Original Article

A Trainable Hearing Aid Algorithm Reflecting Individual Preferences for Degree of Noise-Suppression, Input Sound Level, and Listening Situation

Sung Hoon Yoon[1],* · Kyoung Won Nam[2],* · Sunhyun Yook[3] · Baek Hwan Cho[4] · Dong Pyo Jang[2] · Sung Hwa Hong[5]
In Young Kim[1,2]

[1]Institute of Biomedical Engineering, Hanyang University, Seoul; [2]Department of Biomedical Engineering, Hanyang University, Seoul; [3]Department of Medical Device Management & Research, Samsung Advanced Institute for Health Science & Technology, Sungkyunkwan University, Seoul; [4]S/W Solution Lab, Samsung Advanced Institute of Technology, Yongin; [5]Department of Otolaryngology-Head and Neck Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

Objectives. In an effort to improve hearing aid users' satisfaction, recent studies on trainable hearing aids have attempted to implement one or two environmental factors into training. However, it would be more beneficial to train the device based on the owner's personal preferences in a more expanded environmental acoustic conditions. Our study aimed at developing a trainable hearing aid algorithm that can reflect the user's individual preferences in a more extensive environmental acoustic conditions (ambient sound level, listening situation, and degree of noise suppression) and evaluated the perceptual benefit of the proposed algorithm.

Methods. Ten normal hearing subjects participated in this study. Each subjects trained the algorithm to their personal preference and the trained data was used to record test sounds in three different settings to be utilized to evaluate the perceptual benefit of the proposed algorithm by performing the Comparison Mean Opinion Score test.

Results. Statistical analysis revealed that of the 10 subjects, four showed significant differences in amplification constant settings between the noise-only and speech-in-noise situation ($P<0.05$) and one subject also showed significant difference between the speech-only and speech-in-noise situation ($P<0.05$). Additionally, every subject preferred different β settings for beamforming in all different input sound levels.

Conclusion. The positive findings from this study suggested that the proposed algorithm has potential to improve hearing aid users' personal satisfaction under various ambient situations.

Keywords. Hearing Aid; Classification; Patient Preference; Digital Signal Processing

INTRODUCTION

Hearing impairments were suffered by 360 million people worldwide in 2012, comprising 328 million adults and 32 million children [1]. These numbers have been steadily increasing as the life expectancy increased with advancements in medicine and technology [2]. However, despite the overwhelming number of hearing-impaired people, market surveillance data show that only 1 in 5 adults have adopted hearing aids, and many hearing-

impaired persons still hesitate to purchase hearing aids due to the high-cost and low satisfaction in various noisy situations [3].

When purchasing hearing aids, many people are unsatisfied with the settings prescribed by their audiologists because adjustments made in a quiet clinic environment are often not optimal for real-life use [4-6]. People new to hearing aids may not be aware of what an optimal setting is for them, which would affect the effectiveness of the initial adjustment at the hospital [7]. In such cases, the device user has no choice but to revisit the hospital to readjust the device settings as many times as needed until satisfied with the tuning. However, if there is no relevant hospital or tuning clinic near the user's home or office, repeated visits can incur large time and financial burdens [8]. In addition, due to differences in personal preferences, a setting suitable for one environmental situation may not be suitable for others [5]. Some people prefer to have the directional microphone function enabled when listening to speech sounds in noisy environments, but disabled when listening to music in a music hall or speech sounds in a quiet room. Others prefer to leave the function enabled all the time, regardless of the ambient situation. Additionally, many prefer to adjust the level of noise suppression depending on the variations in the ambient noise, whereas others do not [4]. Further, it has been reported that audiologist-driven settings and patient-driven settings can be quite different [9,10]. For these reasons, a need exists for a hearing-support device that can auto-tune its internal parameters according to the user's personal preferences under various ambient situations, and to meet this need, a trainable or self-fitting hearing aid was introduced [7].

In the initial stage of research on trainable hearing aids, the main focus of training was only the relationship between the intensity of ambient sounds (input sound level [ISL]) and the output volume of the device (amplification constant [AMP]) [7,9,11,12]. Later research revealed that many other factors besides the ISL affects human perception, such as listening situation (LS), type of ambient sound, and internal device factors such as the degree of noise suppression (DNS) [7]. For example, Chalupper et al. [4] demonstrated that the settings of a hearing aid, such as overall gain, compression parameters, noise-reduction strength, and activation/deactivation of the directional microphone, may be sensitive to personal preferences. However, even the recent studies on trainable hearing aids have reflected only one or at most two of these environmental factors into the

device training [4,6,13,14], which limited the user satisfaction under various acoustic conditions. In other words, personal satisfaction of the user would be greatly improved if the device is able to reflect the owner's personal preference in a more expanded environmental acoustic conditions, since the auditory perception is affected by surrounding complex acoustic conditions.

In this study, we propose a trainable hearing aid algorithm that reflects the user's personal preferences in three environmental acoustic conditions (ISL, LS, and DNS), and evaluate the performance of the proposed algorithm using 10 normal-hearing subjects.

MATERIALS AND METHODS

Algorithm implementation scenario

The scenario of the implementation of the proposed algorithm is as follows. Normally, for everyday or first-time use, the user sets the device to normal-operation mode, in which the device adjusts the internal parameters automatically while reflecting the user's preference. Upon encountering an uncomfortable or unoptimized sound in a new environment, the user can train the device by selecting one or both of the two training modes, volume-training and β-training. In volume-training mode, the user can make manual adjustments to the overall output volume via the user interface to add or subtract a certain offset (in 1 dB steps) from the current volume to select an amplification (AMP, dB) value. In β-training mode, the user can select a β value of the beamforming algorithm that fits his or her personal preference. Once training is complete for a specific environment, the user can store the preferred internal parameters selected during training–the outputs of the LS classifier and the ISL detector, which describes the specific environment training was performed in, and the user-selected AMP and β values which are the values the user selected during training–in the database. The algorithm utilizes data gathered during training to better auto-tune to the user's preference. If training is performed in an environment that was trained before (i.e., for the same LS and ISL), the most recent adjustment overrides the previous data. The algorithm in normal-operation mode uses the user-trained database and predetermined interpolation and extrapolation protocols to automatically adjust the AMP, β, and activation or deactivation of beamforming in real-time.

