PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

# Understanding auditory motion perception: the role of temporal fine structure and envelope cues

Michaela WARNECKE[1]; Ruth Y. LITOVSKY[1]

[1] University of Wisconsin-Madison, Waisman Center, 1500 Highland Ave, Madison, WI, 53703, United States

## ABSTRACT

Auditory objects typically involve motion, either because the sound source is in motion, or the listener is. Yet, most psychoacoustic research has focused on investigating sound localization for stationary sources. In normal hearing (NH) listeners, research utilizing "chimera" signals showed that envelope cues are sufficient for robust speech perception, while temporal fine-structure (TFS) cues govern binaural sound lateralization. By contrast, listeners who are deaf and use bilateral cochlear implants (BiCI) do not have access to TFS cues through clinical processors. Recent work has demonstrated poor auditory motion perception in BiCI listeners compared to NH listeners. In this study, we investigate the impact of TFS on auditory motion perception and speech understanding by simulating both stationary and moving sounds consisting of chimera-style speech and noise stimuli in virtual auditory space. Perceptual measures include speech perception, identifying whether a sound is stationary or moving, discriminating the direction of motion and explicitly reporting the absolute range of motion. The results of this study provide a baseline for measurements to be conducted with BiCI listeners. Knowledge of the cues that are not available with clinical bilateral have the potential to provide insight for improved engineering approaches for BiCIs.

Keywords: Auditory Motion, Temporal Fine Structure, localization

## 1. INTRODUCTION

Even though most acoustic signals in our environment are in motion – either because the sound source moves, or we do – most psychoacoustic research investigating parameters of sound perception has been done using stationary stimuli.

Normal-hearing listeners utilize binaural cues to localize sounds in the horizontal plane. These include differences in time of arrival and intensity of sounds at the two ears (Wightman and Kistler, 1992; Blauert, 1997). For stationary sounds, humans show a remarkable sensitivity to changes in sound source locations: the minimum audible angle (MAA), the smallest perceivable difference in location of two stationary sounds, is ~1° for broadband noise stimuli in the frontal hemifield (Mills, 1958; Grantham, 1986). This sensitivity decreases sharply with stimulus frequency and location, up to about 75° thresholds for locations beyond the frontal 60° (Yost, 1974; Moore, 2012).

Measuring the sensitivity to changes in moving sounds is comparatively difficult, because parameters such as velocity and displacement co-vary with the movement and duration of the stimulus. Nonetheless, several studies have aimed at identifying the minimal audible movement angle (MAMA), the smallest perceivable angular range a sound has to move for it to be perceived as moving in the correct direction. Perrot and Marlborough (1989) used a broadband noise on a speaker moving 20°/s and reported MAMA of ~1°, a result that is comparable to the MAA for stationary sounds. Yet, other studies have found more variable results, such as MAMA of 2° to 4° for noise bursts moving at 2.8°/s (Harris and Sargeant, 1971), and noting that MAMA increased consistently, from about 8° to 60° for sound velocities ranging from 90°/s to 360°/s (Perrot and Musicant, 1977). In general, the MAMA appears to be about 2 to 3 times larger than the MAA for horizontal stimuli measured under similar conditions (for a full review, see Carlile and Leung, 2016).

Individuals who have suffered a profound hearing loss can seek to have auditory input through an

[1] mwarnecke@wisc.edu

implantable prosthetic known as a cochlear implant, which stimulates the auditory nerve via electrodes that are inserted into the cochlea. Bilateral cochlear implants (BiCIs) are now provided to a growing number of patients, but BiCI users receive degraded acoustic cues compared to NH listeners, and struggle to achieve similar levels of performance on spatial localization tasks (van Hoesel and Tyler, 2003; Schoen et al., 2005; Litovsky et al., 2009; Aronoff et al., 2010; see also Dorman et al., 2016). On the MAA task, thresholds for BiCI users range from ~3° to 8° at frontal locations and increase sharply to > 45° at peripheral locations (Senn et al., 2005). While the MAMA has not been explicitly tested in cochlear implant users, a recent study investigated the effect of sound duration on auditory motion perception in both normal hearing (NH) and BiCI listeners (Moua et al., 2019). In that study, NH listeners correctly identified stationary and moving sounds, independent of sound duration, as long as the angular movement of the sound was 10° or greater. By contrast, BiCI listeners had trouble discriminating between stationary and moving sounds, especially when sound durations were shorter than 1 second. Moreover, BiCI users were generally biased towards classifying sounds as moving, even when they were stationary.

