2013

Examining the validity and reliability of the Infant-Toddler Meaningful Auditory Integration Scales (IT-MAIS) via Rasch analysis

Anne Denise Schubert

Louisiana State University and Agricultural and Mechanical College

EXAMINING THE VALIDITY AND RELIABILTY OF THE *INFANT-TODDLER MEANINGFUL AUDITORY INTEGRATION SCALES* (*IT-MAIS*) VIA RASCH ANALYSIS

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree
of Master of Arts

in

The Department of Communication Sciences & Disorders

by
Anne Schubert
B.A., University of Michigan, 2011
May 2013

Acknowledgements

I wish to thank my mentor Dr. Brittan Barker for sharing with me her passion for improving assessment and therapy techniques for pediatric cochlear implant users. I admire her for always challenging herself and others to think critically and never settle for 'good enough.' She has played in invaluable role in my growth as a researcher, clinician, and life-long learner.

I would also like to thank Dr. Neila Donovan for patiently and enthusiastically teaching me about Rasch analysis. Thank you to Dr. Hoffman for serving on my committee and sharing his keen insight while establishing this study's methodology.

I thank Aimee Perkins for her words of encouragement, wordsmithing abilities, and generosity. To my kind and hard-working parents, thank you for teaching me to pursue my dreams with determination and a sense of humor. Finally, thank you to the volunteers in the Spoken Language Processing Lab for your support throughout my graduate studies.

Table of Contents

Abstract

In this study, we analyzed the validity and reliability of the *Infant-Toddler Meaningful Auditory Integration Scales* (*IT-MAIS*; Zimmerman-Phillips, Osberger, & Robbins, 2001), an assessment designed to measure listening skills in children ages 0-3 years. The *IT-MAIS* is a caregiver report tool used by speech-language pathologists and audiologists to assess listening skills in children with sensorineural hearing loss (SNHL) pre- and post-cochlear implant (CI). The *IT-MAIS* is widely used; however, it has not undergone thorough psychometric analysis.

Using longitudinal data collected by the University of Iowa Children's Cochlear Implant Program, we analyzed the psychometric properties of the *IT-MAIS* via Rasch analysis, a 1-parameter (1-PL) model of Item Response Theory (IRT; Lord & Novick, 1968). Pre- and post-CI assessments from 23 CI users aged 10 to 36 months were evaluated.

IRT is a form of psychometric analysis that is emerging in the behavioral sciences as a viable alternative to Classical Test Theory for test development and analysis. IRT results are similar, but not identical to classically derived concepts of validity and reliability. Specifically, we analyzed the content and construct validity of the *IT-MAIS*. We found that 2 out of 10 items exceeded misfit criteria, meaning participants did not respond predictably to these 2 items. We also found that the item-difficulty range did not capture the full range of participant ability, especially the higher range of participant ability. Therefore, the *IT-MAIS* may not be assessing higher-level listening skills, particularly in children post-CI. Rasch analysis also revealed that 1 of the 5 rating scale categories was not used predictably, indicating that the rating scale was not used as the test developers (Zimmerman-Phillips, Osberger, & Robbins) intended. To analyze item order relative to sequential development of listening skills, we established an *a priori* item rank order and compared it to item difficulty order established by Rasch analysis. Overall, our

results indicated the *IT-MAIS* did not demonstrate ideal item-level psychometric properties according to Rasch analysis and item order did not reflect sequential development of listening skills. We concluded that the *IT-MAIS* should not be used to assess listening development from pre- to post-CI.

Chapter 1. Précis

It is essential to have an ecologically valid assessment of listening skills. Such an assessment is particularly valuable in order to accurately identify pediatric cochlear implant (CI) candidates. Assessing listening skills is also important for developing appropriate treatment goals and tracking listening development in children post-CI. An assessment that is ecologically valid reflects performance that generalizes to everyday situations; scores from such an assessment are not confined to performance in controlled laboratory or clinical environments. By ensuring the ecological validity of the assessments, we improve professionals' ability to develop appropriate intervention by providing them with realistic evaluations of their clients' listening abilities.

