Improving hearing-aid gains based on automatic speech recognition

DOI:
10.1121/10.0001866

Document Version
Final published version

Link to publication record in Manchester Research Explorer

Citation for published version (APA):
Fontan, L., Le Coz, M., Azzopardi, C., Stone, M., & Füllgrabe, C. (2020). Improving hearing-aid gains based on automatic speech recognition. The Journal of the Acoustical Society of America, 148(3), EL227-EL233. https://doi.org/10.1121/10.0001866

Published in:
The Journal of the Acoustical Society of America

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
1. Introduction

Hearing-aid (HA) fitting is typically a two-stage process during which an initial gain fitting is applied according to a proprietary or device-independent method, followed by behavioral tests with the patient to fine-tune the fitting with the aim of maximizing speech intelligibility and listening comfort.

Fitting methods usually prescribe frequency-dependent amplification on the basis of the patient’s audiometric thresholds. These target gains generally represent a compromise between providing enough amplification to restore audibility and taking into account the patient’s reduced dynamic range in order to avoid presenting sounds at uncomfortable levels. Some fitting methods, such as DSL V5 (Scollie et al., 2005), are based on a loudness-normalization (LN) rationale, aimed at restoring the loudness perception of the patient to that of a normal-hearing listener. Other methods, such as NAL-NL2 (Keidser et al., 2011) and CAM2 (initially referred to as CAMEQ2-HF; Moore et al., 2010b), are based on a loudness-equalization (LE) rationale, aimed at equalizing loudness across frequency bands while maintaining the overall loudness at a comfortable level. While LN-based methods aim at maximizing sound quality and comfort, LE-based methods generally favor speech intelligibility. For example, CAM2 takes into account the general shape of the long-term average speech spectrum (Moore et al., 2008) and defines gains so that, for a given overall loudness, as much as possible of the speech signal falls above the listener’s absolute hearing thresholds. In contrast, the NAL-NL2 gain-prescription rule was defined by using an adaptive computer-controlled process to find optimal gains in terms of loudness and speech intelligibility, as measured by the Speech Intelligibility Index (ANSI, 1997).

Even though initial fittings based on these gain-prescription rules lead to satisfactory outcomes (e.g., Moore and Füllgrabe, 2010; Valente et al., 2018), further fine-tuning is generally required to achieve best results for a given HA user (Moore, 2007). This often involves the use of speech intelligibility tests, which can be time-consuming and, thus, make it impossible to evaluate
many different HA settings (e.g., gain functions) to find their optimal combination in terms of speech intelligibility and listening comfort.

Automatic speech recognition (ASR) techniques represent a novel approach that, without being limited by the patient’s time, motivational and attentional levels, and prior knowledge of the speech material, allows testing of a large number of combinations of HA gains in order to find the gain function yielding the highest speech intelligibility for a given audiometric profile. Moreover, contrary to other models used to evaluate HA settings,1 ASR does not require task-specific reference signals or conditions [see Kollmeier et al. (2016) for a more detailed discussion].

Recent studies demonstrated that ASR can be used to predict trends in speech intelligibility in listeners with age-related hearing loss (ARHL; Fontan et al., 2020; Kollmeier et al., 2016; Schädler et al., 2018). For example, Fontan et al. (2020) used a speech-intelligibility prediction system consisting of a signal-processing algorithm to simulate the effects of ARHL and an ASR system to predict the unaided identification performance of 24 listeners with ARHL for different speech materials (logatoms, words, and sentences). Based on each listener’s audiogram, the speech materials were processed to simulate three perceptual consequences of ARHL (elevation of hearing thresholds, loss of frequency selectivity, and loudness recruitment), and the ASR system was used to compute intelligibility scores. Strong to very strong correlations were observed between human and ASR speech-intelligibility scores.