Structure and operation of the implemented algorithm

Fig. 1 shows the overall schematic of the proposed algorithm. It consists of three sub-blocks: a signal interface block, a user interface block, and a database manager block. The following assumptions were made for simplification: (1) The device contains two user-adjustable algorithms–beamforming, which adjusts the DNS using the β value, and an output volume adjustment, which adjusts the total output volume of the device using AMP;
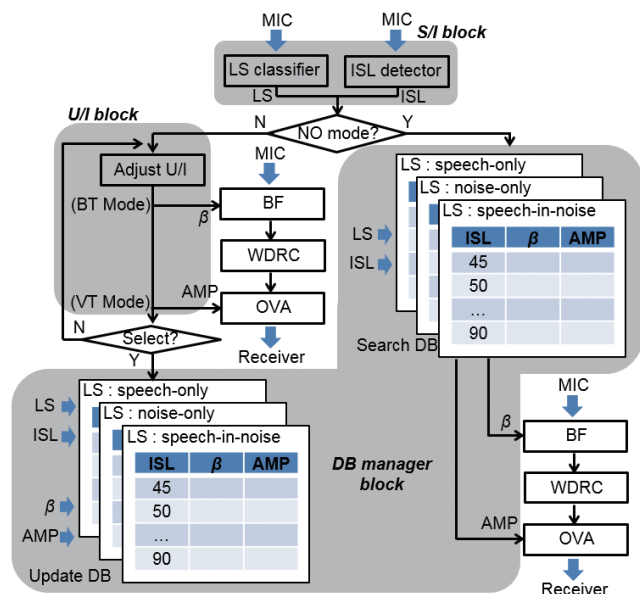
Fig. 1. Overall schematic of the proposed algorithm for a trainable hearing aid that can reflect the variations in all of the ISL, LS, and DNS conditions. MIC, microphone; S/I, signal interface; U/I, user interface; LS, listening situation; ISL, input sound level; DNS, degree of noise suppression; NO, normal operation; BF, beamforming; WDRC, wide dynamic-range compression; OVA, output volume adjustment; AMP, amplification constant in OVA; DB, database.

(2) The wide dynamic-range compression is always active and its channel gains are fixed at purchase; (3) Beamforming is enabled only in speech-in-noise situations and disabled in speech-only and noise-only situations; and (4) The ISL is limited to a 45–90 dB of sound pressure level (dB SPL) range, considering the normal, everyday use presented by Chalupper et al. [4] and Walden et al. [15]. In this study, we used the differential microphone array algorithm suggested by Teutsch and Elko [16] as the beamforming, and the neural network-based LS classifier suggested by Yook et al. [17] that was slightly modified as follows: (1) Additional amplification in front of the classifier to ensure the proper ISL for stable classification performance and (2) Among the four situations classified in the original LS classifier, only three situations–speech-only, noise-only, and speech-in-noise–were enabled (i.e., music-only was disabled).

When external sounds enter the embedded microphones, the average ISL of the most recent 2.5 seconds (which was determined to be the most optimal with respect to accuracy and ability to operate in real-time through repetitive trials) is measured by the ISL detector, and at the same time, the type of ambient situation is determined by the listening situation classifier. When the device is turned on (i.e., proposed algorithm is initiated), initial operation mode is set to normal-operation mode. When the user wants to further train the device, the user needs to change the operation mode to either volume-training or β-training mode. (In this study, we assumed that the user has the means to

freely adjust the operation mode, the AMP, and the β; e.g., a three-step slider and a clickable dial on the surface of the device that enables the user to freely adjust the device setting manually during the training phase, a dedicated remote controller connected to the hearing aid wirelessly that allows manual adjustment, or a dedicated computer or smartphone application for this purpose). In volume-training mode, users can adjust the AMP to their preferences in different ambient listening situations in their normal daily lives. When an adjustment is made, the device records the value of the new AMP, the ambient situation detected by the listening situation classifier, and the ambient sound level detected by the ISL detector into the database. In β-training mode, users can select the DNS to their liking by adjusting the β among the three preset values (1.76, 2.65, and 3.47) [18]. As in volume-training mode, the β value is recorded into the database along with the current LS and ISL values. When reverted back to normal-operation mode after training, the overall training-related functions are disabled, and the device automatically adjusts its internal parameters (AMP, beamforming on/off, and β) in accordance with the training data. In order to perform auto-tuning in normal-operation mode, the proposed algorithm monitors the outputs of the ISL detector and LS classifier in real time to search the database for a matching condition. If an exactly-matched condition exists, the stored value of the parameters are used to auto-tune; however, if one does not exist, (1) the AMP is determined by linear interpolation and extrapolation using the neighboring values in the database, and (2) β is determined by selecting the closest value between the two adjacent values in the database. Since parameter outputs are very sensitive to ISL, preventative measures are taken in normal-operation mode to ensure correct parameter outputs are implemented when the listening environment changes abruptly. Specifically, an ISL change of more than 5 dB is considered abrupt, and for such instants, implementation of the auto-tune parameter output is skipped one interval.

Subject recruitment and test sound preparation for evaluation

Ten normal-hearing subjects (7 males and 3 females; age range, 21 to 30 years; median age, 25 years; mean age, 25.5 years) were recruited in accordance with the method approved by the Institutional Review Board of Hanyang University (HYI-15-011-3). Written consents to participate were acquired from all participants before the experiments, and they were compensated for their participation.

The sound files for evaluation were recorded using a combination of 40 clean speech signals, 30 babble noise signals, and 30 car noise signals [19]. Three types of sound files were recorded to simulate the three listening situations: speech-only, noise-only, and speech-in-noise. The recording took place in a nonreverberant chamber located at Hanyang University using a KEMAR mannequin (Type 45BA; G.R.A.S. Sound & Vibration, Holte, Denmark) with two artificial ears (Type KB0060; G.R.A.S.

