

# Implications of High-Frequency Cochlear Dead Regions for Fitting Hearing Aids to Adults With Mild to Moderately Severe Hearing Loss

Robyn M. Cox,<sup>1</sup> Jani A. Johnson,<sup>1</sup> Genevieve C. Alexander<sup>2</sup>

**Objectives:** A cochlear dead region (DR) occurs at a given frequency when there is a loss of normal functioning of inner hair cells tuned to that frequency. It has been suggested that existence of high-frequency DRs has implications for hearing aid fitting, and that the optimal amount of high-frequency gain is reduced for these patients. However, the data supporting this suggestion has been obtained using listeners with severe or profound hearing loss. It is uncertain whether these results would apply for listeners with the mild to moderately severe hearing loss that is more typical of hearing aid users. This investigation used laboratory and field measurements to examine the effects of reduced high-frequency gain in typical hearing aid users with high-frequency DRs compared with matched users without DRs.

**Design:** The study was a double-blinded, nonrandomized intervention design with 18 matched pairs of adult subjects. In each pair, one subject had high-frequency DRs (usually at one or two test frequencies) and the other subject had no DR. Each subject was fitted unilaterally with a hearing aid having two active programs. One program used a National Acoustics Laboratories (NAL) prescription target for average speech (NAL condition). The other program was identical to NAL except for reduced gain in the high frequencies (low-pass [LP] condition). Outcomes included aided speech recognition in quiet and noise measured in the laboratory, ratings of speech understanding in daily life, and final preference for the NAL or LP program.

**Results:** In laboratory testing, speech recognition in quiet was significantly better when using the NAL program. This result was seen for subjects with DRs and without DRs. When listening in noise, speech recognition was significantly better when using the NAL program for subjects without DRs. For subjects with DRs, results were equivalent for the NAL and LP programs. In daily life, the NAL program received significantly higher ratings for speech understanding, and this result was seen for subjects with and without DRs. When asked about their preference for using the NAL or LP program in daily life, subjects did not always choose the program they had rated as giving better speech understanding, but their preference was not associated with having DRs. About one-third of subjects preferred the LP program. The reason most frequently given for preferring the LP program was that the NAL was too loud.

**Conclusions:** Overall, adult hearing aid wearers with mild to moderately severe hearing loss benefitted from high-frequency gain whether or not they had DRs. In laboratory testing, but not in daily life, subjects with DRs tended to obtain less benefit than those without DRs. However, provision of NAL high-frequency gain never resulted in poorer performance in either group. These results suggest that identification of high-frequency DRs at one or two frequencies does not call for any a priori modification of the target hearing aid prescription for listeners with mild to moderately severe hearing loss.

(*Ear & Hearing* 2012;33;573–587)

<sup>1</sup>School of Communication Sciences and Disorders, University of Memphis, Memphis, Tennessee; and <sup>2</sup>James H. Quillen VAMC, Johnson City, Tennessee.

## INTRODUCTION

Contemporary evidence-based hearing aid prescription methods call for considerable high-frequency gain when the patient's audiogram reveals high-frequency hearing loss (Moore et al. 1999a; Byrne et al. 2001; Seewald et al. 2005). Sometimes, the prescribed amount of high-frequency gain is difficult to achieve in an actual hearing aid fitting because of problems with sound quality or acoustic feedback. Nevertheless, realization of prescribed gain levels remains the standard that most practitioners attempt to meet when possible.

In spite of the evidence in favor of high-frequency amplification, its benefits began to be re-evaluated about a decade ago. This occurred after the publication of several investigations suggesting that the amount of amplification recommended by existing prescriptive methods might not provide expected benefits for some patients (Ching et al. 1998; Hogan & Turner 1998; Turner & Cummings 1999; Amos & Humes 2001). The evidence from these studies implied that there is considerable individual variation in optimal high-frequency gain for individuals whose high-frequency thresholds exceed 55 to 60 dB HL.

It is known that damage to inner hair cells (IHCs) is often present at frequencies where the hearing threshold exceeds about 50 dB HL. However, the extent of IHC damage is variable across individuals (Schuknecht 1993; Killion & Niquette 2000). Some investigators have explored how speech perception is affected by a complete loss of IHC functioning across a limited region of the basilar membrane (Lippmann 1996; Shannon et al. 2002). This kind of cochlear damage has become known as a cochlear dead region (DR) (Moore et al. 1999b). Moore (2001a) suggested that the presence of cochlear DRs might be an explanation for studies that reported a lack of expected benefit from high-frequency gain in some individuals. A DR occurs at a given frequency when there is a total loss of normal functioning of IHCs located at the region tuned to that frequency on the basilar membrane in the cochlea. However, loss of IHC function does not always result in total inability to detect a sound that would normally be detected by that region. Instead, the sound might be detected by the spread of excitation to an adjacent region where IHCs are still functioning. Moore suggested that decreased ability to benefit from high-frequency amplification might be the result of a DR in that frequency region.

Psychophysical tuning curves (PTCs) are a traditional method for identifying DRs in laboratory research (Moore et al. 2000). However, measurement of PTCs is too time consuming to be viable as a clinical test. The Threshold Equalizing Noise (TEN) test was developed by Moore et al. to allow rapid identification of DRs in a clinical setting (Moore et al. 2000; Moore et al. 2004). Although there has been some debate about the relative accuracy of PTCs and the TEN test in identifying DRs,

and varying rates of agreement have been reported (Moore et al. 2000; Summers et al. 2003; Malicka et al. 2010), both test procedures are widely accepted as valid. Further, because only the TEN test is suitable for clinical application, the published research exploring the implications of DRs for hearing aid fitting has relied solely on the TEN test to identify DRs.

The theory underpinning the TEN test is fully described in the aforementioned publications and in others (Moore 2001a, 2001b). In practice, the test calls for the measurement of unmasked and TEN-masked pure tone thresholds. The standard criterion (proposed by Moore et al. 2000) for diagnosing a DR at the test frequency is a finding that the masked threshold exceeds, by at least 10 dB, both the patient's unmasked threshold and the masked threshold that are expected for listeners with normal hearing.

Moore et al. explored speech recognition by listeners with severe to profound high-frequency hearing loss and extensive high-frequency DRs (Vickers et al. 2001; Baer et al. 2002). They concluded that, for subjects with this type of impairment, amplification could be useful for frequencies up to about 1.7 times the lower edge frequency ( $F_e$ ) of the DR. However, amplification of frequencies above  $1.7 \times F_e$  was not beneficial and sometimes was detrimental for speech understanding. Although the subjects in these studies had greater hearing loss than most hearing aid wearers, the findings suggested that the presence of a DR might have implications for hearing aid prescriptions. Consequently, it has been proposed that when high-frequency DRs are present, gain for frequencies above the lower edge of the DR could be reduced relative to that called for by modern generic prescription methods (Moore 2001b). In theory, this reduction of high-frequency gain would not negatively affect speech understanding and might actually improve it. At the same time, less high-frequency gain could diminish problems with acoustic feedback and sound quality. Most recently, Vinay and Moore (2007a) reasoned that the recommended gain reduction could be beneficial for persons with DRs at frequencies below 3.0 kHz but not when DRs occur at or above about 3 kHz, because  $1.7 \times F_e$  would approximate the upper frequency limit of typical hearing aids (about 5.0 kHz). Thus, Vinay and Moore (2007a) suggested that when hearing aid fitting is under consideration, DRs should be measured at 2.0 kHz and below if audiogram thresholds exceed about 60 dB HL at those frequencies.

The implications of high-frequency cochlear DRs for hearing aid fitting have also been of interest to other investigators and practitioners. Their recommendations often have been more sweeping than those of Moore et al. Some have proposed that there is a need to limit high-frequency gain for most or all patients with moderate or worse high-frequency hearing thresholds (Turner 1999; Yanz 2002). Padilha et al. (2007) recommended that amplification be "avoided" at frequencies where there are DRs. On the other hand, several investigators have reported that they were unable to determine any degree of high-frequency hearing loss above which it was not helpful to provide compensatory amplification, although DRs were not specifically measured in these studies (Turner & Henry 2002; Plyler & Fleck 2006; Horwitz et al. 2008).

Cox et al. (2011) measured DRs in about one-third of patients with the mild to moderately severe hearing loss that is typical of hearing aid wearers. However, unlike the subjects studied by Moore et al. (Vickers et al. 2001; Baer et al. 2002), these listeners usually yielded a positive TEN test result at only

one or two test frequencies (similar findings were reported by Hornsby & Dundas, 2009). In the laboratory, the subjects of Cox et al. scored significantly better on a speech recognition test when more high-frequency audibility was provided, regardless of their DR status. Cox et al. recommended further investigation, including field trials with typical hearing aid patients to determine the most effective approach to high-frequency amplification when these types of individuals have DRs.

Relatively few studies have been reported in which performance with hearing aids was tested for individuals with diagnosed DRs. Mackersie et al. (2004) used a matched-pairs design with a threshold-matched control group in a laboratory test to examine consonant identification by hearing aid wearers with and without DRs. The subjects had steeply sloping high-frequency hearing loss and relatively normal low-frequency thresholds. The DRs were identified with the TEN test using the standard DR criterion of 10 dB excess masking. However, subjects in the DR group were required to have positive DR results across three or more contiguous test frequencies. Four of the eight DR subjects had DRs extending upwards from 2.0 kHz and the rest had DRs extending upwards from 3.0 kHz. In this laboratory study, the hearing aid's bandwidth extended through 6.3 kHz. The results of this experiment indicated that subjects with and without DRs performed equivalently in quiet and low-noise conditions. However, in high-noise conditions, subjects without DRs improved whenever additional high-frequency audibility was provided, whereas subjects with DRs reached an asymptote and did not further improve when high frequencies more than 100% above the edge of the DRs were made audible.

Preminger et al. (2005) recruited 49 individuals who were current wearers of bilateral hearing aids and who had at least two hearing thresholds in each ear poorer than 50 dB HL and no thresholds poorer than 80 dB HL. Unaided speech recognition was measured using the Quick Speech in Noise (QSIN) test materials (Killion et al. 2004), comparing standard lists with lists having a high-frequency emphasis above 1.0 kHz. In addition, subjective aided performance was measured using the Abbreviated Profile of Hearing Aid Benefit questionnaire (Cox & Alexander, 1995). The TEN test was used with these listeners to determine the presence or absence of DRs, however, 15 dB of excess masking was required to identify a DR. Thus, these investigators used a stricter criterion to interpret the results of the TEN test than the one proposed by Moore et al. (2000). With this strict criterion, 29% of the hearing aid wearers tested positive for DRs. It was noted that subjects with DRs tended to have more hearing loss than those without DRs. High-frequency DRs were observed in 13 ears. Although subjects without DRs performed better overall on the QSIN test than those with high-frequency DRs, the two groups did not differ in the benefit obtained from increased high-frequency audibility in the QSIN test.