The aim of the present study was to explore if the same speech-intelligibility prediction system could be used to improve HA gains for individual listeners. This was done by first generating for each listener a number of alternative gain functions based on the gain fitting rule CAM2 (i.e., variations around the CAM2 prescription), amplifying speech tokens using those functions, and determining their intelligibility using ASR. Note that the ASR system normalized the intensity of the input signals as a function of their maximum energy, with the energy being defined as the sum of the square magnitude of the signal contained in the 16-ms analysis windows used by the ASR system. Consequently, ASR performance was not affected by the changes in overall intensity that occur when switching between different gain functions. The gain function yielding the highest intelligibility score (henceforth referred to as “OPRA” gains, where OPRA stands for “objective prescription rule based on ASR”) and CAM2 gains were then used to produce amplified speech tokens that were presented to participants with ARHL for identification. Subjective ratings of speech pleasantness were also collected to verify that potential improvements in terms of speech intelligibility would not be obtained at the expense of reduced perceived quality of the speech signals.

2. Method

2.1 Participants

Twenty-four (10 female) right-handed, older (age range: 63–82 yr; mean age = 72.3 yr; standard deviation, SD = 5.5) native French speakers were recruited from a HA dispensing center in Montauban (France). Participants were selected for (i) being experienced HA users, having been fitted (for the most part according to the NAL-NL2 fitting rule; Keidser et al., 2011) in the test (i.e., right) ear for at least 7 months (on average 3 yr, 1 month), and wearing their HAs for at least 5 h per day (on average 10.2 h according to data logging); (ii) having symmetrical (≤ 10 dB interaural difference in PTA for frequencies of 0.5, 1, 2, and 4 kHz) sloping high-frequency mild-to-moderate hearing losses, not exceeding 75 dB hearing level (HL) at 4 kHz in the test ear; and (iii) having air-bone gaps ≤ 10 dB at 0.5, 1, 2, and 4 kHz in the test ear. Air- and bone-conduction audiometry was conducted with calibrated standard audiometric equipment in a sound-treated room in the HA dispensing center. Air-conduction pure-tone audiograms are shown in Fig. 1.

None of the participants reported any neurological disorder, and all passed the Cognitive Disorders Examination (CODEX), a validated short cognitive screen for dementia (Belmin et al., 2007). Prior to participation in the 1-h test session, each participant provided informed written consent. The study was approved by the ethical committee of the Honoré Cave Hospital (Montauban, France).

2.2 Stimuli and procedure

Speech intelligibility was assessed in quiet for the two speech materials most frequently used by HA audiologists in France for speech audiometry (Rembaud et al., 2017):

(i) Words consisting of disyllabic masculine nouns, each preceded by the French masculine definite article (e.g., “le soldat” - “the soldier”; Fournier, 1951).
ii) Sentences taken from the French version of the Hearing in Noise Test (HINT; Vaillancourt et al., 2005) using simple words and a single assertive clause (e.g., “Le camion est rouge.” - “The truck is red.”). Stimuli were calibrated recordings distributed by the Collège National d’Audioprothèse (2007). They were spoken by a single adult male native French speaker and recorded using a 44.1-kHz sampling rate and 32-bit quantization. They were presented to the participant seated in the same sound-treated room as used for the audiometric assessment, using a laptop running a custom presentation program written in Python, a Presonus Audiobox 44 VSL external sound card (Baton Rouge, LA), and the right earpiece of Sennheiser HD650 headphones (Wedemark, Germany).

In the “unaided” condition, stimuli were presented unprocessed at 60 dB sound pressure level (SPL).

In the CAM2 condition, based on each participant’s audiogram and assuming an input level of 60 dB SPL, the gain for 12 center frequencies was computed using the CAM2B-v2 fitting software (Cambridge Enterprise, University of Cambridge, UK). Using a software loudness model of impaired hearing, the algorithm aims to produce a flattened specific loudness pattern over the frequency range from 0.5 to 4 kHz, which is the most important for speech perception (ANSI, 1997). For a detailed description, see Moore et al. (2010b). The gains were implemented in a HA simulator (Moore et al., 2010a), using five processing channels equally spaced on the ERB-number scale (Glasberg and Moore, 1990). The 12-point insertion gain response was implemented as a Finite Impulse Response filter before separation into channels. Two dynamic range compressors were implemented in series in each channel. The first performed the bulk of the prescribed wide dynamic range compression (WDRC) function, while the second was configured as a fast-acting limiter whose compression threshold tracked at a fixed offset above the running mean level measured in the first compressor. The limiter function was therefore only activated by occasional peaks (< 1% of the time) in the channel speech signal. The attack times were 200, 100, 100, 100, and 100 ms, while the release times were 2000, 1500, 1200, 1000, and 1000 ms for channels 1 to 5, respectively. The channel compression thresholds were 47, 44, 40, 35, and 45 dB SPL, respectively. For an input level of 60 dB SPL and speech in quiet, channel 5 would be activated mainly by the peaks within the channel. A small error in the programming of the HA insertion gain meant that the insertion gains used were suitable for a 55 dB SPL input. This meant that the actual replay level per channel was up to 3.3 dB higher than prescribed, depending on the channel compression ratio.

