Single-sided deafness cochlear-implant perception and simulation: localization and spatial-masking release

Coral Dirks[1] , Andrew J. Oxenham[2], Peggy B. Nelson[1]
[1]Department of Speech-Language Hearing Sciences [2]Department of Psychology, University of Minnesota

Background

Single-sided deafness (SSD) refers to a condition in which people have normal or near-normal hearing in one ear and little or no residual hearing in the opposite ear. With few exceptions, most SSD patients experience substantial deficits in localizing sounds in the horizontal plane; their errors are generally greatest when sounds originate on the poorer ear side, and sounds presented on the poorer ear side tend to be localized to the better ear side (Van Wanrooij & Van Opstal, 2004). SSD also affects a person's ability to understand speech in noisy environments in at least three ways: (1) loss of "better ear" when speech is on the poorer ear side, (2) loss of binaural squelch, and (3) reduced ability to use perceived spatial differences between target speech and masker, especially when target speech and masker are similar, leading to "informational masking" (Freyman et al., 1999). Recently some SSD patients have received a cochlear implant (CI) in their poorer ear, primarily to relieve tinnitus. Preliminary research suggests that these patients' localization accuracy improves from 30° to 15° error (Arndt et al., 2011) and that their speech reception thresholds (SRTs) significantly decrease (improve) ~1 dB in certain spatial configurations after implantation (Blasco & Redleaf, 2014). However, little is know about the specific cues SSD patients with CIs use to localize sound and improve speech understanding in spatial environments. Moreover, a recent study suggests that the spatial configurations in which researchers measured SSD/CI patients' SRTs do not to produce optimal performance (Culling, Jelfs, Talbert, Grange, & Backhouse, 2012).

In this pilot study, normal-hearing (NH) listeners completed localization and speech understanding in noise tasks in sound field with and without masking noise in one ear to simulate profound, unilateral sensorineural hearing loss. The NH listeners also completed a speech-in-noise task over headphones in which silence (to simulate SSD) or vocoded stimuli (to simulate a cochlear implant) were presented to one ear. We tested NH listeners to:

- Identify the listening conditions in which we expect to observe the maximum benefit of the implant,
- Obtain baseline measures to serve as 'upper limits' for the benefit we may expect to observe in patients with SSD and CIs, and
- Obtain estimates of the benefit we may expect to observe in patients with SSD and CIs in the speech in noise task.

The results will provide a benchmark for potential improvement when a CI is combined with contralateral normal hearing.

General Methods

University of Minnesota Multi-Sensory Perception Laboratory (MSP):

- Sound-treated booth with capacity to simulate complex, real-world acoustic environments via 48-speaker array
- In this simulation, we used 19 speakers located along the front horizontal plane from -90° to 90° azimuth (adjacent speakers separated by 10°).

Latin Squares Design: For each experiment, we used a Latin squares design to balance conditions across subjects.

Adaptive Speech Understanding in Noise (SPIN) Experiment

Hypothesis: In NH listeners, listening with two ears (binaurally) will improve speech understanding in noise performance when speech and noise are spatially separated. These improvements will extend beyond the better-ear advantage and encompass the benefits of perceived separation associated with informational masking.

Listeners: 6 normal-hearing (NH), native speakers of American English.

Task: While fixating gaze toward 0° azimuth, listeners were presented with a sentence in a masker and verbally repeated the sentence they heard to an examiner in the sound booth.

Methods/Stimuli: 2 masker types X 3 spatial configurations X 2 listening conditions = 12 blocks
TARGET SPEECH: Harvard IEEE sentences spoken by a female talker (lists 61-72); fixed at rms level of 55 dB SPL
MASKERS:

1. Time-reversed, two-talker female babble (TTB), taken from the target sentences
2. Gaussian noise spectrally shaped to the long-term average spectrum of target sentences (SSN)

SPATIAL CONFIGURATIONS:

1. Speech and masker co-occurring from 0° azimuth (S0N0)
2. Speech at 60° azimuth and masker at -60° azimuth (S+60N-60)
3. Speech at -60° azimuth and masker at 60° azimuth (S-60N+60)