Gordo and Iorio (2007) reported a laboratory study of 15 subjects with high-frequency DRs ( $F_e$  between 1.0 and 1.5 kHz) and 15 subjects without DRs. Each subject was fit bilaterally with programmable eight-channel hearing aids. Program 1 provided amplification from 0.1 to 8.0 kHz. Program 2 provided no gain at frequencies above 2.56 kHz. Speech recognition scores in quiet and noise were measured with each program for monosyllabic words and for sentences. For subjects without DRs, performance was significantly better with Program 1, whereas, for subjects with DRs, performance was significantly better with

Program 2. The authors concluded that when DRs are present at high frequencies, avoiding amplification in this frequency range gives better results. However, the two groups in this study were not matched for hearing loss. Compared with subjects without DRs, subjects with DRs had substantially worse thresholds at high frequencies.

It is noteworthy that the three laboratory investigations, in which subjects with and without DRs used individually fitted hearing aids, all involved subjects with the mild to moderately severe hearing loss that is typical of hearing aid wearers. This is in contrast with the severely or profoundly impaired subjects studied by Moore et al. (Vickers et al. 2001; Baer et al. 2002). The investigations with hearing aid wearers have not yielded consistent conclusions regarding the implications of high-frequency DRs for high-frequency gain in hearing aid fittings. Mackersie et al. (2004) and Preminger et al. (2005) both found that hearing aid wearers benefited from high-frequency amplification under some conditions, regardless of their DR status. In addition, Mackersie et al. noted that, compared with a matched control group, listeners with DRs did not show additional benefit from a broadband signal when the noise level was high, but they did not perform worse. In contrast, Gordo and Iorio (2007) found that hearing aid wearers with DRs did not benefit from high-frequency amplification and were actually penalized by it, whereas a comparison group without DRs (with substantially better hearing thresholds) did benefit from high-frequency amplification. Hence, there is still uncertainty about the most effective approach to hearing aid fitting for patients with mild to moderately severe hearing loss who have high-frequency DRs versus the optimal fitting for patients with similar hearing loss who do not have DRs. One approach to resolving this uncertainty is to conduct an investigation that includes both laboratory and field trial elements in which patients with and without DRs use hearing aids with prescribed and reduced high-frequency gain. This article reports the results of such a study.

We asked whether performance was affected when high-frequency gain was reduced for typical hearing aid patients who tested positive for high-frequency DRs on the TEN test. Each research question compared the effects of a change in high-frequency gain for a group of subjects positive for DRs with the effects of the corresponding gain change for a matched group of subjects negative for DRs.

1. When listening in the laboratory to speech in quiet, do subjects perform better, equally, or worse when fit with amplification using an evidence-based prescription method compared with amplification with reduced high-frequency gain, and does the answer to this question differ for listeners with DRs compared with a matched control group without DRs?
2. When listening in the laboratory to speech in noise, do subjects perform better, equally, or worse when fit with amplification using an evidence-based prescription method compared with amplification with reduced high-frequency gain, and does the answer to this question differ for listeners with DRs compared with a matched control group without DRs?
3. When listening in daily life to speech in noise, do subjects rate their performance better, equally, or worse when fit with amplification using an evidence-based prescription method compared with amplification with reduced

high-frequency gain, and does the answer to this question differ for listeners with DRs compared with a matched control group without DRs?

4. Considering overall listening in daily life, do subjects prefer amplification using an evidence-based prescription method or amplification with reduced high-frequency gain, and does the answer to this question differ for listeners with DRs compared with a matched control group without DRs?

## MATERIALS AND METHODS

For this nonrandomized intervention study, hearing-impaired adults were allocated to one of two groups. Subjects with DRs were assigned to the experimental group. Subjects without any DRs were assigned to the control group. Each individual in the experimental group was matched with an individual in the control group. Each subject was provided with a unilateral hearing aid fitting. The hearing aid was programmed with two frequency responses: the first (the National Acoustics Laboratories [NAL] program) was based on the NAL-NL1 (NAL-Nonlinear, Version 1, Byrne et al. 2001) real-ear targets for 70 dB speech input; the second (the low-pass [LP] program) was identical to the first up to 1.0 kHz with gain smoothly rolled off at higher frequencies. Because some studies have suggested that a period of use is needed before listeners achieve their final performance with a new frequency response (see Turner et al. 1996, for a review), a 2-week acclimatization period was allowed before laboratory testing of aided speech recognition in quiet and in noise. Then, over the next 2-week period, each subject rated the two programs in their daily listening environments. Lastly, laboratory testing of aided speech recognition in quiet and in noise was repeated, and each subject was interviewed about which hearing aid program they preferred in their daily lives. Subjects were paid for their participation. For each subject, the entire protocol required four test sessions.

Subjects were blinded to the specific purpose of the study and to the differences between the hearing aid programs. They were told that the study was looking at different types of hearing aid fittings for different types of hearing losses. In addition, the scoring of laboratory speech-understanding tests was blinded in the following way: the researcher who administered the test scored it as usual. At the same time, the subject's spoken responses were digitally recorded. These recordings were later scored by a second researcher who was blinded to the subject and hearing aid program. Any scoring discrepancies were resolved by a third listener.

### Participants

All of the subjects with DRs were recruited from individuals identified in an earlier study of DRs in typical hearing aid patients (Cox et al. 2011). Typical hearing aid patients were defined as follows: (1) adults with bilateral flat or sloping sensorineural hearing impairments, and (2) hearing thresholds of 60 to 90 dB HL for at least part of the frequency range from 1 to 3 kHz, and (3) thresholds no better than 25 dB HL for frequencies less than 1 kHz. In addition, the subjects reported a fairly active lifestyle, good physical and mental health, and English as their primary language. To be eligible for the present

study, subjects were also required to be capable of doing speech intelligibility ratings. Exclusion criteria were as follows: history of ear surgery, chronic middle or outer ear pathology, evidence of retro-cochlear involvement, and known psychiatric or neurological disorder.

The TEN (HL) test was used for the diagnosis of DRs (Moore et al. 2004). Cox et al. (2011) provide a detailed description of the methods of test administration and scoring. Briefly, stimuli were presented monaurally via an ER-3A earphone coupled to the ear with a compressible foam plug. Unmasked and TEN-masked thresholds were determined for pure tones at the following test frequencies: 0.50, 0.75, 1.0, 1.5, 2.0, 3.0, and 4.0 kHz. The standard threshold-testing protocol with 5 dB ascending steps was used until the threshold zone was reached. Then the test was completed using 2 dB steps. All the TEN test data were obtained in a single test session. Cox et al. provide a discussion of the similarities of their results compared with that of previous studies, as well as a review of the evidence regarding the reliability of the TEN test. On the basis of the standard criterion of 10 dB excess masking, Cox et al. determined that 53 of 170 typical hearing aid candidates had a DR for at least one test frequency from 0.5 to 4.0 kHz. Some of these individuals were existing hearing aid users and some were considering amplification for the first time.

To be eligible for the experimental group in the study reported in this paper, it was required for one or more DRs to be present in the better ear at or above 1.0 kHz and no DRs to be present at frequencies below 1.0 kHz. Of the 53 subjects with DRs identified by Cox et al. there was a pool of 33 that met these requirements and all of them were eligible to participate in the study. However, of this group of 33, a matched control could be identified for only 21 individuals. These 21 pairs of subjects were enrolled in the study. Although 21 pairs of subjects began the field trial protocol, 3 pairs were dropped from the study because one member of the pair did not complete the protocol. Eighteen matched pairs completed the protocol and their data are presented in this article.

Individuals in the matched control group were recruited from the subjects of Cox et al. (2011) who were not diagnosed with DRs in either ear, or from volunteers who did not participate in the Cox et al. study. Each control subject was matched with an experimental subject according to age (within 10 years), degree of hearing loss (average threshold across 0.5, 1.0, and 2.0 kHz within 5 dB), slope of audiogram from 0.5 to 4.0 kHz (within 5 dB/octave), and aided ear (right/left), although aided ear did not match for two pairs. In each group there were 13 fittings in the right ear and 5 in the left ear. There were 24 men and 12 women. Eleven pairs were the same gender, 7 pairs were not. The experimental subjects' ages ranged from 54 to 88 years, with a mean age of 72 years (SD = 11.14). The control subjects' ages ranged from 56 to 85 years, with a mean age of 74 years (SD = 8.34). Twelve experimental and ten control subjects were current hearing aid wearers when they began the study. The remainder of the subjects had not previously worn hearing aids. Figure 1 depicts the mean test ear audiogram for the experimental and control groups. On average, subjects in both groups had sloping mild to severe hearing impairment in the ear fitted with a hearing aid for this study.

The experimental group was diverse with regard to the distribution of DRs across the frequency range of 1.0 to 4.0 kHz. This distribution is documented in Table 1. For 8 of the experimental subjects, the lowest DR measured was located between 1.0 and

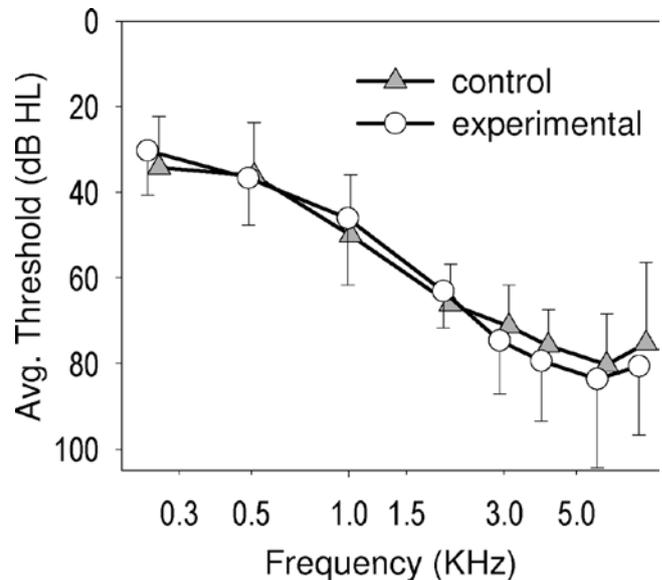


Fig. 1. Mean test ear audiograms for the experimental and control groups. Bars depict 1 standard deviation (SD).  $N = 18$  per group.

2.0 kHz. On the basis of the hypothesis of Vinay and Moore (2007a), reduction of high-frequency gain relative to prescribed levels would be more beneficial for these 8 individuals than for the 10 whose DRs were measured at or above 3.0 kHz. This issue was considered in the data analyses. Ten of the experimental subjects had DRs in one frequency only. This is consistent with the finding of Cox et al. (2011) that most of their typical hearing aid wearers who were positive for DRs were positive at only one test frequency. A large proportion of subjects with DRs in only one frequency was also reported by Hornsby and Dundas (2009). This issue was considered in the data analyses.

