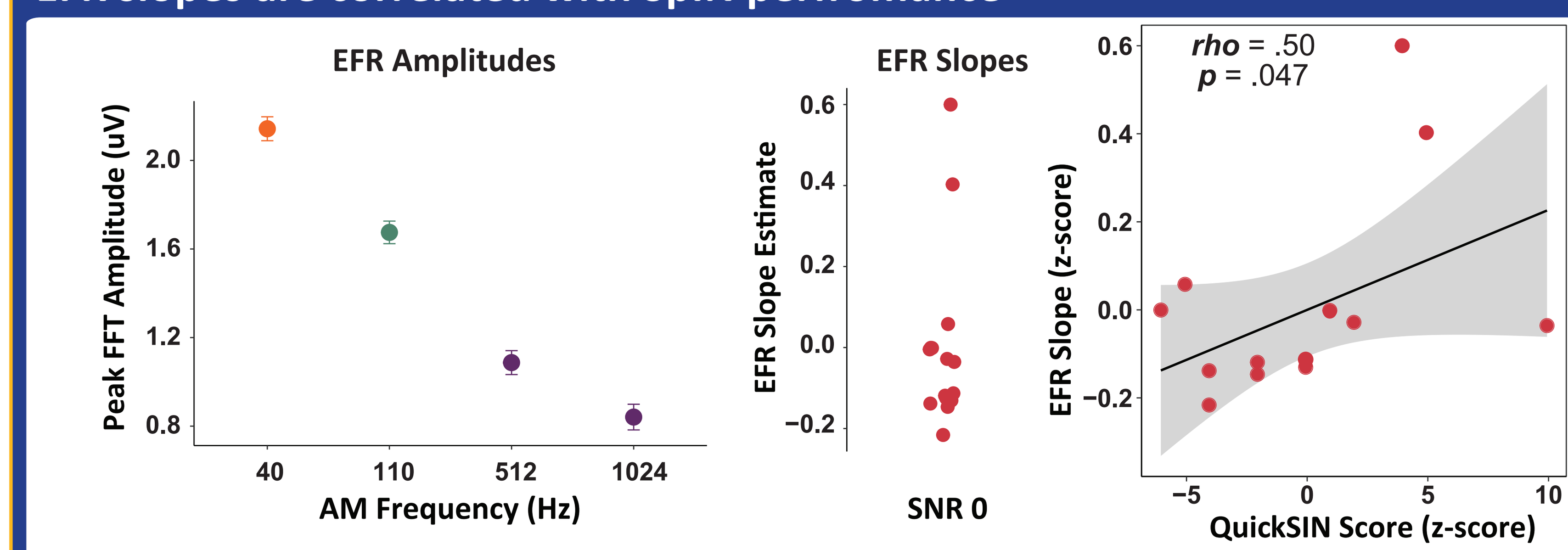
Pupil ~ ot1*SNR + (ot1|Subject:SNR)

