Interpreting the efficacy of frequency-lowering algorithms

By Francis Kuk, Denise Keenan, Jane Auriemmo, Petri Korhonen, Heidi Peeters, Chi Lau, and Bryan Crose

Despite a long history of research and commercial efforts,[1] hearing aids with frequency-lowering algorithms have become popular only recently. Their lack of commercial success may be attributed in part to the immaturity of analog technology when these devices were introduced such that artifacts were plentiful. But insufficient training provided to the wearers of such devices, unrealistic expectations, and inadequate means to evaluate their efficacy are equally important contributors to the limited acceptance for this technology.

Widex re-introduced the concept of linear frequency transposition in its Inteo hearing aid in 2006 under the name Audibility Extender.[2] Since then, we have explored various avenues to better understand how such a feature can be fitted[3,4] and its use facilitated.[5] Just as important, we also studied (and developed) research tools that may be optimal for evaluating such an algorithm. Our effort led us to report on the efficacy of such an algorithm in a simulated hearing loss,[6] in an open-tube fitting,[7] in children,[8] and in adults in quiet and in noise.[9] We have learned that demonstrating the efficacy of a frequency-lowering algorithm is not a straightforward matter. We would like to share our experience in this paper.

WHAT IS AN ADEQUATE EVALUATION?

Any evaluation must begin with the assumption that the optimal settings have been verified. That is, the hearing aid settings are indeed providing the *potential* for the added audibility. Verification does not guarantee that the Audibility Extender (AE) algorithm is helpful and/or produces meaningful changes. It simply verifies that the algorithm results in "optimal" acoustic changes that are needed for improved performance. If the settings are not optimal, one cannot draw any conclusion on the efficacy of the processing algorithm. See Kuk et al.[3] for a brief description on how to verify the AE settings using the SoundTracker.

An adequate evaluation informs the clinician (and patient) how changes in the acoustic condition (i.e., from using the AE) affect the patient's performance. However, performance is not an all-or-nothing phenomenon. While one may not notice any substantive changes on one task, changes may occur on other tasks. Indeed, one lesson we have learned is that a battery of tests varying in the cognitive/integrative demand is necessary to evaluate the efficacy of the AE.

THE EVALUATION BATTERY

The auditory system is capable of handling tasks of varying complexity. Auditory signals must be detectable (or audible) before they can be discriminated. However, being discriminable does not guarantee that they can be identified. Thus, evaluation needs to be inclusive and multi-leveled. Otherwise, an incomplete picture of the efficacy may result.

Detection

Detection is the most fundamental aspect of hearing. It reflects the wearer's ability to sense the presence or absence of a sound. An example of a detection task is the aided sound field threshold. It measures the softest level at which a patient can hear the intended sound.[10] Almost all the studies that evaluated the efficacy of frequency-lowering algorithms reported an improvement in aided thresholds. For example, Figure 1 shows the improvement in aided thresholds with the use of AE over the "no-AE" or master program in school-aged children.[8] Note that without the AE, five children did not respond at 4000 Hz, and the average aided threshold was 70 dB HL even with the master program (left). On the other hand, with the AE, every child (n=10) responded with an average threshold at 35 dB HL at 4000 Hz (right).

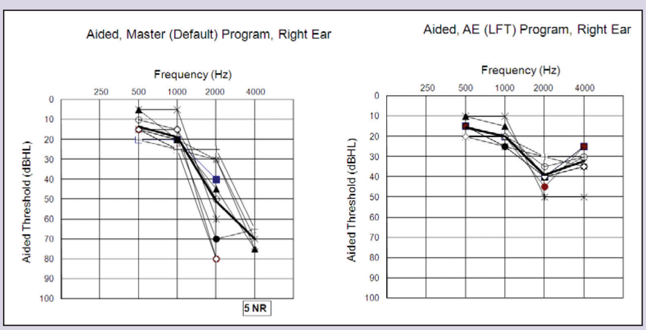

Figure 1. *Individual and averaged aided sound-field thresholds with the master program (left) and the AE program (right) on ten hearing-impaired children. NR stands for no response (from Auriemmo et al.,[7]).*