Collectively, these studies suggest that BiCI listeners are disadvantaged compared to NH listeners in terms of sound localization and sound motion identification. Given that most natural sounds are not stationary, understanding what contributes to such deteriorated auditory motion perception is of great interest and importance.

The auditory input received by CI listeners is generally degraded compared to that of NH listeners, in both temporal and spatial resolution: a total of eight independent channels are stimulated at the same time, each of which amplitude-modulates an electrical pulse train (Rubinstein, 2004). This means that the temporal fine-structure (TFS) of the original sound is discarded. In fact, Senn et al. (2005) show that none of their participants with BiCIs discriminated between interaural fine structure differences, indicating that no TFS information was transmitted via cochlear implants. Despite not having access to acoustic TFS, BiCI users show good speech perception in quiet (e.g. Fishman et al., 1997; Friesen et al., 2001; Litovsky et al., 2009). This is not surprising, as natural speech is remarkably robust to this kind of signal degradation: studies with NH listeners in quiet have shown that only about four to eight channels of spectral *envelope* information are necessary to retain 85% to 100% speech perception (Smith et al., 2002).

While the envelope of a speech signal may be sufficient for speech perception, its temporal fine structure has been shown to impact sound source lateralization (e.g. Smith et al., 2002). However, this lateralization cue appears to be limited to lower frequencies (< 2.5 kHz; Wightman and Kistler, 1992; see also Brungart and Simpson, 2008).

Given that the low frequency TFS of a signal aids in perceptually locating a sound source, we explored whether removing low frequency TFS from a moving signal impacts auditory motion perception of NH listeners. If so, it would indicate that low frequency TFS of a signal may be one cue that influences auditory motion perception. Further, it could potentially explain why BiCI users, who do not have access to low frequency TFS, show poor performance when asked to identify components of moving sounds (Moua et al., 2019).

Additionally, we were interested in learning more about how envelope cues affect auditory motion perception: Human speech sounds are complex, containing not just frequency modulations, but also envelope modulations across time. To understand what parts of speech might influence auditory motion perception, we utilized a modified version (see Methods section 2.4) of the Chimeras introduced by Smith et al. (2002). This allowed us to create a set of stimuli which tested (1) the influence of signal envelope, and (2) the influence of low frequency TFS on auditory motion perception.

## 2. METHODS

### 2.1 Participants

Nine listeners (8 female) between 19 to 32 years of age participated in this study. All listeners had normal hearing assessed audiometrically and were naïve to the study's experimental design and purpose. All experimental procedures followed the regulations set by the National Institutes of Health and were approved by the University of Wisconsin's Human Subject Institutional Review Board.

## 2.2 Psychoacoustic testing

All testing was done in a sound booth (2.9 m x 2.74 m x 2.44 m; Acoustic Systems, Austin, TX, USA) with inner walls covered in acoustic foam (Pinta Acoustics, Minneapolis, MN, USA). Participants sat in a chair facing a horizontal array of 37 loudspeakers (Cambridge SoundWorks, North Andover, MA, USA; TDT Technologies, Alachua, FL, USA), which covered spatial locations from -90° (left) to +90° (right) in 5° steps. Four infrared cameras (OptiTrack, Natural Point Inc., Corvallis, OR, USA) were mounted on the ceiling of the sound booth to continuously monitor reflective markers attached to a small, custom-built handheld laser pointer, which was used to indicate the location of the perceived sound. On each trial, participants pressed a small touch screen (OnLap 1303, GeChic, Taichung City 403, Taiwan) to start the trial. Subsequently, a sound (moving or stationary) played at a location along the horizontal speaker array. Sounds were moved in virtual auditory space by applying vector-based amplitude panning (VBAP) to the stimuli (Pulkki, 1997). The sound was calibrated at 65 dBA SPL using a sound level meter (System 824, Larson Davis, Depew, NY, USA). Participants were instructed to face forward and keep their head still during sound presentation. After stimulus offset, participants could move their head. They were instructed to use the pointer to indicate the perceived sound location (for stationary sounds), or to trace the sound from its perceived start location to its perceived end location (for moving sounds). After indicating the sound location/tracking, the frontal touch screen displayed a matrix of all possible words that were part of the stimulus set (see below). Participants could either select a word from the matrix or select a button to indicate they did not recognize the word. After submitting their answer, a new trial could be started by pressing the touch screen.