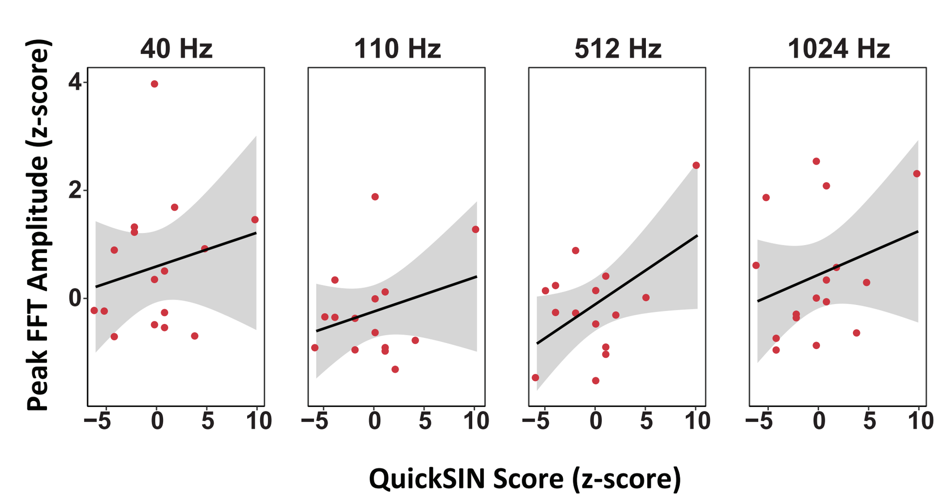


Averaged pupillary slopes demonstrated a change in listening effort associated with SpIN performance. While SpIN peformance remained near ceiling for 5 dB SNR, listening effort increased.

Individual slope estimates did not show significant correlations with SpIN intelligibility scores at 0 dB SNR.

EFR slopes are correlated with SpIN perfromance





A similar GCA was fit for Peak FFT Amplitudes across AM rates for each Montage using a first-order orthogonal polynomial (ot1):
PeakFFTAmp ~ ot1*Montage + (ot1|Subject:Montage)

EFR slopes at Fz-R montage were correlated with individual SpIN performance at 0 dB SNR, though indivudal EFR amplitudes were not.

This suggests the EFR slope is a better metric for bottom up sensory processing as it includes auditory processing at multiple stations, and likely minimizes variability due to experiemental conditions.

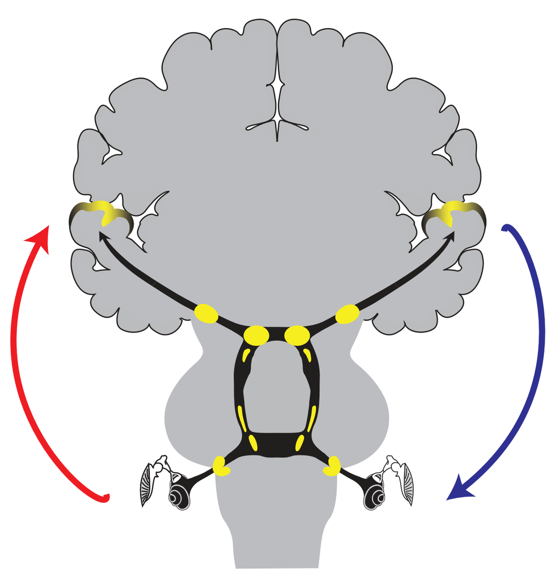
Optimal model for SpIN intelligibility includes inetractions between bottom-up and top-down measures

Model	Adjusted-R^2
SpIN ~ EFR	.133
SpIN ~ Pupil	.203
SpIN ~ EFR + Pupil	.371

+.070
+.168

A linear regression model of speech intelligibility with both EFR slope and Pupil slope as predictors explained the greatest variability in SpIN performance compared to models with just EFR or Pupil alone.

