Spatial Speech Enhancement: Thrive in More Challenging Environments

Jumana Harianawala, Au.D. and Ben Waite, MS

Noisy environments present the hearing aid user with many challenges. Background noise amplified through a hearing aid may be perceived as excessively loud and even uncomfortable. Attending to a target talker in the presence of interfering talkers and babble noise can be difficult and cause cognitive fatigue. It is no surprise then that dissatisfaction with hearing aid performance in noise is a chief complaint amongst hearing aid users (Kochkin, 2000; Hickson, Clutterbuck, Khan 2010; Hong, Oh, Jung, Kim, Kang, & Yeo, 2014; Johnson, Xu, & Cox, 2016) and a strong determinant for reduced or discontinued use of hearing aids (McCormack & Fortnum, 2013; Hickson, Meyer, Lovelock, Lampert, & Khan, 2014; Aazh, Prasher, Nanchahal, & Moore, 2015; Bennett, Laplante-Lévesque, Meyer, & Eikelboom, 2017).

Essentially all modern hearing aids use some form of directional processing and digital noise reduction (DNR) to combat these issues. Both treatments reduce the amount of noise presented to the hearing aid user, but in substantively different ways. Whereas directional processing attenuates sounds originating in the rear (including speech) with the intent of accentuating speech originating in the front of the user, DNR systems attenuate sounds that have characteristics typical of noise regardless of the direction of arrival. The result is that directional processing improves speech intelligibility for front-facing speech, while DNR systems improve comfort (Mueller, Weber, & Hornsby, 2006; Palmer, Bentler, & Mueller, 2006) and reduce listening effort (Bentler, Wu, Kettle, & Hurtig, 2008; Sarampalis, Kalluri, Edwards, & Hafter, 2009) while preserving, but not improving, speech intelligibility.

The simultaneous suppression of noise and preservation of desired speech is at the heart of the DNR problem, and the efficacy of a DNR system depends largely on how well it separates the noise in an acoustical scene from the desired speech signal. If the system fails to identify noise as noise, it won’t attenuate anything and thus offers no relief to the hearing aid user; if the system incorrectly identifies target speech as noise, it will be attenuated, and intelligibility will ultimately suffer.

Single microphone DNR systems take the signal from a single microphone as their input. Starkey’s Voice iQ2, a proprietary form of fast-acting single microphone noise reduction, exploits known differences in the temporal characteristics of speech and noise in order to separate them. Simply put, the level of a speech signal changes over short (syllabic) time durations, whereas the level of a noise signal tends to remain constant in the short term (i.e., noise tends to be “stationary”). Thus, time-frequency regions where the signal is relatively stationary are attenuated, while dynamic speech-like signals are presented to the hearing aid user with the prescribed gain. Clinical investigation of Voice iQ2 has demonstrated improvements in comfort, sound quality, and willingness to accept background noise (Pisa, Burk, & Galster, 2010). Independent studies have also demonstrated that listeners experience decreased listening effort when using Voice iQ2 while listening in noise (Desjardin & Doherty, 2014).

While Voice iQ2 has undoubtedly improved the hearing aid user’s experience in stationary background noise, environments where the noise is more speech-like, and thus less stationary, continue to be a problem. Consider the background noise encountered in a noisy restaurant or bar. Interfering speech from the left and right sits atop a fluctuating noise floor comprised of dozens of speakers in a reverberant space. The restaurant, like so many environments encountered in daily life, is full of speech-like noise, and single microphone DNR systems reliant on differences in the speech and noise parts of the acoustic environment
struggle to keep up. To do more effective noise reduction in these environments, we must improve the input to the noise reduction system.