In the OPRA condition, for each participant, the 12 frequency-specific gains prescribed by CAM2 were split into four frequency ranges (0.125–0.5 kHz, 0.75–1.5 kHz, 2–4 kHz, 6–10 kHz). Within each frequency range, the CAM2 gains were systematically varied by 0, ±3, or ±6 dB and applied to each of the three frequencies in a given frequency range, implemented with smooth transitions between step changes in gain. This resulted in 625 (=5^4) gain functions, including the original CAM2 gain function. The compression ratios were identical for all gain functions and corresponded to those used in the CAM2 condition. Each of the gain functions was implemented in the HA simulator to amplify 50 disyllabic words (Fournier, 1951) that were not used in the identification task performed by the participants. The processed speech was
then used as input to a hearing-loss simulator implemented in MATLAB, mimicking some of the perceptual consequences of ARHL (Nejime and Moore, 1997). In the present study, only loss of audibility and loudness recruitment were simulated since the simulation of loss of frequency selectivity (which is also typically associated with hearing loss and leads to the “smearing” of spectral components) has been shown to lead to weaker associations between human and machine intelligibility (see Table III in Fontan et al., 2020). The severity of the degradation imposed by the hearing-loss simulator was dependent on the participant’s audiometric thresholds. The amplified and then degraded stimuli were finally fed into an ASR system run on the OSIRIM platform (http://osirim.irit.fr/site/en), which is a cluster of 928 central processing units and 28 graphical processing units. The acoustic and language models used and the general functioning of the ASR system are described in Fontan et al. (2020). The ASR system used a lexicon consisting of 6491 disyllabic words [corresponding to the “large” lexicon described in Fontan et al. (2020)]. The range of ASR performance (i.e., maximum word recognition rate to minimum word recognition rate) for all 625 gain functions tested for each participant was, on average, 33.7 percentage points (SD = 14.9). In each of the four frequency ranges used, increases in ASR performance were observed for both positive (i.e., +3 or +6 dB) and negative (i.e., –3 or –6 dB) gain changes. As expected, given the use of the intensity-normalization process in the ASR system, all uniform variations of CAM2 gain functions (i.e., –3, +3, –6, or +6 dB at all frequencies) yielded the same ASR performance as that obtained with the original CAM2 gain functions.

The amplified speech obtained using the gain function yielding the highest machine intelligibility was then presented to the older hearing-impaired (OHI) participants for identification. Figure 2 illustrates the absolute and relative differences between OPRA and CAM2 gains. Compared to CAM2 gains, OPRA tended to provide more amplification for frequencies up to and including 4 kHz (+5.0 dB for frequencies ≤ 0.5 kHz, and +2.9 dB for frequencies > 0.5 kHz and ≤ 4 kHz), while providing less amplification at higher frequencies (–3.4 dB for frequencies > 4 kHz).

Participants were tested in all three listening conditions (unaided, CAM2, and OPRA), using both sets of speech materials. For each set of speech material, participants first completed the test in the unaided condition and then in the two aided conditions. The participants were pseudo-randomly allocated to one of the four possible orders (2 aided conditions × 2 speech materials) to obtain the same number of participants (N = 6) for each order. Prior to data collection in each of the test conditions, practice was provided in the form of 10 words or 10 sentences that were not used in the test phase.

Participants reported back verbally the words they had heard and were instructed to guess in case they were uncertain. No feedback was provided. For both sets of speech materials, words correctly identified were scored manually by the experimenter. Word intelligibility in each listening condition was computed as the average identification performance for 50 words, corresponding to five of the 10-word lists (Lists 1–5, 7–11, or 13–17) developed by Fournier (1951). This number of words has been shown to yield satisfactory reproducibility of behavioural intelligibility outcomes (Moulin et al., 2016). Sentence intelligibility in each listening condition was computed as the average word-intelligibility score for one of the 20-sentence HINT lists (List 1, 2, or 3; Vaillancourt et al., 2005). Lists were counterbalanced across listening conditions.

