CHAPTER TEN

Electroacoustic Verification Measures with Modern Hearing Instrument Technology

Susan D. Scollie and Richard C. Seewald

Introduction

Modern-day hearing instruments offer a wide variety of signal processing options, including wide-dynamic range compression (WDRC), multiple channels, selectable compression speeds, dual/directional microphones, and noise detection. Similarly, modern-day hearing instrument test systems offer a wide variety of test signals, varying in their bandwidth, spectral shape, level, and temporal characteristics. The process of verifying the performance of a given hearing instrument for a specific infant or child is therefore complicated by the process of matching appropriate test signal characteristics to the signal processing characteristics of the hearing instrument, in order to obtain an accurate test. In this chapter, we will review some current issues in electroacoustic verification of modern-day hearing instruments and we will outline some evidence-based strategies for verification with modern hearing instrument technology.

The Purpose of Electroacoustic Verification

Recent consensus statements from expert audiologists have suggested that verification is necessary to evaluate whether the hearing instrument is providing appropriate levels of audibility for speech and appropriate levels of output limiting for loud sounds (Pediatric Working Group 1996). As shown in figure 1, these measures generally address two main purposes.

First, verification measures should allow the clinician to know, as accurately as possible, that the hearing instrument fitting will facilitate auditory development. This is accomplished in three steps. First, the clinician must choose a prescriptive formula that has been shown to provide appropriate speech intelligibility, loudness perception, comfort, and real-world benefit in previous evaluations of children with comparable hearing. Second, prescriptive targets for hearing instrument performance are generated using the selected prescriptive formula, for specific types and levels of test signals. Finally, the clinician uses the same types and levels of test signals to measure the performance of the hearing instrument, and sets the instrument to meet the prescriptive targets as closely as possible. Through this process, the experimental treatment (i.e., the prescribed frequency response and level) that has been shown to provide benefit is systematically replicated by the clinician for each infant or young child fitted. This allows a reasonable expectation that auditory development will be facilitated through amplification.

A second main purpose of verification is to provide a reasonable substitute for immediate feedback from the hearing instrument user. Older children and

Two purposes of electroacoustic verification:

- To know, as accurately as possible, that the hearing aid fitting will ***facilitate*** auditory development.
- To provide a reasonable ***substitute*** for immediate feedback from the wearer.

Figure 1. Purposes of electroacoustic verification in pediatric hearing instrument fitting.

Address correspondence to: Susan D. Scollie, M.Cl.Sc., Research Audiologist, Child Amplification Laboratory, National Centre for Audiology, University of Western Ontario, London, ON, Canada, N6G 1H1

Characteristics of a good verification protocol:

- It tells us how the hearing aid processes *speech.*
- It tells us the levels at which *output is limited.*
- It is an *efficient, reliable, and valid* procedure.
- It can be used with *infants.*
- It is *meaningful.*

Figure 2. Characteristics of a verification protocol with universal applicability in pediatric hearing instrument fitting.

adults can communicate with the clinician to indicate if the sound quality and loudness of the fitting are acceptable, through informal communication and/or formal testing of speech perception, loudness perception, or preference. For infants and young children, a substitute for this type of testing is to verify that the hearing instrument provides a prescribed frequency response and level that has been shown to provide comfort, acceptable loudness, and/or preference in past studies. This allows a reasonable expectation that the child's auditory experiences when using the hearing instrument will not be aversive.

For either of these purposes to be accomplished, a sufficiently accurate and detailed, yet clinically feasible, verification protocol must be employed. The characteristics that define such a protocol are listed in figure 2, and will be discussed separately in the following sections.

Verification of Speech Processing Versus Output Limiting

It is essential that verification procedures separately evaluate hearing instrument performance for amplification of speech and limiting of maximum output. Speech processing and output limiting are determined by independent factors in hearing instruments. Therefore, any one test of hearing instrument performance for speech inputs will not provide information about the instrument's maximum output and vice versa. This implies that a minimum of two verification tests is required. One test should assess hearing instrument performance for speech-level inputs, while the other should assess the maximum output of the instrument. It is well established that the maximum output must be assessed only with a high-level (e.g., 90 dB SPL or greater), narrowband test signal (Pediatric Working Group 1996; ASHA 1998). In contrast, speech-level performance can be evaluated using either narrowband or broadband stimuli, provided appropriate steps are taken to ensure test accuracy. These steps will be described later in this chapter.

Efficiency, Reliability and Validity

Evaluation of speech processing and output limiting is preferably conducted with procedures that are clinically efficient, reliable, and valid. Efficiency is determined by the time taken to complete the procedure. Typically, electroacoustic measurements are very rapid, ranging from about 1 to 15 seconds in length once test equipment has been set up. Reliability is determined by the difference that is expected by two consecutive tests of the same type. In infants and young children, probe microphone measurements have test-retest reliability of about 2 dB, when carried out by an experienced clinician (Sinclair et al. 1996). This was not significantly different from the test-retest reliability for the same real ear measurement in adults (Sinclair et al. 1996). Validity reflects the measurement's ability to accurately determine hearing instrument performance without errors due to signal processing, signal type, or noise. To summarize, good verification protocols should incorporate measures with high levels of efficiency, reliability and validity.

Compatibility with Infants

In the present era of newly developing universal infant hearing screening programs, development of clinical protocols must take into consideration a wider range of ages and developmental levels than was previously relevant. Even five years ago, hearing instrument fitting prior to 6 months of age was a relatively rare occurrence. Now, many clinicians consider hearing instrument fitting at three to four months of age to be relatively routine. Therefore, the design of an optimal verification protocol must take into account the unique characteristics of the infant population. Measurement procedures that are not compatible with the skill set or capabilities of infants must be modified.

Meaningful Verification

When possible, hearing instrument verification should be designed to provide as much meaning as

possible. This requires the use of test procedures that predict functional benefit with hearing instruments. Test procedures should therefore evaluate hearing instrument performance against measures of residual auditory capacity. Valid comparisons of unaided versus aided hearing, and calculations of the Speech Intelligibility Index (ANSI 1997a) are more meaningful than are simple decibel estimates of hearing instrument gain. Graphical displays that show the relationships between unaided hearing and aided hearing instrument output are more meaningful than graphical displays of hearing instrument output alone. Several authors have described verification schemes that may be used as counseling tools, provided that meaningful graphical displays are employed (Harrison 2000; Stelmachowicz 2000). These electroacoustic-based counseling tools are particularly useful and appropriate in infant-toddler hearing instrument fitting, because they can serve as predictive substitutes for subjective feedback from the hearing instrument user.

General Classification of Verification Procedures

Historically, two types of hearing instrument verification procedures have been applied with infants and children and can be classified as (1) audiometric-based and (2) electroacoustic-based. As we evaluate the relative strengths and limitations of each class of verification procedures, it is important to keep in mind the purpose of this stage of the hearing instrument fitting process. At the verification stage, the question that we are attempting to answer is: To what extent does the measured or predicted real-ear performance of the hearing instrument under consideration conform to the desired performance characteristics for the infant/child we are fitting with amplification? Thus, before entering into the verification stage it is assumed that the clinician has determined, on some theoretical basis, what the desired performance characteristics are for a given infant/child. Without such specification of desired performance characteristics, it is simply impossible to answer the question that is being asked at the verification stage. Simply stated, one enters into the verification stage knowing what they want the hearing instrument to do in theory (i.e., electroacoustic selection) and then applies a set of measures (i.e., verification) to determine the extent to which the desired real-ear hearing instrument performance has been achieved. The relative strengths and limitations associated with both audiometric-based and electroacoustic-based procedures, for pediatric applications, are considered within the sections to follow. For a more extensive treatment of this topic, the reader is referred to a chapter by Seewald, Moodie, Sinclair and Cornelisse (1996).

