# Proving the classifier right: Using feature activation to validate accurate environmental detection

Jennifer Groth, MA; Tao Cui, AuD.

## Abstract

Automatic programs in hearing aids are commonly fit to hearing aid users as they can potentially provide benefit of complex features without effort on the part of the user. Their performance depends on accurate identification of acoustic environments as well as accurate application and control of advanced features. Previous work has shown varying accuracy in environmental classification across premium hearing aid brands. A new environmental classification and feature control system called AcoustIQ™ has been introduced in the Beltone Illuminate hearing aids. This paper reports on a technical study that builds on earlier studies of environmental classification accuracy by assessing accuracy of application of directionality and gain features in 3 types of prototypical sound scenes. The AcoustIQ system showed the highest accuracy of all systems tested with 88% greater accuracy than the best performing of the other systems.

Today's cars have many automatic features that operate based on weather, road conditions and traffic. A simple example is windshield wipers that activate and adjust based on detected precipitation on the windshield. Automatic features are meant to make driving easier and safer, but drivers can find them annoying and even dangerous when they activate at the wrong time or startle the driver with invalid alerts and sudden automatic actions. Systems that slam on the brakes unexpectedly due to shadows or overpasses are a great example of a feature intended to keep drivers safe, but that too often annoys them unnecessarily or even causes a perilous situation. Both the detection and identification of driving conditions as well as the strategy for activating features automatically are crucial elements in how successfully the system supports drivers in real life.

Like cars, hearing aids also deploy systems for identification and control to automatically apply sound processing features. These systems rely primarily on acoustic information picked up by the hearing aid microphones to analyse the listening conditions. Based on this analysis, the hearing aids automatically take actions intended to support the hearing aid wearer in real life listening environments. This requires not only accurate classification of the acoustic environments but also accurate application of beneficial features for these environments. This is more important than ever before as hearing aids are commonly fit with an automatic program, and these programs are increasingly opaque in terms of how they apply advanced features. Not only that, evidence for the appropriateness of the settings is remarkably lacking.[1]

The Beltone hearing aid system responsible for environmental analysis and automatic application and control of features is AcoustIQ. Beginning with the Beltone Illuminate, AcoustIQ delivers real time accurate sound sensing and engagement of the right features at the right time, unlocking broader access to the full sound environment.

Some investigators have assessed the accuracy of environmental classification in hearing aids by exposing hearing aids to defined sound scenes and then reading out their datalogging to see how the system categorized the sound scenes.[2-4] Groth et al[3] demonstrated the highest accuracy in environmental classification for the Beltone system compared to other hearing aid brands, with an overall accuracy of 91%. Key findings were that all systems could identify quiet and speech only environments with a high degree of accuracy, and that accuracy decreased in more complex sound scenes. Systems that included automatic music programs showed the lowest levels of accuracy overall.

There are some complications involved in using datalogging to assess and compare accuracy of environmental classification. One is that the datalogging of some systems explicitly labels environments without indicating what features were activated while the environment was detected,

while other systems data log the time spent in defined automatic programs where the name of the program suggests what the identified environment was. The labels for the prototypical environments and automatic programs are different across systems, making them difficult to compare. And different systems also have different numbers of environments or automatic programs meaning that decisions must be made about how to map them all to a common subset of environments. Perhaps the most critical limitation of studies like this is the assumption that sound scene classification that matches human judgments corresponds to the hearing aids behaving in a beneficial way in response to those scenes. In other words, the basic assumption that accurate environmental analysis leads to optimum activation of features is not proven by this approach.

Another way to evaluate accuracy of environmental classification was reported by Husstedt et al.[5] Their method involved presentation of different sound scenes to hearing aids with a continual comparison of the output to a reference to see how the output of the hearing aids changed depending on the sound scene. Their design was intended to infer what environment the hearing aid was detecting and what programs were automatically activated. Their results indicated that datalogging of environments or programs isn't enough to draw conclusions about accuracy because knowing how the environmental classification identifies the environment does not guarantee that features are applied optimally. Their work showed that hearing aids make mistakes in activating programs in cases where humans agree and have no trouble in recognizing the environment. For example, one oddity they noted was that a device they tested identified piano music at a moderate presentation level as speech: "This is astonishing, since the test signal does not include any voices and no human would classify this track as speech." We had a similar finding in our study[2] involving a vacuum cleaner test signal. Four devices consistently classified this sound as music, which we also found astonishing.