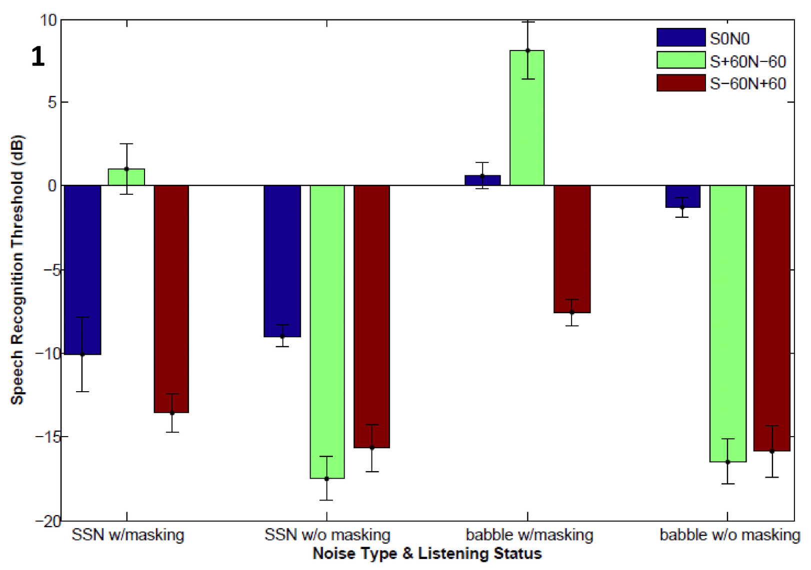
LISTENING CONDITIONS: With and without masking noise presented to the right ear (to simulate SSD)

- Examiner judged whether the entire sentence was repeated correctly
- Noise level increased with each incorrect response and decreased with each correct response
- Listeners performed the task with and without masking noise in their **right** ears to simulate profound unilateral hearing loss

Data Analysis:

- Estimated SRT = average of the last even number of "turning points" in signal-to-masker ratio; "turning points" are the signal-to-masker ratios (SMRs) at which the direction of the noise level changed from louder to softer or softer to louder.

Adaptive SPIN Experiment : Results/Discussion



- When hearing on the simulated deaf side is restored, we observe improvements in speech-to-masker ratio between 18 and 25 dB when we present target speech and a masker on opposite sides **(Fig. 1)**.
- The 7-dB increase in the effect between SSN and TTB reflects the release from informational masking in the babble condition, based on perceived location, over and beyond the effects of classical binaural interactions (Freyman et al., 1999).

Figure 7: Speech reception thresholds measured using an adaptive procedure for six NH listeners whose right ears were masked in binaural and monaural listening conditions

Localization Experiment

Hypothesis: NH listeners will integrate ILD and ITD cues in the stimulus fine structure and envelope across ears will approximately equal accuracy. When masking noise is presented to one ear (to simulate SSD), errors will increase and responses will be biased toward the unmasked ear.

Listeners: 2 NH, native English speakers

Task: While fixating gaze toward 0° azimuth, listeners listened to one of six types of stimuli randomly presented from one of the 19 speakers in the front horizontal plane. After each stimulus presentation, listeners indicated on a computer screen which speaker they heard the sound coming from.

Methods/Stimuli

- Stimuli was presented at nominal overall rms level of 60 dB SPL and roved by ±10 dB around the nominal level.
- Each stimulus type was played 5 times from each of 19 speakers in the front horizontal plane.
- Listeners completed this task with and without masking noise presented one ear to simulate SSD.
- Tonal stimuli: total duration = 300 ms; 100 ms raised-cosine onset and offset ramps
- TEN noise added to high-frequency complex tone conditions; presented from all speakers in the front horizontal plane at a level per equivalent rectangular bandwidth (ERB) sufficient to mask distortion

Localization Experiment Cont'd

Stimuli	Type of Cues Tested
300-Hz pure tone	Low-frequency ITDs in stimulus fine structure
Unmodulated high-frequency complex tone comprised of pure tones >1500 Hz	High-frequency ILDs
Same high-frequency complex tone modulated by 80-Hz sinusoid with 100% modulation depth	High-frequency ILDs and ITDs in the stimulus envelope
Unfiltered NU-6 words	Combination of ITDs and ILDs in stimulus envelope and fine structure
NU-6 lowpass-filtered at 640 Hz	Low-frequency ITDs in stimulus fine structure
Nu-6 highpass-filtered at 1500 Hz	High-frequency ILDs and ITDs in stimulus envelope

Table 1. Stimulus type and cues tested in the localization experiment.