Audiometric-Based Verification Procedures

For many years, sound field aided threshold measures (see figure 3) have been the primary means for verifying hearing instrument performance with infants and children. Over time, clinicians have used the results of sound field aided threshold testing for a variety of purposes. For example, sound field

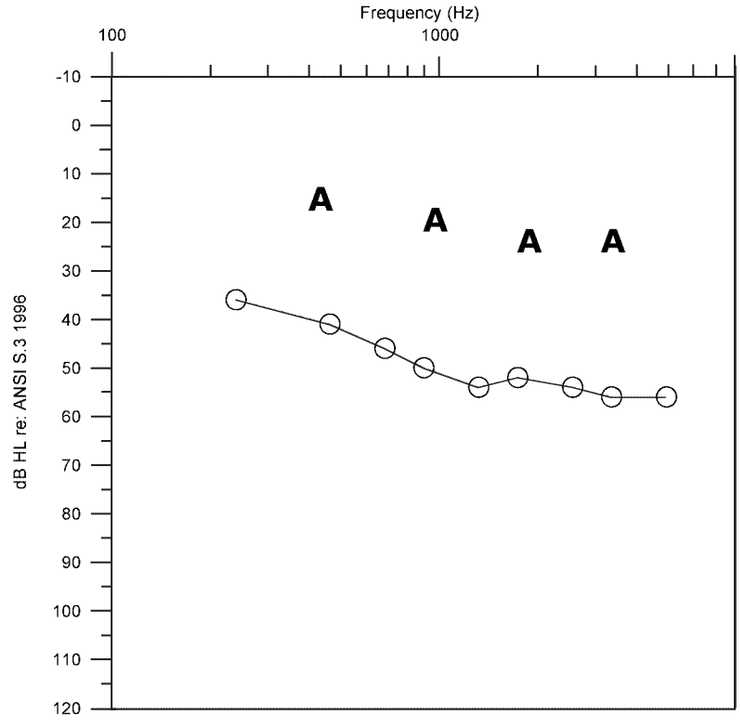

Figure 3. Example of an aided audiogram.

aided thresholds have been used to (1) directly measure some quantity of interest (e.g., detection level by frequency); (2) compare aided performance under different amplification conditions (e.g., performance with hearing instrument "A" versus performance with hearing instrument "B"); and (3) predict the levels at which amplified speech will be received within the aided condition. In reality, the verification of aided performance with infants and children, using aided threshold measures, is often approached with some unique combination of all of the above purposes simultaneously.

Assuming that a valid measurement of aided thresholds can be obtained, the results of these measurements provide clinicians with an estimate of the level of sound that can be detected within the audiometric test environment and have at least one advantage relative to the alternative electroacoustic-based procedures. Specifically, the results of sound field aided threshold testing do provide some information about the infant's/child's ability to detect the presence of sound within the aided condition. Thus, these measure do provide some information about *auditory performance* with amplification.

However, as we consider using sound field aided threshold measures for the purpose of verifying the real-ear performance of a hearing instrument for an infant or child, we need to acknowledge several limitations that are known to be associated with these measures as well as any potential threats to the validity of the results we obtain. The first and perhaps most obvious limitation of these measures for the purposes of the fitting-related decision-making is that a reliable behavioral response is required. As pediatric audiologists who work with infants know, this is not a trivial requirement. Second, even when reliable aided thresholds can be measured, we need to acknowledge that the real-ear frequency response of the hearing instrument under consideration is being sampled at a very limited number of frequencies. Third, when sound field aided thresholds are used exclusively at the verification stage, several important electroacoustic characteristics of the instrument(s) under consideration cannot be quantified, including, for example, the input/output characteristics and output limiting characteristics of the hearing instrument across frequencies. This is rapidly becoming a more serious limitation of sound field aided threshold testing as we work with a new generation of instruments that embody more sophisticated automatic signal process capabilities.

As noted earlier, one popular use of sound field aided thresholds has been to use them in predicting the levels at which amplified speech will be received within the aided condition. It has been known for many years that this application of sound field aided thresholds will be invalid under some conditions (e.g., Schwartz and Larson 1977). The primary problem relates to the fact that aided thresholds are measured in response to audiometric test signals that are low in level relative to the SPLs associated with average conversational speech. Consequently, the sound field aided thresholds do not provide a valid estimate of how a given hearing instrument will operate in response to everyday speech inputs. Thus, in many cases the sound field aided thresholds will tend to overestimate the available gain, and consequently, the levels at which amplified speech will be received by the infant/child within the aided condition (Schwartz and Larson 1977; Seewald, Hudson, Gagné and Zelisko 1992; Stelmachowicz and Lewis 1988). This is most certainly a problem for instruments which apply any form of non-linear signal processing.

Over the years, two studies have been reported in which the sound field aided threshold-based approach to predicting amplified speech sensation levels has been compared with alternative electroacoustic-based procedures (Schwartz and Larson 1977; Seewald et al. 1992) The results of both studies confirmed that under certain electroacoustic conditions, sound field aided thresholds provide unrealistically high estimates of the sensation levels at which speech will be received through amplification. On the basis of their findings, Schwartz and Larson concluded (more than 20 years ago!) that "the traditional sound-field audiogram is inappropriate for determining useable amplification of severe and profoundly hearing-impaired children" (p. 406). Furthermore, Schwartz and Larson strongly advised pediatric audiologists to consider alternative electroacoustic-based procedures (e.g., Erber 1973) to the traditional sound field aided audiogram when measuring hearing instrument performance for pediatric applications.

Electroacoustic-Based Procedures

Electroacoustic approaches to verification use well-defined signals to generate precise measurements of hearing instrument response by frequency. Hearing instrument response may be measured in various formats, either in the real ear or in a 2cc coupler. Regardless of the specific procedure, electro-

acoustic verification is highly efficient and reliable, and is easily modified to assess either speech processing or output limiting. Recently however, the validity of pairing electroacoustic tests with published prescriptive formulae has been questioned, at least for some classes of hearing instruments. Specifically, the effects of multiple channels of compression, fast or slow compression action, and noise detection have been raised as possible confounds that threaten the validity of electroacoustic test signals (Kuk and Ludvigsen 1999). In addition, some electroacoustic measurement formats are inherently more meaningful and/or more compatible with the infant population than others. In summary, electroacoustic-based verification is likely more efficient, reliable and valid than audiometric verification but specific electroacoustic procedures may have limited validity, meaningfulness, and clinically feasibility with some devices and/or age groups. These limitations will be discussed below.

Measurement Format Limitiations

As stated above, electroacoustic measures can be carried out in either the real-ear or the 2cc coupler. Measurement formats within each of these locations may be defined in either gain or sound pressure level (SPL). Early applications of probe microphone technology developed an electroacoustic analog to functional gain, termed "insertion gain". This measurement format is shown in figure 4. The lower curve is the gain produced by the open ear canal, termed the Real Ear Unaided Gain (ANSI 1997b). The top curve is the gain produced by the hearing aid for a given test signal, measured near the tympanic membrane, and termed the Real Ear Aided Gain (ANSI 1997b). The difference, in dB, between these two measurements is termed Real Ear Insertion Gain, shown by the middle curve in figure 4.

The REIG measurement format was developed as an electroacoustic analog to functional gain (Mueller 1992). The open ear gain was selected as the electroacoustic analog to the unaided audiogram, partially because protocols for assessing functional gain recommended that unaided hearing be assessed with open ears in sound field (Skinner 1988). However, this is rarely the case in current clinical practice. Consequently, this calls into question the validity of using an open ear resonance to represent the unaided condition.

A further limitation that is associated with insertion gain measures is that they are expressed in units (i.e., dB REIG) that cannot be directly compared to audiometric thresholds or loudness discomfort levels (LDLs), either in dB HL or dB SPL. Therefore, it is not possible to directly judge the adequacy of a frequency response by comparing it to thresholds and/or LDLs. These limitations are also true for any measurement of coupler gain or REAG. Therefore, while gain measures are certainly accurate and reliable, they are not as potentially meaningful as other electroacoustic formats.