## Procedure

The research protocol was approved by the Institutional Review Board of the University of Memphis.

**Session 1: Preliminary Data** • Case history information, hearing aid experience (yes/no), and audiograms were obtained. The protocol was reviewed and the consent form was signed. If an experimental subject had DRs in both ears, the test ear was selected as the better-hearing ear, or, if the loss was equivalent in the two ears, the subject chose the test ear. For individuals in the control group, the test ear was their better hearing ear or the loss was equivalent in the two ears, and the test ear audiogram matched that of their counterpart in the experimental group. An impression of the test ear was taken, and a skeleton ear mold with adjustable vent was ordered.

**Session 2: Hearing Aid Fitting** • Each subject was fit unilaterally with a hearing aid. Unilateral fittings were chosen to avoid any ambiguity when subjects switched between programs during the field trial ratings. If two devices had been used, subjects would have been required to switch back and forth between programs for each device separately to make the ratings, and it was felt that the potential for errors in pairing the ratings with the programs would be higher under these conditions. The study predated the wide availability of bilaterally synchronized hearing aids. Thirty-five subjects wore Starkey Axent II

**TABLE 1. Test frequencies with DRs (kHz) for subjects in the experimental group**

Subject	Test Frequency With DRs (kHz)				
	1.0	1.5	2.0	3.0	4.0
E04	X		X	X	X
E14		X			
E03		X	X		
E15			X		
E16			X		
E02			X		X
E07			X		X
E05			X	X	X
E11				X	
E12				X	
E10				X	X
E13				X	X
E09					X
E08					X
E06					X
E17					X
E19					X
E20					X

behind-the-ear style hearing aids. These were midlevel technology hearing aids with four channels of compression and eight bands for gain adjustment across frequencies. The 36th subject used a Starkey Destiny 1200 hearing aid, which had eight channels of compression and 12 bands for gain adjustment across frequencies (the Destiny hearing aid was a successor to the Axent II and had somewhat more effective feedback management, which this subject needed). To maintain the contrast between the NAL and LP programs during their use in diverse listening situations, the decision was made to use linear processing rather than wide dynamic range compression. Essentially, linear processing was achieved by setting the compression ratios at, or close to, 1.0. The volume control was disabled for both programs. The ear mold was appropriately vented for each subject. There were no open fittings. The feedback management algorithm was activated. Directional microphone and noise reduction features were present in the hearing aids but were disabled for the study. All these hearing aid settings were consistent with the current best practice evidence (Larson et al. 2000; Bentler 2005; Davies-Venn et al. 2009). Half of the subjects were fit with the NAL program in the hearing aid's Program 1 (P1). The other half were fit with the LP program in P1.

Each hearing aid was fit with two active programs. One program is referred to as the *NAL program*. The other program, referred to as the *LP program*, was equal to the NAL program below 1.0 kHz. At higher frequencies the LP gain was rolled off smoothly with the details determined by the estimated  $F_e$  of the lowest frequency DR for that subject. There are no published guidelines for reducing high-frequency gain to optimize the fitting when DRs are present. Laboratory studies using filters can achieve precise cutoff frequencies and very steep slopes for high-frequency rejection, but this is not feasible in real hearing aids. For this study using real hearing aids, the rules for

high-frequency gain reduction were based on the consideration of the frequency of the lowest measured DR and the need to manipulate the frequency gain function in a manner that would produce a reasonable difference between the two test conditions in the desired frequency region without generating undesirable artifacts by overprocessing the signal. For the 15 experimental subjects whose lowest positive DR was at frequencies of 2.0 kHz or higher, gain was decreased by 2 to 3 dB at 2 kHz, 5 dB at 2.5 kHz, 10 dB at 3 kHz, and then smoothly rolled off to 0 dB gain above 6 kHz. For the 3 experimental subjects whose lowest positive DR was at 1.0 or 1.5 kHz, LP gain was decreased by 10 dB at 2 kHz and surrounding frequencies were adjusted as described earlier. Each subject in the control group was fit in the same manner as his or her match in the experimental group.

Objective verification of the fitting was accomplished using the Audioscan Verifit real ear probe microphone system (Versions 2.2 and 2.4). The NAL program was fit first. Because the hearing aid's processing was essentially linear, the goal was to match the prescription for average level speech. The NAL-NL1 prescription and real ear aided response (REAR) targets for speech at 70 dB SPL were used to fulfill this goal. Although the NAL-NL1 prescription was developed for fittings using nonlinear processing, the target for average speech is essentially the same as the target generated using the earlier widely validated NAL prescription for fittings using linear processing (NAL-RP, Byrne et al. 2001). Speech was presented at 70 dB SPL and the hearing aid was adjusted so that the measured REAR was within  $\pm 5$  dB SPL of the prescription targets from 0.25 kHz through at least 4.0 kHz. The maximum power output (MPO) was measured using a series of tone bursts swept across the frequency range at 85 dB SPL. If the subject noted any discomfort for the tone bursts, MPO was adjusted for comfort as necessary. The subject then rated the loudness and quality of speech presented at 65 dB SPL. If the amplified speech was not comfortably loud and of acceptable quality, the hearing aid programming was adjusted as needed and real ear measurements (REAR and MPO) were reobtained. Most subjects required some reduction of gain at the highest frequencies relative to the NAL prescription. Next, the LP program was adjusted to equal the fitted NAL program up to 1.0 kHz and to roll off smoothly at higher frequencies as described previously. Then, the loudness of the NAL and LP programs was compared by switching between them while the subject rated the loudness of speech presented at 65 dB SPL. The overall gain of the LP program was adjusted, if necessary, until the NAL and LP programs were judged to be equally loud. Most subjects did not require much adjustment. REAR for speech at 70 dB SPL, and MPO, were measured for the LP program. REARs for soft speech (50 dB SPL) were also obtained for both programs. Figures 2 (experimental group) and 3 (control group) depict the mean fitted REARs for speech at 70 dB SPL for the NAL and LP programs compared with the mean NAL-NL1 prescription for speech at 70 dB SPL. Similar patterns of real ear differences between the NAL and LP programs were seen for MPO and soft speech measures.

Subjective verification procedures were carried out in a typical (not sound-treated) room. These measures included loudness ratings of everyday sounds and the subject's own voice. In addition, the Ling Six-Sound test (Ling & Ling 1978) was administered in a soft voice to assess the audibility of soft speech. The hearing aid fittings, including both programs, were judged by the subjects to be acceptable for

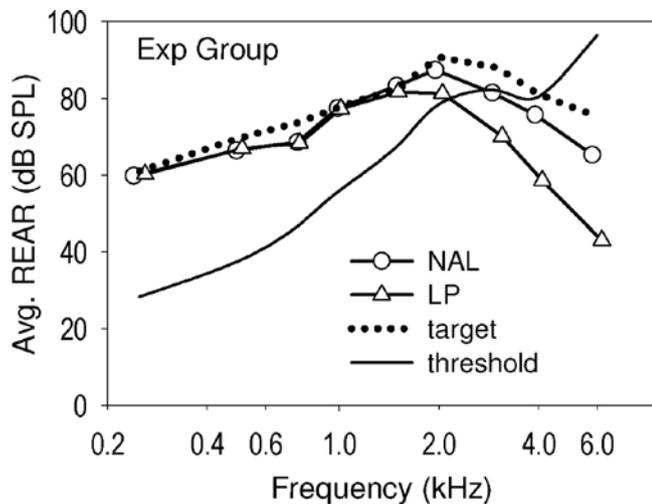


Fig. 2. Mean fitted REAR for speech at 70 dB SPL for the NAL and LP programs, compared with the mean NAL-NL1 prescription for speech at 70 dB SPL. Data are shown for the experimental group. NAL-NL1, National Acoustics Laboratories-Nonlinear, Version 1; LP program, low-pass program.

long-term use. Instructions for the use and care of the hearing aid during the 2-week acclimatization period were reviewed with the subjects at this time and extra hearing aid batteries were provided. Subjects were also given a brochure that reinforced this information and reminded them about the details of the research protocol. To ensure that both programs (NAL and LP) were used equivalently throughout the study, subjects were instructed to use P1 on odd-numbered dates and P2 on even-numbered dates.

**Acclimatization Period** • Subjects were given 2 weeks to acclimatize to the hearing aid. This provided new users with time to become adjusted to amplified sound, while those with previous experience used this time to acclimatize to any differences between the experimental hearing aid and their own instruments. To ensure reasonable compliance with the protocol, subjects agreed to wear the hearing aid for at least 4 hr each day (as noted in the following, they reported wearing the hearing aids actually longer than this). Each subject was interviewed by telephone 2 to 3 days after the fitting to see how he or she was progressing, to ensure compliance with the wearing schedule, and to make sure that there were no fit or sound quality issues with the hearing aid. If the subject reported any problems, an appointment was scheduled to adjust the hearing aid and review the procedures. Real ear responses were remeasured if any changes were made to the hearing aid programs (final fittings are shown in Figs. 2 and 3). After this, no further fine-tuning adjustments were allowed for the duration of the study.

**Session 3: Speech Recognition Measurement and Field Trial Training** • *Speech recognition tests were performed in a sound-treated room.* Speech and noise were presented from a head-sized loudspeaker located 1 m in front of the subject. Recorded speech and noise were routed through a GSI-61 audiometer to an amplifier and the loudspeaker. The non-test ear was plugged and muffed. The test began with the hearing aid set to P1. Half the subjects began with speech testing in quiet, and the other half began with testing in noise. Control subjects were tested using the same test lists and order as their counterparts in the experimental group.

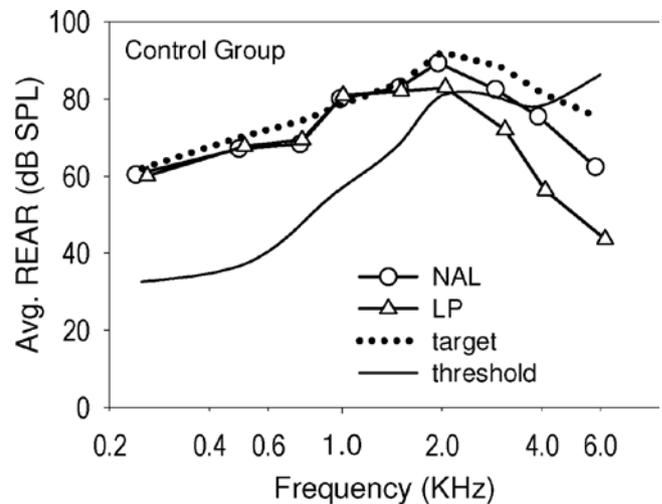


Fig. 3. Mean fitted REAR for speech at 70 dB SPL for the NAL and LP programs, compared with the mean NAL-NL1 prescription for speech at 70 dB SPL. Data are shown for the control group. NAL-NL1, National Acoustics Laboratories-Nonlinear, Version 1; LP program, low-pass program.