Bear in mind that improvement in aided threshold does not mean that patients can hear the 4000-Hz tone as a 4000-Hz tone. They hear the 4000-Hz tone as a lower frequency, 2000 Hz in the case of the AE. It will be a different tone with other frequency-lowering

Kitchen:	29. Phone ring (in other room)	53. Hearing aid held in hand
1. Dripping faucet	30. Phone button tones	54. Snap closures
2. Gas stove ignition	***Family Room:***	***Sounds in Public:***
3. Microwave buttons	31. Squeaky furniture	55. Elevator bell
4. Microwave alarm	32. Door hinge	56. Coins jingling
5. Cracking ice	33. Ticking clock	57. Wet shoes on tile
6. Aluminum foil	34. Coo-coo clock	58. Shopping carts
7. Candy wrappers	35. Adjusting window blinds	59. Instruments (ex: piccolo)
8. Whistling tea kettle	36. Fire crackling in fireplace	60. Cash register printing receipt
9. Plastic bag/wrap	37. Door bell	61. Music in stores over intercom (doctor's office/waiting area)
10. Dripping faucet	38. Door locking	62. High heel shoes on hard floor
Dining:	39. Rain on the roof	***Sounds outdoors:***
11. Silverware	40. Various light switches	63. Cracking ice
12. Spoon stirring a drink	a. Pull chain	64. Rustling leaves
13. Ice clinking in a glass	b. Standard flip switch	65. Birds
14. Glasses clinking for a toast	***Children:***	66. Crickets
15. Tapping glass with a spoon	41. Child's squeaky toy	67. Whistling
16. Hitting chop-sticks together	42. Child's toy that plays music or beeps	68. Wind chimes
17. Soda fizzing	***Pets:***	69. Bicycle bell
18. Other people eating	43. Pet toe nails on tile	***Car sounds:***
19. Chair scraping the floor	44. Pet collar tags	70. Car turn signal
Office:	45. Whining dog	71. Left key in ignition with door open (warning ding)
20. Keyboard buttons	46. Meowing cat	72. Door locking
21. Rustling paper	47. Pet toys with bell or squeak	73. Screeching tires
22. Paperclips on hard surface	***Personal:***	74. Checking air in tires (hiss)
23. Clicking a pen	48. Clothes rustling	75. Washing window or mirror
24. Scissors	49. Brushing hair	76. Shaking keys
25. Stapler	50. Jewelry (ring on hard surface, noisy bracelet, etc.)	
26. Computer sounds	51. Watch alarm	
27. Computer mouse click	52. Tapping fingernails	
28. Phone ring (in same room)		

Figure 2. *Widex developed this checklist to be used to sensitize hearing aid wearers' attention to some sounds in their daily environments.*

schemes. The important point is that *information* carried by the 4000-Hz tone is now decoded by the neurons responsible for the 2000-Hz (or other lower frequencies) auditory fibers.

An improvement in sound detection can also be assured when patients report increased awareness of sounds in their daily environments. Often patients comment on a "fuller range of sounds," "feeling less fatigue," and "increased comfort" with the use of the AE program. These are reflections of the benefits of frequency lowering as more sounds become available. These benefits should be documented as indications of improved sound detection. This level of benefit may be immediately measurable with the use of the AE.

We have developed a checklist to measure improvement in sound awareness. This list (Figure 2) identifies 76 sounds that typically occur in everyday listening situations. After the fitting and verification of the AE (or any other hearing aids), we give this checklist to the patient and ask him or her to check off the sounds that are audible without the hearing aid, with the hearing aid in the master program, and with the hearing aid in the AE program. A difference in the number of sounds heard between the unaided and the aided conditions reflects the improvement offered by that hearing aid condition. Furthermore, using this checklist can help heighten the patient's awareness to sounds in general (and to the AE in particular).