## 2.3 Stimuli and Experimental Design

Speech tokens were composed of 56 two-syllable spoken word recordings (TVM corpus: Helfer and Freyman et al., 2009). The duration of words averaged 508 ms. A total of six stimuli were created for each of the words: 1. Original speech (OS); 2. Spectrally-matched noise (SMN); 3. Speech chimera (SC); 4. Noise chimera (NC); 5. Filtered speech chimera (SC_f); and 6. Filtered noise chimera (NC_f). These six conditions fall into one of three categories: Original stimuli (OS, SMN), Chimera stimuli (SC, NC), and filtered Chimera stimuli (SC_f, NC_f). Further, each stimulus condition contains either speech (OS, SC, SC_f) or noise (SMN, NC, NC_f) in its envelope. To create chimeras of each word, we utilized a modified version of Smith et al.'s (2002) Chimera-generating approach. Briefly, a recording of speech and a spectrally-matched noise token were each split into eight frequency bands spanning from 80 to 8000 Hz. Subsequently, for each frequency band, the envelope and TFS of both input signals were extracted and exchanged, such that the envelope of one signal was convolved with the TFS of the other, and vice versa. A summation across frequency bands resulted in two, multiple-band chimeras. A second set of chimeras was bandpass-filtered between 2.5 kHz and 8 kHz using a 6th order Butterworth filter to remove low frequency TFS.

Each stimulus condition was tested four times: twice stationary, and each once while it was moving horizontally to the left or the right. Sounds moved a 10° angular range. Participants listened to equal numbers of stationary/moving sounds, as well as an equal number of the six conditions. All of the variables were randomized during testing.

In order to estimate each participant's threshold for what constituted a stationary vs. a moving sound, we tested stationary localization prior to the main experiment. The stimuli for this testing were made from a different set of words, but created in the same way as stimuli for the main experiment. As such, the *stationary-only* testing served to (1) get the participants familiarized with the sound stimuli, and (2) test the effect of low frequency TFS on stationary localization.

## 2.4 Data Analysis

Overall, two main components were analyzed for this data set: (1) the impact of low frequency TFS on stationary localization and auditory motion perception (2) the impact of the envelope cue (Speech/Noise) on auditory motion perception.

Raw data were summarized across participants and conditions. Analysis for statistical significance was done in JMP (SAS), and all analyses were performed across stimulus category (Original, Chimera, filtered Chimera) and envelope type (Speech/Noise). We entered participant ID as a random variable, unless otherwise stated. A mixed model method was used and, when needed, posthoc analyses were performed with Tukey HSD comparisons.

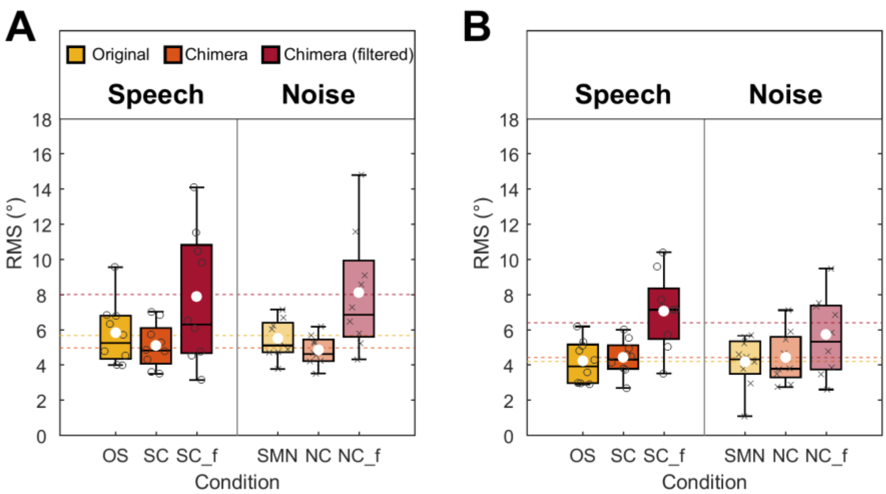
All data graphs plot the variable of interest (y-axis) per condition (x-axis). Importantly, data graphs are further grouped by their envelope content: conditions that contained speech in their envelope (OS, SC, SC_f) are plotted on the left, while conditions with noise in their envelope (SMN, NC, NC_f) are plotted on the right. The category that each condition falls into (see above) is indicated by color.