A final concern with insertion gain is that the aided real ear measurements are taken while the hearing instrument wearer is seated in front of a loudspeaker. For accurate estimation of insertion gain, the individual should remain the same distance

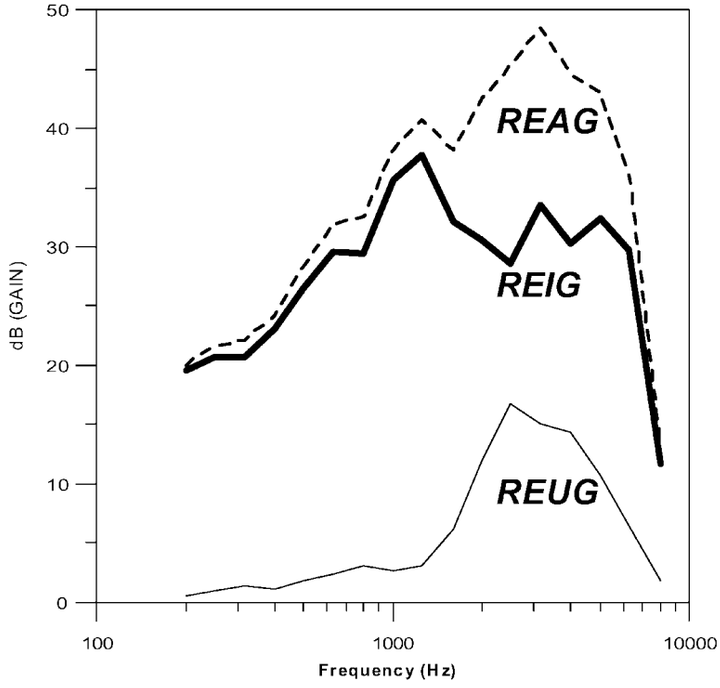

Figure 4. Example of an insertion gain verification measurement.

from the speaker for both REUG and REAG measurements, and not vocalize or turn his/her head. These requirements are simply not feasible with the infant and toddler population, and this has lead to relatively infrequent use of REIG/REAG/REAR measurements in pediatric hearing instrument fitting (Arehart, Yoshinaga-Intano, Thomson, Gabbard and Brown 1998).

Improvements to Measurement Format

One measurement format that has been proposed to facilitate meaningful verification of children's hearing aids is the "SPLogram" (Seewald 1995). Intended as a replacement for the aided audiogram, this measurement format uses direct comparisons of thresholds, unaided speech levels, and aided speech levels to facilitate informed evaluation of the adequacy of real-ear hearing instrument performance. An example of an unaided SPLogram is shown in figure 5.

This graph is a plot of real-ear sound pressure level (SPL) by frequency. Right ear thresholds are plotted as circles (○), and predicted Loudness Discomfort Levels (LDL) (Seewald 1991) are plotted as asterisks (*). The area between the thresholds and the LDLs is referred to as the "dynamic range" or "auditory area". These data points have been converted, for the purposes of graphic presentation, into the dB SPL scale (re: the tympanic membrane) via the application of the Real Ear to Coupler Difference (RECD) and Reference Equivalent Threshold Sound Pressure Levels (RETSPL: ANSI 1996) in an HL to SPL transform (Scollie, Seewald, Cornelisse and Jenstad 1998; Seewald and Scollie 1999).

It can be seen in figure 5 that by using this format, the unaided thresholds and LDLs may be directly compared to the levels associated with the unaided speech spectrum, shown as a solid dark line, and/or to normal hearing sensitivity, shown by the lowermost line on the SPLogram. Unaided speech is represented by the one-third octave band levels of running speech, averaged over several sentences, and spoken at an overall level of 70 dB SPL re: the free field (Cox and Moore 1988). Also, the dashed lines located 12 dB above and 18 dB below the unaided long-term average speech spectrum (LTASS) display the peaks and valleys of the unaided speech envelope per frequency (Cox, Mateisch and Moore 1988). For the purposes of graphic presentation, the free field values for the unaided LTASS have been transformed to values in the real ear via the application of an age-appropriate free-field to tympanic membrane transfer function (Kruger 1987). This unaided SPLogram display can be an important tool in counseling normal-hearing caregivers of infants and young children with hearing loss, particularly when the normal hearing sensitivity line is used as a reference point.

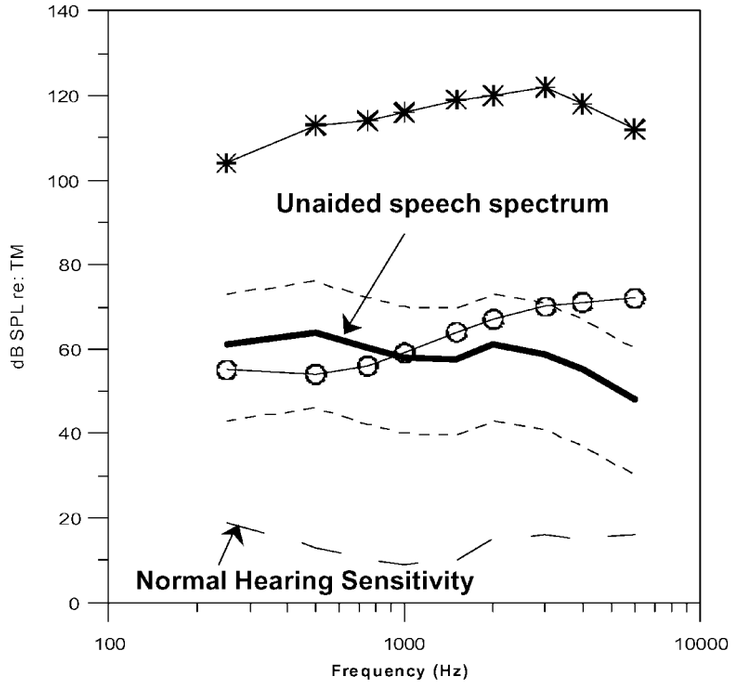

Figure 5. Example of an unaided SPLogram.

Interpretation of an SPLogram display is similar to the calculation of the Articulation Index or Speech Intelligibility Index (SII: ANSI 1997a), in that contributions to intelligibility are expected to begin when the peaks of speech are at threshold, and are expected to reach maximum values when the peaks of speech are 30 dB above threshold. In SII calculations, some frequency regions, usually in the

higher frequencies, are weighted to be more important than other frequency regions. Some available software systems combine the SPLogram display with AI-type calculations to provide numeric guidance for interpretation (Stelmachowicz, Lewis, Kalberer and Cruetz 1994). As with any audibility-based evaluation, expected limitations to the intelligibility of the aided speech signal, such as reduced cochlear integrity, distortion, noise, reverberation, or other limiting factors (Moore 1996; Ching, Dillon and Byrne 1998) should always be taken into consideration in interpreting the audibility displayed in the SPLogram format.

The Situational Hearing Aid Response Profile (SHARP) uses both SPLograms and Audibility Index calculations to provide objective assessment of hearing instrument candidacy and performance, based on electroacoustic measurements (Stelmachowicz et al. 1994). A SHARP analysis of the unaided listening condition was completed for the unaided thresholds shown in figure 3, for the following listening conditions: average conversational speech at 1 meter (60 dB SPL) and 4 meters (48 dB SPL), raised voice at 1 meter (70 dB SPL), and a classroom teacher speaking from 7 meters away (61 dB SPL). The SHARP analysis is shown in figure 6. As the expected input level of speech drops from 70 dB SPL to 48 dB SPL, progressively less of the speech envelope is above threshold, as shown by the shaded regions on each SPLogram. Also, audibility is primarily available in the low frequency regions, thereby limiting the contribution that can be made by the important high frequency regions. Therefore, even though a substantial bandwidth and sensation level of speech is available in the 70 dB condition (i.e., the top right plot in figure 6), the Audibility Index value is only .47 out of a possible