Discrimination

Discrimination refers to the ability to tell if two sounds are different from each other. This usually involves a "same-different" paradigm in which the wearer is presented with two sounds and asked to judge if they are the same or different. A variant of this is seen in tasks that ask the wearer if the singular form of a word (e.g., cat) is different from its plural form (e.g., cats). A discrimination task does not require the wearer to identify the meaning of the sound, simply to indicate if the sounds are different from each other.

While good performance on a discrimination task reflects the patient's ability to use the newfound audibility (from amplification and frequency lowering) to discern differences between sounds, it does not guarantee that the person can identify either of the sounds or use that difference for higher processing. As evidence, some earlier research found improvement in discrimination tasks but not on identification ability.[11] For that reason, we focused our effort on the higher level of processing/testing (i.e., identification) instead of on discrimination tasks.

Discrimination tasks often use speech stimuli and are therefore sometimes called "speech discrimination tests." However, in clinical audiology, speech-identification/-recognition tests such as the W-22 and NU-6 have also been referred to as speech-discrimination tests. For that reason, it could be confusing to call a discrimination task using speech materials a "speech-discrimination test." Clinicians must understand the nature of the evaluation task and not assume that all speech tests require identification skills, since they may require only discrimination ability.

Identification

The ability of hearing aid wearers to identify sounds requires sufficient audibility of the necessary cues that are *keys* to identifying the target sound. In addition, the wearer must be able to match the sound(s) to his/her repertoire. This can become an issue in evaluating the efficacy of frequency-lowering devices in that the improved audibility may be "foreign" to the wearer and delay identification of the sound.

Simple sounds (i.e., those with simple spectral and temporal structures) are easier to relearn and identify. But more complex sounds such as speech may be more difficult to identify even though audibility is guaranteed. Sometimes, a subjective change in sound quality accompanies the improvement in sound identification. There are many considerations when using speech-identification tests as a measure of the efficacy of frequency lowering.

CONSIDERATIONS IN CURRENT SPEECH TESTS

Think high-frequency sounds

Frequency-lowering algorithms work to ensure the audibility of *high-frequency sounds* that are either inaudible or unusable when amplified with conventional means. Given today's advances in hearing aids, such as a broader bandwidth and active feedback cancellation, the hearing loss in the high frequencies must be severe to profound for frequency lowering to provide unequivocal benefit. In today's terminology, the high frequencies should be "dead" so the recoding of the high-frequency sounds is the only way to gain access to them. This means the evaluation materials must include sufficient information in the high frequencies to provide a basis for evaluating the efficacy of frequency lowering.

High-frequency content varies by gender

The typical speech tests now used in the clinic are monosyllabic words presented in quiet or sentence tests presented in noise. Although these tests may be suitable for a general evaluation of word comprehension, they may not be optimal for the evaluation of frequency-lowering devices for several reasons.

The first reason is that most of these tests are spoken by a male speaker (although female versions of some tests are available). In general, male speakers tend to have a lower fundamental frequency than female speakers. This results in a spectrum with restricted output in the high frequencies. Figure 3 compares the spectrum of the /s/ phoneme spoken by a male and a female. One can see for the male, the dominant energy is around 5000 Hz, whereas for the female it is around 8000 Hz. This means that a patient with aidable hearing up to 4000 Hz may be able to identify the /s/ phoneme when spoken by a male, but have more difficulty with /s/ when it is spoken by a female. Thus, speech-identification tests using a male speaker may not reflect the benefits of a frequency-lowering algorithm. Therefore, a female version may better demonstrate the algorithm's efficacy.

High-frequency consonants occur less frequently

A second reason for not using typical speech tests is the frequency of occurrence of high-frequency speech sounds. Popular speech tests such as the NU-6[12] and the W-22[13] are phonetically balanced. This means that the frequency with which these phonemes occur in the test reflects their occurrence in real life. While this is reasonable for a speech test designed to reflect daily communication difficulty, it is by definition limiting sounds that do not occur as frequently as others.