## 3. RESULTS

In this study, we examined how low frequency temporal fine structure and envelope cues impact auditory motion perception. To do so, 9 participants listened to six different stimuli (OS, SMN, SC, NC, SC_f, NC_f), which were presented both as moving and as stationary.

### 3.1 Stationary Sound Localization

Prior to the main experiment, each participant performed a stationary-only test to reveal their stationary localization threshold (see Section 2.3). Individual and group data for RMS localization errors are shown in Figure 1A. Category (Original, Chimera, filtered Chimera) is indicated by color. Across conditions, RMS errors ranged from 3.16° to 14.81°. We observed a main effect of stimulus category ($F_{2,40} = 12.5$; $p < 0.0001$), and posthoc analyses confirmed significantly larger RMS errors in conditions with bandpass-filtered Chimera stimuli (mean = 8°; SEM = 0.8°) compared to unfiltered stimuli (Original: mean = 5.6°, SEM = 0.34°, t = 3.68, $p = 0.0019$; Chimera: mean = 4.9°, SEM = 0.4°, t = -4.78, $p < 0.001$).




s

Figure 1 – Stationary localization RMS errors. **(A)** RMS errors (y-axis) collected during the stationary-only phase across participants (individual markers) along with median (black line) and mean (white circle) per condition (x-axis). Mean values per category (legend) are indicated by dashed lines. Data are organized by whether a given condition had speech (left, circle) or noise (right, cross) in the envelope. **(B)** same as (A) for correctly identified stationary trials collected during the main experiment.

During the main experiment, participants were presented with both moving and stationary stimuli, and asked to discriminate between these two categories of motion. When participants correctly identified stationary sounds as stationary, RMS error patterns were the same as in the stationary-only testing phase (Fig. 1B): RMS errors ranged from 1.1° to 10.42° across conditions. We observed a main effect of stimulus category ($F_{2,40} = 20.9$; $p < 0.0001$), and posthoc analyses confirmed larger RMS errors in conditions with bandpass-filtered Chimera stimuli (mean = 6.4°; SEM = 0.52°) compared to unfiltered stimuli (Original: mean = 4.2°, SEM = 0.3°, t = 5.88, $p < 0.0001$; Chimera: mean = 4.4°, SEM = 0.28°, t = -5.28, $p <0.0001$).

### 3.2 Auditory Motion Perception

To evaluate how well participants could distinguish between stationary and moving sounds, *d'* was used as a sensitivity measure. Figure 2A plots *d'* for individuals and group data for each condition.

Sensitivity was lower in conditions that had speech in their envelopes (Fig. 2A left; mean = 0.45, SEM = 0.07), compared to conditions that had noise in their envelopes (Fig. 2A, right; mean = 0.86, SEM = 0.14). Statistical analyses revealed that this sensitivity difference was significant ($F_{1,40} = 11.61$, $p = 0.0015$). Further, we found a significant effect of category ($F_{2,40} = 6.75$; $p = 0.003$), and posthoc analyses revealed that the sensitivity was significantly lower for filtered stimuli (mean = 0.35, SEM = 0.08), compared to unfiltered Chimera stimuli (mean = 0.76, SEM = 0.143; t = 2.74, $p = 0.024$), or Original stimuli (mean = 0.87, SEM = 0.18; t = -3.49, $p = 0.0034$). We also found a significant interaction between category and envelope type ($F_{2,40} = 4.44$, $p = 0.018$).