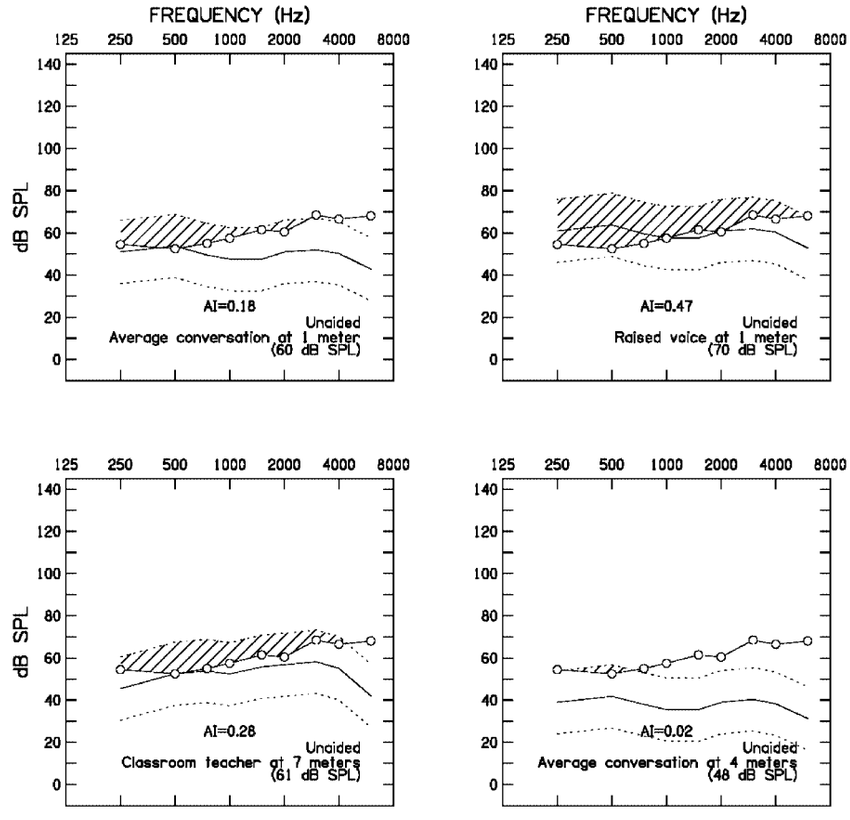

Figure 6. Example of various unaided SPLograms generated by the Situational Hearing Aid Response Profile (SHARP) software package. Each panel displays the audibility available, without hearing aids, in a specific listening condition.

score of 1.0. Poorer values are noted at the lower level listening conditions, with almost no speech audible (AI = .02) from a distance of 4 meters away or more unless a raised vocal effort is used. This SHARP profile would agree with parental concerns of poorer speech and language development than would be expected of normal-hearing peers, and inconsistent responses to environmental sounds, especially those from a distance. In school-aged children, difficulty in classroom listening tasks would be expected. Clearly, this child would be a candidate for some form of intervention.

One important component of intervention would be an amplification system in the form of personal hearing instruments. As shown in figure 7, prescriptive targets for hearing instrument performance can be generated for soft, average, and loud level speech signals, and plotted within the SPLogram format. The common plotting format for both unaided information and prescriptive targets for hearing instrument performance provides an important link between assessment, selection and verification.

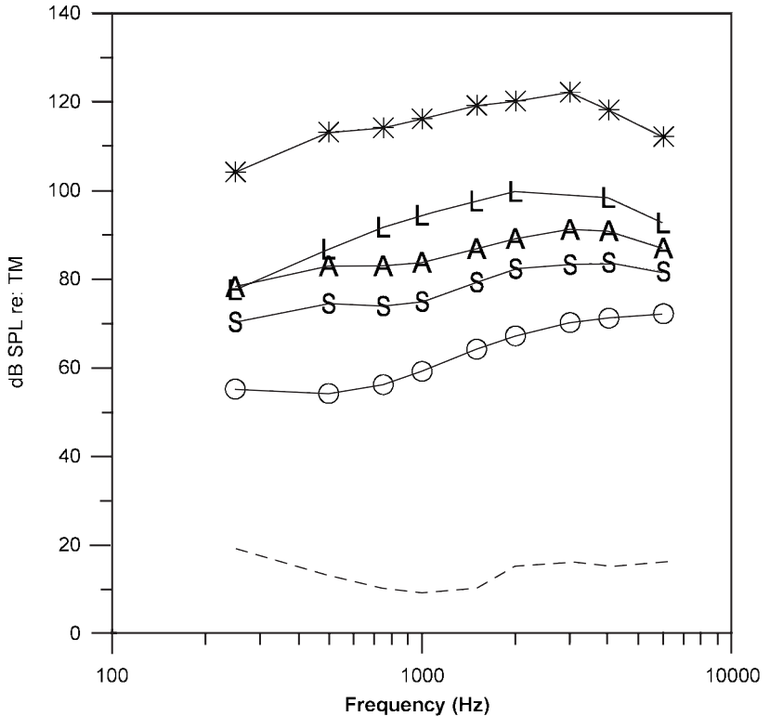

Figure 7. Example of prescriptive targets for soft (S), average (A), and loud (L) level speech inputs, plotted in an SPLogram format.

Verification measures, as stated above, cannot be easily carried out with infants and young children using conventional probe microphone protocols. Particularly, the requirement for the child to sit up and still during measurement can be problematic. An alternative to this procedure has been proposed by Moodie, Seewald and Sinclair (1994) and validated in two separate studies (Seewald, Moodie, Sinclair and Scollie 1999; Munro and Hatton 2000). Rather than directly measuring hearing aid performance at the eardrum, the real-ear performance of the hearing instrument can be predicted from coupler measures, using an acoustic transformation procedure. This acoustic transform involves the measurement of the infant/child's RECD using appropriate procedures (Munro and Salisbury 2001). The RECD, although it is a real-ear procedure, is measured with an insert phone transducer rather than a probe microphone system's loudspeaker. For this reason, it does not require the client to sit up and face any particular direction. This allows the RECD to be measured even when the baby is sleeping in the mother's arms, or if a toddler is turning his or her head during the measurement.[1] In summary, predicted real-ear responses may be accurately derived from coupler measurements using the RECD, and plotted on an SPLogram. This facilitates use of electroacoustic verification techniques with the pediatric population, and interpretation of verification measures in a meaningful context.

An example of a hearing instrument, fitted to the threshold values shown in figure 3, will be described. Previous studies have shown improved benefit with low level, low ratio compression systems, compared to linear amplification (Jenstad, Seewald, Cornelisse and Shantz 1999; Jenstad, Pumford, Seewald and Cornelisse 2000). Therefore, a wide-dynamic range circuit type was pre-selected for fitting, in a BTE style

[1] For advice on measuring the RECD with infants and young children, see Bagatto (2000).

with a soft earmold to accommodate changes in ear size and prevent injury to the pinna (Pediatric Working Group 1996). These targets have an average compression ratio of 2:1 across the frequency range of 250 through 6000 Hz, with compression ratios per frequency ranging from 1.8 to 2.5:1. The instrument under consideration is a digital device, with multiple channels of compression. Although these features are not essential for a basic hearing aid fitting, they are expected to provide the following advantages in the fitting process. Multiple channels of compression tend to allow more flexibility in the frequency response shape of the hearing instrument. This can be helpful, especially with changes over time due to factors such as ear infections, ear canal growth, earmold changes, and, more rarely, fluctuations in hearing sensitivity.

For this example, the hearing aid was set to provide the electroacoustic performance shown in figure 8. Measurements of the instrument's maximum output response (i.e., thin solid line) and speech processing (i.e., heavy solid line and dashed lines) are shown. It may be noted that the maximum output of the hearing aid does not exceed LDLs at any frequency. The estimates of the aided LTASS and aided peaks and valleys of speech were measured with running conversational speech, in one-third octave bands. The level of the amplified speech, per frequency, is similar to the prescribed frequency response to approximately 4000 Hz. For simplicity, only the frequency response for a 70 dB SPL speech input is shown. Other speech input levels can be assessed using further verification measures, or with further verification plus Articulation Index calculations, as described in the Aided SHARP analysis below.

For this SHARP analysis, listening situations specific to infancy have been chosen, to further evaluate the fitting of a wide dynamic range compression (WDRC) instrument shown in figure 9. Specifically, this SHARP analysis evaluates aided audibility for average conversational speech from 1 and 4 meters distance, as well as average conversational speech received by the infant's near ear when held in cradle and hip positions by a caregiver. Across these situations, the Aided Audibility Index ranges from .80 to .94. The shaded regions on each SHARP SPLogram provide a graphical depiction of the level and shape of the aided signal in each targeted situation. Overall, the SHARP analysis indicates that the hearing instrument, with WDRC, will adequately amplify speech information from a wide range of azimuths and distances to the infant's ear canal. This hearing instrument fitting should facilitate communication development through a broad exposure to auditory stimuli in the infant's environment.