Welling et al[6] took a clinical approach to examining how hearing aid settings line up with listening environments but reached a similar conclusion. They found that acoustic information doesn't always reflect how people perceive a listening situation. They fit test participants with a hearing aid that had multiple programs including an automatic program and had them use Ecological Momentary Assessment via an app to collect perceptual data. They were interested in whether there was a relationship between the acoustic environment detected and what type of program was selected. Their results did not support systematic relationships between satisfaction with selected programs and acoustic environment information. Their results can in part be explained by the fact that listener intent is context dependent. The acoustic environments that were identified may have aligned with how humans would characterize them, but the listening goals of participants in the real world could easily be different than what the environment indicates they "should" be. Environments with music are a good example of how this could play out. A hearing aid user dining with friends at a restaurant where background music is playing might rather focus on the conversation at the table rather than the music. But the environmental classification and steering system would not know this and might apply adjustments to enhance the music. This strongly implies that successful application of features in an automatic program hinges on making their benefit available to users without interfering with their listening goals in the moment.

Because an essential part of AcoustIQ is how sound processing features are controlled due to environmental classification, we wanted to build on our earlier work on accuracy of environmental classification with an assessment of how accurate automatic activation of features is across hearing aid brands. We focused on how each system applies directionality and gain adaptations because these features are universal among hearing aid brands. The goal of the study was to evaluate accuracy of how these features are activated in prototypical sound scenes in the automatic programs of premium hearing aid brands.

## Methods

### Hearing aids and programming

This was a technical study with no human participants. A pair of the Beltone Illuminate 17 RIE hearing aids and the newest premium RIE hearing aids from five other brands were included. Each pair of hearing aids was programmed in the respective fitting software as a binaural fitting to ensure that any features relying on wireless communication or streaming between the devices could be activated. Each set of hearing aids had two programs. One program was omnidirectional with advanced features disabled. This program served as the baseline for determining whether directionality was active and whether gain changes had occurred in response to the stimuli. The second program in each case was the default automatic program. Both programs were set to 15 dB linear gain to allow focus on features other than compression that could affect gain.

## Sound scenes and recordings

Sound scenes processed through the hearing aids were recorded for analysis offline. For the recordings, each set of hearing aids was fitted to a mannikin with miniature DPA 4060 microphones sealed in the ear canals with putty to better capture the effects of the signal processing. The mannikin was placed in the center of a 5-speaker setup in a quiet office environment. Speakers were located 1.5m from the mannikin at 0, 45, 135, 225 and 315 degrees azimuth.

Three types of sound scenes were used: voices only, speech-in-noise and noise only. These scenes were also used in our previous technical studies of hearing aid environmental classification accuracy based on datalogging.[2,3] The first of these datalogging studies included a subjective

classification of the sound scenes as a human perceived reference that the datalogging results could be compared to as a gold standard for accuracy. Using these scenes in the current study allowed us to establish reasonable success criteria for accurate feature engagement. Table 1 provides an overview of the eight total sound scenes and presentation details. Presentation levels for all scenes except voices only were high with low signal-to-noise ratios (SNR). We designed the setup including presentation levels to create sound scenes where user intent in a real-world setting would very likely align with expectations for feature activation. Based on each brand's product-related literature, it was expected that these conditions would result in maximum engagement of noise management features in the scenes with noise.

| Sound scene | Environment | Presentation details |
|---|---|---|
| Voices only | Speech in quiet | 65 dB SPL speech from 0° |
| Café | Speech in noise | 75 dB SPL speech from 0°; 75 dB SPL noise from 45, 135, 225 and 315° |
| Living room | Speech in noise | 75 dB SPL speech from 0°; 75 dB SPL noise from 45, 135, 225 and 315° |
| Party | Speech in noise | 75 dB SPL speech from 0°; 75 dB SPL noise from 45, 135, 225 and 315° |
| Station | Speech in noise | 75 dB SPL speech from 0°; 75 dB SPL noise from 45, 135, 225 and 315° |
| Super-market | Speech in noise | 75 dB SPL speech from 0°; 75 dB SPL noise from 45, 135, 225 and 315° |
| Vacuum cleaner | Noise | 75 dB SPL from all 5 speakers |
| Hand mixer | Noise | 75 dB SPL from all 5 speakers |

Table 1. Sound scenes and presentation levels.