Sound & Vibration) placed in the center of the chamber. Two behind-the-ear type hearing aid mock-ups, shaped identically to the commercialized device (Canta7; GNReSound A/S, Ballerup, Denmark), were mounted on the two artificial ears to record the sound in separate files. Four identical speakers (HS50M; Yamaha Co., Hamamatsu, Japan) were each positioned 120 cm from the mannequin, surrounding it, in the 0°, 90°, 180°, and 270° directions. The four speakers were calibrated before the recording using a commercial sound-level meter (2250 Light; Brüel & Kjær Co., Naerum, Denmark) to make sure the outputs were the same for all the speakers. The originating direction of sound during the recording for each listening situation were as follows: (1) speech-only: clean speech from 0°; (2) noise-only: identical noises from 0°, 90°, 180°, and 270°; and (3) speech-in-noise: clean speech from 0° and identical noises from 90°, 180°, and 270°. The input signal-to-noise ratio of the mixed signals was fixed at 0 dB. For the recordings of each of the three situations, the overall ISL was adjusted from 45 to 90 dB in 5 dB steps. A total of 12 test sets were recorded for algorithm verification (PS_1 and PS_2), subjective training (TS_1), and subjective evaluation (TS_2) as listed below (10 sound clips per set; no duplication or overlap between CS_1–CS_5, BN_1–BN_3, and CN_1–CN_2).

- PS_1 contained one test set consisting of seven 1,000-Hz pure tone signals with 60, 65, 70, 75, 80, 85, and 90 dB SPL.
- PS_2 contained one test set consisting of six 1,000-Hz pure tone signals with 63, 67, 73, 77, 83, and 87 dB SPL.
- TS_1 contained three test sets, each consisting of one clean speech set (CS_1) and one babble noise set (BN_1): one speech-only set (ES_1), one noise-only set (BN_1), and one speech-in-noise set (CS_1+BN_1).
- TS_2 contained seven test sets, each consisting of three clean speech sets (CS_2, CS_3, and CS_4), two babble noise sets (BN_2 and BN_3), and two car noise sets (CN_1 and CN_2): one speech-only set (CS_2), two noise-only sets (BN_2 and CN_1), and two speech-in-noise sets (CS_3+BN_3 and CS_4+CN_2).

Evaluation of the implemented algorithm

Before evaluating the perceptual benefit of the proposed algorithm, we verified the proper operation of the implemented blocks of the algorithm: first, the performance of the ISL detector was evaluated with PS_1; second, the interpolation performance of the database manager was evaluated by training the algorithm with PS_1 in volume-training mode and evaluating the algorithm using PS_2 in normal-operation mode; and lastly, the performance of the LS classifier was evaluated using both TS_1 and TS_2.

Next, we evaluated the performance of the overall algorithm: first, an examiner trained the algorithm with TS_1 and adjusted the AMP and β values to his preferred setting for each test sound (each sound was repeatedly played until the examiner completed training). Then, the algorithm was switched to normal-operation mode, and TS_1 and TS_2 signals were sequentially entered into the signal interface block in random order, and the automatically-adjusted AMP and β values were recorded. Each sound clip was repeated until five parameter outputs were recorded, and then the next clip was played.

After the overall verification process of the algorithm was completed, the Comparison Mean Opinion Score (CMOS) test [20] was performed with 10 subjects to compare the differences in perceptual benefits among the three different parameter settings.

- Fixed setting (T1 setting; representing a traditional hearing aid): fixed amplification constant and β value regardless of the variations in input signals (both the ISL detector and LS classifier were disabled).
- User-trained setting without LS classifier (T2 setting; representing a traditional trainable hearing aid): automatically adapting the user-defined amplification constants and β values to the variations in the input sound level (only the ISL detector was enabled, and the LS classifier was disabled).
- User-trained setting with LS classifier (T3 setting; representing the proposed algorithm): automatically adapting the user-defined amplification constants and β values to the variations in both the input sound level and the listening situation (both the ISL detector and the LS classifier were enabled).

TS_2 sounds were utilized to record test sounds for the CMOS test, and wide dynamic-range compression in Fig. 1 was disabled (bypassed) during recording because the test was performed using normal-hearing subjects only.

To record test sounds for the CMOS test, the algorithm must be trained prior to the recording as the recording would be done with the auto-tuned internal parameters based on the trained data. Each subject was given thorough instructions regarding the experiment before being positioned in the middle of the nonreverberant chamber wearing a headset. Training for T1, T2, and T3 setting differed from each other: for the T1 setting: (1) Training-related algorithms were all disabled; (2) The β value was preset to 3.1 [18]; (3) Each subject listened to a 1,000-Hz pure tone signal through the headset (70 dB SPL) and was asked to manually adjust the AMP to his or her preferred sound level (comfortable level); and (4) The subject-selected AMPs were recorded for all subjects. For the T2 setting: (1) The LS classifier was disabled (and the database manager was also adjusted to accommodate the T2 setting); (2) Each subject was asked to train (i.e., manually adjust the AMPs and β values to his or her preferred setting for each sound) the algorithm using TS_1; and (3) The subject-selected parameters for each sound were recorded for all subjects. For the T3 setting: (1) All algorithms in Fig. 1 were enabled; (2) Each subject was asked to train the algorithm using TS_1; and (3) The subject-selected parameters for each sound were recorded for all subjects. During training, each sound clip was repeated until the subject selected his or her preferred parameter

setting. After the training phase, the subject was given a break while the examiner recorded test sounds (output signals of the algorithm in normal-operation mode when TS_2 signals were entered into the algorithm with the recorded parameter settings for each subject) for the CMOS test.

After recording the test sounds, the subject returned to the original position in the chamber and wore the headset again to proceed with the CMOS test. During the CMOS test, identical test sounds with two different settings (either T1 and T3, or T2 and T3) were played sequentially, and the subject was asked to score the relative quality of the second sound (comparison audio) compared to the first sound (reference audio), in the range of –3 to +3. For instance, if the comparison audio sounded much better than the reference audio, a score of 3 would be given or a –2 if the comparison audio was worse than the reference audio. Objective preference between two settings can be measured by comparing identical audio in different settings. The order of test sounds (from TS_2), and the selection of reference and comparison audio (e.g., T1 after T3 or T3 after T1), were all randomized (the total scores of the three parameter settings represent the total quality score of each setting). To minimize the expected learning effects from the duplicated use of test sounds (although it may not be as significant for a quality test compared to an intelligibility test), the order was carefully arranged to prevent same test sound from repeating within 5 trials. In addition, to reduce physical and psychological fatigue of the subject, resting intervals of 5 seconds were given between tests, and an additional resting interval of 2 minutes every 25 trials.