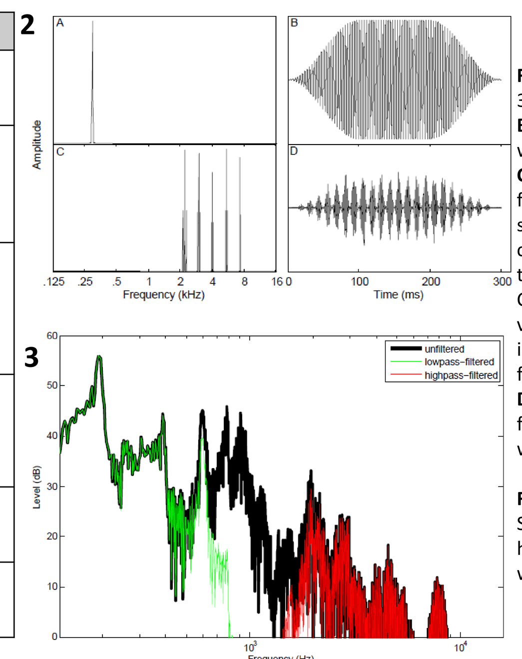


Figure 2 (upper panel). A. 300 Hz pure tone spectrum. **B.** 300 H pure tone waveform. **C.** Modulated high frequency complex tone spectrum. The frequency components correspond to the CFs of standard MED-ED CI map. The sidebands visible here are not present in the unmodulated high-frequency complex tone. **D.** Modulated high-frequency complex tone waveform.

Figure 3 (lower panel). Spectra of unfiltered, highpass, and lowpass NU-6 words.

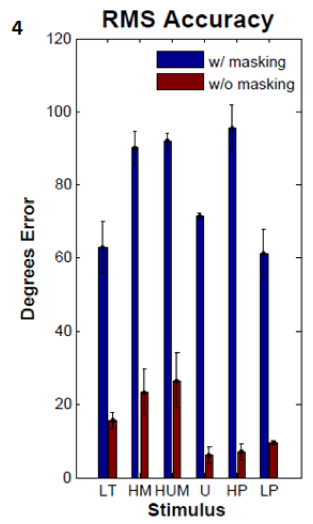
Data Analysis:

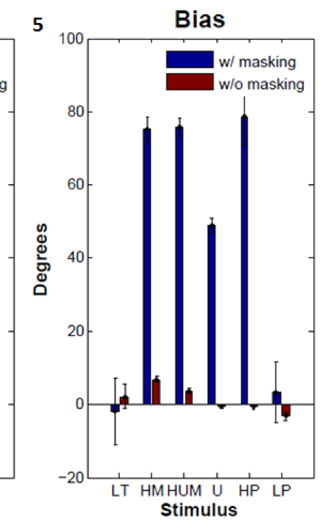
Root Means Square (RMS) Error: the square root of the mean squared difference in degrees between sound source location and perceived sound location; combined estimate of systematic bias and response variability

Systematic Localization Bias: the mean (signed) difference between the actual and perceived angles for each speaker; separates the variability from the systematic response bias to determine whether the responses are systematically biased

Standard Deviation at Mean Response Location: estimated standard deviation of the responses around the estimated mean response location; partials out systematic response bias to determine variability around the perceived location