Aided speech recognition in quiet was measured using the Computer-Assisted Speech Perception Assessment (CASPA; Mackersie, Boothroyd & Minniear 2001). The CASPA stimuli consist of 20 lists of digitally recorded consonant-vowel-consonant words spoken by a female talker. Each list contains 10 words and uses the 30 vowels and consonants that occur most commonly in English consonant-vowel-consonants. Five CASPA lists were presented for each of the NAL and LP programs. The list order was randomly assigned and the hearing aid memory order was interleaved as follows: P1 (2 lists), P2 (5 lists), P1 (3 lists). One practice list was administered before the test. Stimuli were presented at a 65 dB SPL (a typical level for speech in quiet). Subjects were instructed to repeat the words and were encouraged to guess. Responses were scored using the CASPA computer software. The software can provide several types of scores (percent correct of vowels, consonants, phonemes, and words). In this study, we elected to use the score for total consonants correct (initial plus final consonants).

Aided speech recognition in noise was measured using the Bamford-Kowell-Bench Speech In Noise (BKB-SIN test (Ety-moticResearch 2005)). In this test, sentences spoken by a male talker are presented with a masker of 4-talker babble. Each list contains 10 sentences and 31 scoring words. The test comprises pairs of lists that have been equated for difficulty. When a list is presented, the signal to noise ratio (SNR) worsens by 3 dB for each sentence so that each list covers the signal to noise ratio range of +21 dB to -6 dB. In this study, two pairs of lists were presented for each of the NAL and LP programs. The order of list pairs was randomly assigned and the hearing aid memory order was interleaved as follows: P1 (1 list pair), P2 (2 list pairs), P1 (1 list pair). One practice list was administered before the test. Sentences were presented at a 73 dB SPL (a "loud but OK" level, as recommended for this test). Subjects were instructed to repeat each sentence and were encouraged to guess. The score for each list was the number of key words correct in each sentence, summed over all 10 sentences. The total score was the average number of words correct across all lists.

Subjects were trained in the procedures to be used in the 2-week field trial. The field trial was designed to quantify speech understanding under a variety of everyday listening conditions for both the NAL and the LP programs. The rating protocol was based on the procedure used by Preminger et al. (2000). To perform a rating, subjects selected an occasion when they were listening to speech and could understand some but not all of the words. They were asked to listen for a few minutes, switching between P1 and P2. Next, they recorded an estimate of the percentage of speech they understood with each program, and some basic information about the listening situation. This procedure was to be completed twice each day for 14 days, resulting in 28 ratings of each program by each subject. In addition, a checklist was completed daily during this 2-week period to track the extent and variety of the subject's daily listening experiences and to record the daily hours of hearing aid use. The checklist entries were used to verify that the trial was completed according to the protocol.

During session 3 after speech recognition testing, methods for completing the daily ratings and checklists were demonstrated in the laboratory, and the subject was required to show that he or she could perform ratings using the procedure. As a further precaution, the subject was telephoned early in the field trial to assure that there were no problems with completing the ratings. Each subject was given a booklet with daily rating sheets and daily listening checklists, and subjects were reminded of the wearing schedule for the P1 and P2 hearing aid programs. The field trial was conducted during the next 2 weeks.

**Session 4: Speech Recognition Measurement and Preference Interview** • The aided speech recognition measures performed in session 3 were repeated in session 4. In this session, testing began with the hearing aid set to P2, and different lists of words and sentences were used from the CASPA and BKB-SIN tests.

The daily ratings and listening activity checklists completed during the 2-week field trial were collected. The subject was interviewed about his or her preference for the P1 or P2 hearing aid program in everyday listening. The interview included nine verbally delivered questions covering wearing preference for the following: understanding speech in quiet, understanding speech in noise, hearing best over long distances, best sound of own voice, best sound quality, best loudness, best localization, least tiring, and most comfortable sound. After these questions, the subject was asked for his or her preference for using P1 or P2 overall. The subject was then asked to provide (in his or her own words) the three most important reasons for his or her choice of P1 or P2 overall.

## RESULTS

Three types of outcome data were obtained: laboratory speech understanding, real-world ratings of amplified speech, and patient preferences. Statistical analyses were performed using SPSS Version 16.

### Laboratory Speech Understanding

Speech understanding in quiet was measured in the laboratory in sessions 3 and 4 using the CASPA test. These data were used to address the first research question: When listening in

the laboratory to speech in quiet, do subjects perform better, equally, or worse when fit with amplification using an evidence-based prescription method compared with amplification with reduced high-frequency gain, and does the answer to this question differ for listeners with DRs compared with a matched control group without DRs?

Possible CASPA scores for consonants correct ranged from 0 to 100 for each program. Figure 4 depicts the results obtained with the CASPA test. Each symbol illustrates the mean score (both sessions) for the NAL and LP conditions for one subject. If the subject received the same mean score for both programs, their symbol lies on the diagonal. A symbol above the diagonal indicates a higher score for the NAL program and one below the diagonal indicates a higher score for the LP program. Subjects in the control group are depicted with filled triangles, whereas experimental group subjects are represented with open symbols. Experimental subjects with a DR at only one frequency are depicted using squares. Experimental subjects with DRs at two or more frequencies are depicted using circles. Further, the subgroup of individuals in the experimental group whose lowest DRs were measured at 1.0 to 2.0 kHz are further delineated using an 'X' inside the open symbol. The arrow points to the data of subject E04 who is further discussed in the subsequent paragraphs.

Most of the symbols in Figure 4 are either above the diagonal or very close to it. Considering the size and direction of the differences between NAL and LP scores, there is no clear trend among the experimental subjects suggesting different results for those with a DR at one test frequency versus those with DRs at more than one test frequency. Similarly, there is no clear trend

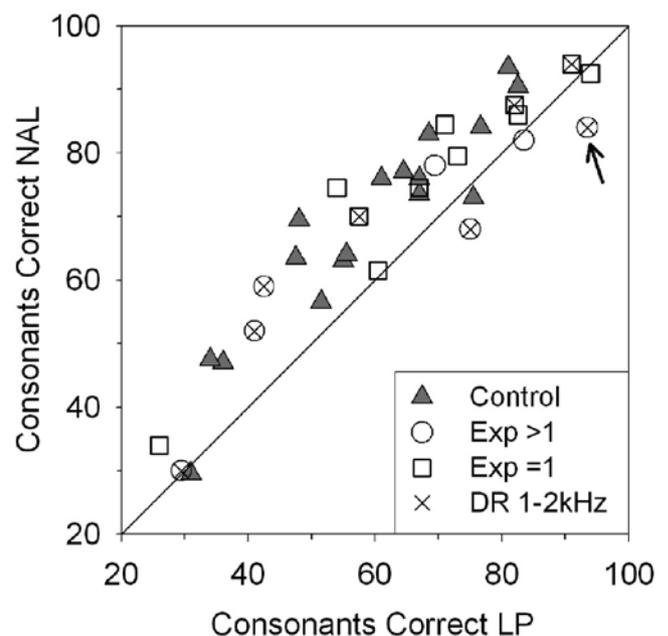


Fig. 4. CASPA scores for the NAL and LP programs for each subject. Subjects in the control group are depicted with filled triangles. Experimental group subjects are depicted with open squares (subjects with DR at one test frequency) or open circles (subjects with DRs at two or more test frequencies). The subgroup of individuals in the experimental group, whose lowest DR was measured at or below 2.0 kHz, are identified with an 'X' inside the open symbol. The arrow points to the data of subject E04. CASPA, Computer-Assisted Speech Perception Assessment; NAL, National Acoustics Laboratories; LP, low pass; DR, dead region.

suggesting different results for those whose lowest DR was in the 1.0 to 2.0 kHz range. For the experimental group, the mean score obtained using the LP program (66.3, SD = 21.0) was less than that using the NAL program (71.8, SD = 18.4). Similarly for the control group, the mean score obtained using the LP program (56.4, SD = 19.1) was less than that obtained using the NAL program (66.6, SD = 18.3).

Statistical questions about the CASPA data were addressed using a three-way mixed model analysis of variance (ANOVA). The variables were session (3 and 4), program (NAL and LP), and group (experimental and control). The main effect of session was not statistically significant,  $F(1, 34) = 0.78, p = 0.38$ . The main effect of group was not statistically significant,  $F(1,34) = 1.42, p = 0.24$ . The main effect of program was statistically significant,  $F(1,34) = 43.96, p < 0.001$ . Also, the Program  $\times$  Group interaction was marginally statistically significant,  $F(1,34) = 4.09, p = 0.051$ , indicating that the superiority of the NAL program was different for the two groups. Post hoc pairwise comparisons (least significant differences) exploring the Program  $\times$  Group interaction revealed that the two groups did not yield significantly different scores for the NAL program ( $p = 0.409$ ) or for the LP program ( $p = 0.148$ ). Further, the NAL program scored significantly higher than the LP program for both the experimental group ( $p = 0.003$ ) and the control group ( $p < 0.001$ ). These results support the pattern observed in Figure 4: when listening to speech in quiet, the NAL program gave better performance than the LP program, and this result was seen for both the experimental and control groups. However, the NAL-LP mean difference was larger for the control group than for the experimental group.

Speech understanding in noise was measured in the laboratory in sessions 3 and 4 using the BKB-SIN test. These data were used to address the second research question: When listening in the laboratory to speech in noise, do subjects perform better, equally, or worse when fit with amplification using an evidence-based prescription method compared with amplification with reduced high-frequency gain, and does the answer to this question differ for listeners with DRs compared with a matched control group without DRs?

Possible BKB-SIN mean scores for words correct ranged from 0 to 31 for each program. Figure 5 depicts the results obtained with the BKB-SIN test. Each symbol illustrates the mean score (both sessions) for the NAL and LP conditions for one subject. If the subject received the same mean score for both programs, the symbol lies on the diagonal. A symbol above the diagonal indicates a higher score for the NAL program and one below the diagonal indicates a higher score for the LP program. Subjects in the control group are depicted with filled triangles, whereas experimental group subjects are represented with open symbols. Experimental subjects with a DR at only one frequency are depicted using squares. Experimental subjects with DRs at two or more frequencies are depicted using circles. Further, the subgroup of individuals in the experimental group whose lowest DRs were measured at 1.0 to 2.0 kHz are further delineated using an 'X' inside the open symbol. The arrow points to the data of subject E04 who is further discussed in the subsequent paragraphs.