Each recording started with several seconds where a white noise test signal at 65 dB SPL was presented with the hearing aids in the omnidirectional program. This part of the recording served as the baseline for the analysis. The white noise was presented 5 times, once from each speaker surrounding the mannikin. Following this, the sound scene was presented, and the hearing aids were immediately switched to the automatic program. The sound scene played for 4 minutes to allow the hearing aids to adapt to the scene. Immediately when the 4-minute sound scene ended, the white noise test signal was again presented once at each speaker location without switching the program.

## Data analysis

All recordings were high-pass filtered with a cutoff frequency 100 Hz to remove low frequency interference. Audacity® digital audio workstation was used for analysis of the recordings.

Feature activation was evaluated by comparing the baseline response to the white noise test signal in the omnidirectional program at each speaker location to the final response to the test signal following the sound scene presentation in the automatic program. To simplify analysis, overall hearing aid output levels rather than spectral analyses were used.

Many features in hearing aids beyond the basic amplification system can affect gain. For example, all types of noise reduction features reduce gain. Automatic volume control features may either reduce or increase gain. The output of the hearing aid contains the net effect of these features and it is not possible to disentangle them in the automatic program where multiple features may be enabled simultaneously. In this paper, we call the sum of features affected by gain controlling features "gain agents."

The effect of gain agents was calculated by subtracting the final output level at 0° azimuth from the initial level at 0°.

To detect whether directional microphones were activated by the sound scene, and to get an impression of how strong the directional setting was, we defined a directional benefit based on comparing the hearing aid output at 0° azimuth with the output at 45° off-axis in the omnidirectional versus the automatic program. The directional benefit was calculated for each ear side as follows:

**Right side** = (final level at 0°- final level at 45°)
- (initial level at 0°- initial level at 45°)
**Left side** = (final level 0° - final level 315°)
- (initial level 0°- initial level 315°)

The reason for using the closest speaker locations at 45° and 315° as the basis for the calculations was to better differentiate strong directionality with narrower beamwidths. This method, plus taking the overall output level across the bandwidth above 100 Hz makes the directional benefit values shown in this paper smaller than would be typical for conventional measures of directionality such as Directivity Index or Front-Back Ratio. Therefore, the directional benefit we report here should not be interpreted to align with perceptual speech-in-noise test results with any of the tested hearing aids.

## Determining accuracy

Clinical guidelines for management of adults with hearing loss provide clear guidance in fitting and verifying hearing aids in accordance with a prescription for gain-frequency response and output. They also provide recommendations for hearing aid selection that include directional microphones, noise reduction, and other advanced sound processing features.[7,8] What they don't provide is guidance on how advanced hearing aid features should be applied or adjusted. Hearing Care Professionals tend to rely on hearing aid manufacturer defaults and recommendations regarding how to fine-tune or adjust settings for sound processing features other than gain.[9] The lack of specific guidance and clinical consensus in this area makes sense considering the variety and complexity of what is commercially available in hearing aids and how quickly technology updates are introduced. The criteria for accuracy were therefore defined according to evidence-based expectations to how gain agents and directionality would activate automatically in response to the different sound scenes.

There is a copious body of literature demonstrating that directionality helps speech understanding in noisy situations,[10,11] as well as literature supporting the use and benefit of digital noise reduction in noisy scenes.[10,12] In addition to benefit, some studies indicate differing preferences for gain, noise reduction and directionality in varying listening environments and listening contexts.[13-15] Reference to this literature allowed us to define the generous criteria outlined in Table 2. Our intention was to increase the fairness of the evaluation by setting a low bar for what the hearing aids would be expected to apply.

| | Voices only | Speech in noise | Noise |
|---|---|---|---|
| **Gain agents** | None or mild increase | Mild or more decrease | Moderate or more decrease |
| **Directional benefit** | None or mild | Moderate or more | Moderate or more |

Table 2. Criteria for accuracy of feature activation

We defined the specific criteria in determining accuracy as follows:
**None = 0**
**Mild = 1-2 dB**
**Moderate = 2-3 dB**
**Strong = >3 dB**

The rationale for assessing accuracy in each sound scene type considered the specific setup including direction of target speech and levels of speech and noise. For voices against a quiet background, audibility and natural sound quality are important. It is unnecessary to provide noise management features except to reduce low-level background or device internal noise. Very mild directionality to replicate pinna cues can be desirable but otherwise directionality disturbs the spatial hearing cues that contribute to naturalness. For the noise only scenes – especially when the noise is loud as in the current study - listening comfort is paramount. Noise management features that are strong enough to provide a noticeable decrease in the bothersome noise matter. Finally, the speech-in-noise scenes as presented in this study require an enhanced focus on the target speech in front while also adding to listening comfort with decreased gain for the noise. The decrease in gain was expected to be less than for the noise scenes for some brands depending on the complexity of their digital noise reduction features. Some noise reduction features may disengage when speech is present while others act simultaneously with speech.