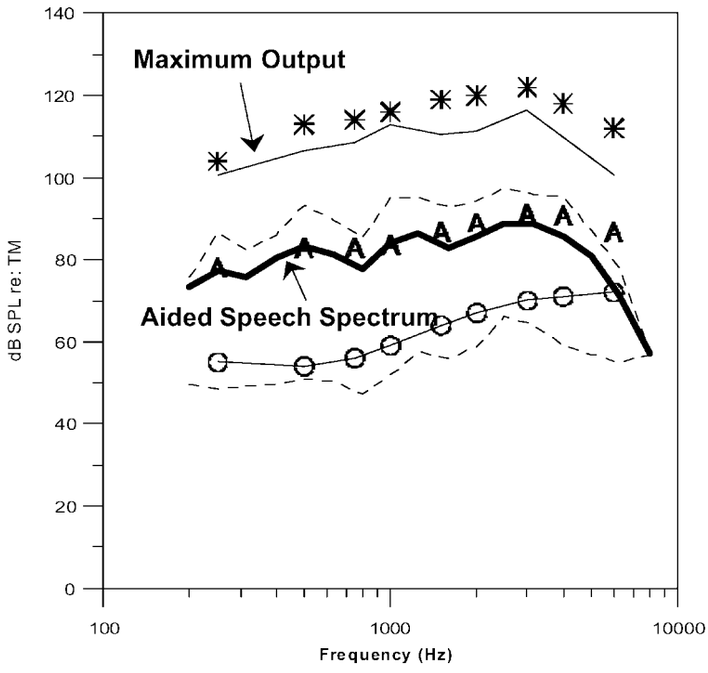

Figure 8. Example of a digital hearing aid, set to targets and verified in the SPLogram format.

In summary, SPLogram verification strategies, in combination with measures of the infant's RECD, fulfill the requirements of a universal pediatric verification protocol shown in figure 2. However, it should be noted that the counseling utility of the SPLogram approach depends upon the plotting of aided real-world stimuli, such as speech. Although it is possible to create an SPLogram that plots the results of tests using artificial stimuli such as aided pure tone sweeps, this is somewhat less meaningful, and therefore somewhat less useful for counseling and hearing aid orientation, than a

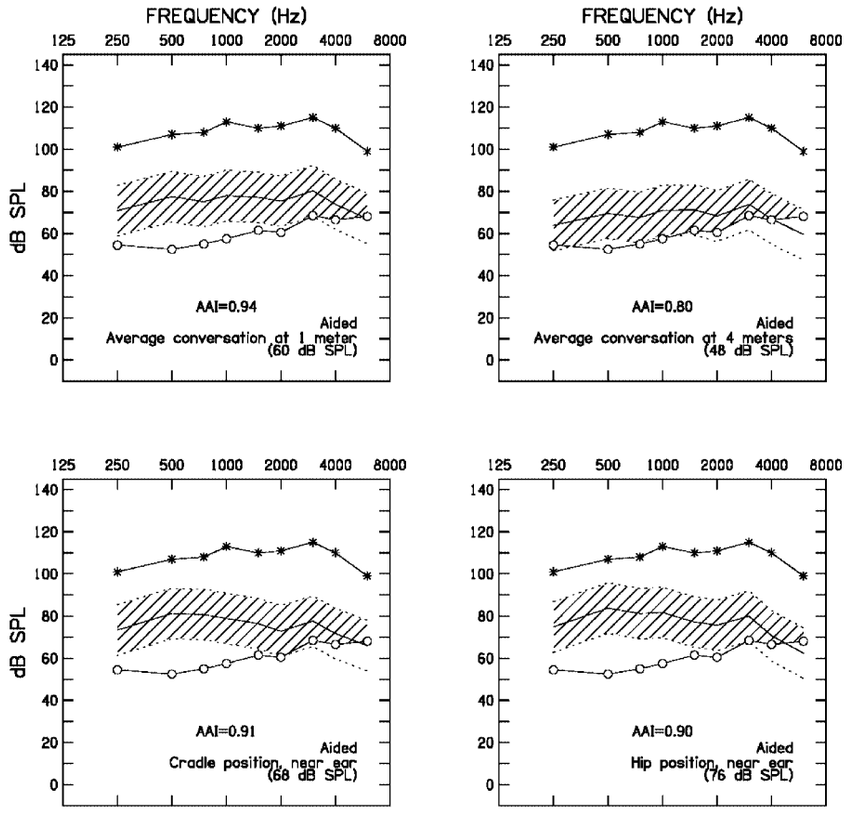

Figure 9. Example of various aided SPLograms generated by the Situational Hearing Aid Response Profile (SHARP) software package. Each panel displays the audibility available, with hearing aids, in a specific listening condition.

speech-based SPLogram. The challenge, then, is to derive estimates of the levels and shape of the aided speech signal from clinical non-speech test signals.

Test Signal Limitations

In most clinically available test systems, electroacoustic verification relies upon synthetic test signals to evaluate the speech processing and output limiting of a hearing instrument. There are many cited instances in which a clinical test signal has failed to accurately estimate the performance of a hearing instrument for a real-world test signal (Stelmachowicz, Lewis, Seewald and Hawkins 1990; Stelmachowicz, Kopun, Mace and Lewis 1996). The authors of some recent publications have hypothesized that such instances are more likely to occur with digital signal processing instruments due to a greater number of compression channels, lower compression thresholds, a wider range of compression speeds, and the presence of noise detection processing (Kuk and Ludvigson 1999).

In response to these concerns, a variety of newly developed and increasingly complex test signals have been described (Frye 2000; Groth 2001; Lesiecki, Majest and Redinger 2001; Robson 2001; Sinclair, Cole and Pumford 2001). These test signals typically employ some combination of spectral weighting, bandwidth, frequency modulation, duration modulation, and amplitude modulation to mimic the acoustic properties of speech. It is hypothesized by the developers of these signals that a more speech-like synthetic signal will be more likely to provide an accurate estimate of speech processing (Cole and Sinclair 1998; Frye 2000).

For the clinician, a plethora of hearing aid signal processing options interacts with an abundance of test signals to create an apparently complex set of decisions that must be made in order to ensure an accurate verification protocol. We have recently com-

pleted two studies that attempt to evaluate these interactions (Scollie and Seewald [in review]; Scollie, Joyce and Seewald [in review]). These studies used two samples of hearing instruments. In the first sample, 41 hearing instruments were selected and adjusted to typical use settings. The 41 devices were distributed among slightly sloping moderate, moderately-severe, and severe to profound hearing losses, and included a wide range of circuitry, including 17 digital devices. In the second sample, two digital hearing aids were set to an extreme range of settings (i.e., minimum and maximum compression settings), generating 38 test cases. The accuracy for tests of both maximum output and speech-level amplification were evaluated for these instruments using two clinically available test systems (i.e., Fonix 6500, Audioscan RM500).

Maximum Output Measures

The maximum output responses of 41 hearing aids were measured with a 90 dB SPL pure tone sweep delivered from the Fonix 6500, and with the "MPO" test from the Audioscan RM500. The results of these tests were compared to the aided peak levels of speech presented at 83 dB SPL and 70 dB SPL, measured on the same set of hearing aids. Peak levels of speech were defined as the 99th percentile of the amplitude distribution per one-third octave band, measured in 126 msec intervals. Comparisons of the peaks from the 70 dB SPL speech signal and peaks from the 83 dB SPL speech signal were made, and the higher of the two was chosen to represent the peak level of speech. In some cases, the peaks from the 70 dB SPL signal were higher. The 83 dB SPL signal was shaped to represent shouted vocal effort, which has a distinct mid-frequency emphasis (Olsen 1998). Therefore, it was possible for the extreme-frequency peaks from the 70 dB SPL speech to exceed those of the 83 dB speech, despite the lower overall level. For example, this occurred in 90% of cases at 250 Hz, but in 0% of cases at 1000 Hz.