Localization Experiment: Results/Discussion





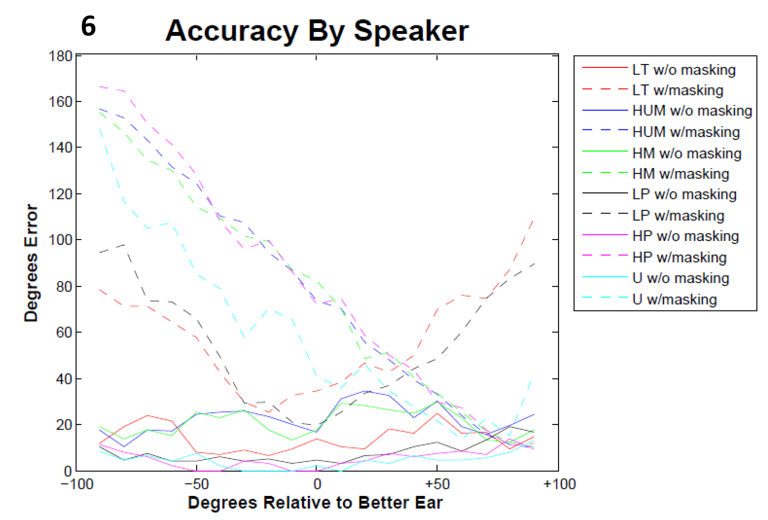


Figure 4: RMS Accuracy with standard error bars for two NH listeners, where LT represents low tone, HM represent modulated high-frequency complex tone, HUM represents unmodulated high-frequency complex tone, U represents unfiltered words, HP represents high-pass words, and LP represents low-pass words.
Figure 5: Response bias with standard error bars for the same two NH listeners, where positive bias represents error in the direction of the better ear and negative bias represents error in the direction of the poorer ear.
Figure 6: RMS Accuracy by speaker location where the lines represent the six different types of stimuli. Blue lines represent average performance across listeners in masked conditions and red lines represent average performance across listeners in conditions unmasked conditions. Positive degrees represent degrees on the better ear side and negative degrees represent degrees on the masked ear side.

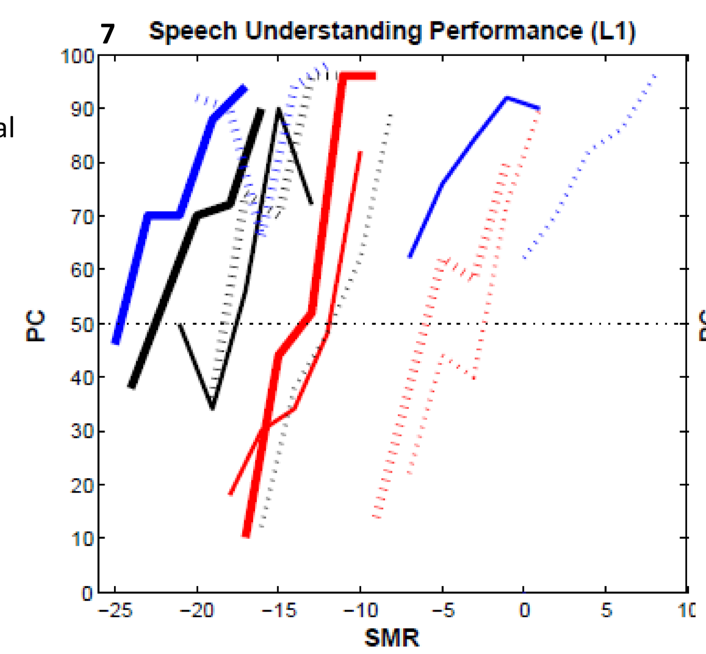
- Monaural errors are 3-10 times larger than binaural localization errors **(Fig. 4)**.
- Errors around the perceived sound location appear to be more variable when one ear is masked **(Figs. 4 & 5.)**
- Monaural errors and bias are larger in conditions where listeners had access to ILDs only, ILDs + ITDs in the stimulus envelope, and all ITD/ILD information **(Figs. 4 & 5)**.
- Errors are biased toward the better ear in conditions where listeners had access to ILDs only, ILDs + ITDs in the stimulus envelope, and all ITD/ILD information. **(Fig. 5.)**
- When one ear was masked, listeners tended to perceive sounds as coming from 0° azimuth in the low-pass/low-tone conditions, producing a "v-shaped" pattern of error, or as coming from the speaker at 90° azimuth on the better ear side, producing a downward sloping pattern of error **(Fig. 6.)** This may explain why listeners appear to perform better monaurally in the low-tone/low-pass monaural listening conditions **(Figs. 4 & 5.)**