Statistical analysis

In this study, repeated-measures analysis of variance (RM-ANOVA) test was performed to determine whether there was a significant differences in each subject's training result among the three listening situations: speech-only, noise-only, and speech-in-noise. Post-hoc Tukey honest significant difference test was further applied only when significant difference ($P<0.05$) was observed.

Table 1. Results of the ISL detection and ISL interpolation tests

ISL detection test (dB SPL)		ISL interpolation test (dB SPL)	
Actual SL	Detected SL	Input SL	Output volume
90	90.19	87	87.26
85	85.30	83	83.82
80	81.59	77	79.73
75	78.49	73	76.09
70	72.50	67	68.16
65	65.27	63	63.27
60	60.27	-	-

Error rate for ISL detection test and ISL interpolation test was in the range of 0.2%–4.6% and 0.3%–4.2%, respectively.
ISL, input sound level; SPL, sound pressure level; SL, sound level.

RESULTS

Table 1 provides the results of the ISL detection test and ISL interpolation test. The implemented ISL detector could detect the intensity of the actual input signals with a 0.2%–4.6% error rate, and in addition, it could also interpolate the volumes of the input signals as intended with a 0.3%–4.2% error rate.

Fig. 2 shows the AMP variation patterns when both training and verification were performed using TS_1. The AMPs in normal-operation mode followed the trained values in all situations
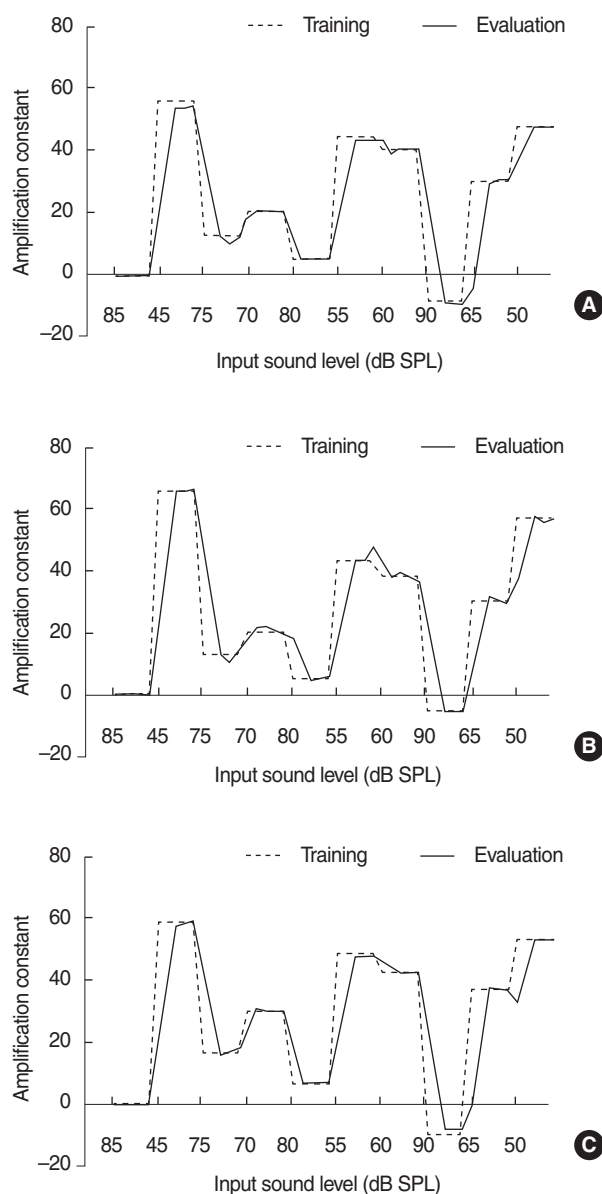
Fig. 2. Amplification constant (AMP) variation patterns when both training and verification were performed using TS_1 (unit: dB): (A) Speech-only. (B) Noise-only. (C) Speech-in-noise. Dotted line indicates data collected during training. Solid line indicates amplification output by the proposed algorithm. SPL, sound pressure level.

with some amount of time delay, and parameter interpolation also operated properly when required. Errors between the trained and adjusted AMPs were 0.01–2.71 dB, 0–4.23 dB, and 0–3.33 dB for speech-only, noise-only, and speech-in-noise respectively.

Fig. 3 shows the AMP variation patterns when training was performed using TS_1 and verification using TS_2. The AMPs in normal-operation mode also followed the trained values in all situations, and interpolation operated properly. Errors between the trained and adjusted AMPs were 0.06–4.37 dB, 0.02–4.99 dB, and 0–11.20 dB for speech-only, noise-only, and speech-in-noise respectively.

Table 2 shows the user-selected AMP and β values during the training phase of the subjective test. RM-ANOVA results showed significant difference existed among the three listening situations for all ISL (in the order from 45 dB to 90 dB, *F*-value is 8.00, 15.81, 19.99, 14.75, 17.67, 11.58, 9.86, 9.06, 6.14, and 4.70 respectively and *P*<0.05), and further, post-hoc analysis indicated that, in speech-in-noise conditions, subjects select significantly higher AM*P* values compared to speech-only and noise-only in all ISL levels except 90 dB where this was the case only in noise-only (*P*=0.008, Bonferroni adjustment).