For each type of sound scene and for all sound scenes combined, accuracy was determined simply by counting the times that the feature was activated as expected, dividing by the total number of sound scenes (of each type and in total) and multiplying by 100 to give a result in percent. For example, there was only one voices only sound scene, because previous datalogging studies have demonstrated that speech in quiet is identified correctly with at or near 100% accuracy across hearing aid brands. Thus, there were only 3 possible outcomes for voices only: 0% (neither gain agents nor directional benefit criteria met), 50% (either gain agents or directional benefit criteria met) and 100% (both gain agents and directional benefit criteria met.) There were 2 noise scenes which meant there were twice as many possible outcomes as for voices only. And for speech in noise, there were 5 sound scenes, increasing the possible accuracy outcomes fivefold.

## Results and Discussion

The overall accuracy of feature engagement per hearing aid brand is shown in Figure 1 while the 3 types of sound scenes are broken down in terms of accuracy in Figure 2. The AcoustIQ system was overall the most accurate with 88% better accuracy than the best performing other brand. Results reflected those of earlier studies of environmental classifiers based on datalogging where quiet or voices only environments were identified with the greatest accuracy. Like in those studies, activation of features was the most accurate for the voices only sound scene. Three brands, including the AcoustIQ-driven Beltone Illuminate, showed 100% accuracy for voices only. Results with AcoustIQ were in fact at 100% accuracy for all 3 sound scene types. Performance in feature activation for noise and speech in noise sound scenes for the other brands was mixed and without a consistent pattern.

These results suggest that environmental classification as indicated by datalogging doesn't tell the whole story. In other words, knowing what type of acoustic environment the hearing aid system identified doesn't necessarily align well with how you might believe features are applied in an automatic program. This is the same conclusion drawn by Husstedt et al[5] and reinforces the importance of accuracy both in environmental classification as well as automatic decision-making on application of features.

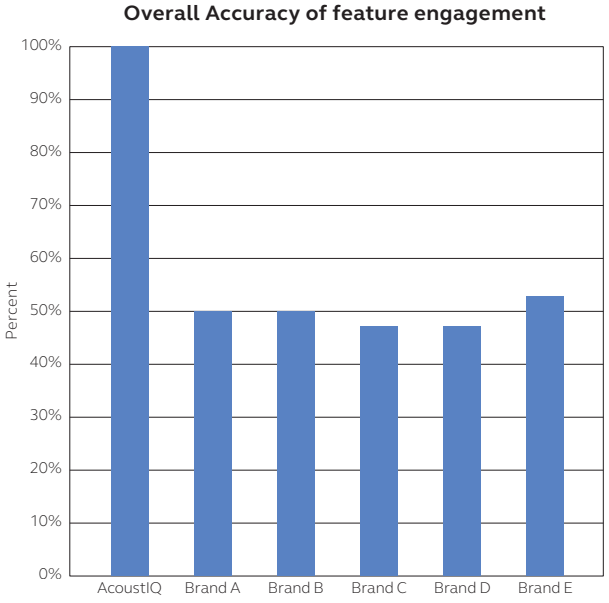



Figure 1. Overall accuracy of engagement of directionality and gain agents across hearing aid brands.

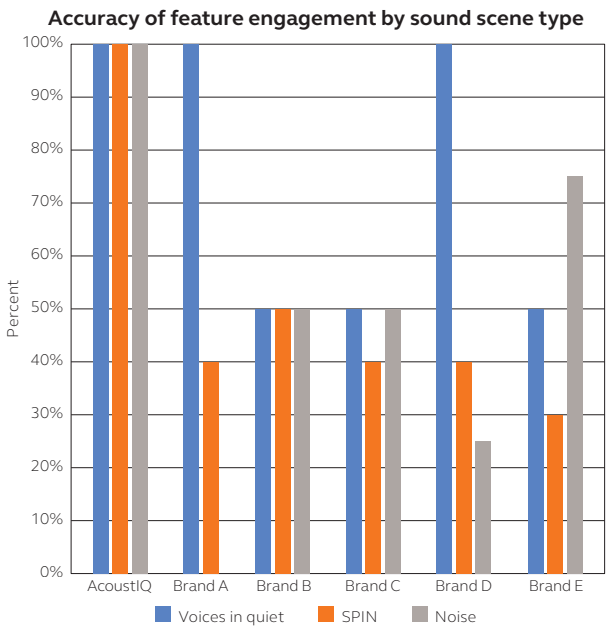



Figure 2. Accuracy of feature engagement according to the type of sound scene for each brand.