Best fit model: SpIN ~ EFR + Pupil
$F(2,12) = 5.126$, $p = .025$

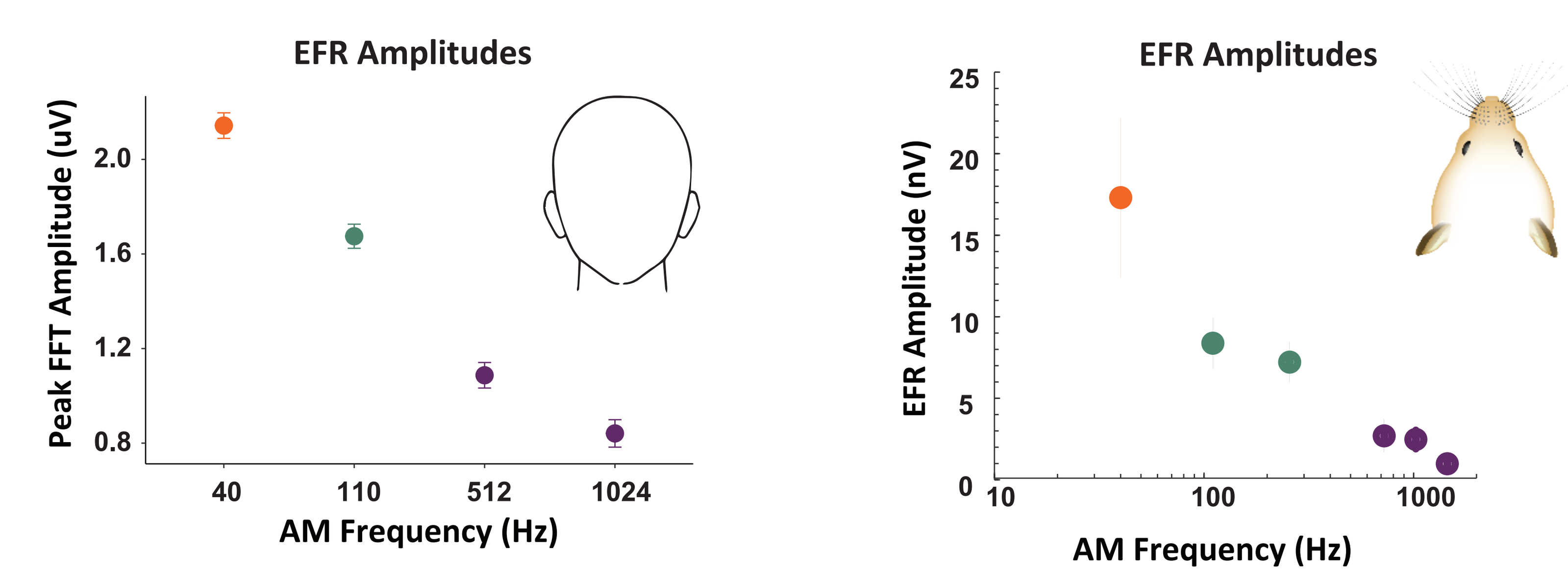


This model indicates that interplay between bottom-up (EFRs) and top-down (listening effort) correlates contribute to SpIN intelligibility in younger adults with normal hearing thresholds.

Summary

- There is significant individual variability in SpIN performance in listeners with normal audiograms
- EFRs to increasing modulation frequencies can be used to emphasize central vs. peripheral generators.
- Decrease in SpIN performance is associated with increase in listening effort, with some subliminal changes that are not reflected in performance.
- EFR slopes are correlated with SpIN performance, suggesting a role for the transformation of auditory information along the ascending auditory pathway.
- EFR slopes may be a better metric for bottom-up processing and/or central gain compared to individual EFR amplitudes.
- Optimal model for SpIN intelligibility includes bottom-up measures of sensory coding and top-down measures of listening effort.

Future Directions



Ongoing work is related to replicating these results in an animal model of low-frequency hearing, the Mongolian gerbil. Changes in EFR slopes can be measured to identical stimuli in the gerbil, and can be compared with responses to carefully induced pathological conditions such as noise-induced cochlear synaptopathy.

Acknowledgements and References

This work was supported by NIDCD-P50DC015817 (DP) and NIDCD-T32DC011499 (trainee: JRM.). The authors thank Jessica L'Heureux and Jennifer Klara for administrative and technical support.

1. Bregman, A. S. (1990). Auditory Scene Analysis: The Perceptual Organization ofSound. Cambridge, MA: MIT Press. 2. Shinn-Cunningham, B. (2017). Cortical and Sensory Causes of Individual Differences in Selective Attention Ability Among Listeners With Normal Hearing Thresholds. Journal of Speech, Language, and Hearing Research, 60, 2976–2988. 3. Parthasarathy, A., Hancock, K. E., Bennett, K., Degruttola, V., & Polley, D. B. (2020). Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing. ELife, 9, e51419. 4. 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. The Journal of the Acoustical Society of America, 116, 2395. 5. Mirman, D. (2014). Growth curve analysis and visualization using R. CRC Press.