At the end of each aided listening task, the participants were asked to rate the pleasantness of the speech tokens on a 11-point Likert scale, ranging from 0 (“Very unpleasant/artificial”) to 10 (“Very pleasant/natural”).
3. Results

The top panel of Fig. 3 shows intelligibility scores for words and sentences for each of the three listening conditions. In the unaided condition, word-intelligibility scores were widely distributed, ranging from 0 to 92% (median: 46%). Unaided sentence-intelligibility scores were higher (median: 68%), and their distribution was bimodal, with five participants scoring below 10% and the others scoring above 38% (maximum: 96.9%).

For both sets of speech materials, intelligibility increased markedly with HA amplification, leading to ceiling effects in the two aided conditions. Median scores were marginally higher for OPRA gains than for CAM2 gains: 97% vs 96%, and 99.5% vs 97.1% for words and sentences, respectively. According to Wilcoxon signed-rank tests, the difference between gain prescriptions was significant for words ($Z = 217, p = 0.003$; two-tailed) and sentences ($Z = 247, p = 0.001$; two-tailed).

The bottom panel of Fig. 3 shows the subjective ratings of speech pleasantness for words and sentences in the two aided conditions. Ratings were lower in the CAM2 condition (median: 7.0 for both words and sentences) than in the OPRA condition (median: 8.3 and 8.5 for words and sentences, respectively). According to Wilcoxon signed-rank tests, this difference was significant for words ($Z = 220, p < 0.001$; two-tailed) and sentences ($Z = 221, p = 0.002$; two-tailed).

4. Summary and discussion

This study provides evidence that a reference-free ASR system can be used to improve individually prescribed HA gains. Based on the audiograms of 24 OHI participants, two of the perceptual consequences of ARHL (elevation of hearing thresholds and loudness recruitment) were simulated, and an HA simulator was used to find the HA gain function yielding highest ASR performance for each participant. A possible concern was that any increase in the level at the output of the HA simulation would result in better ASR performance, by compensating for the “noise” induced by the ARHL simulation and thus improving the “signal-to-noise ratio.” This would have led to the maximization of the amplification gains (i.e., +6 dB at all frequencies) for each participant. However, uniform gain variations yielded the same outcome as the original CAM2 gain functions, showing that the ASR performance was not directly affected by the presentation level at the output of the HA simulation. Indeed, as the ASR system normalized the energy of the input signals, its performance was only affected by changes in the relative distribution of the signal energy as a function of frequencies, which impacts the probability for speech phones to be correctly identified. In this study, such spectral distortions included the effects of the ARHL simulation, as well as non-uniform changes in the HA gain functions aiming at (partially)
compensating these effects. As a result, the optimal variations of CAM2 gains found for each participant implied both gain increases and gain decreases in each of the four frequency ranges used in this study (i.e., 0.125–0.5 kHz, 0.75–1.5 kHz, 2–4 kHz, and 6–10 kHz). Significant difference in ASR performance was observed between OPRA gains and the gains prescribed by the CAM2 fitting rule (meanCAM2: 30.0%; meanOPRA: 43.3%; t[23] = 10.5, p < 0.001; two-tailed).

OPRA gains yielded significantly higher human word- and sentence-intelligibility scores than the gains recommended by the CAM2 fitting rule. Even though these differences were statistically significant, they were small, possibly due to near-perfect performance of most participants in both aided conditions. In future studies, ceiling effects in human performance could be avoided by using a lower presentation level, which, however, would be less representative of standard conversational levels. Employing an adaptive procedure to track speech recognition thresholds in quiet is not an option, given the use of signal-intensity normalization techniques in ASR systems, making them insensitive to the overall level of the speech signal. Another possible way to avoid ceiling effects is to use linguistically more challenging speech materials than those used in the present study, or to present the target speech in interfering background noise. The applicability of ASR to the prediction of speech-in-noise intelligibility is currently under investigation.