Most of the symbols in Figure 5 are slightly above the diagonal. Considering the size and direction of differences between NAL and LP scores, there is no clear trend among the experimental subjects suggesting different results for those with a DR at one test frequency versus those with DRs at more than

one test frequency. Similarly, there is no clear trend suggesting different results for those whose lowest DR was in the 1.0 to 2.0 kHz range. For the experimental group, the mean score obtained using the LP program (18.1, SD = 4.3) was less than that obtained using the NAL program (18.7, SD = 3.7). Similarly for the control group, the mean score obtained using the LP program (17.0, SD = 4.0) was less than that using the NAL program (18.7, SD = 3.6).

Statistical questions about the BKB-SIN data were addressed using a three-way mixed model ANOVA. The variables were session (3 and 4), program (NAL and LP), and group (experimental and control). Exploration of the data characteristics revealed that there were significant deviations from the normal distribution. These issues were corrected by applying a log10 transformation to the BKB-SIN data, and the ANOVA used the transformed data. The main effect of session was not statistically significant,  $F(1, 34) = 0.05, p = 0.824$ . The main effect of group was not statistically significant,  $F(1, 34) = 0.301, p = 0.587$ . The main effect of program was statistically significant,  $F(1, 34) = 20.33, p < 0.001$ . Also, the Program  $\times$  Group interaction was significant,  $F(1,34) = 6.24, p = 0.017$ . The interaction was explored using post hoc pairwise comparisons, and these indicated that the NAL program scored significantly higher than the LP program for the control group ( $p < 0.001$ ), but not for the experimental group ( $p = 0.164$ ). As seen with the CASPA test, the two groups did not yield significantly different scores for the NAL program ( $p = 0.963$ ) or for the LP program ( $p = 0.324$ ).

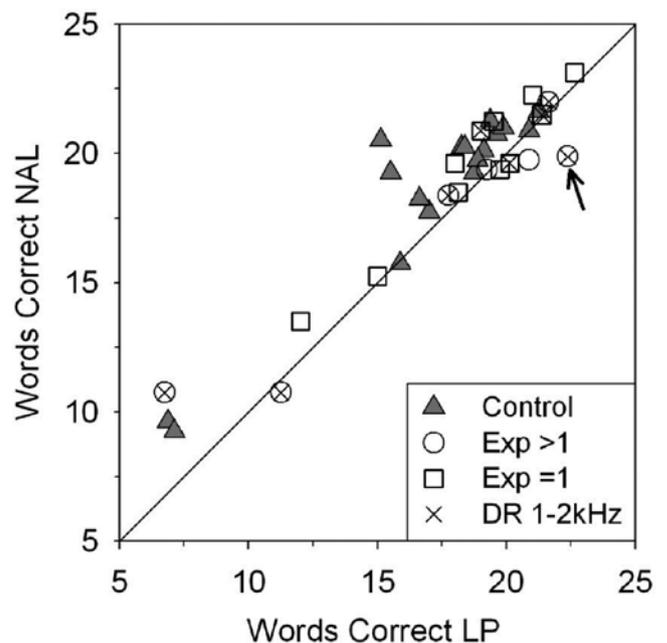


Fig. 5. BKB-SIN scores (untransformed) for the NAL and LP programs for each subject. Subjects in the control group are depicted with filled triangles. Experimental group subjects are depicted with open squares (subjects with DR at one test frequency) or open circles (subjects with DRs at two or more test frequencies). The subgroup of individuals in the experimental group, whose lowest DR was measured at or below 2.0 kHz, are identified with an 'X' inside the open symbol. The arrow points to the data of subject E04. BKB-SIN score, Bamford-Kowell-Bench Speech In Noise; NAL, National Acoustics Laboratories; LP, low pass; DR, dead region.

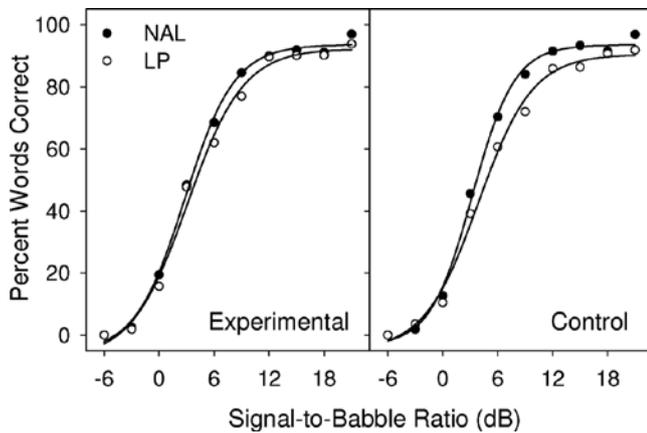


Fig. 6. Performance intensity functions for the NAL and LP programs for the experimental and control groups. NAL, National Acoustics Laboratories; LP, low pass.

An additional analysis was carried out using the data for speech recognition in noise. Because each BKB-SIN list includes sentences presented at 10 different signal to noise ratios, it was possible to construct performance intensity (PI) functions for scores obtained using the LP and NAL programs. Figure 6 depicts the PI functions obtained by fitting the mean score at each SNR using a four-parameter sigmoid function. It is noteworthy that, for both groups, the PI function for the NAL program is steeper than the one for the LP program. This pattern shows that, at a given SNR, speech tended to be more intelligible when amplification provided access to more high-frequency cues. However, the difference between the PI functions is more apparent for the control group than for the experimental group, which is consistent with the Program  $\times$  Group interaction reported earlier.

To further assess the magnitude of the differences between speech recognition scores obtained with the NAL and LP programs, standardized effect sizes (Cohen's  $d$ ) were computed for each group for speech in quiet and in noise (Cohen 1988). The effect sizes are shown in Table 2. The smallest effect size ( $d = 0.13$ ) was seen when individuals with DRs listened to speech in noise. The largest effect size ( $d = 0.55$ ) was seen when individuals without DRs listened to speech in quiet. It is noteworthy that the effect sizes are smaller overall for the listeners with DRs than for those without DRs, and smaller overall for speech in noise than for speech in quiet.

### Real-World Ratings of Amplified Speech

The field trial protocol called for each subject to rate their understanding of speech twice per day in real-world environments. These data were used to address the third research question: When listening in daily life to speech in noise, do subjects rate their performance better, equally, or worse when fit with amplification using an evidence-based prescription method compared with amplification with reduced high-frequency gain, and does the answer to this question differ for listeners with DRs compared with a matched control group without DRs?

Subject E13 experienced repeated difficulties in following the rating protocol, so the ratings from E13 and C13 were excluded from the analyses, leaving a total of 17 pairs for the field trial

TABLE 2. Standardized effect sizes (Cohen's  $d$ ) for differences in speech understanding between the NAL and LP programs for laboratory tests of speech in quiet (CASPA test) and speech in noise (BKB-SIN test)

	Experimental	Control
Speech in quiet	0.28	0.55
Speech in noise	0.13	0.42

Data are given for the experimental and control groups. CASPA, Computer-Assisted Speech Perception Assessment; NAL, National Acoustics Laboratories; LP, low pass.

data. Ratings occurred for 14 consecutive days and subjects also completed a checklist each day. Checklists recorded the subject's hearing aid use that day and documented the different types of listening situations they had experienced (both rated and unrated). Subjects in the experimental group reported mean hearing aid use of 10.5 hr/day ( $SD = 3.24$ ). Subjects in the control group reported mean hearing aid use of 9.2 hr/day ( $SD = 2.83$ ). The difference in reported use between the two groups was not significant,  $t(32) = 1.25$ ,  $p = 0.22$ . Subjects from both groups reported experiencing about six different listening situations per day, on average, including the rated situations.

In each rated situation, the subject recorded the estimated percentage of words he or she understood when using each hearing aid program. The protocol called for 28 ratings of each program, but not every subject submitted this number of valid ratings. Some ratings were missing because the subject could not identify an appropriate listening situation on that day. Other ratings were not used because they were not completely reported (e.g., no indication of talker type or background noise). A total of 891 ratings comprised the final data. The number of ratings provided by each subject ranged from 17 to 29 with a mean of 26. For each subject, a single rating for each program was determined by averaging all the ratings for that program. Figure 7 depicts the pairs of ratings for each subject. Each symbol illustrates the mean ratings for the NAL and LP conditions for one subject. If the subject allotted the same mean rating to both programs, the subject's symbol lies on the diagonal. A symbol above the diagonal indicates a higher rating for the NAL program and one below the diagonal indicates a higher rating for the LP program. Subjects in the control group are depicted with filled triangles, whereas experimental group subjects are represented with open symbols. Experimental subjects with a DR at only one frequency are depicted using squares. Experimental subjects with DRs at two or more frequencies are depicted using circles. Further, the subgroup of individuals in the experimental group whose lowest DRs were measured at 1.0 to 2.0 kHz are further delineated using an 'X' inside the open symbol. The arrow points to the data of subject E04 who is further discussed in the subsequent paragraphs.

Most of the symbols in Figure 7 are either very close to the diagonal or above it. Considering the size and direction of differences between the NAL and LP ratings, there is no clear trend among the experimental subjects suggesting different results for those with a DR at one test frequency versus those with DRs at more than one test frequency. Similarly, there is no clear trend suggesting different results for those whose lowest DR was in the 1.0 to 2.0 kHz range. For the experimental group, the mean rating of the LP program (78.6,  $SD = 10.57$ ) was less than the mean rating of the NAL program (85.14,  $SD = 9.44$ ). Similarly for the control group, the mean rating for the LP program

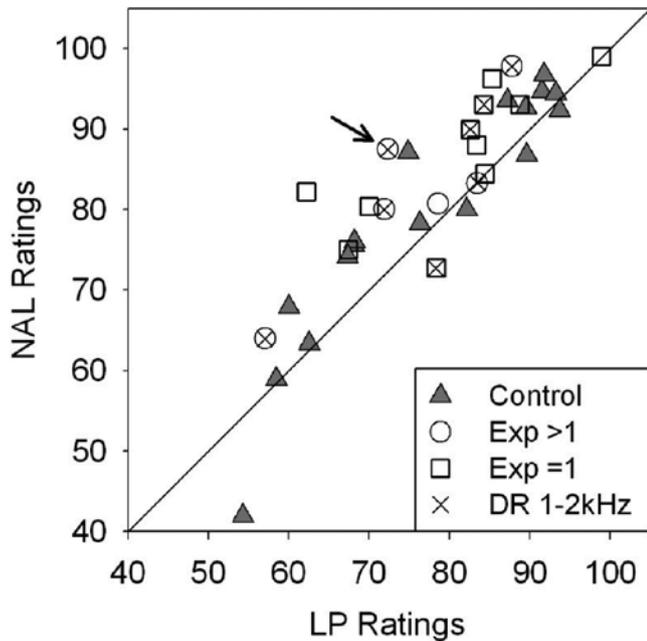


Fig. 7. Mean rating of daily life speech understanding using NAL and LP programs for each subject. Subjects in the control group are depicted with filled triangles. Experimental group subjects are depicted with open squares (subjects with DR at one test frequency) or open circles (subjects with DRs at two or more test frequencies). The subgroup of individuals in the experimental group, whose lowest DR was measured at or below 2.0 kHz, are identified with an 'X' inside the open symbol. The arrow points to the data of subject E04. NAL, National Acoustics Laboratories; LP, low pass; DR, dead region.