This is the case for fricatives, which tend to be higher in frequency and to occur less frequently than other sounds. For example, Denes showed that the frequency of occurrence of the /n/ was 11% in real-life whereas for fricatives like /s/ and /ʃ/ it was 6.9% and 1.1%, respectively.[14] Consequently, the percentage of these phonemes in a phonetically balanced test will be relatively low. Even though the wearer may benefit from lowering the higher-frequency consonant sounds (such as /s/, /ʃ/ etc.), the fact that they are a small percentage of the total number of sounds in the typical speech test could mask such a benefit in the overall score. This may make any observable benefit appear insignificant.

The likely improvement in speech-identification scores may be estimated by low-pass filtering the speech materials. Figure 4 shows the absolute performance from a group of normal-hearing individuals on a nonsense syllable test as a function of cut-off frequencies. Both male and female speech materials were included. As one can see, the change in performance from a broadband condition to a low-pass condition with a cutoff at 4000 Hz changed the overall consonant score by less than 5%. This magnitude of difference is within the typical test-retest variability of most speech tests.

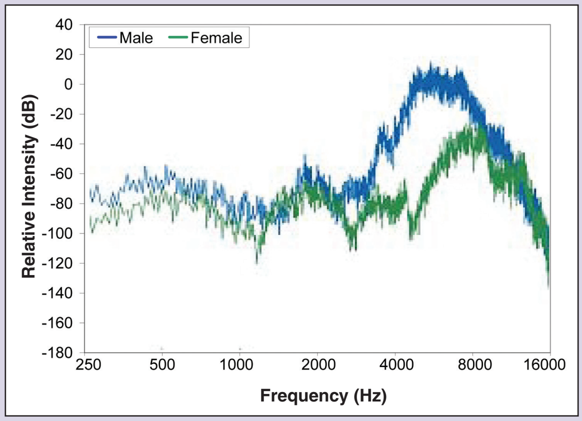

Figure 3. *Spectral difference between male and female production of the /s/ phoneme.*

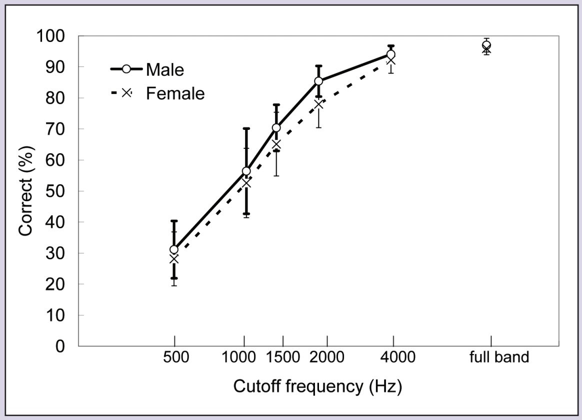

Figure 4. *Absolute consonant-identification scores as a function of low-frequency cutoff.*

With a cutoff at 2000 Hz, the male materials showed an 11% drop whereas the female materials showed a 16% drop. This means that unless the wearer has a substantial degree of hearing loss at or below 2000 Hz or unless the test evaluates only high-frequency consonant sounds, conventional speech tests are unlikely to reveal a substantial degree of improvement in identification score from frequency lowering.

Benefits may be offset

Because frequency lowering alters the natural spectral content of the input, it is possible that different phonemes are affected differently by its actions (e.g., Braida et al.[1]). This suggests that the increased identification of some phonemes may be offset by the potential decreased identification of others. The net result would be no or little improvement in speech understanding. No improvement may mean an ineffective algorithm or it may mean an effective algorithm in which the positive changes are offset by the negative changes. The latter case would suggest that the frequency lowering can be beneficial. One need only identify the sounds that may benefit and those that may not in order to set proper expectations. The information is also helpful in developing better algorithms. Unfortunately, current speech tests do not offer that level of analysis.