The peaks of amplified speech were compared to the measured maximum output from the two clinical tests. Results of this comparison are shown in figure 10, with 500, 1000, 2000, and 4000 Hz shown as panels (a) through (d), respectively. An accurate test of maximum output should be located at or above the level of amplified real-world signals (i.e., any point above the diagonal line in each panel in figure 10). For this study, the following tolerance level for underestimation of maximum output was used: the measured maximum output must be no less than 3 dB below the speech peaks, on average, from 500 to 4000 Hz. It was found that no hearing instruments fell outside of this tolerance region for the RM500 MPO test signal. However, one instrument fell outside of this tolerance region for the pure tone test using the Fonix 6500 system at 90 dB SPL. This particular device had an active noise reduction processor that could not be disabled during testing. It was found that the noise reduction processor affected the frequency response of the hearing aid at all test levels including testing performed at 90 dB SPL. Overall, the assessment of maximum output is successful in describing the ceiling of hearing aid output in the majority of cases.

Assessment of Frequency Response for Speech

An important issue in electroacoustic verification relates to the accuracy in estimating the frequency response of amplified speech using non-speech clinical test signals. This issue was also evaluated experimentally, using the two samples of hearing aids described above (Scollie and Seewald [in review]; Scollie, Joyce and Seewald [in review]). At the 70 dB SPL test level, both the Audioscan Swept and Fonix Composite Noise signals were found to predict the aided frequency response for real speech to within 5 dB RMS for approximately 80% of tested cases. Error patterns for both test signals were somewhat predictable. Overall, the Fonix Composite Noise was more likely to underestimate aided speech output (figure 11a), while the Audioscan Swept signal was more likely to overestimate aided speech output (figure 11b). Digital hearing aids with fixed, active noise reduction processing tended to be inaccurately tested by the Composite Noise signal, likely due to the steady-state presentation of this signal. For this reason, it was recommended that noise reduction processing be disabled prior to testing with steady-state Composite Noise. Further, the speech output from high-power instruments tended to be overestimated by the Audioscan Swept signal. However, the results of our studies suggest that both signals provide sufficient accuracy for clinical use in estimating the aided frequency response of speech.

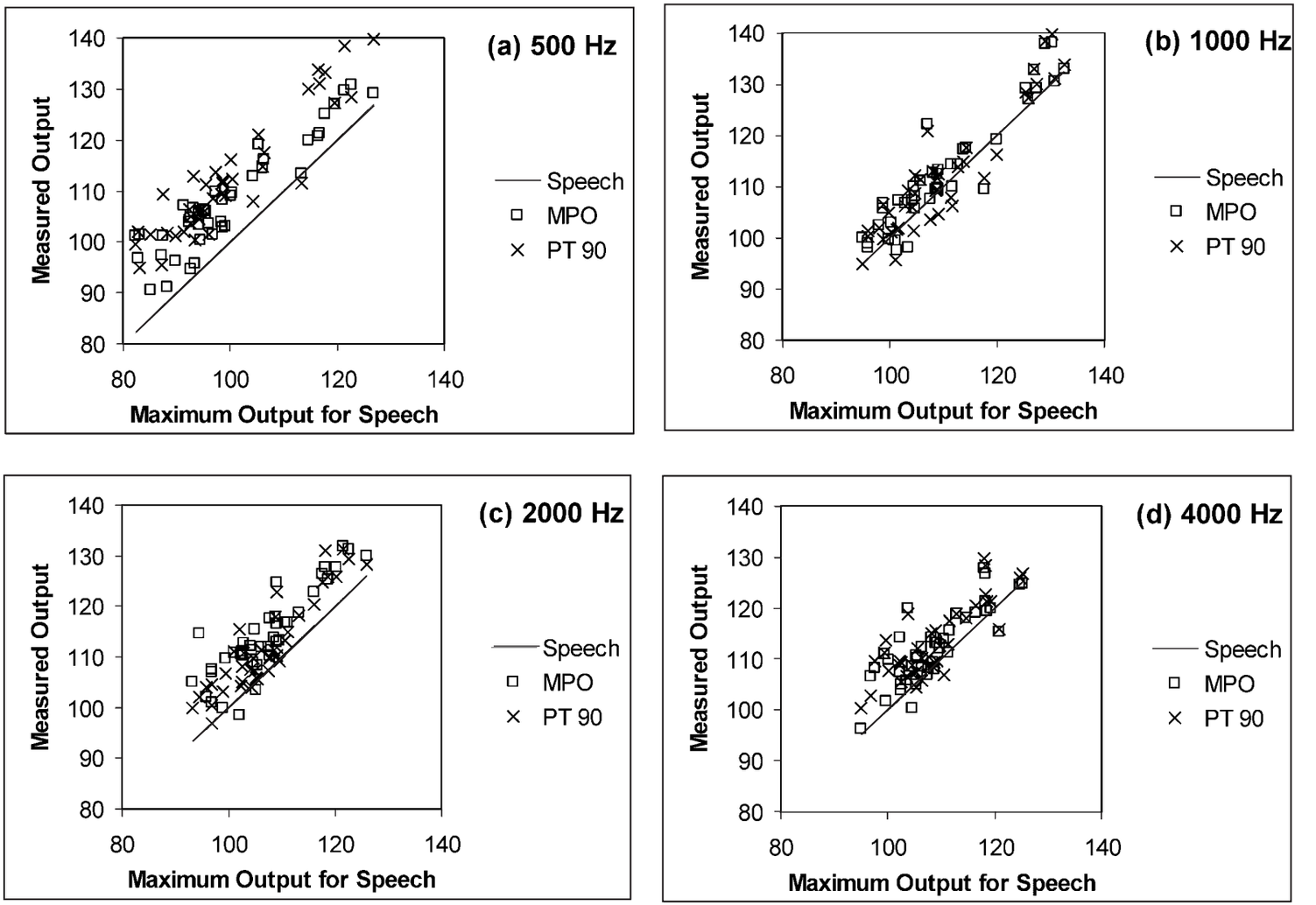

Figure 10. Comparison between two clinical measures of hearing instrument maximum output and the peaks of speech, from 41 hearing aids. Each panel represents data gathered at one frequency.

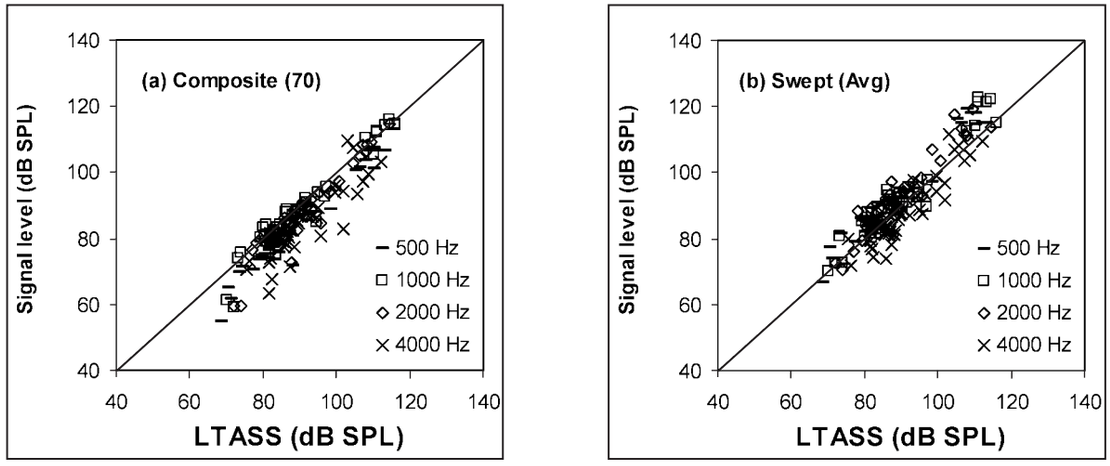

Figure 11. Comparison between two clinical measures of frequency response and the Long-Term Average Speech Spectrum (LTASS), of 41 hearing aids. Each panel represents data gathered from one clinical test signal.

Improvements to Test Signal Accuracy

In general, a more speech-like test signal will yield a more accurate estimation of frequency response for aided speech. The acoustic properties that vary from test signal to test signal include: a) spectral shape; b) bandwidth; c) level; d) changes in level over time; e) changes in frequency content over time; and f) crest factor. The test signals evaluated in the reported studies tended to match some, but not all, acoustic properties of speech. Therefore, they were not entirely accurate in their estimations of the aided frequency response of speech. It is reasonable, although untested, to assume that more recently developed test signals, which incorporate more/all of the acoustic features of speech, will yield a higher degree of accuracy in predicting the amplified levels for real speech inputs. However, it is also possible to improve upon the test accuracy of signals that are not entirely speech-like through the application of correction procedures.