Tables 3 and 4 give the results of the CMOS test for T1–T3 and T2–T3, respectively. In both tables, higher score indicate
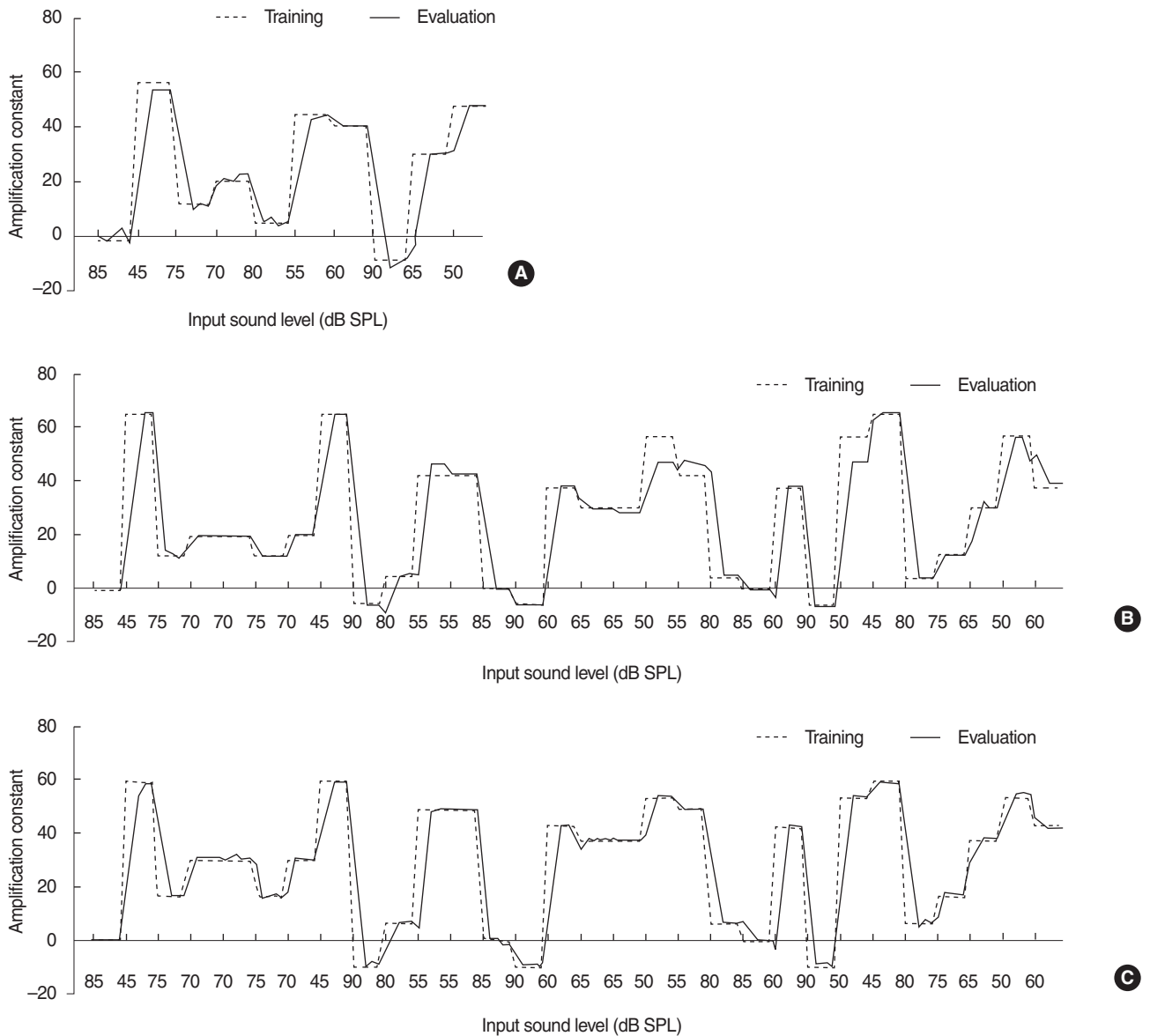

Fig. 3. Amplification constant (AMP) variation patterns when training was performed using TS_1 and verification using TS_2 (unit: dB): (A) Speech-only. (B) Noise-only. (C) Speech-in-noise. Dotted line indicates data collected during training. Solid line indicates amplification output by the proposed algorithm. SPL, sound pressure level.

Table 2. User-selected AMP and β values in the subjective test (in training phase)