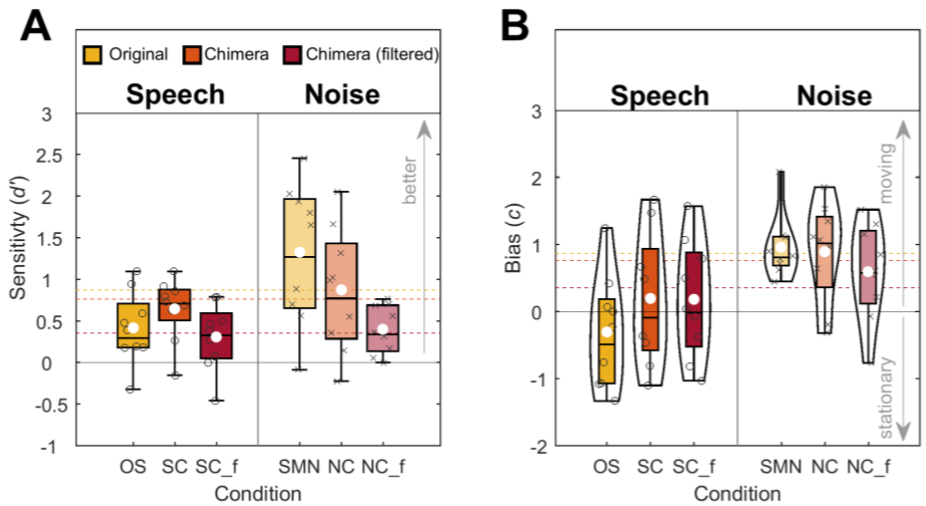



Figure 2 – Sensitivity and Bias estimates. **(A)** Sensitivity (y-axis) across participants (individual markers) along with median (black line) and mean (white circle) per condition (x-axis). Mean values per category (legend) are indicated by dashed lines. Data are organized by whether a given condition had speech (left, circle) or noise (right, cross) in the envelope. Better sensitivity is indicated by larger values (higher *d'*). **(B)** Same as (A) but for Bias estimates. Additionally, violin plots add probability density function of the data distribution. Larger vs. smaller values indicate bias towards classifying sounds as moving vs. stationary, respectively.

The difference in sensitivity across envelope types and category led us to evaluate whether participants showed a bias towards perceiving certain stimuli as stationary or moving. Figure 2B plots the bias for individuals and group data for each condition. Overall, bias values were lower in conditions that had speech in their envelopes (mean = 0.02, SEM = 0.17), compared to conditions which had noise in their envelopes (mean = 0.81, SEM = 0.12). Statistical analyses revealed that this difference was significant ($F_{1,40} = 50.42$, $p < 0.0001$). We did not find a main effect of category, but report a significant interaction between stimulus category and envelope type ($F_{2,40} = 5.06$, $p = 0.0109$).

Overall, participants correctly identified the motion of sounds (stationary vs. moving) 62.4% of the time. Figure 3 plots the individual and group data as mean percent (± SEM) for correctly identifying stationary sounds (Fig. 3A) or moving sounds (Fig. 3B), and the absolute angle that participants tracked for moving sounds (Fig. 3C). Overall, participants more often identified stationary trials as stationary in conditions that had speech in the envelope (Fig. 3A, left; mean = 56.7%, SEM = 5.7%), compared to conditions with noise in the envelope (Fig. 3A, right; mean = 38.3%, SEM = 4.8%). Statistical analyses confirmed that this difference was significant ($F_{1,40} = 26.7$, $p < 0.0001$).