While the statistically significant improvements in speech intelligibility with OPRA gains are probably not clinically significant due to their small size, the simultaneously obtained pleasantness judgements revealed markedly higher preference scores for speech amplified according to OPRA. This can be explained by the fact that, compared to CAM2, OPRA called for less amplification at high (> 4 kHz) and more amplification at low (≤ 4 kHz) frequencies, which is known to lead to higher judgements of pleasantness (e.g., Moore et al., 2011).

It should be noted that in the current study the speech material was presented in quiet at a single level. Hence, it remains to be shown whether OPRA benefits would also be observed for other presentation levels for speech in quiet and for speech in noise.

Also, this preliminary study was limited to the comparison of CAM2 and OPRA gains. Further work is warranted to compare OPRA outcomes—in terms of both machine and human speech intelligibility—to the outcomes obtained with other speech-intelligibility prediction systems, which are often less complex, such as the Speech Intelligibility Index used to define NAL-NL2 prescription gains (Keidser et al., 2011).

It is possible that ASR-based fitting could be further improved by refining the gain prescription, using narrower frequency ranges, and varying the compression ratio in each of these ranges. The present study aimed to avoid excessive computational costs by testing a single set of fixed compression ratios and by splitting the 12 frequency-specific CAM2 gains into four frequency ranges that were systematically varied to determine OPRA gains. Increasing processing power and/or processing optimization would allow for the use of a higher number of frequency ranges and compression settings (i.e., compression ratios, attack and release times of the compressor), as well as longer speech items (ideally produced by different speakers of both genders), which might improve the certainty of OPRA settings and the generalizability of the associated beneficial outcomes. These benefits might be further increased by using other (e.g., more physiologically inspired) models of human peripheral auditory processing in the HL simulator and/or as the front end of the ASR system itself (e.g., Holmberg et al., 2007). Finally, the effects of normalizing input signals based on their loudness instead of their maximum energy should be explored.

In conclusion, the results obtained in this study provide proof of concept that ASR techniques can be used to improve, at no “experimental cost” to the HA user, the gains recommended by an established fitting rule, such as CAM2. Promisingly, this approach has led to the improvement of both speech intelligibility and speech pleasantness.

Acknowledgments

The experiments presented in this paper were carried out using the OSIRIM platform (http://osirim.irit.fr/site/en) that is administered by the Institut de Recherche en Informatique de Toulouse (IRIT) and supported by the French National Center for Scientific Research (CNRS), the Occitanie Region, the French Government, and the European Regional Development Fund (ERDF). The authors are indebted to Etienne Revol-Buisson for helping with participant recruitment and providing test facilities and audiometric equipment. The authors thank Professor Brian C. J. Moore and three anonymous reviewers for their helpful comments on previous versions of this manuscript. Some of the work was conducted while L.F. was a visiting fellow at the Institute of Advanced Studies at Loughborough University (UK). M.A.S. was funded by the National Institute for Health Research (NIHR) Manchester Biomedical Research Centre. C.F. acted as a scientific consultant for Archean Technologies.
References and links

1For an overview, see the special issue of the International Journal of Audiology on model-based concepts and assessment of HA technologies (2018, vol. 57 sup. 3).
2All participants wore receiver-in-the-ear HAs, bar one who wore behind-the-ear HAs.