Figure 3 shows the average directional benefit and effect of gain agents per brand for each category of sound scenes. The adaptations in the voices only scene were smaller than in the other types of scenes although a few results were unexpected considering that this scene is easy for all systems to classify. For example, Brand B showed a dramatic volume increase exceeding 6 dB in the overall output which might be jarring for some users in an automatic program. In contrast, Brand C significantly reduced overall volume which could negatively impact audibility. While either of these adaptations might be preferred by individual hearing aid users in specific contexts, the job of an automatic program is to satisfy user needs and preferences broadly. Another unusual outcome was that Brand E appeared to be applying directionality in the voices only scene that was as strong or stronger than in the scenes with noisy backgrounds. And finally, Brand A was assessed at 0% accuracy for the Noise scenes. It is worth noting that this is one of the brands mentioned previously that had consistently classified the "Vacuum cleaner" sound scene as music in a previous study. If we look at the directional benefit and gain agent results for each of the noise scenes separately for Brand A, it is likely that music enhancing settings were applied to the vacuum cleaner sound scene in the current study. For the hand mixer scene, Brand A showed a mild directional benefit and a mild gain decrease while for the vacuum cleaner scene, there was no directional benefit and a mild gain *increase*.

Overall, the results of this study mirror those of earlier environmental classifier accuracy analyses. Earlier studies showed that the more complex the acoustic environment, the less likely that the environment is classified consistently with human judgments. The current work adds to this by demonstrating that automatic application of features does not always meet expectations based on the acoustic environment.

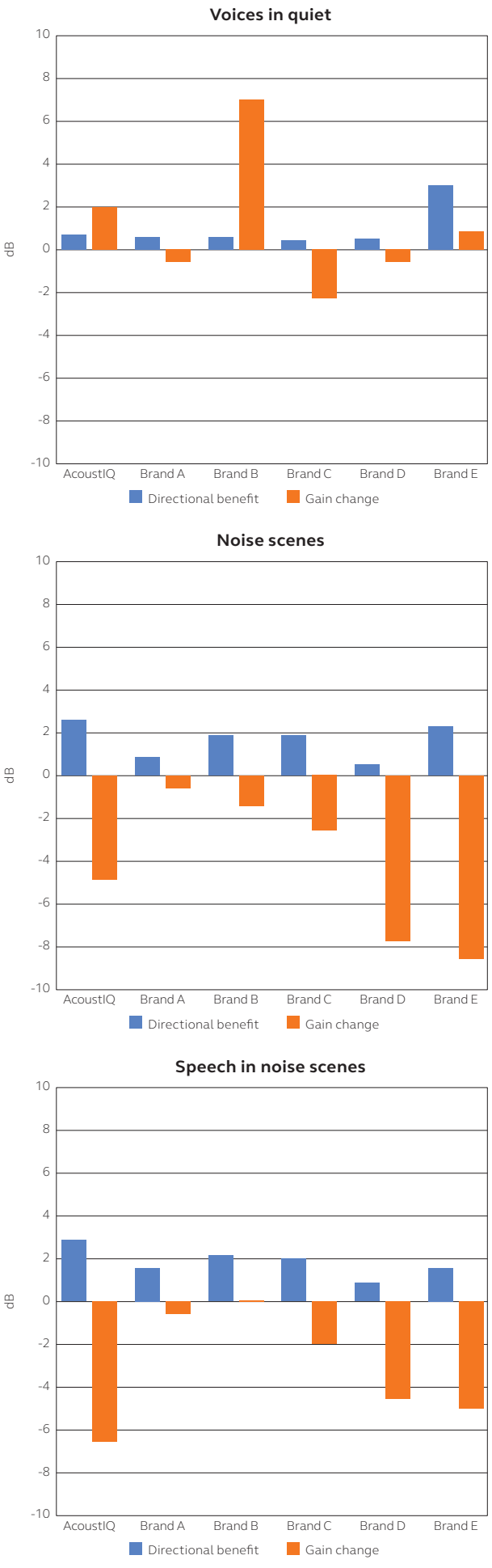



Figure 3. Average directional benefit and effects of gain agents for the three sound scene types per brand.