(76.98, SD = 13.76) was less than the mean rating for the NAL program (79.67, SD = 15.16).

Statistical questions about the rating data were addressed using a two-way mixed model ANOVA. The variables were program (NAL and LP) and group (experimental and control). The main effect of program was statistically significant,  $F(1,32) = 20.6$ ,  $p < 0.001$ . The main effect of group was not statistically significant,  $F(1, 31) = 0.734$ ,  $p = 0.398$ . The Program  $\times$  Group interaction approached but did not quite reach statistical significance,  $F(1,32) = 3.6$ ,  $p = 0.068$ . Because of its a priori interest to this investigation, the Program  $\times$  Group interaction was explored using post hoc pairwise comparisons. These indicated that the NAL program scored significantly higher than the LP program for the experimental group ( $p < 0.001$ ), and the result for the control group approached but did not quite reach statistical significance ( $p = 0.07$ ). As seen for both the laboratory speech perception tests, the two groups did not yield significantly different scores for the NAL program ( $p = 0.216$ ) or for the LP program ( $p = 0.700$ ).

Standardized effect sizes (Cohen's  $d$ ) were computed to assess the magnitude of the differences between the NAL and LP programs in estimated speech understanding in daily life. These effect sizes were  $d = 0.65$  for listeners with DRs and  $d = 0.19$  for listeners without DRs.

### Patient Preferences

After all data were collected, each subject was interviewed about his or her experiences using the two programs in the experimental hearing aid. The interview covered three topics in the following order: (1) Program preference for each of the nine

criteria, (2) Which of the two programs he or she would prefer for overall everyday use, and (3) Up to three reasons for the expressed overall preference. These data were used to address the fourth research question: Considering overall listening in daily life, do subjects prefer amplification using an evidence-based prescription method or amplification with reduced high-frequency gain, and does the answer to this question differ for listeners with DRs compared with a matched control group without DRs?

Across both groups 23 subjects preferred the NAL program and 11 subjects preferred the LP program. When asked how certain they were about their preference, 31 subjects were either "reasonably certain" or "very certain." The three remaining subjects were "slightly uncertain." None was "very uncertain." The major concern for this investigation was whether subjects with DRs were more likely than their partners without DRs to prefer the LP program. To explore this issue, an analysis of the data was performed using the McNemar test for matched pairs. There were nine pairs in which both individuals preferred the NAL program and three pairs in which both preferred the LP program. For four pairs, the control subject preferred the LP program whereas the experimental subject preferred the NAL program. There was one pair in which the control subject preferred the NAL program and the experimental subject preferred the LP program. The results of the McNemar test indicated that the preference for the LP program over the NAL program was not significantly different for the two groups ( $p = 0.375$ ). Thus, although most of the subjects of both groups preferred the NAL program overall, a noteworthy number of individuals preferred the LP program, and this preference was not associated with the presence of DRs.

Because the presence of DRs did not seem to contribute to subjects' overall preferences for the NAL or LP program, it was of interest to further explore the data in an attempt to gain some insight into the basis for the preferences. For these explorations, the two groups were combined. One possibility was that subjects preferred the program that they rated as best for speech understanding in daily life. An assessment of this possibility is shown in Figure 8. The figure replicates the rating data from Figure 7 but the symbols now depict each subject's overall preference for the NAL or LP program. If each subject's preference was consistent with his or her daily life ratings, all the symbols above the diagonal would be filled circles and all those below the diagonal would be open circles. Many subjects did choose the program that they rated as highest for speech understanding in challenging daily life situations. However, the presence of five open circles above the diagonal and a filled circle below the diagonal shows that not all subjects made this decision.

In a further attempt to appreciate the basis for subjects' preferences, we examined the relationships between their overall preference and their responses to the nine interview questions about preference relative to specific criteria. These data are summarized in Figure 9. For each criterion, the figure depicts the number of subjects who chose the NAL program, who had no preference, and who chose the LP program. In addition, the contingency coefficient computed between the preference for that criterion and the overall preference across the 34 subjects is displayed on the right side of Figure 9. This coefficient is an indicator of the strength of the association between the two variables. There are several noteworthy aspects of Figure 9: (1) The two criteria with the strongest relationship to overall preference were sound quality and speech in noise. (2) The NAL program

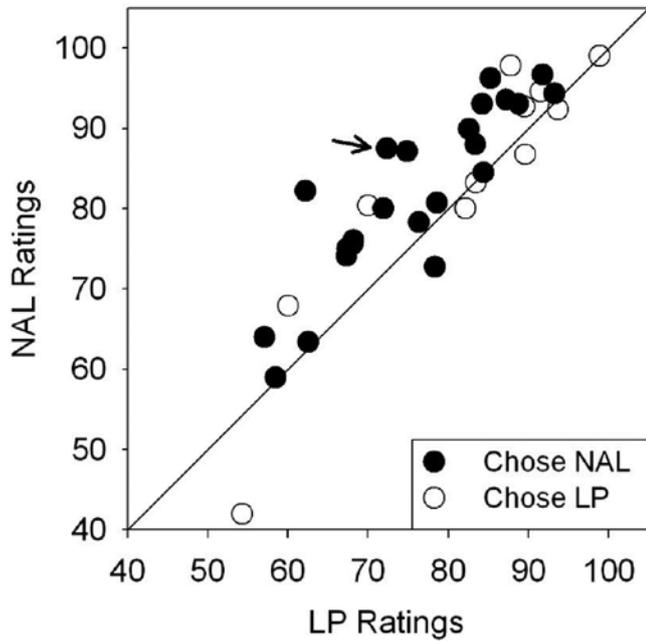


Fig. 8. Rating data from Figure 7 but with the symbols showing each subject's overall preference for the NAL or LP program. The arrow points to the data of subject E04. NAL, National Acoustics Laboratories; LP, low pass.

was substantially preferred over the LP program for speech in quiet, but this preference was not strongly related to the overall preference. (3) The LP program was preferred more often than the NAL program for only two criteria—own voice and less tiring.

At the end of the interview, subjects were asked to give reasons for their preference in their own words. The reasons given were subjected to a content analysis to derive overall themes (Krippendorff 2004) and the results are shown in Figure 10. There are clear differences in the themes identified as reasons for choosing each program. The most frequently volunteered reason for choosing the NAL program related to greater speech clarity, whereas this theme was not often invoked by subjects who preferred the LP program. The most frequently volunteered reason for choosing the LP program was that the other program (i.e., NAL) was too loud, whereas this idea was very rarely mentioned by subjects who preferred the NAL program. Issues related to sound quality and softness of the other program were mentioned about equally by subjects with both preferences.

**DISCUSSION**

We explored the effects of reducing high-frequency amplification for typical hearing aid patients who tested positive for high-frequency DRs at one or more test frequencies. This type of adjustment has sometimes been recommended in a nonspecific way that calls for reducing high-frequency gain in general for patients with high-frequency DRs. Other researchers have recommended more targeted gain reductions (Vickers et al. 2001; Baer et al. 2002) and gain reductions that would apply only to patients with DRs at or below 2.0 kHz (Vinay and Moore, 2007a). Our research design involved 18 matched pairs of subjects with each pair comprising one person without measured DRs and the other person with at least one measured DR at frequencies between 1.0 and 4.0 kHz. The number and frequencies of DRs among the experimental

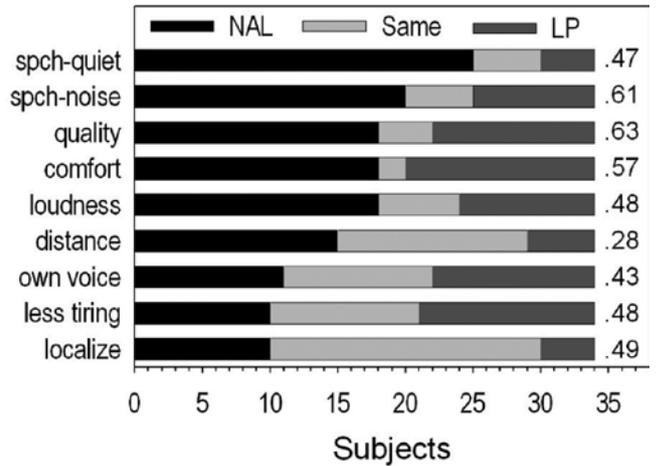


Fig. 9. Summary of responses to the nine interview questions about preference relative to specific criteria. For each criterion, the figure depicts the number of subjects who chose the NAL program, who had no preference, and who chose the LP program. The contingency coefficient between the subject's preference for that criterion and the subject's overall preference, which was computed across the 34 subjects, is displayed on the right side. NAL, National Acoustics Laboratories; LP, low pass.

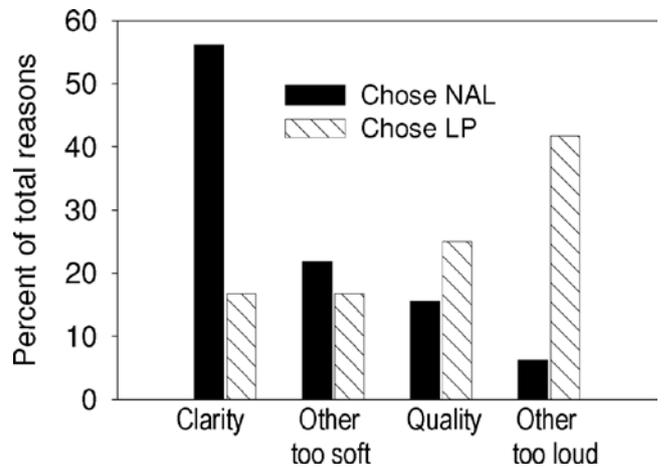


Fig. 10. Overall themes derived from reasons given by subjects for their preference for the NAL or LP program. Themes are shown separately for subjects who preferred the NAL program and subjects who preferred the LP program. NAL, National Acoustics Laboratories; LP, low pass.

subjects were representative of that seen by Cox et al. (2011) in a large group of typical hearing aid patients. In the experimental group, 8 subjects had their lowest measured DRs between 1.0 and 2.0 kHz, and the remaining 10 had their lowest measured DRs at 3.0 or 4.0 kHz. Eleven subjects had a DR at only one test frequency and only two subjects had DRs at more than two test frequencies. Objective measures of speech understanding were obtained in quiet and in noise in a laboratory setting. Subjective ratings of speech understanding were obtained in daily life environments. Last, each subject was interviewed about which program they preferred to use in the particular circumstances of their daily lives.