ANSI. (1997). ANSI S3.5-1997, Methods for the calculation of the speech intelligibility index. (American National Standards Institute, New York).
Belmin, J., Pariel-Madjessi, S., Surun, P., Bentot, C., Feteana, D., Lefebvre des Noettes, V., Onen, F., Drunat, O., Trivallie, C., Chassagne, P., and Golmard, J.-L. (2007). “The cognitive disorders examination (Codex) is a reliable 3-minute test for detection of dementia in the elderly (validation study on 323 subjects).” Presse Med. 36(9), 1183–1190.
Collège National d’Audioprothèse. (2007). Précis d’audioprothèse—Tome II [Reference manual for hearing-aid specialists—Part II]. (Elsevier-Masson, Paris, France).
Fontan, L., Cretin-Maitezn, T., and Füllgrabe, C. (2020). “Predicting speech perception in older listeners with sensorineural hearing loss using automatic speech recognition,” Trends Hear. 24, 1–16.
Fournier, J. E. (1951). Audiométrie vocale: les épreuves d’intelligibilité et leurs applications au diagnostic, à l’expertise et à la correction prothétique des surdités [Speech audiometry: speech-intelligibility tests and their application for the diagnosis, survey and hearing-aid rehabilitation of hearing losses]. (Maloine, Paris, France).
Glasberg, B. R., and Moore, B. C. J. (1990). “Derivation of auditory filter shapes from notched-noise data,” Hear. Res. 47(1–2), 103–138.
Holmberg, M., Gelbart, D., and Hemmert, W. (2007). “Speech encoding in a model of peripheral auditory processing: Quantitative assessment by means of automatic speech recognition,” Speech Commun. 49(12), 917–932.
Keidser, G., Dillon, H., Flax, M., Ching, T., and Brewer, S. (2011). “The NAL-NL2 prescription procedure,” Audiol. Res. 1, 88–90.
Kollmeier, B., Schälsker, M. R., Warzybok, A., Meyer, B. T., and Brand, T. (2016). “Sentence recognition prediction for hearing-impaired listeners in stationery and fluctuation noise with FADE: Empowering the attenuation and distortion concept by Plomp with a quantitative processing model,” Trends Hear. 20, 233121651665579.
Moore, B. C. J. (2007). Cochlear Hearing Loss: Physiological, Psychological and Technical Issues, 2nd ed. (Wiley, Chichester, England).
Moore, B. C. J., and Füllgrabe, C. (2010). “Evaluation of the CAMEQ2-HF method for fitting hearing aids with multi-channel amplitude compression,” Ear Hear. 31(5), 657–666.
Moore, B. C. J., Füllgrabe, C., and Stone, M. A. (2010a). “Effect of spatial separation, extended bandwidth, and compression speed on intelligibility in a competing-speech task,” J. Acoust. Soc. Am. 128, 360–371.
Moore, B. C. J., Füllgrabe, C., and Stone, M. A. (2011). “Determination of preferred parameters for multichannel compression using individually fitted simulated hearing aids and paired comparisons,” Ear Hear. 32(5), 556–568.
Moore, B. C. J., Stone, M. A., and Füllgrabe, C. (2008). “Spectro-temporal characteristics of speech at high frequencies, and the potential for restoration of audibility to people with mild-to-moderate hearing loss,” Ear Hear. 29, 907–922.
Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (2010b). “Development of a new method for deriving initial fittings for hearing aids with multi-channel compression: CAMEQ2-HF,” Int. J. Audiol. 49(3), 216–227.
Moulin, A., Bernard, A., Tordella, L., Vergne, J., Gisbert, A., Martin, C., and Richard, C. (2016). “Variability of word discrimination scores in clinical practice and consequences on their sensitivity to hearing loss,” Eur. Arch. Oto-Rhino-L. 274, 2117–2124.
Nejime, Y., and Moore, B. C. J. (1997). “Simulation of the effect of threshold elevation and loudness recruitment combined with reduced frequency selectivity on the intelligibility of speech in noise,” J. Acoust. Soc. Am. 102(1), 603–615.
Rembaud, F., Fontan, L., and Füllgrabe, C. (2017). “L’audiométrie vocale en France: état des lieux,” Cahiers de l’Audition 6, 22–25.
Schälsker, M. R., Warzybok, A., and Kollmeier, B. (2018). “Objective prediction of hearing aid benefit across listener groups using machine learning: Speech recognition performance with binaural noise-reduction algorithms,” Trends Hear. 22, 233121651876895.
Scollie, S., Seewald, R., Cornelisse, L., Moodie, S., Bagatto, M., Laurnagaray, D., Beaulac, S., and Pumpford, J. (2005). “The Desired Sensation Level multistage input/output algorithm,” Trends Amplif. 9(4), 159–197.
Vaillancourt, V., Larocbe, C., Mayer, C., Basque, C., Nali, M., Erks-Brophy, A., ... Giguère, C. (2005). “Adaptation of the HINT (hearing in noise test) for adult Canadian francophone populations,” Int. J. Audiol. 44(4), 358–361.
Valente, M., Oeding, K., Brockmeyer, A., Smith, S., and Kaliogeri, D. (2018). “Differences in word and phoneme recognition in quiet, sentence recognition in noise, and subjective outcomes between manufacturer first-fit and hearing aids programmed to NAL-NL2 using real-ear measures,” J. Am. Acad. Audiol. 29(8), 706–721.