For trials with moving sound, the opposite pattern was observed: participants more often correctly identified moving trials as moving in conditions with noise in their envelope (Fig. 3B, right; mean = 85%, SEM = 3.13%), compared to speech in the envelope (Fig. 3A, left; mean = 57%, SEM = 5.19%). This difference was statistically significant ($F_{1,40} = 48.7$, $p < 0.0001$), and there was a significant interaction between the stimulus category and envelope type ($F_{2,40} = 5.76$, $p = 0.0063$). Further, participants also correctly identified the direction of motion (left/right) significantly more often in conditions that had noise in the envelope (Fig. 3B right, transparent; mean = 72.3%, SEM = 3.57%), compared to those with speech in the envelope (Fig. 3B left, transparent; mean = 38.3%, SEM = 3.5%; $F_{1,40} = 92.6$, $p < 0.0001$). Here, too, the interaction between the stimulus category and envelope type was significant ($F_{2,40} = 11.5$, $p < 0.0001$).

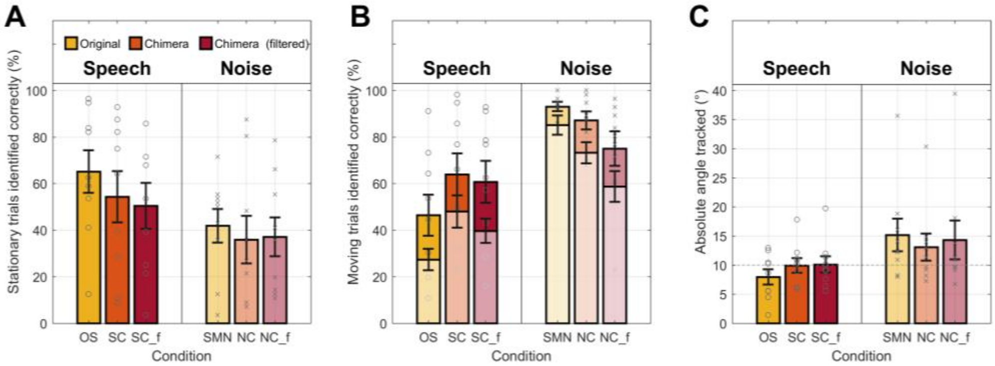



Figure 3 – Percent correct for different measures of auditory motion perception. **(A)** Mean percent correct (± SEM, y-axis) per condition (x-axis) across participants (individual markers) for identifying stationary trials as stationary. **(B)** Mean percent correct (± SEM, y-axis) per condition (x-axis) across participants (individual markers) for identifying moving trials as moving. Additionally, transparent data bars indicate how often participants correctly identified the direction of motion. **(C)** Mean absolute angle tracking (± SEM, y-axis) per condition (x-axis) across participants (individual markers) for sounds that moved 10° (dashed line).

How well could participants track the absolute angle that the sound moved? All moving sounds had an angular range of 10°, and when stimuli had speech in the envelope, participants tracked the angle of the moving sound significantly more accurately (mean = 9.33°, SEM = 0.75°) compared to when there was noise in the envelope (mean = 11.55°, SEM = 0.66°; $F_{1,37} = 17.3$, $p = 0.0002$). There was no effect of category.

## 4. DISCUSSION

In everyday life we are constantly surrounded by moving sounds. Yet, the cues that may help us identify, locate, and evaluate the direction of moving sounds are not well understood. In this study, we aimed to understand how auditory motion perception is influenced by (1) low frequency temporal fine structure of a signal, and (2) envelope cues.

Overall, the low frequency TFS affected both stationary localization and the ability to distinguish between stationary and moving sounds. All participants performed well when localizing stationary sounds (Fig. 1). Importantly, when we removed low frequency TFS from the stimulus (SC_f, NC_f) sound localization accuracy degraded (Fig. 1, Chimera (filtered)). This result indicates that low frequency TFS impacts precise sound source localization, and confirms that our manipulation of bandpass-filtering the chimera stimulus was effective. Further, this finding is generally in line with previous research showing that sound source lateralization depends on temporal alignment of signal TFS, but not so much on envelope (Smith et al., 2002).