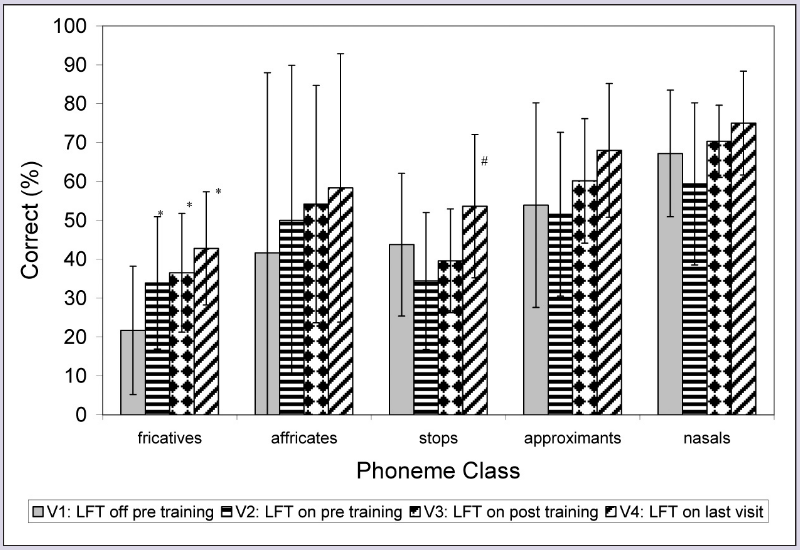

Figure 5. *Absolute identification scores on the ORCA-NST for five phoneme classes at four hearing aid intervals: v1 – master or no AE (LFT) at initial fit; v2 – AE at initial fit; v3 – AE at one month use; and v4 – AE at 2 month use. The magnitude of the standard deviation is included (from Kuk et al.[9]).*

AN ANALYTIC NONSENSE SYLLABLE TEST

In view of such a difficulty, we devised a nonsense bisyllable test (ORCA-NST), which has a consonant-vowel-consonant-vowel-consonant (CVCVC) format. This is an open-set test that contains all 23 English consonants, with each appearing at least once in the initial, medial, and final word positions (unless prohibited). There are a total of 115 items on the list (in random order). A female and a male speaker version of the test, in full (115 items) and shortened (32 items), are available. A detailed description of the test, including normative data, is found in Kuk et al.[15] One advantage is that at the end of the test a display of the scores for each phoneme class is immediately available.

We evaluated the efficacy of the AE on eight hearing-impaired adults with a severe to profound high-frequency hearing loss using the ORCA-NST.[9] Figure 5 shows the performance on each phoneme class (fricatives, affricates, stops, nasals, approximants) at four hearing aid time intervals (visit 1 - master or no AE, visit 2 - AE baseline, visit 3 - AE at 1 month with training, and visit 4 - AE at 2 months without training). One can see that for fricatives and affricates the benefit of the AE was immediate. There was a 10% improvement in fricative identification even at the initial fit (without any training).

Furthermore, the benefit in fricative and affricate identification improved over time.

For other phoneme classes, notably stops, there was an initial decline in identification, which reverted itself over time. For example, there was a 10% decrease of stop consonants over the master program at initial fit. That decrease narrowed over time. After 2 months, the identification of the stop consonants with the AE program was higher than with the master program. This was also true for the other consonants. This shows that a speech test that allows analysis of phoneme errors could offer more insights into how frequency lowering may have affected speech perception at the fundamental level.

The ORCA-NST was developed primarily for research purposes, but the findings from using that test allowed us to formulate a simple nonsense syllable list for clinical evaluation of frequency-lowering algorithms. While this list (see Figure 6) does not provide all the information available from the ORCA-NST, it does offer a quick sampling of the speech sounds most likely to be affected by frequency lowering.

The list includes only voiced and voiceless fricatives and stops appearing at the initial and final word positions. These phonemes are included because they showed the most changes in response to frequency lowering in our studies. Minimally, the unvoiced syllables should be administered live-voice or in recorded form at a soft level (e.g., 30 dB HL) and at a conversational level (50 dB HL). Testing should be done with the wearers' own hearing aids (or unaided) as well as with the frequency-lowering aids with and without the algorithm activated. Testing should be conducted at the initial fitting, as well as over time (1 month, 3 months, and 6 months) to document any acclimatization.