**SPATIAL SPEECH ENHANCEMENT**

Starkey’s new *Spatial Speech Enhancement* (SSE) feature uses the acoustical inputs from both hearing aids to better separate speech from noise, and thus improve attenuation of noise at no cost to the desired speech input. This is facilitated by high-fidelity ear-to-ear audio streaming via a near-field magnetic induction (NFMI) radio introduced with the *Thrive platform with Livio AI and Livio hearing aids*. SSE compares ipsilateral audio input with the streamed contralateral audio input. When these inputs are similar, they are treated as a desired input, and when they are different, they are treated as unwanted noise. This treatment has a dual effect. First, random stationary noise is treated as noise, much like it is by a single microphone DNR system. Second, and in contrast to a single-microphone method, coherent inputs (e.g. speech) from the side, which are uncorrelated at the left and right ears, are also treated as noise.

Consider again the restaurant scenario; you are attending to a talker directly in front of you while interfering talkers converse at a table to your right. To a single microphone noise reduction system, the interfering speech to the right looks identical at the left and right ears. To SSE however, the interfering speech from the right looks different at the right and left ears. The interfering speech is altered as it travels from the right to the left ear (around the listeners head) resulting in level and phase differences. SSE uses these differences to separate the unwanted noise from the desired signal and is thus able to relieve the hearing aid user from interfering speech.

The net-effect is that SSE is more successful than single-microphone DNR techniques in attenuating speech-like noise, and in improving the signal-to-noise ratio (SNR) for speech in speech-like noise (including interfering speech). The effect of SSE on speech-like noise is illustrated in Figure 1. Panel (A) shows, as the speech-like noise input grows louder, SSE’s efficacy relative to the single-microphone technique improves. Panel (B) illustrates a similar effect, but this time with a speech-in-speech-like-noise mixture. As the signal-to-noise ratio of this mixture degrades (i.e., as the noise level grows relative to the level of the target speech), SSE’s ability to improve SNR at the output of the hearing aid (i.e., its ability to attenuate noise while preserving the target speech) grows relative to that of the single-microphone DNR technique.

**Figure 1:** Comparison of electroacoustic results achieved by Single Microphone Noise Reduction techniques (Voice IQ2) and Binaural Microphone Noise Reduction (Spatial Speech Enhancement). Panel (a) shows the reduction for speech like noise alone and Panel (b) shows signal to noise ratio (SNR) improvement for a speech in speech like noise mixture.
NOISE REDUCTION AND DIRECTIONAL PROCESSING REVISITED

As previously noted, modern hearing aids employ DNR systems in concert with directional processing to enhance the user experience in noisy environments, with each approach having unique goals and offering unique benefits. SSE represents a departure from this paradigm in that it acts both as a classical DNR system which attenuates stationary noises, while also providing a directional benefit by suppressing coherent signals from the side. This connection with directional processing motivates an integrated approach to noise reduction and directional processing.

The automatic enabling and disabling of directional processing in hearing aids is, like single microphone DNR techniques, ubiquitous. At lower input levels, directional processing is turned off (omnidirectional) or moderated (directional only in the high frequencies), while at higher input levels, broadband directional processing is enabled. This allows for optimal sound clarity and access at low input levels and focus on the look direction at higher input levels. With the introduction of SSE, a similar approach is applied to noise reduction. At higher input levels, SSE is enabled, and at lower input levels, single microphone noise reduction is enabled.

With the Thrive platform, the single microphone noise reduction feature has been reformulated specifically to pair with SSE. It employs the theoretical underpinnings of SSE but does not require streamed contralateral audio to work. Accordingly, it does not offer the same benefits relative to speech-like noise but provides identical performance relative to stationary noise. This facilitates a seamless switching between SSE at higher input levels (thus identified in Inspire as “Speech in Loud Noise”) where it is most beneficial, and single noise reduction (or “Speech in Noise”) at lower input levels where sound clarity and situational awareness are most important.

HEARING REALITY SOUND ENHANCEMENT

Noise reduction and directional processing are managed by Starkey’s Inspire X fitting software. The Sound Manager screen seen in Figure 2 provides for adjustment of Speech in Loud Noise and Speech in Noise. The strength of Speech in Noise and Speech in Loud Noise can be adjusted via the Speech in Noise control in the Sound Enhancement screen. The rationale for modulating noise reduction strength via the Speech in Noise control is consistent with past iterations of noise reduction features; some users like a lot of noise reduction, while others prefer more natural audio. The Speech in Loud Noise control enables and disables SSE.