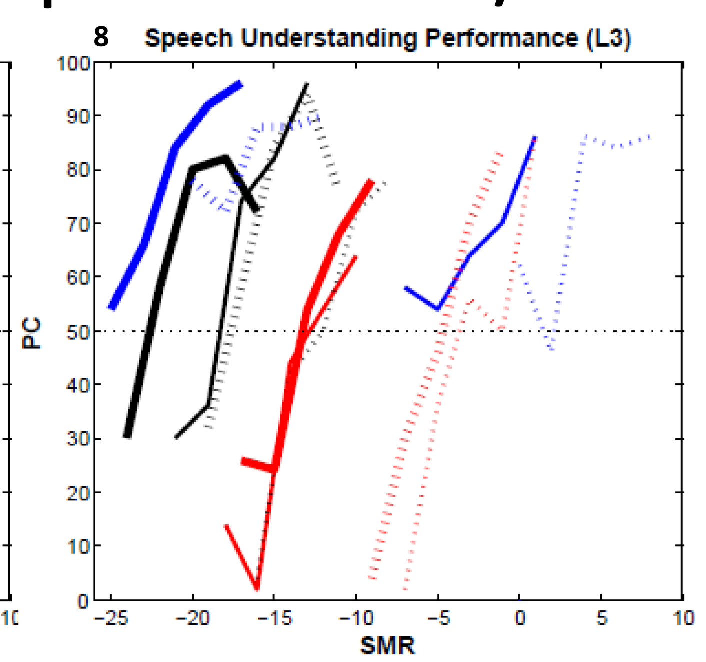
Sound Field SPIN Experiment

Hypothesis: see Adaptive SPIN experiment
Listeners: 2 NH, native English speakers; L1 = control from localization exp. and L3 = control who did not complete localization exp.
Task: While fixating gaze toward 0° azimuth, the listeners listened to a sentence presented in noise then typed what they heard onto a computer screen. When the listeners were satisfied with their responses, they clicked a button to save their responses to a data file.
Methods/Stimuli: 2 noise types X 3 spatial configurations X 2 listening conditions X 5 SMRs = 60 blocks

- Same as the Pilot SPIN Experiment except the noise was fixed at 5 SMRs: SRT, SRT-2, SRT-4, SRT-6, SRT-8, where SRT = average SRT measured in the Pilot Adaptive SPIN Experiment
- **Right** ear masked in each listener during monaural listening conditions to simulate SSD

Sound Field SPIN Experiment: Results/Discussion





Figures 7 & 8: Proportion of key words correctly identified in a list as a function of signal-to-masker ratio in two NH listeners. The right ear was masked in L1 and L3.

LINE WIDTH: thick = without masking; thin = with masking presented to **right** ear
LINE TYPE: solid = SSN, dotted = babble
LINE COLOR: red = S0N0, blue = S+60N-60, black = S-60N+60