Subject	LS	Input sound level (dB)									
		45	50	55	60	65	70	75	80	85	90
1	SO	−11.95	−19.66	−26.71[a)]	−33.76	−36.78	−39.46	−41.81	−43.83[a)]	−46.85	−48.29
	NO	−17.99	−21.34	−25.37[b)]	−30.74	−32.97	−40.47[b)]	−42.70	−50.20	−52.55	−56.58
	SIN	−8.59	−13.62	−16.31	−21.34	−24.36	−29.40	−30.07	−34.09	−40.81	−48.19
	β	3.47	2.65	2.65	2.65	2.65	3.47	2.65	2.65	2.65	2.65
2	SO	−23.02	−30.07	−33.76	−37.45	−42.15	−44.83	−47.85	−57.92	−63.29	−66.64
	NO	−2.89	−12.28	−18.99[b)]	−21.01	−30.07	−31.75[b)]	−38.46[b)]	−46.17	−51.88	−57.92
	SIN	−8.93	−13.96	−16.64	−21.34	−27.05	−31.07[b)]	−33.76	−41.48	−47.85	−53.89[a)]
	β	3.47	3.47	3.47	3.47	3.47	3.47	3.47	2.65	3.47	3.47
3	SO	−13.40	−27.67	−30.20	−29.10	−36.80	−36.80	−43.40	−43.40	−57.70	−48.00
	NO	−23.32	−23.32	−34.32[b)]	−34.32	−47.11	−47.11[b)]	−53.51	−60.65	−60.65	−60.65
	SIN	−10.03	−10.03	−21.03	−21.03	−21.03	−19.80	−38.50	−30.80	−52.80	−52.80
	β	3.47	2.65	3.47	2.65	3.47	1.76	3.47	1.76	3.47	3.47
4	SO	−15.30	−20.00	−25.03	−33.09	−36.78	−37.79	−42.48	−45.17	−54.68	−54.68
	NO	−4.90	−11.28	−15.97[b)]	−18.99	−23.69	−29.73[b)]	−33.76[b)]	−35.77	−41.14	−46.51
	SIN	−2.55	−5.23	−10.94	−15.30	−25.37	−30.07	−35.77	−40.13	−45.17	−47.85
	β	2.65	1.76	1.76	1.76	3.47	3.47	3.47	3.47	2.65	2.65
5	SO	−6.30	−15.10	−21.70	−25.00	−28.30	−30.50	−31.90	−38.50	−41.01	−52.09
	NO	−8.70	−17.50	−20.80[b)]	−24.30	−28.70[b)]	−39.70[b)]	−41.90	−47.40	−49.60	−50.70
	SIN	5.68	0.26	−4.21	−11.91[b)]	−17.41	−20.71	−20.71	−31.71	−33.91	−43.81
	β	2.65	2.65	2.65	2.65	2.65	2.65	1.76	2.65	1.76	2.65
6	SO	−13.96	−20.34	−26.38	−27.72	−33.42	−36.44	−40.81	−45.17	−51.54	−54.23
	NO	−1.88	−9.26	−15.30[b)]	−18.99	−26.38	−29.40	−34.09	−39.80[b)]	−45.00	−60.00
	SIN	1.14	−4.90	−9.60[b)]	−15.97	−19.66	−25.70	−30.07	−33.76	−38.46	−44.83
	β	3.47	2.65	2.65	2.65	3.47	2.65	2.65	3.47	2.65	2.65
7	SO	−0.61	−6.11	−13.81	−21.51	−29.21	−34.71	−36.91	−39.11	−43.51	−42.41
	NO	−11.00	−18.70	−24.20[b)]	−29.70	−30.80	−34.10[b)]	−36.30[b)]	−37.40	−41.80	−44.00
	SIN	9.90	2.20	−1.10	−6.60	−12.10	−15.40	−18.70	−25.30	−33.00	−34.10
	β	2.65	3.47	2.65	2.65	2.65	1.76	3.47	2.65	2.65	1.76
8	SO	−10.36	−13.66	−20.78	−25.21	−29.15	−29.15	−37.27	−37.27	−40.22	−43.18
	NO	−13.42	−17.82[b)]	−20.02[b)]	−26.62	−31.02	−34.32[b)]	−39.76[b)]	−41.96	−50.33	−50.33
	SIN	−7.99	−10.92	−16.37	−21.51	−21.52	−28.66	−35.55	−41.95	−45.15	−50.31
	β	3.47	3.47	2.65	3.47	2.65	3.47	3.47	2.65	3.47	2.65
9	SO	−10.39	−16.99	−21.39	−24.69	−27.99	−29.09	−31.29	−41.19	−37.76	−37.27
	NO	−16.21	−23.91	−30.51[b)]	−38.21	−43.71[b)]	−49.21[b)]	−53.61	−59.11	−59.11	−61.31
	SIN	−7.00	−9.71	−15.86	−19.55	−23.49	−28.90	−35.55	−36.53	−39.98	−43.42
	β	3.47	3.47	2.65	3.47	3.47	3.47	3.47	3.47	3.47	3.47
10	SO	−1.09	−9.89	−20.89	−21.99	−27.49	−28.59	−31.89	−35.19	−40.72	−43.42
	NO	−8.61	−11.91	−17.41[b)]	−22.91	−26.21	−29.51	−31.71[b)]	−35.01	−37.21	−40.51
	SIN	15.15	4.56	−3.80	−6.02	−7.74	−12.17	−19.80	−19.55	−28.90	−30.91
	β	3.47	3.47	3.47	3.47	2.65	2.65	2.65	2.65	3.47	2.65

Repeated-measures analysis of variance (RM-ANOVA) revealed significant differences among the three listening situations for all input sound level (ISL), and further post-hoc analysis indicated that, in speech-in-noise conditions, subjects select significantly higher AMP values compared to speech-only and noise-only in all ISL levels except 90 dB where this was the case only in noise-only.

AMP, amplification constant; LS, listening situation; SO, speech-only; NO, noise-only; SIN, speech-in-noise.

[a)]Not actually NO, but classified as NO. [b)]Not actually SO, but classified as SO.

[a)] and [b)] represent misclassified cases.

DISCUSSION

preference of T3 setting over the other setting being compared to. The T3 setting showed higher CMOS scores than the T1 setting for all subjects and listening situations. In contrast, the T2 setting and the T3 setting showed almost equivalent CMOS scores; that is, the preferences were different among the subjects.

In the subjective test, all subjects showed significant differences in their AMP settings ($P<0.05$) in accordance with the variations in the LS condition under the same ISL (usually between noise-only and speech-in-noise). Also, the user-selected AMP

Table 3. Results of the CMOS test between T1 setting and T3 setting using 10 speech-only sounds, 20 noise-only sounds, and 20 speech-in-noise sounds (TS_2)

Subject	1	2	3	4	5	6	7	8	9	10
Speech-only	2.20±0.79	1.00±1.56	0.90±1.37	1.10±1.29	1.00±1.94	1.00±1.49	1.70±1.16	1.50±1.43	2.00±0.47	1.80±1.48
Noise-only	1.30±0.98	0.15±1.93	0.55±0.94	0.15±1.14	0.30±1.30	−0.10±1.29	1.30±0.86	0.45±1.23	1.30±0.98	0.60±1.76
Speech-in-noise	0.40±1.23	0.35±1.76	0.50±1.24	0.40±1.23	0.55±1.73	0.20±1.44	0.65±0.88	0.40±1.79	0.80±1.24	0.80±1.28

Values represent the CMOS scores (higher score indicate the amount of preference of T3 setting over the other setting).
CMOS, Comparison Mean Opinion Score.