![Figure 2: Inspire X Sound Manager and Sound Enhancement screens showing controls for Speech in Noise (SIN) and Spatial Speech Enhancement (SSE) features.](image-url)
Default settings are designed to provide an optimal balance of noise reduction and audio quality, but a vital service of the fitting professional is to adjust that balance to the preference of the individual patient.

**VALIDATION STUDY**

**Participants and Fittings**
A laboratory study compared hearing aid wearers’ preference for SSE and Voice iQ2 when listening to speech in complex noisy environments. Fifteen individuals (seven men and eight women) ranging in age from 54 to 87 years (mean age of 72 years) participated in the study. On average, all participants had bilateral, symmetrical, mild-to-severe sensorineural hearing loss (Figure 3). All participants were experienced full-time users of bilateral hearing aids.

The hearing aids were fit to Starkey Hearing Technologies’ proprietary fitting formula, e-STAT (Scheller & Rosenthal, 2012). SSE and Voice iQ2 were compared at their default settings (up to 10 dB of noise reduction). The microphone mode was set to omnidirectional to allow for the study of SSE in isolation from directional processing. All participants were fit bilaterally with appropriately sized power domes. Real-ear verification was performed, and fine-tuning adjustments were made to ensure audibility. The two memories were matched in all respects except for the difference in noise reduction algorithms, SSE and Voice iQ2.

**Preference Test**
Participants’ preference between the SSE and Voice iQ2 was evaluated via a paired comparison task. Participants were seated in a sound-treated booth in the center of eight-speaker array. Target speech was presented from a loudspeaker directly in front of the participant (0° azimuth) at 73dBSPL and a diffuse babble noise at 70dBSPL was presented from seven other loudspeakers surrounding the participant (±45, ±90, ±135, 180°). In addition to diffuse babble, interfering speech signals at 70dBSPL were also presented from the loudspeakers on either side (±90°) of the participant. Four different speech-in-noise configurations were tested:

- **Diffuse babble only (no target speech)**
- **Target speech in diffuse babble**
- **Target speech in diffuse babble + 1 interfering talker**
- **Target speech in diffuse babble + 2 interfering talkers**

Participants blinded to the current setting switched between SSE and Voice iQ2, for each speech and noise configuration and indicated their preference. Each comparison was made twice by every participant.

**Results**
Figure 4 displays the preference (as percentage) for SSE and Voice iQ2. The Wilcoxon Signed-Ranks Test indicated that preference for SSE was significantly higher than preference for Voice iQ2, Z = 3.327, p < 0.001. The preference for SSE over Voice iQ2 was greater than would be expected due to chance (p < 0.5), for all but the target speech in diffuse babble configuration. Given there is only one speech like talker in the target speech in diffuse babble configuration, as discussed earlier,
both SSE and Voice iQ2 can similarly differentiate target speech from diffuse noise resulting in similar SNR improvement. For more complex noise environments with interfering talkers, as demonstrated by these results, the SSE is largely preferred over Voice iQ2.

CONCLUSION

DNR algorithms are known to improve comfort in noisy environments. In less complex environments with a single talker in diffuse babble noise, single microphone DNR systems such as Voice iQ2 provide sufficient noise reduction. However, for more complex environments with interfering talkers, single microphone DNR systems are insufficient. For these environments, Spatial Speech Enhancement is preferred over Voice iQ2. The study presented in this article clearly illustrates a preference by hearing aid wearers for Spatial Speech Enhancement available in Livio and Livio AI hearing aids over Voice iQ2. SSE offers the hearing aid user the benefit of minimizing the negative effects of distracting talkers and background noise in even the most challenging environments.

REFERENCES


The authors would like to acknowledge Ivo Merks for his contributions.

Figure 4: Participant preference (shown as percentage) for SSE & Voice iQ2 per noise configuration collapsed across all 15 participants. Graph shows preference for SSE vs. Voice iQ2 as stacked bar.