Soundfield SPIN Exp: Results/Discussion Cont'd

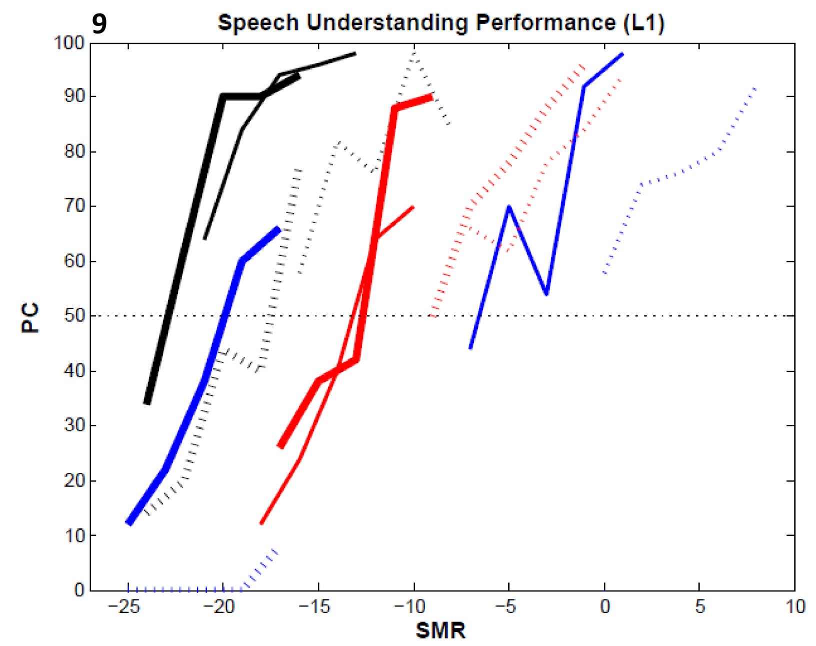
- Listeners performed best when speech and noise were spatially separated, except in the monaural listening conditions when speech was presented on the right.
- For a given noise type, the SRTs corresponding to the binaural, spatially separated listening conditions are not equal, as we had expected. However, those SRTs are smaller (better) than SRTs for the given noise type in the binaural, co-located listening conditions.
- The largest improvements in SRT appear to occur when hearing is "restored" to (masking noise removed from) the right ear in SSN and babble. Grossly, this improvement is larger for babble than for SSN.

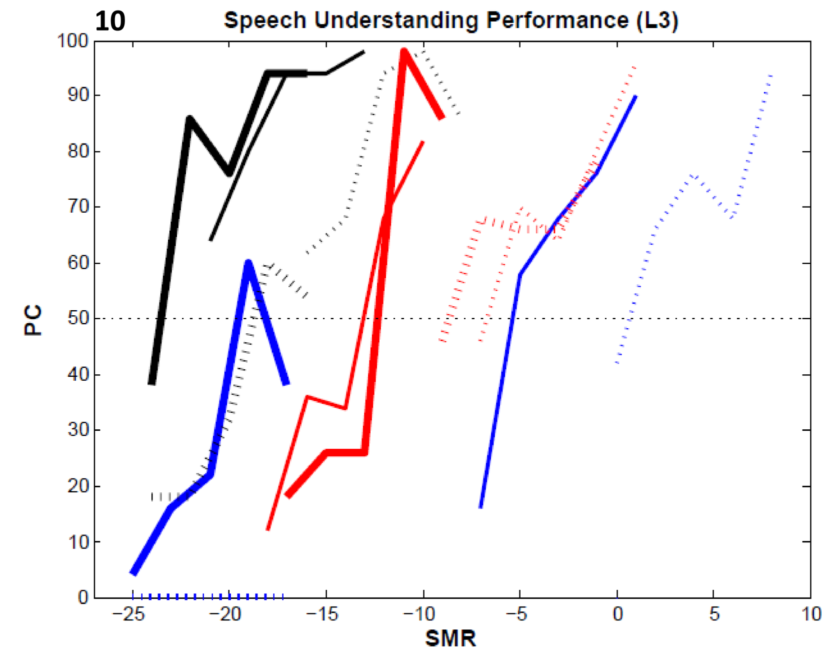
Vocoded SPIN Experiment

Hypothesis: Listening to unprocessed stimuli in one ear and vocoded stimuli in the other ear (to simulate a CI) will improve speech understanding in noise performance, relative to monaural listening, when speech and noise are spatially separated. These improvements will extend beyond the better-ear advantage and encompass the benefits of perceived separation associated with informational masking.
Task: Over headphones, NH listeners listened to sentences presented in noise then typed the sentences they heard onto a computer screen. When the listeners were satisfied with their responses, they clicked a button to save their responses to a data file.
Listeners: 2 NH, native English speakers; L1 and L3 from Sound Field SPIN Experiment
Methods/Stimuli:

- Same as SF SPIN task except that listeners completed this task with no stimuli or tone vocoded stimuli presented to their **right** ears
- Spatial environment simulated using head-related transfer functions from a mannequin head (KEMAR)
- Vocoder channel CFs = CFs of standard MED-EL CI map.