We further evaluated participants' abilities to distinguish between stationary and moving sounds. Sensitivity scores (Fig. 2A) were degraded when stimuli did not contain low frequency TFS. Recently, Moua et al. (2019) tested auditory motion perception for simulated stationary and moving sounds in both NH and BiCI listeners. Their results showed that BiCI listeners had more difficulty discriminating between stationary and moving sounds compared to NH listeners. In that study, NH listeners' access to low frequency acoustic TFS may have contributed to better performance compared to that of BiCI listeners. CI speech processors discard low frequency TFS which means that binaural TFS cues are also absent (for review see Kan and Litovsky, 2015). Our results show that removing access to low frequency TFS impacts the ability to discriminate between stationary and moving sounds in NH listeners. As such, BiCI listeners may benefit from low frequency TFS cues to more accurately localize stationary sounds, and to be able to improve their discrimination of stationary and moving sounds.

The present study also showed that envelope cues strongly influence auditory motion perception. When participants listened to stimuli with speech in the envelopes, their ability to discriminate between stationary and moving sounds was significantly worse than when stimuli had noise in their envelopes (Fig. 2A). In fact, the significant interaction between stimulus category and envelope type shows that, as stimuli envelopes change from containing speech to containing noise, discriminability increased; in the case of original stimuli (OS, SMN) the increase was more than twofold on average. This finding suggests that moving sound stimuli with a flat noise envelope are easier to classify correctly. This interpretation is corroborated by the finding that stimuli with noise envelopes were generally perceived as moving, compared to stimuli with speech envelopes (Fig. 2B).

It was not just the correct identification of motion that was impacted by envelope type, but also the correct motion direction: stimuli with noise in their envelope showed better direction discrimination, which suggests that the noise envelope aids in receiving a clearer percept of the moving sound altogether. In fact, for *both* stationary (Fig. 3A) and moving sounds (Fig. 3B), stimuli that had a speech envelope reached about only 57% correct identification performance, indicating that this envelope cue makes it generally difficult to distinguish motion (cf. Fig. 2A).

Overall motion sensitivity scores were low (mean = 0.66, SEM = 0.08), indicating that 10° movement was difficult to distinguish from a non-moving sound. Future work can address this issue with an easier task, using sounds that move in larger increments.

Measurements of sound motion detection and direction give a good indication of auditory motion perception, however, we were also interested in learning about whether participants can accurately trace the distances that sounds moved. In this study, participants were able to track the angular motion well (Fig. 3C), and estimates were more accurate when stimuli had a speech envelope. These angular distance estimates stem from trials on which participants correctly identified a moving sound as moving, and only ~ 38% of trials with a speech envelope fulfilled that requirement. By contrast, an average of 72% of trials with noise envelope contributed to the estimate for angular distance tracked.

These results outline important factors in auditory motion perception. Specifically, learning that motion perception is degraded for stimuli with speech envelopes aids in designing future experiments that evaluate motion and speech perception. It is important to understand the cues that underlie auditory motion perception, not just for NH listeners, but also for listeners with hearing impairments. Research in NH listeners has shown that discriminating sound motion is highly dependent on the frequency content, stimulus duration, and movement velocity (for a full review see Carlile and Leung, 2016). The current results show that sound envelope may be another important factor impacting sound motion perception. This is an important finding because some stimulus features (such as envelope) can be confidently represented by CIs, and perceived by BiCI users, whereas other features (such as TFS) are not represented by CIs with fidelity. This further underscores the importance of TFS for auditory motion perception; here we have shown that low frequency TFS impacts the discriminability of stationary and moving sounds. The ability to discriminate between stationary and moving sounds is essential, especially in everyday life, where moving sounds can indicate fast approaching danger.

## 5. CONCLUSIONS

In this experiment, we evaluated whether low frequency TFS or signal envelopes impact auditory motion perception. Our results show that low frequency TFS degrades stationary localization, and the ability to distinguish between stationary and moving sounds. We further find that envelope cues strongly influence auditory motion perception. Specifically, a speech envelope makes it more difficult to distinguish stationary from moving sounds, and biases the listener towards categorizing a sound as stationary. By contrast, a noise envelope biases a listener towards categorizing sounds as moving.

## ACKNOWLEDGEMENTS

The authors thank the participants for their time and patience. The research was funded by NIH-NIDCD (R01DC8083) and NIH-NICHD (U54 HD090256).