MAKING THE RIGHT COMPARISON

This brings up an important detail that must not be overlooked. Some studies showed frequency-lowering algorithms resulting in substantial improvement in speech scores. However, the comparison was made to wearers' own hearing aids and not to the frequency-lowering hearing aids with the frequency-lowering algorithm deactivated. Although the comparison to the wearers' own hearing aids may be impressive, it fails to control for differences in other aspects of the hearing aids.

For example, there may be a difference between the instruments in overall bandwidths, number of channels, compression parameters, distortion levels, and other algorithms such as noise reduction, directional microphones, etc. A difference in any of these parameters could lead to a substantial difference in performance that is *not* attributable to the frequency-lowering algorithm per se. To determine if the algorithm works, the

CV (voiced)	CV (unvoiced)	VC (voiced)	VC (unvoiced)
bee	pee	eeb	eep
dee	tee	eed	eet
gee	kee	eeg	eek
vee	fee	eev	eef
THee (ð)	thee (θ)	eeTH (ð)	eeth (θ)
zee	see	eez	ees
zhee (ʒ)	shee (ʃ)	eezh (ʒ)	eesh (ʃ)
dgee (dʒ)	chee (tʃ)	eedg (dʒ)	eech (tʃ)

Figure 6. *Simplified high-frequency syllable test: voiced and voiceless fricatives and stops that may be used to evaluate frequency-lowering algorithms.*

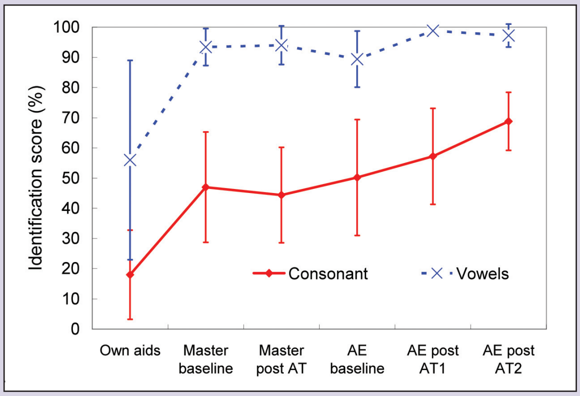

Figure 7. *Comparison of identification scores (vowels – blue dashes; consonants – red line) on a nonsense syllable test between subjects' own aids, Inteo Master and Inteo AE program over time. (from Auriemmo et al.[8]).*

comparison must be made on the same hearing aid with the frequency-lowering algorithm turned "on" and "off."

An example that demonstrates the importance of the comparison condition is seen in Auriemmo et al.[14] In that study, identification performance on a nonsense syllable test was made between the children's own digital hearing aids and the Inteo hearing aid with transposition "off" (master) and "on" (AE). Figure 7 shows the identification scores at a 30-dB HL presentation level over time. If one compares the performance of the children's own aids (first interval) with the AE at baseline (fourth interval), one will conclude that AE improves consonant and vowel identification by over 30%. On the other hand, if one compares the own aid with the Inteo hearing aid in the master's baseline setting (second interval), one sees almost the same 30% improvement. This means that the noted improvement between the AE at baseline and the subjects' own aids was not the result of AE alone. Rather, it was because of a hearing aid (Inteo) better than the subjects' own instruments.

The benefit of the AE alone was minimal when one compared the AE baseline (fourth interval) and the master (second and third intervals). The benefit of the AE did not become apparent until after the children had worn the AE program for 3 weeks (post AT1). After 6 weeks of AE use (post AT2), the average overall improvement in consonant score from the master program was 20%!

EFFECT OF TRAINING AND/OR ACCLIMATIZATION

One observation from Figure 5 was that while using the AE resulted in an immediate improvement with fricatives, stops showed a 10% decrease on the nonsense syllable test during the initial comparison. This confirms speculation that the actions of the AE could result in confusion of speech cues.