Table 4. Results of the CMOS test between T2 setting and T3 setting using 10 speech-only sounds, 20 noise-only sounds, and 20 speech-in-noise sounds (TS_2)

Subject	1	2	3	4	5	6	7	8	9	10
Speech-only	0.80±1.32	0.80±1.32	−0.30±1.25	0.20±1.03	0.20±1.55	−1.00±0.82	2.10±0.74	0.10±0.99	−0.50±1.43	2.10±0.57
Noise-only	1.35±1.09	−0.25±0.97	−0.05±0.83	−0.05±0.76	0.80±0.89	−0.45±1.05	1.45±1.00	−0.10±1.45	−0.35±1.27	1.15±1.23
Speech-in-noise	0.20±1.11	−0.30±1.38	0.00±0.65	−0.30±0.66	−0.75±1.16	−0.45±1.05	−0.05±1.05	−0.15±0.75	−0.20±1.24	−0.55±1.15

Values represent the CMOS scores (higher score indicate the amount of preference of T3 setting over the other setting).
CMOS, Comparison Mean Opinion Score.

values in speech-in-noise were generally higher than those in noise-only and speech-only situations under the same ISL, which implies that the listener prefers to reduce the volume of the device when there is no interesting sound (noise-only) or there is no acoustic interference (speech-only) to reduce the perceptual discomfort from the noise. In contrast, when an interesting sound is mixed with environmental interference, the listener usually prefers to increase the volume of the device to increase the intelligibility of the interesting sound while enduring the perceptual annoyance from the loud interference. This preference, though it was derived from normal-hearing subjects, demonstrates the high potential for LS classification to benefit hearing-impaired subjects. Furthermore, when considering the user-selected β values in speech-in-noise situations under different ISL condition, the participants all preferred different β settings (e.g., subjects 1, 2, 6, 8, 9, and 10 selected 2.65 and 3.47, subject 5 selected 1.76 and 2.65, and subjects 3, 4, and 7 selected 1.76, 2.65, and 3.47) with no apparent pattern. Additionally, participants sometimes preferred different β values as the ISL condition changed, which further demonstrates the necessity of user-adjustable noise-reduction algorithm (adjusting the DNS) for the trainable hearing aid. These results demonstrate that all ISL, LS, and DNS conditions can affect the results of training and as a result, affect the performance (personal satisfaction) of the device in real-life situations, which justifies the incorporation of all of those factors in the device operation.

In the CMOS test, the T3 setting showed significantly improved sound quality over the T1 setting [21]; however, the T3 setting showed no significant difference compared to the T2 setting despite the difference in parameter setting between T2 and T3 settings. These results may be associated with the fact that the LS classifier used in the current study occasionally misclassified noise-only (babble) situations as speech-only situations (Table 2), which may have introduced error to the results. In addi-

tion, the output of the LS classifier varied depending on when the user selection was made during the training phase (e.g., an identical babble noise could be classified as noise-only or speech-only, depending on the timing of the user selection), which greatly affected the classification accuracy among the participants. The CMOS test results between T2 and T3 would be improved by enhancing the performance of the LS classifier [5]. For example, looking at only the speech-only situation (which was least affected by the misclassification) in Table 4, T3 showed improved CMOS score (on average) in a majority of the subjects, which indicates the potential benefit of the proposed algorithm. In the current study, the proposed T3 setting resulted in CMOS scores 1.42, 0.60, and 0.51 (on average) higher than the T1 setting, in speech-only, noise-only, and speech-in-noise situations, respectively. The proposed T3 setting also resulted in scores 0.45 and 0.35 higher than the T2 setting, in speech-only and noise-only situations, respectively (scores were lower only in the speech-in-noise situation), which is consistent with the report by Chalupper et al. [4], which suggests that the classifier control training achieved more comfortable loudness ratings (in about 72%–88% of situations) than did the gain control training (in about 68%–82% of situations) on average.

There are three scenarios that may cause abnormal operation of the proposed algorithm in normal-operation mode. First, when the ambient sound level changes abruptly (by more than 5 dB), the output of the ISL detector may not detect the change as quickly because it uses the most recent 2.5 seconds of data to calculate (i.e., the input sound data from the previous track is temporarily dominant over the new one). As a result, the ISL detector may display a sound level with less than 5 dB difference (even though the change in environment was permanent, not temporary), and hence the algorithm assumes the environment did not change and proceed to implement incorrect internal parameters instead of refraining from an update for one in-

terval as intended. For example, note the errors in the 1st recording of 65 dB shown in Fig. 2A, as well as the errors in the 1st recording of 65 dB and the 1st recording of 50 dB shown in Fig. 3A. Second, the LS classifier may misclassify the category of the input sounds during training or during normal-operation mode, which would negatively influence the database, thereby negatively affecting the auto-tuned internal parameters. For example, as shown in Figs. 2B, 3B, the 55 dB babble noise sound was misclassified as speech-only during the training phase (in other words, the parameters for 55 dB noise-only were recorded as 55 dB speech-only, which resulted in an absence of data for 55 dB in the noise-only database). As a result, the parameters for 55 dB babble noise was applied to 55 dB speech-only in normal-operation mode, and the parameters for the 55 dB babble noise-only situation were determined by interpolation of the surrounding values, 50 dB and 60 dB noise-only parameters. The errors in 55 dB (babble noise) and 50 dB (car noise) shown in Fig. 3B were also caused by this scenario. Third, the first and second error-inducing scenarios may occur simultaneously, as shown in Fig. 3C: the 1st recording of 80 dB (babble noise) and the 1st recording of 75 dB (car noise).

There are several limitations in the current study. First, during the subjective evaluations, we considered only one user-controllable noise reduction algorithm: β-adjustable beamforming. However, in commercial devices, additional algorithms may be required to achieve environment-adaptive adjustment, such as single-microphone-based noise-reduction algorithms using spectral subtraction and Wiener filtering [17,22]. To modify the currently implemented algorithm to have a more conventional hearing aid algorithm structure, it is necessary to include additional user-controllable algorithms, such as single-microphone-based noise-reduction, to better reflect users' personal preferences for various degrees of ambient noise suppression and speech distortion. Second, wide dynamic-range compression was bypassed during the verification process because only normal-hearing subjects participated in the subjective test. However, in actual hearing aids, the use of wide dynamic-range compression is indispensable because it (with an additional fitting formula [23]) compensates for the threshold shift in the audiogram caused by damage to the inner and outer hair cells, which may also affect the personal parameter preferences set during the training phase [4,13]. For an implemented algorithm to be suitable for hearing-impaired persons, wide dynamic-range compression with appropriate band gains for individual hearing-impaired subjects should be added before the output volume adjustment, and then, the procedures described in the current study should be performed. In fact, compression learning achieved the largest comfortable loudness ratings, in about 76%–91% of situations, in the report of Chalupper et al. [4]. Third, the amount of time spent in the training phase was relatively short; continuous training of weeks or months after the initial training would be ideal to more accurately verify the ef-