**Speech Understanding in the Laboratory and in the Real World**

Most subjects achieved better speech understanding in a quiet laboratory when they used the NAL program rather than

the LP program (Fig. 4). This difference was statistically significant for both groups of listeners. However, the magnitude of improvement obtained with additional high-frequency cues, measured by the standardized effect size, was small for the listeners with DRs ( $d = 0.28$ ) but moderate for those without DRs ( $d = 0.55$ ). In other words, additional high-frequency cues were useful for both groups of listeners, but they were more useful for those who did not have DRs.

In the laboratory test, most subjects achieved better speech understanding in noise when they used the NAL program rather than the LP program (Fig. 5). However, the differences between the two programs were generally less in noise than in quiet. The difference was statistically significant for listeners without DRs but not for listeners with DRs. Consistent with this result, the standardized effect of additional high-frequency gain on speech understanding in noise was very small for the listeners with DRs ( $d = 0.13$ ) but still moderate for those without DRs ( $d = 0.42$ ). In other words, as expected, additional high-frequency cues in noise were quite useful for listeners without DRs. For subjects with DRs, additional high-frequency cues were of minimal value when listening to speech in noise, however, they did not reduce performance (Fig. 6).

In real-world listening environments where they could understand some but not all of the words, it was seen that subjects tended to give a higher rating to their own speech understanding when they listened using the NAL program (Fig. 7). The difference achieved or approached statistical significance for both subject groups, in contrast with the laboratory result in noise where only the control group showed a significant improvement with the NAL program. Further, in real-world listening, the advantage conferred by added high-frequency cues, measured by the standardized effect size, was much larger for the listeners with DRs ( $d = 0.65$ ) than for those without DRs ( $d = 0.19$ ). Perhaps the authenticity of combining natural speech, real environments, and visual cues in the real-world ratings was responsible for this outcome. In any event, this finding, which would not have been predicted from the laboratory test results in noise, underscores the importance of conducting effectiveness research with patients functioning in their everyday environments. It cannot validly be assumed that performance seen in a laboratory setting will also be observed in daily life.

Vickers et al. (2001) and Baer et al. (2002) examined the benefits of high-frequency amplification in quiet and noise, respectively, for listeners having severe to profound hearing loss with and without high-frequency DRs. Both studies determined that, in that group of listeners, those with DRs often did not benefit from increased audibility of high-frequency speech cues, and both suggested that this finding had important implications for hearing aid fitting for individuals with high-frequency DRs. In addition, Vinay and Moore (2007a) recommended that gain adjustments for DRs would be most appropriate for listeners with DRs measured at or below 2.0 kHz. Our study was undertaken to explore the application of the recommendations of Moore et al., as well as more sweeping recommendations from others, by attempting to implement them in fittings with typical hearing aid users (with mild to moderately severe hearing loss) wearing real hearing aids. Cox et al. (2011) found that, although the TEN(HL) test was positive for DRs in about one-third of typical hearing aid users, the DRs were usually found in only one or two test frequen-

cies. Thus, the experimental subjects in the present study did not have the extensive contiguous DRs that were characteristic of the subjects studied by Moore et al. Generally, the results of our study indicated that typical hearing aid users wearing real hearing aids did benefit from increased audibility of high-frequency speech cues in most situations, and never performed more poorly with them. Thus, the recommendation to reduce high-frequency gain for persons with high-frequency DRs does not seem to apply to this type of patient.

The results of our study combining laboratory and field data generally reinforce and extend the studies of Mackersie et al. (2004) and Preminger et al. (2005). Both previous studies used laboratory tests to explore the implications of DRs for hearing aid fitting with subjects who were typical hearing aid users, and both reached the same conclusions as we report, namely, that hearing aid wearers benefited from high-frequency amplification regardless of their DR status. Mackersie et al. also noted that in high noise levels, high-frequency gain conferred less benefit for listeners with DRs than for those without DRs. Our laboratory data are consistent with this finding (Fig. 6). However, this trend was not seen when data were collected in real-world environments.

Our study outcomes are not consistent with those of Gordo and Iorio (2007), who found that hearing aid wearers with DRs actually performed more poorly with high-frequency amplification than without it. The differences between our study outcomes and those of Gordo and Iorio probably resulted from a combination of audiogram differences and hearing aid fitting differences. Given that our study used a matched-group design, that is, subjects with hearing loss that is typical of hearing aid wearers, and real hearing aid fittings used in real-world environments, our study offers high ecological validity.

It is reasonable to ask whether this outcome is dependent on the number of test frequencies that are positive for DRs. Much of the previous research on this topic used subjects who had DRs that were contiguous at three or more test frequencies, whereas only subjects E04 and E05 in our study fulfilled that requirement (Table 1). Of these two individuals, subject E04 had the more extensive DRs, and this subject's results are indicated with an arrow in Figures 4, 5, 7, and 8. E04 performed noticeably better with the LP program than with the NAL program in the laboratory tests (Figs. 4 and 5) but better with the NAL program in the daily life trial (Fig. 7), and ultimately selected the NAL program as preferred (Fig. 8). Thus, subject E04's laboratory results support a recommendation to reduce high frequency gain when there are extensive DRs, however, in real-world listening this subject preferred the NAL program and performed better with it.

It is also noteworthy that DRs in three or more test frequencies is an unusual finding among typical hearing aid candidates. In Cox et al. (2011), 83% of ears with measured DRs had DRs at one to two test frequencies, and 17% had DRs at three to four test frequencies. The pattern of multiple DRs in the experimental subjects in the present study (89% had DRs at one to two test frequencies and 11% had DRs at three to four test frequencies) was similar to that seen in Cox et al. In summary, our research indicates the following points: (1) measured DRs at three or more high frequencies are rarely seen in typical hearing aid candidates, and (2) typical hearing aid patients who have DRs in one or two high-frequency regions are still likely to benefit from high-frequency gain in daily life.

## IMPLICATIONS FOR CLINICAL PRACTICE

The implications of this research are limited to patients with adult-onset hearing loss similar to that of typical hearing aid wearers (mild to moderately severe extent and flat or sloping configuration). The laboratory and real-world outcomes of this study indicate that a positive result for the TEN test at one or two frequencies between 1.0 and 4.0 kHz does not call for an a priori modification in that patient's target prescription for amplification: A patient with one or two high-frequency DRs seems likely to benefit from high-frequency gain in much the same way as one without DRs. Whether the same implication can be drawn for persons who have DRs at three or more high frequencies is not certain because there were only two in this research. In addition, our data do not allow any statement about the implications for hearing aid fitting of extensive DRs in severe to profound hearing loss, or low-frequency DRs (however, see Summers 2004; Vinay & Moore 2007b; Vinay et al. 2008).

The results of this study continue to strengthen the evidence base that supports the use of a generic gain prescription similar to the NAL method when the goal is to maximize speech perception (Mueller 2005). Figures 2 and 3 illustrate the average differences in high-frequency gain between the NAL and LP programs. It is striking that, when audibility limitations imposed by thresholds are taken into account, these differences do not seem to be large. Nevertheless, they resulted in perceptible and significant differences in speech understanding for both groups of subjects. Clearly, even small differences in high-frequency gain can have consequences for speech understanding in everyday life.

On the other hand, it is also evident that speech understanding, although a major contributor to decisions, was not the only important consideration in everyday listening. About one-third of the subjects expressed a preference for the LP program at the end of the study. Only one of these gave the LP program substantially higher ratings for speech understanding in daily life, and several gave the NAL program substantially higher ratings (Fig. 8). Furthermore, the presence or absence of DRs was not related to the preference for the LP program in daily life.

It would be valuable to understand why some subjects apparently preferred the program that did not yield the best speech understanding. Examination of Figure 9 tends to support conventional wisdom that sound quality and speech understanding in noise are the primary elements in preference decisions, because these individual criteria had the strongest relationship to overall preference. However, examination of Figure 10 yields another important insight: the dominant theme for preferring the LP program was that the NAL program was judged to be too loud. This was unexpected because, during the verification/fine-tuning process, the MPO of each program was adjusted using current best-practice procedures to avoid discomfort, and every subject judged the two programs to be equally loud for conversational speech. Nevertheless, although the two programs were equally loud for a low-frequency dominated signal (speech) it seems likely, in retrospect, that they were not equally loud for other environmental sounds that contained more high-frequency energy. In addition, although both programs were adjusted for loudness comfort during the fitting process, the maximum output for the NAL program was higher than that for the LP program in the high frequencies. This difference could also

contribute to the overall impression of the NAL program being louder, or too loud, for some environmental sounds. Because the levels and the relationship between the NAL and LP programs were fixed in this study, it is not certain that individual preferences for the LP program would have been observed if those subjects had been able to control the volume or if further fine-tuning had been allowed.

## Other Considerations

This investigation was a nonrandomized clinical trial. Because it is not possible to randomly assign subjects to groups in this sort of study, there is always the possibility that the experimental and control groups were not equivalent in some important but unappreciated way(s), despite the matching criteria. However, we have not identified any threats to the validity of the outcomes.

Each subject wore one hearing aid in this study because switching programs on a single device could be accomplished more accurately by subjects during the daily life ratings. Would the result have been different if subjects had worn two hearing aids? It seems quite likely that the results of providing more (or less) high-frequency cues in one ear of a listener would be similar to the results of making the same change in both ears of that same listener. However, it is unknown whether binaural interactions would be different for the experimental and control groups and whether this might affect the relative benefits of high-frequency cues. In addition, it is noteworthy that unilateral listening was assured in the laboratory tests by plugging and muffing the unaided ear whereas the participation of the unaided ear in the daily life ratings is unknown (however, the sensitivity of the unaided ear was either equal to or worse than that of the aided ear).

It is also of interest to consider whether the activation of the directional microphone or the digital noise reduction features would have affected the results of the study. These features were disabled to make sure that there would be no post hoc uncertainty about whether they might have changed the relationship between the NAL and LP programs depending on the dynamic environments encountered by the subjects during the field trial. If directional processing had been activated, it might have promoted better speech understanding in any given listening environment. However, this effect would apply equally to the two programs and is unlikely to have changed the result. Enabling the digital noise reduction feature would not have altered the speech-understanding results. On the other hand, activation of digital noise reduction in noisy non-speech environments might have reduced the loudness difference between the two programs by reducing high-frequency output more for the NAL program than for the LP program. It is conceivable that this could have reduced the overall preference for the LP program by some subjects who found the NAL program too loud.

In conclusion, although hearing aid fittings in this study used current best-practice fitting procedures, the gain adjustments made to accommodate each subject's DRs were relatively crude. Exactly how much gain reduction to use, where to place it, and how to shape it were issues that were determined partly by the consideration of existing recommendations and partly by the technical functioning of contemporary real hearing aids. It is not known whether a different result would be obtained if the

frequency gain function was more precisely customized based on the location of the measured DRs.