On the other hand, an examination of the performance of the AE after a month of use (with training) shows improvement in identification of *all* phoneme classes. This includes sounds such as fricatives where there was initial benefit from the AE and sounds such as stops where the initial effect of the AE was negative. Indeed, some of the errors on stop consonants were corrected at the 1-month evaluation interval. Wearing the AE for an additional month resulted in an improvement in stop phoneme identification beyond the master program. This shows that the initial confusion at the phoneme level may be transitory in nature.

Whether or not the noted change reflects a change in neural/cortical reorganization from the difference in acoustic cues remains unclear. It is also unclear what the exact time course of this change is or if it is critical that training be provided. In Kuk et al. 2009,[8] self-paced, interactive training was provided only for the first month of AE use and no additional training was provided for the second month of AE use—and yet performance continued to improve.[9]

Clinical implications

The observation that performance with the AE improved over time has important clinical implications. First, it stresses the importance of experience (or time) with the AE to fully realize its potential.

A second implication relates to the initial acceptance for frequency lowering. Because AE moves the high frequencies to a lower frequency area, the initial sound quality may be unnatural. A common clinical response when a patient has a negative reaction is to readjust the hearing aids to a more pleasant sound quality or to de-activate frequency lowering to preserve the original speech. Unfortunately, this is the exact opposite of what needs to be done.

What clinicians and hearing aid wearers must realize is that the unnaturalness resulting from the additional cues will become a natural part of the wearer's sound repertoire over time. The sound quality will become more acceptable and speech understanding will improve, but only if the wearer continues to receive the additional cues by using the AE at the assigned settings. Reducing or eliminating the transposition because the patient doesn't like the initial sound quality misses the opportunity of improving the patient's speech skills further. The clinician, once satisfied with the fitting through its verification, should counsel the patient to use the recommended settings for maximum benefit.

A third clinical implication is the timing of the efficacy studies. Our studies showed that AE benefits are best observed after wearers had worn the devices for at least 1 month. This means

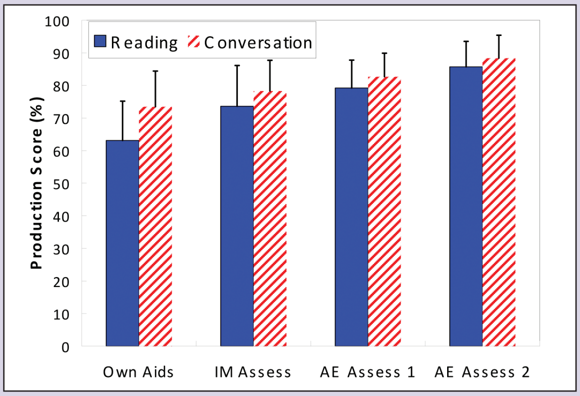

Figure 8. */s/ and /z/ production accuracy on a reading task and a conversation task with the children's own hearing aids, Inteo master (IM assess), AE at 3 weeks (assess 1), and AE at 6 weeks (assess 2) (from Auriemmo et al.[8]).*

that the evaluation of AE efficacy is an on-going process and cannot be completed at the initial fitting.

SPEECH PRODUCTION

The benefit of frequency transposition extends beyond the perceptual realm. In Auriemmo et al., benefits with the AE were seen in the accuracy of the children's articulation of the /s/ and /z/ phonemes in a contextual background (a reading task and a conversation task).[8] Figure 8 shows the articulation accuracy of the /s/ and /z/ production in a reading task and in a conversational task with the children's own hearing aids (own aids), the Inteo Master program (IM assess), and the AE at 3 weeks post-fitting (AE assess 1), and the AE at 6 weeks post-fitting (AE assess 2).

The increased audible cues on the fricative sounds provided the children the necessary cues to monitor their own speech production on high-frequency consonant sounds. Thus, in the clinical evaluation of AE, one should not restrict oneself to direct, immediate evaluation measures, but should also look at indirect measures that may also depend on the availability of audibility cues. A measure on speech-production skills in children is a worthwhile and necessary measure.