fect of user training [24]. In addition, according to previous reports, the final parameter settings can be affected by the starting point of the training [22]. The effect of such initial parameter settings on the proposed algorithm should also be investigated in future works. Furthermore, in the current study, each subject performed the training phase with a one-signal-for-each-condition protocol. In real circumstances, the personal preferences in a specific condition can differ from time to time depending on the variations in emotional and circumstantial states of the device user. To deal with such ambiguous situations, in the current study, the algorithm was designed to override the previous parameter settings with the most recent parameter setting for the same exact environment. We based this on the assumption that the device user would adapt and become more familiar with the device, and as this happens, preference settings would also adapt. However, this overriding protocol may not always be appropriate in all situations; therefore, for example, it may be more effective to save all duplicated user selections for the same condition in the database, and then determine the output parameters by one of the following methods: (1) averaging the saved values or (2) giving different weights to the saved values in accordance with the recorded time, and averaging the weighted values. These approaches should also be considered in future studies. Finally, due to limited number of hearing-impaired patients available to participate in this study, normal-hearing persons participated in this study instead, which is the limitation of this study. Though this is less than ideal, the positive results observed with normal-hearing participants are highly likely to affect hearing-impaired persons in a similar way.

In conclusion, as the experimental results showed, the preferred settings for each of the ISL, LS, and DNS conditions were different among subjects, which justifies the incorporation of more environmental factors during the training phase for more dedicated device individualization. Based on the result of this study, it is very likely for greater personal satisfaction to be achieved under various ambient situations by reflecting individual preference to a more extensive acoustic conditions to further personalize hearing aid control.

CONFLICT OF INTEREST

ACKNOWLEDGMENTS

REFERENCES

1. World Health Organization. WHO global estimates on prevalence of hearing loss: morality and burden of diseases and prevention of blindness and deafness WHO, 2012. Geneva: World Health Organization; 2012.

2. McPherson B. Innovative technology in hearing instruments: matching needs in the developing world. Trends Amplif. 2011 Dec;15(4): 209-14.

3. Kochkin S. MarkeTrak VII: customer satisfaction with hearing instruments in the digital age. Hearing J. 2005 Sep;58(9):30.

4. Chalupper J, Junius D, Powers T. Algorithm lets users train aid to optimize compression, frequency shape, and gain. Hearing J. 2009 Aug;62(8):26.

5. Lamarche L, Giguere C, Gueaieb W, Aboulnasr T, Othman H. Adaptive environment classification system for hearing aids. J Acoust Soc Am. 2010 May;127(5):3124-35.

6. Taylor B. Advanced user control optimizes real-world listening preferences. Hearing J. 2011 Mar;64(3):26.

7. Dillon H, Zakis JA, McDermott H, Keidser G, Dreschler W, Convery E. The trainable hearing aid: what will it do for clients and clinicians? Hearing J. 2006 Apr;59(4):30.

8. Kochkin S. MarkeTrak VII: obstacles to adult non-user adoption of hearing aids. Hearing J. 2007 Apr;60(4):24-51.

9. Dreschler WA, Keidser G, Convery E, Dillon H. Client-based adjustments of hearing aid gain: the effect of different control configurations. Ear Hear. 2008 Apr;29(2):214-27.

10. Boymans M, Dreschler WA. Audiologist-driven versus patient-driven fine tuning of hearing instruments. Trends Amplif. 2012 Mar;16(1): 49-58.

11. Zakis JA, Dillon H, McDermott HJ. The design and evaluation of a hearing aid with trainable amplification parameters. Ear Hear. 2007 Dec;28(6):812-30.

12. Keidser G, Dillon H, Convery E. The effect of the base line response on self-adjustments of hearing aid gain. J Acoust Soc Am. 2008 Sep; 124(3):1668-81.

13. Keidser G, Alamudi K. Real-life efficacy and reliability of training a

14. Nielsen JB, Nielsen J, Larsen J. Perception-based personalization of hearing aids using Gaussian processes and active learning. IEEE/ ACM Trans Audio Speech Lang Process. 2015 Jan;23(1):162-73.

15. Walden BE, Surr RK, Grant KW, Van Summers W, Cord MT, Dyrlund O. Effect of signal-to-noise ratio on directional microphone benefit and preference. J Am Acad Audiol. 2005 Oct;16(9):662-76.

16. Teutsch H, Elko GW. First- and second-order adaptive differential microphone arrays. In: Proceedings of the 7th International Workshop on Acoustic Echo and Noise Control; 2001 Sep 10–13; Darmstadt, Germany. 35-8.

17. Yook S, Nam KW, Kim H, Hong SH, Jang DP, Kim IY. An environment-adaptive management algorithm for hearing-support devices incorporating listening situation and noise type classifiers. Artif Organs. 2015 Apr;39(4):361-8.

18. Schaub A. Digital hearing aids. New York (NY): Thieme Medical Publishers Inc.; 2008.

19. Buchler M, Allegro S, Launer S, Norbert D. Sound classification in hearing aids inspired by auditory scene analysis. EURASIP J Appl Signal Process. 2005 Jan;18:2991-3002.

20. International Telecommunication Union. P.800: methods for subjective determination of transmission quality [Internet]. Geneva: International Telecommunication Union; c2008 [cited 2016 Jun 1]. Available from: https://www.itu.int/rec/T-REC-P.800-199608-I/en.

21. Ricketts TA, Hornsby BW. Sound quality measures for speech in noise through a commercial hearing aid implementing digital noise reduction. J Am Acad Audiol. 2005 May;16(5):270-7.

22. Mueller HG, Hornsby BW, Weber JE. Using trainable hearing aids to examine real-world preferred gain. J Am Acad Audiol. 2008 Nov-Dec;19(10):758-73.

23. Byrne D, Dillon H, Ching T, Katsch R, Keidser G. NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures. J Am Acad Audiol. 2001 Jan;12(1):37-51.

24. Pierce JL, Kostova T, Dirks KT. The state of psychological ownership: integrating and extending a century of research. Rev Gen Psychol. 2003 Mar;7(1):84-107.

hearing aid. Ear Hear. 2013 Sep;34(5):619-29.