## Conclusion

Results of this study support that accuracy of environmental classification does not guarantee automatic control of advanced features that aligns with evidence-based expectations. Just as automatic features in a car should not unduly alarm, restrict or endanger drivers in getting where they need to go, automatically applied features in hearing aids should provide benefit while not distracting the hearing aid user and allowing them to maintain control of what they want to listen to. Both environmental analysis and the automatic control are critical in achieving this. The Beltone AcoustIQ system for environmental classification and feature application is highly accurate and more accurate than the systems in other premium hearing aids. This allows automatic control of advanced processing in Beltone hearing aids that supports hearing aid users in their daily listening environments rather than taking control from them.

## References


1. de Graaff F, Huysmans E, Ket JC, Merkus P, Goverts ST, Leemans CR, Smits C. Is there evidence for the added value and correct use of manual and automatically switching multimemory hearing devices? A scoping review. International Journal of Audiology. 2018 Mar 4;57(3):176-83.

2. Cui T, Groth J. How accurate are environmental classifiers in hearing aids?. Retrieved from AudiologyOnline. 2017.

3. Groth J, Kirkwood B, Cui T, Hougaard O, Dieu L. Empowering hearing aid users with directional technology designed for the real world. Beltone white paper. 2023.

4. Yellamsetty A, Ozmeral EJ, Budinsky RA, Eddins DA. A comparison of environment classification among premium hearing instruments. Trends in Hearing. 2021 Mar;25:2331216520980968.

5. Husstedt H, Wollermann S, Tchorz J. A method to analyse and test the automatic selection of hearing aid programs. In Proceedings of the International Symposium on Auditory and Audiological Research 2017 Dec 14 (Vol. 6, pp. 143-150).

6. Welling J, Krug AE, Fischer RL, Schinkel-Bielefeld N. What helps or hinders hearing aid wearers to change hearing aid programs. Jahrestagung der Deutschen Gesellschaft für Audiologie. 2020.

7. Valente M. Guideline for audiologic management of the adult patient. Audiology Online. 2006.

8. Hearing Aid Fitting for Adult and Geriatric Patients. Audiology Practice Standards Organization. 2021. Available at https://www.audiologystandards.org/standards/current.php.

9. Anderson MC, Arehart KH, Souza PE. Survey of current practice in the fitting and fine-tuning of common signal-processing features in hearing aids for adults. Journal of the American Academy of Audiology. 2018 Feb;29(02):118-24.

10. Bentler RA. Effectiveness of directional microphones and noise reduction schemes in hearing aids: A systematic review of the evidence. Journal of the American Academy of Audiology. 2005 Jul;16(07):473-84.

11. Jespersen CT, Kirkwood BC, Groth J. Increasing the effectiveness of hearing aid directional microphones. In Seminars in Hearing 2021 Aug (Vol. 42, No. 03, pp. 224-236). Thieme Medical Publishers, Inc..

12. Lakshmi, M.S.K., Rout, A. and O'Donoghue, C.R., 2021. A systematic review and meta-analysis of digital noise reduction hearing aids in adults. Disability and Rehabilitation: Assistive Technology, 16(2), pp.120-129.

13. Walden BE, Surr RK, Cord MT, Dyrlund O. Predicting hearing aid microphone preference in everyday listening. Journal of the American Academy of Audiology. 2004 May;15(05):365-96.

14. Walravens E, Keidser G, Hickson L. Consistency of hearing aid setting preference in simulated real-world environments: Implications for trainable hearing aids. Trends in Hearing. 2020 Jun;24:2331216520933392.

15. Christensen JH, Rumley J, Gil-Carvajal JC, Whiston H, Lough M, Saunders GH. Predicting individual hearing-aid preference from self-reported listening experiences in daily life. Ear and Hearing. 2024 Sep 1;45(5):1313-25.

Worldwide Headquarters
Beltone A/S
Lautrupbjerg 7
DK-2750 Ballerup
Denmark
Tel.: +45 4575 1111
beltone.com

CVR no. 55082715

© 2024 GN Hearing A/S. All rights reserved. Beltone is a trademark of GN Hearing A/S.

M201751GB-26.06-Rev.A