CONCLUSIONS

We have gained significant knowledge on how to evaluate the benefits of frequency-lowering algorithms. When evaluating such an algorithm, it is important first to verify that the AE settings are optimal before pursuing any additional evaluative measures. It is also recognized that performance improves over time, and the wearer's initial reaction must be interpreted with caution. When validating benefit, it is important to recognize that benefit is hierarchical.

Our limited, conventional view of benefit that examines only speech-perception measures must be expanded to include evaluative measures on detection, discrimination, identification, and speech production to be comprehensive.

Francis Kuk, PhD, is Director of Audiology at Widex/ORCA-USA. **Denise Keenan**, MA; **Chi Lau**, PhD; and **Heidi Peeters**, MA, are Research Audiologists at ORCA. **Jane Auriemmo**, AuD, is Manager of the Pediatric Program at Widex USA. **Petri Korhonen**, MS, and **Bryan Crose**, BS. are Research Engineers at ORCA. Readers may contact Dr. Kuk at fkuk@widexmail.com or at fkuk@aol.com.

REFERENCES

1. Braida L, Durlach I, Lippman P, et al.: *Hearing Aids: A Review of Past Research of Linear Amplification, Amplitude Compression, and Frequency Lowering*, ASHA Monographs number 19. Rockville, MD: American Speech-Language-Hearing Association, 1979.
2. Kuk F, Korhonen P, Peeters H, et al.: Linear frequency transposition: Extending the audibility of high frequency information. *Hear Rev* 2006;13(10):42-48.
3. Kuk F, Keenan D, Peeters H, et al.: Critical factors in ensuring efficacy of frequency transposition I: individualizing the start frequency. *Hear Rev* 2007a;14(3):60-67.
4. Auriemmo J, Kuk F, Stenger P: Criteria for evaluating linear frequency transposition in children. *Hear J* 2008;61(4):50-54.
5. Kuk F, Keenan D, Peeters H, et al.: Critical factors in ensuring efficacy of frequency transposition II: facilitating initial adjustment. *Hear Rev* 2007b;14(4):90-96.
6. Korhonen P, Kuk F: Use of linear frequency transposition in simulated hearing loss. *JAAA* 2008; 19(10):639-650.
7. Kuk F, Peeters H, Keenan D, Lau C: Use of frequency transposition in thin-tube, open-ear fittings. *Hear J* 2007c;60(4):59-63.
8. Auriemmo J, Kuk F, Lau C, et al.: Effect of linear frequency transposition on speech recognition and production in school-age children. *JAAA* 2009; 20(5):289-305.
9. Kuk F, Keenan D, Korhonen P, Lau C: Efficacy of linear frequency transposition on consonant identification in quiet and in noise. *JAAA* 2009; 20(8):465-479.
10. Kuk F, Ludvigsen C: Reconsidering the concept of the aided threshold for nonlinear hearing aids. *Trends Amplif* 2003;7(3):77-97.
11. Rees R, Velmans M: The effect of frequency transposition on the untrained auditory discrimination of congenitally deaf children. *Brit J Audiol* 1993; 27:53-60.
12. Tillman TW, Carhart R: *An Expanded Set for Speech Discrimination Utilizing CNC Monosyllabic Words: Northwestern University Auditory Test No. 6. Technical Report No. SAM-TR-66-55.* San Antonio: USAF School of Aerospace Medicine, Brooks Air Force Base, 1966.
13. Hirsh IJ, Davis H, Silverman SR, et al.: Development of materials for speech audiometry. *J Sp Hear Dis* 1952;17:321-337.
14. Denes PB: On the statistics of spoken English. *J Acoust Soc Am* 1963;35(6):892-904.
15. Kuk F, Lau C, Korhonen P, et al.: Development of the ORCA Nonsense Syllable Test. 2010 (submitted).