Scollie, Joyce and Seewald (in review) evaluated several correction procedures, ranging from simple addition of gain to multiple regression involving measures of compression ratio and release time. Two of these procedures offer immediate improvement and are feasible for clinical use. These procedures fall into two general categories: procedures for use with pure tone sweeps or Composite Noise, and procedures for use with the Audioscan RM500 test signals.

Pure Tone Sweeps or Composite Noise

Test signals have specific spectral shapes, and are typically either speech-weighted (i.e., Composite Noise) or flat (i.e., pure tone sweeps). Prescriptive targets are generally calculated assuming a specific input shape, often assuming a published speech spectrum (Cox and Moore 1988; Cornelisse, Gagné and Seewald 1991). If the spectral shape of the test signal does not match the spectral shape of the published spectrum used in calculating targets, error can be introduced into the target matching process. As described below, it is possible to reduce this error by applying a correction.

Slight differences in spectral shape generally do not produce large differences in hearing aid gain, for low to moderate test levels (Scollie, Joyce and Seewald [in review]; Stelmachowicz et al. 1996). That is, the gain for speech will be similar to the gain for a test signal, even if the measured responses in SPL are slightly different. Therefore, if the gain from the test signal is added to the unaided speech spectrum, the spectral shape of the test signal is essentially removed. This tends to reduce error due to spectral shape differences between a test signal and speech. This procedure will be referred to as a "gain correction" in the paragraphs below.

The gain correction approach has been evaluated for both pure tone sweeps and for Composite noise, in a sample of 41 hearing aids (Scollie, Joyce and Seewald [in review]). For pure tone sweeps, 93–95% of

test cases were accurately estimated using the gain correction, for 50 and 70 dB test levels. For Composite Noise, 88–93% of test cases were accurate with the gain correction, at the same test levels. For both pure tones and Composite noise, the gain correction also provided improvement to test accuracy on a second set of digital hearing aids with extreme compression ratio settings. Two practical considerations apply in using the gain correction: (1) the gain correction is successful in digital hearing aids with noise reduction processing only if this processing is disabled prior to making gain measurements, and all other signal processing parameters (i.e., shaping, compression ratio, etc.) remain at user settings; (2) the gain correction, or a similar correction procedure, is currently implemented in DSL and SHARP software systems. In DSL, this correction is applied to gain measures prior to plotting on the Speech SPLogram.

In summary, the gain correction provides improved test accuracy for test signals that have a different spectral shape than the presumed speech spectrum. We can use the gain correction as follows. First, evaluate frequency shaping with measures of Composite Noise gain at a test level of 70 dB SPL. Then, assess gain for low-level inputs using Composite Noise at a test level of 50 to 55 dB SPL, and measure maximum output with pure tone sweeps at a test level of 90–100 dB SPL. In combination, these tests would constitute a comprehensive and accurate test of electroacoustic performance, provided that noise reduction processing was not active during testing.

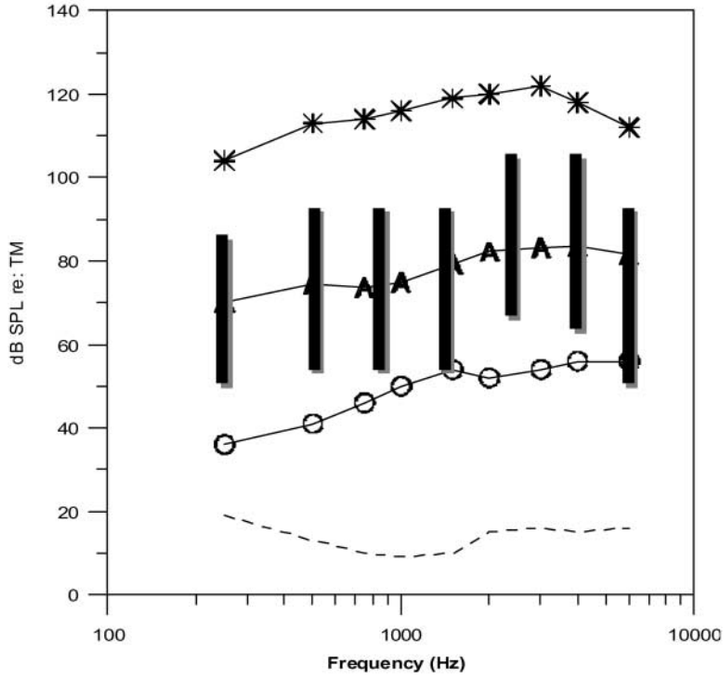

Figure 12. Verification of hearing aid performance using the Audioscan RM500 Dynamic test signal.

Audioscan RM500 Test Signals

As noted above, the Swept signal from the Audioscan sometimes overestimates the aided levels of speech. In this set of studies it was found that a gain correction did not offer any improvement for this signal. This is because there were no discrepancies between the spectral shape of the signal and the speech spectrum used by Scollie and Seewald (in review). However, a second signal is available in the RM500 system, which attempts to estimate the peaks and valleys of running amplified speech. For this 'dynamic' test signal the peaks and valleys of amplified speech are estimated by the tops and bottoms of the vertical bars (see figure 12). It is assumed that these bars surround the aided Long Term Average Speech Spectrum (LTASS), and that the LTASS is likely located at the midpoint to the 75% point of the bars. The hearing instrument fitting shown in figure 12 would be judged to be "on target" for average speech inputs since approximately the vertical bars fall approximately 50% above and 50% below the targets for average speech.

Scollie, Joyce and Seewald (in review) evaluated the accuracy of using the midpoint of the Dynamic bars to estimate the aided speech signal. The effectiveness of this correction was found to vary with test signal level. At the nominally "soft" test level (i.e., 53 dB SPL), the Swept signal was more accurate (90% of cases) in hearing instruments with moderate compression ratio settings. When compression settings were extreme, or if testing was conducted at the "average" or "loud" test levels, the midpoint of the Dynamic signal was a more accurate predictor of real speech output

than the Swept signal. From our findings with this test system, it seems reasonable to fit the hearing aid by surrounding the average speech targets with the bars of the Dynamic/Average test, and to assess amplification for low-level signals using the Swept/Soft signal. Combined with the "MPO" evaluation of maximum output, this would constitute a comprehensive and accurate test of electroacoustic performance.

Summary

Hearing instrument fitting with infants and young children presents many challenges. For young infants, hearing instrument prescriptive targets will be based upon electrophysiologic estimates of auditory thresholds and the acoustic properties of the ear canal, with little or no real contribution from behavioral audiometric testing. Similarly, evaluation of the appropriateness of a given hearing aid will be based on electroacoustic assessments of hearing aid function, with very little contribution from behavioral responses to hearing instrument output. To the extent possible, electroacoustic verification should be paired with systematic observation of infant responses to sound by caregivers (Harrison 2000). Also to the extent possible, electroacoustic verification should be both meaningful and accurate, and should reflect hearing aid processing for both speech-level inputs and for output limiting. It is possible, and necessary, for the clinician to select test signals and procedures that ensure accurate electroacoustic verification, even for infants and for higher technology hearing aids.

References

American National Standards Institute. 1996. Specifications for audiometers, *ANSI* S3.6–1996. New York: Acoustical Society of America.

American National Standards Institute. 1997a. Methods for calculation of the speech intelligibility index, *ANSI* S3.5–1997. New York: Acoustical Society of America.

American National Standards Institute. 1997b. Methods of measurement of real-ear performance characteristics of hearing aids, *ANSI* S3.46–1997. New York: Acoustical Society of America.

Arehart, K.H., Yoshinaga-Itano, C., Thomson, V., Gabbard, S.A., and Brown, A.S. 1998. State of the states, the status of universal newborn hearing screening, assessment, and intervention systems in 16 states. *American Journal of Audiology* 7:1–14.

ASHA Ad Hoc Committee on Hearing Aid Selection and Fitting. 1998. Guidelines for hearing aid fitting for adults. *American Journal of Audiology* 7(1): 5–13.

Bagatto, M. 2000. Optimizing your RECD measurements. *Hearing Journal* 54(9): 32, 34–46.