Vocoded SPIN Experiment: Results/Discussion





Figures 9 & 10: Proportion of key words correctly identified in a list as a function of signal-to-masker ratio in two NH listeners. The right ear was masked in L1 and L3.

LINE WIDTH: thin = without stimuli presented to **right** ear; thick = with vocoded stimuli presented to **right** ear
LINE TYPE: solid = SSN, dotted = babble
LINE COLOR: red = S0N0, blue = S+60N-60, black = S-60N+60

- NH listeners tended to perform better when speech and noise were spatially separated than when they co-occured at 0° azimuth, except in the monaural listening conditions when speech was presented on the right. It is unclear whether this is true when speech was presented on the right side, noise was presented on the left side, and vocoded stimuli was sent to the right ear.
- When speech and noise were spatially separated, listeners performed better binaurally when speech was closer to the left, "normal-hearing" ear than the right, vocoded ear.
- The largest improvements in SRT appear to occurred in SSN when speech was presented on the right side and hearing was "restored" via vocoding to the right ears. Grossly, this improvement is slightly smaller than that observed in the SF SPIN Exp.
- It is difficult to tell whether statements in the point above are true in babble and whether the improvement from monaural to binaural listening is larger in SSN or babble.

CONCLUSIONS

- In the localization task, presenting masking noise to one ear in our NH listeners effectively simulated SSD; our listeners' struggled to identify where sounds originated and tended to perceive sounds on in their unmasked, "better" ear.
- With access to hearing in both ears, our NH listeners used ILD and ITD cues in the stimulus envelope and fine structure equally well to localize sound. We expect that a CI can improve localization for patients with SSD via high frequency ILD cues and via ITD cues in the temporal envelope, but not via ITD cues in the low-frequency stimulus fine structure.
- In the speech and noise tasks, it appears that presenting speech and noise at 60 degrees on either side of the head will capture the largest benefit of using an implant in SSD patients; NH listeners demonstrated between 18 and 25 dB improvement in SRT when masking noise was removed from the ear and speech was presented on the masked ear side.
- NH listeners experienced an additional 7 dB release from informational masking beyond 18 dB release from energetic masking when speech originated on the masked, "poorer" ear side. We hypothesize that the benefits of the implant in patients with SSD will extend beyond the better ear effect to encompass the benefits of perceived separation associated with informational masking.
- When we stimulated listening through a CI using vocoded stimuli over headphones, NH listeners appeared to perceive spatial differences in speech and noise and use those differences to selectively attend to speech in noise. However, the benefit of listening through a combination of acoustic and simulated electric hearing were smaller than listening to unprocessed stimuli.
- We plan to extend this tests to patients with SSD and cochlear implants, and compare their results with age-matched NH listeners.

References

Arndt, S., Aschendorff, A., Laszig, R., Beck, R., Schild, C., Kroeger, S., … Wesarg, T. (2011). Comparison of pseudobinaural hearing to real binaural hearing rehabilitation after cochlear implantation in patients with unilateral deafness and tinnitus. *Otology & Neurotology 32*(1), 39–47.

Blasco, M. a, & Redleaf, M. I. (2014). Cochlear implantation in unilateral sudden deafness improves tinnitus and speech comprehension: meta-analysis and systematic review. *Otology & Neurotology*, 35(8), 1426-1432.

Culling, J. F., Jelfs, S., Talbert, A., Grange, J. A, & Backhouse, S. S. (2012). The benefit of bilateral versus unilateral cochlear implantation to speech intelligibility in noise. *Ear and Hearing*, *33*(6), 673–82.

Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *Journal of the Acoustical Society of America*, *106*, 3578–3588.

Van Wanrooij, M. M., & Van Opstal, A. J. (2004). Contribution of head shadow and pinna cues to chronic monaural sound localization. *Journal of Neuroscience 24*, 4163–4171.

Acknowledgments

McKnight Professorship (AJO)
American Auditory Society Mentored Student Poster Award (CD)