## REFERENCES

1. Wightman FL, Kistler DJ. The dominant role of low-frequency interaural time differences in sound localization. The Journal of the Acoustical Society of America. 1992; 91(3):1648-61.
2. Blauert J. Spatial hearing: the psychophysics of human sound localization. MIT press; 1997.
3. Mills AW. On the minimum audible angle. The Journal of the Acoustical Society of America. 1958; 30(4):237-46.
4. Grantham DW. Detection and discrimination of simulated motion of auditory targets in the horizontal plane. The Journal of the Acoustical Society of America. 1986; 79(6):1939-49.
5. Yost WA. Discriminations of interaural phase differences. The Journal of the Acoustical Society of America. 1974; 55(6):1299-303.
6. Moore BC. An introduction to the psychology of hearing. Brill; 2012.
7. Perrot DR, Marlborough K. Minimum audible movement angle. J. Acoustic. Soc. Am. 1989; 85:1773-5.
8. Harris JD, Sergeant RL. Monaural/binaural minimum audible angles for a moving sound source. Journal of Speech and Hearing Research. 1971; 14(3):618-29.
9. Perrott DR, Musicant AD. Minimum auditory movement angle: Binaural localization of moving sound sources. The Journal of the Acoustical Society of America. 1977; 62(6):1463-6.
10. Carlile S, Leung J. The perception of auditory motion. Trends in hearing. 2016; 20:2331216516644254
11. van Hoesel RJ, Tyler RS. Speech perception, localization, and lateralization with bilateral cochlear implants. The Journal of the Acoustical Society of America. 2003; 113(3):1617-30.
12. Schoen F, Mueller J, Helms J, Nopp P. Sound localization and sensitivity to interaural cues in bilateral users of the Med-El Combi 40/40+ cochlear implant system. Otology & Neurotology. 2005;26:429-37.
13. Litovsky RY, Parkinson A, Arcaroli J. Spatial hearing and speech intelligibility in bilateral cochlear implant users. Ear and hearing. 2009; 30(4):419.
14. Aronoff JM, Yoon YS, Freed DJ, Vermiglio AJ, Pal I, Soli SD. The use of interaural time and level difference cues by bilateral cochlear implant users. The Journal of the Acoustical Society of America. 2010; 127(3): EL87-92.
15. Dorman MF, Loiselle LH, Cook SJ, Yost WA, Gifford RH. Sound source localization by normal-hearing listeners, hearing-impaired listeners and cochlear implant listeners. Audiology and Neurotology. 2016; 21(3):127-31.
16. Senn P, Kompis M, Vischer M, Haeusler R. Minimum audible angle, just noticeable interaural differences and speech intelligibility with bilateral cochlear implants using clinical speech processors. Audiology and Neurotology. 2005; 10(6):342-52.
17. Moua K, Kan A, Jones HG, Misurelli SM, Litovsky RY. Auditory motion tracking ability of adults with normal hearing and with bilateral cochlear implants. The Journal of the Acoustical Society of America. 2019; 145(4):2498-511.
18. Rubinstein JT. How cochlear implants encode speech. Current opinion in otolaryngology & head and neck surgery. 2004; 12(5):444-8.
19. Fishman KE, Shannon RV, Slattery WH. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. Journal of Speech, Language, and Hearing Research. 1997; 40(5):1201-15.
20. Friesen LM, Shannon RV, Baskent D, Wang X. Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. The Journal of the Acoustical Society of America. 2001; 110(2):1150-63.
21. Smith ZM, Delgutte B, Oxenham AJ. Chimaeric sounds reveal dichotomies in auditory perception. Nature. 2002; 416(6876):87.
22. Brungart DS, Simpson BD. Effects of temporal fine structure on the localization of broadband sounds: potential implications for the design of spatial audio displays. ICAD 2008. Paris.
23. Pulkki V. Virtual sound source positioning using vector base amplitude panning. Journal of the audio engineering society. 1997; 45(6):456-66.
24. Helfer KS, Freyman RL. Lexical and indexical cues in masking by competing speech. The Journal of the Acoustical Society of America. 2009; 125(1):447-56.
25. Kan A, Litovsky RY. Binaural hearing with electrical stimulation. Hearing Research. 2015; 322:127-37.