Ching, T.Y.C., Dillon, H., and Byrne, D. 1998. Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification. *Journal of the Acoustical Society of America* 103(2): 1128–1140.

Cole, W., and Sinclair, S.T. 1998. The Audioscan RM500 speechmap/DSL fitting system. *Trends in Amplification* 3(4): 125–139.

Cornelisse, L.E., Gagné, J.P., and Seewald R.C. 1991. Ear level recordings of the long-term average spectrum of speech. *Ear and Hearing* 121: 47–54.

Cox, R.M., Mateisch, J.S., and Moore, J.N. 1988. Distribution of short-term rms levels in conversational speech. *Journal of the Acoustical Society of America* 84(3): 1100–1104.

Cox, R.M., and Moore, J.N. 1988. Composite speech spectrum for hearing aid gain prescriptions. *Journal of Speech and Hearing Research* 31(1): 102–107.

Erber, N.P. 1973. Body-baffle and real-ear effects in the selection of hearing aids for children with deafness. *Journal of Speech and Hearing Disorders* 38: 224–231.

Frye, G. 2000. Testing digital hearing instruments: The basics. *Hearing Review* 7(8): 20, 22, 24, 26–7.

Groth, J. 2001. Influence of advanced hearing aid technology on choice of signal for probe microphone measures. *Trends in Amplification* 5(2): 69–72.

Harrison, M. 2000. How do we know we've got it right? Observing performance with amplification. In R. Seewald (ed.), *A sound foundation through early amplification: Proceedings of an international conference* (pp. 119–140). Stäfa, Switzerland: Phonak AG.

Jenstad, L.M., Seewald, R.C., Cornelisse, L.E., and Shantz, J. 1999. Comparison of linear gain and wide dynamic range compression hearing aid circuits: Aided speech perception measures. *Ear and Hearing* 20(2): 117–26.

Jenstad, L.M., Pumford, J., Seewald, R.C., and Cornelisse, L.E. 2000. Comparison of linear gain and wide dynamic range compression hearing aid circuits II: Aided loudness measures. *Ear and Hearing* 21(1): 32–44.

Kruger, B. 1987. An update on the external ear resonance

in infants and young children. *Ear and Hearing* 8(6): 333–336.

Kuk, F.K., and Ludvigsen, C. 1999. Variables affecting the use of prescriptive formulae to fit modern nonlinear hearing aids. *Journal of the American Academy of Audiology* 10(8): 458–465.

Lesiecki, W. Majest, B., and Redinger, B. 2001. Universal REM-Link and ICRA signals – High-tech fitting tools for high-tech hearing instruments. *Trends in Amplification* 5(2): 73–76.

Moodie, K.S., Seewald, R.C., and Sinclair, S.T. 1994. Procedure for predicting real-ear hearing aid performance in young children. *American Journal of Audiology* 3(1): 23–31.

Moore, B.C. 1996. Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. *Ear and Hearing* 17(2):133–61.

Mueller, H.G. 1992. Terminology and procedures. In H.G. Mueller, D.B. Hawkins, and J.L. Northern (eds.), *Probe microphone measurements: Hearing aid selection and assessment* (pp. 41–66). San Diego, CA: Singular Publishing Group.

Munro, K.J., and Hatton, N. 2000. Customized acoustic transform functions and their accuracy at predicting real-ear hearing aid performance. *Ear and Hearing* 21: 59–69.

Munro, K.J., and Salisbury, V.A. 2001. Is the real-ear to coupler difference acoustic transform independent of the measurement earphone? Poster presentation, A Sound Foundation through Early Amplification: An International Conference, Chicago, IL.

Olsen, W.A. 1998. Average speech levels and spectra in various speaking/listening conditions: A summary of the Pearson, Bennett, and Fidell (1977) Report. *American Journal of Audiology* 7(2) :1–5.

Pediatric Working Group. 1996. Amplification for infants and children with hearing loss. *American Journal of Audiology* 5(1): 53–68.

Robson, H. 2001. Using a variety of signals to improve real-ear testing. *Trends in Amplification* 5(2): 77–79.

Scollie, S.D., Joyce, M.J., and Seewald, R.C. In review. Evaluation of electroacoustic test signals II: Development and cross-validation of correction factors. *Ear and Hearing.*

Scollie, S.D., and Seewald, R.C. In review. Evaluation of electroacoustic test signals I: Comparison to amplified speech. *Ear and Hearing.*

Scollie, S.D., Seewald, R.C., Cornelisse, L.E. and Jenstad, L.M. 1998. Validity and repeatability of level-independent HL to SPL transforms. *Ear and Hearing* 19(5): 407–413.

Schwartz, D.M., and Larson, V.D. 1977. A comparison of three hearing aid evaluation procedures for young children. *Archives of Otolaryngology* 103: 401–406.

Seewald, R.C. 1991. Hearing aid output limiting considerations for children. In J. Feigin and P. Stelmachowicz, (eds.), *Pediatric amplification: Proceedings of the 1991 national conference* (pp 19–35). Omaha NE: Boys Town National Research Hospital Press.

Seewald, R.C. 1995. The desired sensation level (DSL) method for hearing aid fitting in infants and children. *Phonak Focus* 20. Stäfa, Switzerland: Phonak AG.

Seewald, R.C., Hudson, S.P., Gagné, J.P., and Zelisko, D.L. 1992. Comparison of two procedures for estimating the sensation level of amplified speech. *Ear and Hearing* 13(3): 142–149.

Seewald, R.C., Moodie, K.S., Sinclair, S.T., and Cornelisse, L.E. 1996. Traditional and theoretical approaches to selecting amplification for infants and young children. In F.H. Bess, J.S. Gravel and A.M. Tharpe (eds.), *Amplification for children with auditory deficits* (pp. 161–192). Nashville TN: Bill Wilkerson Center Press.

Seewald, R. C., Moodie, K. S., Sinclair, S. T., and Scollie, S. D. 1999. Predictive validity of a procedure for pediatric hearing instrument fitting. *American Journal of Audiology* 8(2):143–152.

Seewald, R.C., and Scollie, S.D. 1999. Infants are not average adults: Implications for audiometric testing. *Hearing Journal 52*(10): 64–72.

Skinner, M.W. 1988. *Hearing aid evaluation.* Englewood Cliffs NJ: Prentice-Hall.

Sinclair, S.T., Beauchaine, K.L., Moodie, K.S., Feigin, J.A., Seewald, R.C., and Stelmachowicz, P.G. 1996. Repeatability of a real-ear-to-coupler difference measurement as a function of age. *American Journal of Audiology* 5(3): 52–56.

Sinclair, S.T., Cole, W.A., and Pumford, J.M. 2001. The Audioscan RM500 real-ear hearing aid analyzer: Measuring for a successful fit. *Trends in Amplification* 5(2): 81–90.

Stelmachowicz, P.G. 2000. How do we know we've got it right? Electroacoustic and audiometric measures. In R. Seewald (ed.), *A sound foundation through early amplification: Proceedings of an international conference* (pp. 109–118). Stäfa, Switzerland: Phonak AG.

Stelmachowicz, P.G., Kopun, J., Mace, A.L., and Lewis, D.E. 1996. Measures of hearing aid gain for real speech. *Ear and Hearing* 17(6): 520–527.

Stelmachowicz, P.G., and Lewis D.E. 1988. Some theoretical considerations concerning the relation

between functional gain and insertion gain. *Journal of Speech and Hearing Research* 31(3): 491–496.

Stelmachowicz, P.G., Lewis, D., Kalberer, A., and Cruetz, T. 1994. Situational hearing-aid response profile (SHARP), Version 2.0: User's manual. Omaha NE: Boys Town National Research Hospital.

Stelmachowicz, P.G., Lewis, D.E., Seewald, R.C., and Hawkins, D.B. 1990. Complex and pure tone signals in the evaluation of hearing-aid characteristics. *Journal of Speech and Hearing Research* 33:380–385.

Acknowledgement

Support for the manuscripts summarized in this chapter was received from the Ontario Rehabilitation Technology Consortium, Etymonic Design, Inc., the Natural Sciences and Engineering Research Council of Canada, Siemens Hearing Instruments, and Phonak AG.