## CONCLUSIONS

On the basis of this research, we recommend against limiting high-frequency gain prescription solely because a patient has DRs in one or two high-frequency regions. Considering the significance tests, the effect sizes (Table 2), and the PI functions (Fig. 6), the results indicate that when listening to highly controlled speech in a quiet laboratory setting, typical hearing aid patients with DRs benefited from high-frequency cues, although not as much as those without DRs. Further, when listening to highly controlled speech in a noisy laboratory setting, typical hearing aid patients with DRs did not benefit significantly from high-frequency cues but did not perform more poorly, whereas those without DRs did benefit. In addition, when listening to everyday speech in typical environments, typical hearing aid patients with DRs benefited from high-frequency cues at least as much as those without DRs. It is important to note that we did not observe any situation in which typical hearing aid patients with DRs were penalized by the provision of high-frequency gain similar to the NAL prescription: either they benefited from more high-frequency gain or they performed no worse with it.

This result does not necessarily imply that all typical hearing aid patients will prefer the additional high-frequency gain suggested by a prescription similar to NAL, even when they are aware that this type of fitting optimizes speech understanding. About one-third of subjects expressed an overall preference for the LP program, which provided less high-frequency gain. The main reason for preferring the LP program overall was an impression that the NAL program was too loud. This outcome reaffirms the primacy of achieving acceptable loudness for all types of sounds during hearing aid fittings.

## ACKNOWLEDGMENTS

This research was supported by funding from National Institutes of Health, National Institute on Deafness and Other Communication Disorders grant R01DC006222: "Optimizing Hearing Aid Fitting for Older Adults." Hearing aids for the study were provided by Starkey Hearing Aids, Minneapolis, Minnesota.

Address for correspondence: Robyn M. Cox, 807 Jefferson Avenue, Memphis, TN 38105, USA. E-mail: robyncox@memphis.edu

Received March 1, 2011; accepted January 25, 2012.

## REFERENCES

- Amos, N. E., & Humes, L. E. (2001). The contribution of high frequencies to speech recognition in sensorineural hearing loss. In D. J. Breebart, A. J. Houtsma, A. Kohlrausch, V. F. Puijs & R. Schoonhoven (Eds). *Physiological and Psychophysical Bases of Auditory Function* (pp. 437–444). Maastricht, The Netherlands: Shaker Publishing, BV.
- Baer, T., Moore, B. C., Kluk, K. (2002). Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. *J Acoust Soc Am*, 112(3 Pt 1), 1133–1144.
- Bentler, R. A. (2005). Effectiveness of directional microphones and noise reduction schemes in hearing aids: A systematic review of the evidence. *J Am Acad Audiol*, 16, 473–484.
- Byrne, D., Dillon, H., Ching, T., et al. (2001). NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures. *J Am Acad Audiol*, 12, 37–51.
- Ching, T. Y., Dillon, H., Byrne, D. (1998). Speech recognition of hearing-impaired listeners: predictions from audibility and the limited role of high-frequency amplification. *J Acoust Soc Am*, 103, 1128–1140.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cox, R. M., & Alexander, G. C. (1995). The abbreviated profile of hearing aid benefit. *Ear Hear*, 16, 176–186.
- Cox, R. M., Alexander, G. C., Johnson, J., et al. (2011). Cochlear dead regions in typical hearing aid candidates: prevalence and implications for use of high-frequency speech cues. *Ear Hear*, 32, 339–348.
- Davies-Venn, E., Souza, P., Brennan, M., et al. (2009). Effects of audibility and multichannel wide dynamic range compression on consonant recognition for listeners with severe hearing loss. *Ear Hear*, 30, 494–504.
- EtymoticResearch. (2005). *BKB-SIN Speech in Noise Test, Version 1.03*. Elk Grove Village, IL: Etymotic Research.
- Gordo, A., & Iorio, M. C. M. (2007). Dead regions in the cochlea at high frequencies: 884 Implications for the adaptation to hearing aids. *Rev Bras Otorinolaringol*, 73, 299–307.
- Hogan, C. A., & Turner, C. W. (1998). High-frequency audibility: benefits for hearing-impaired listeners. *J Acoust Soc Am*, 104, 432–441.
- Hornsby, B. W. Y., & Dundas, J. A. (2009). Factors affecting outcomes on the TEN (SPL) test in adults with hearing loss. *J Am Acad Audiol*, 20, 251–263.
- Horwitz, A. R., Ahlstrom, J. B., Dubno, J. R. (2008). Factors affecting the benefits of high-frequency amplification. *J Speech Lang Hear Res*, 51, 798–813.
- Killion, M. C., & Niquette, P. A. (2000). What can the pure-tone audiogram tell us about a patient's SNR loss? *Hearing J*, 53, 46–53.
- Killion, M. C., Niquette, P. A., Gudmundsen, G. I., et al. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, 116(4 Pt 1), 2395–2405.
- Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology* (2nd ed.). Thousand Oaks, CA: Sage.
- Larson, V. D., Williams, D. W., Henderson, W. G., et al. (2000). Efficacy of 3 commonly used hearing aid circuits: A crossover trial. NIDCD/VA Hearing Aid Clinical Trial Group. *JAMA*, 284, 1806–1813.
- Ling, D., & Ling, A. H. (1978). *Aural Habilitation: The Foundations of Verbal Learning in Hearing-Impaired Children*. Washington, DC: A. G. Bell Association for the Deaf.
- Lippmann, R. P. (1996). Accurate consonant perception without mid-frequency speech energy. *IEEE Trans Speech Audio Processing*, 4, 66–69.
- Mackersie, C. L., Boothroyd, A., Minniear, D. (2001). Evaluation of the Computer-Assisted Speech Perception Assessment Test (CASPA). *J Am Acad Audiol*, 12, 390–396.
- Mackersie, C. L., Crocker, T. L., Davis, R. A. (2004). Limiting high-frequency hearing aid gain in listeners with and without suspected cochlear dead regions. *J Am Acad Audiol*, 15, 498–507.
- Malicka, A. N., Munro, K. J., Baker, R. J. (2010). Diagnosing cochlear dead regions in children. *Ear Hear*, 31, 238–246.
- Moore, B. C. J. (2001a). Dead regions in the cochlea: Diagnosis, perceptual consequences, and implications for the fitting of hearing aids. *Trends Amplif*, 5, 1–34.
- Moore, B. C. J. (2001b). *Dead regions in the cochlea: Implications for the choice of high-frequency amplification*. In R. C. Seewald & J. S. Gravel (Eds.), *A Sound Foundation Through Early Amplification* (pp. 153–166). Stafa, Switzerland: Phonak AG.
- Moore, B. C., Glasberg, B. R., Stone, M. A. (1999a). Use of a loudness model for hearing aid fitting: III. A general method for deriving initial fittings for hearing aids with multi-channel compression. *Br J Audiol*, 33, 241–258.
- Moore, B. C., Glasberg, B. R., Stone, M. A. (2004). New version of the TEN test with calibrations in dB HL. *Ear Hear*, 25, 478–487.
- Moore, B. C., Glasberg, B. R., Vickers, D. A. (1999b). Further evaluation of a model of loudness perception applied to cochlear hearing loss. *J Acoust Soc Am*, 106, 898–907.
- Moore, B. C., Huss, M., Vickers, D. A., et al. (2000). A test for the diagnosis of dead regions in the cochlea. *Br J Audiol*, 34, 205–224.
- Mueller, H. G. (2005). Fitting hearing aids to adults using prescriptive methods: an evidence-based review of effectiveness. *J Am Acad Audiol*, 16, 448–460.
- Padilha, C., Garcia, M. V., Costa, M. J. (2007). Diagnosing cochlear "dead" regions and its importance in the auditory rehabilitation process. *Braz J Otorinolaringol*, 73, 556–561.
- Plyler, P. N., & Fleck, E. L. (2006). The effects of high-frequency amplification on the objective and subjective performance of hearing instrument users with varying degrees of high-frequency hearing loss. *J Speech Lang Hear Res*, 49, 616–627.

- Preminger, J. E., Carpenter, R., Ziegler, C. H. (2005). A clinical perspective on cochlear dead regions: Intelligibility of speech and subjective hearing aid benefit. *J Am Acad Audiol*, *16*, 600–13.
- Preminger, J. E., Neuman, A. C., Bakke, M. H., et al. (2000). An examination of the practicality of the simplex procedure. *Ear Hear*, *21*, 177–193.
- Schuknecht, H. F. (1993). *Pathology of the Ear* (2nd ed.). Philadelphia, PA: Lea & Febiger.
- Seewald, R., Moodie, S., Scollie, S., et al. (2005). The DSL method for pediatric hearing instrument fitting: historical perspective and current issues. *Trends Amplif*, *9*, 145–157.
- Shannon, R. V., Galvin, J. J. III, Baskent, D. (2002). Holes in hearing. *J Assoc Res Otolaryngol*, *3*, 185–199.
- Summers, V. (2004). Do tests for cochlear dead regions provide important information for fitting hearing aids? *J Acoust Soc Am*, *115*, 1420–1423.
- Summers, V., Molis, M. R., Müsch, H., et al. (2003). Identifying dead regions in the cochlea: psychophysical tuning curves and tone detection in threshold-equalizing noise. *Ear Hear*, *24*, 133–142.
- Turner, C. (1999). The limits of high-frequency amplification. *Hearing J*, *52*, 10–14.
- Turner, C., & Cummings, K. J. (1999). Speech audibility for listeners with high-frequency hearing loss. *Am J Audiol*, *8*, 47–56.
- Turner, C., & Henry, B. A. (2002). Benefits of amplification for speech recognition in background noise. *J Acoust Soc Am*, *112*, 1675–1680.
- Turner, C., Humes, L., Bentler, R., et al. (1996). A review of past research on changes in hearing aid benefit over time [see comments]. *Ear Hear*, *17*, 14S–25S.
- Vickers, D. A., Moore, B. C., Baer, T. (2001). Effects of low-pass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies. *J Acoust Soc Am*, *110*, 1164–1175.
- Vinay, B. T., & Moore, B. C. (2008). Speech recognition in noise as a function of highpass-filter cutoff frequency for people with and without low-frequency cochlear dead regions. *J Acoust Soc Am*, *123*, 606–609.
- Vinay, & Moore, B. C. J. (2007a). Prevalence of dead regions in subjects with sensorineural hearing loss. *Ear Hear*, *28*, 231–241.
- Vinay, & Moore, B. C. J. (2007b). Speech recognition as a function of high-pass filter cutoff frequency for people with and without low-frequency cochlear dead regions. *J Acoust Soc Am*, *122*, 542–553.
- Yanz, J. L. (2002). Cochlear dead regions and practical implications for high frequency amplification. *Hearing Review*, *9*, 58–63.