The *Infant-Toddler Meaningful Auditory Integration Scale* (*IT-MAIS*; Zimmerman-Phillips, Osberger, & Robbins, 2001) was one of the first parent report tools made available for assessing CI candidacy and monitoring listening development post-CI in children ages 0 to 3 years. Items included in the *IT-MAIS* are intended to measure 3 underlying principles: vocalization behavior, alerting to sounds, and deriving meaning from sound (Zimmerman-Phillips et. al, 2001). Clinicians, researchers, and CI manufacturers use the *IT-MAIS* to assess the listening skills in young children with hearing loss pre- and post-CI; however, it was developed without rigorous psychometric analysis. Without that psychometric rigor, users have to be cautious about conclusions they make based on the test's findings (validity), and perhaps question the reliability of the respondents' reported pre- and post-CI behaviors (reliability). In the present study, our research questions were as follows:

1. Does the *IT-MAIS* data meet the assumptions for Rasch analysis (i.e. unidimensionality, local independence)?

2. Does the *IT-MAIS* demonstrate item-level psychometric properties, reflecting that the latent trait ("auditory integration") can be measured?

3. Does the *IT-MAIS* demonstrate adequate range of person ability and item difficulty to separate respondents into more than 2 levels?

4. Do the items on the *IT-MAIS* reflect content validity based on *a priori* item hierarchy rankings?

The present study is an initial step in psychometrically analyzing the *IT-MAIS* at the item level. We used Item Response Theory (IRT)—a statistical method of assessing the measurement properties of assessments (Baylor et. al, 2011, Donovan, Rosenbek, Ketterson, & Velozo, 2006)—to determine whether the *IT-MAIS* is an ecologically valid assessment of early auditory skills. The results of this analysis (p. 40) showed that the *IT-MAIS* did *not* demonstrate ideal item-level psychometric properties because: 1) 2 out of 10 items did not fit the construct and were consequently removed from the analysis; 2) the item difficulty hierarchy did not sufficiently measure the full range of respondents' ability; and 3) the respondents did not use the 5-unit rating scale reliably. In the discussion section of this paper (p. 57), we discuss 3 directions for future research: 1) revise the *IT-MAIS* item order and rating scale categories, 2) establish a theoretical foundation and operational definition for "listening development" and 3) analyze the psychometric properties of existing listening assessments, specifically the *LittlEARS Auditory Questionnaire* (*LittleEARS;* Tsiakpini et al., 2004) and the *Parents' Evaluation of Aural/Oral Performance of Children* (*PEACH*; Ching & Hill, 2005a).

In this document, first the literature review (p. 11) will orient the reader to test development via psychometric analysis. Next, there is an overview of the *IT-MAIS* (p. 21) and its current role in assessing listening skills in young CI users. Finally, the results of our Rasch

analysis on the *IT-MAIS* data are presented (p. 40) and followed by a discussion of our findings

(p. 57), including suggestions for future research (p. 65).

Chapter 2: A Review of the Literature

The American Speech-Language-Hearing Association (2013) calls for evidence-based practice in speech-language pathology and audiology. Evidence-based practice is a valuable framework for providing quality, clinical services (Ratner, 2006; Dollaghan, 2007). The goal of evidence-based practice is to provide appropriate care for patients including diagnosis, developing treatment goals, and tracking progress using valid and reliable outcome measures (Roberts, 2008; Edmonds & Donovan, 2012). However, valid and reliable resources are not always available. The *Infant-Toddler Meaningful Auditory Integration Scales (IT-MAIS;* Zimmerman-Phillips, Osberger, & Robbins, 2001) is an example of a tool that is commonly used, although its validity and reliability have not been formally established. Speech-language pathologists and audiologists assess listening skills in children aged 0 to 3 years with sensorineural hearing loss (SNHL) using the *IT-MAIS*. Cochlear implant (CI) clinical research programs specifically use the *IT-MAIS* to determine CI candidacy and track development of listening skills post-implantation (Barker, Kenworthy, & Walker, 2011; Osberger, Zimmerman, & Koch, 2002; Franz, 2002) despite the fact that it lacks validity and reliability. When pediatric CI programs rely on the *IT-MAIS* to make decisions and track progress post CI-stimulation, the lack of validity and reliability makes their findings (regarding functional listening skills) questionable at best. Using findings from the *IT-MAIS* to establish treatment goals and track listening development risks providing inappropriate intervention for children with SNHL. Given that the *IT-MAIS* is lacking psychometric analyses, in the present study we employed Rasch analysis, a 1-PL model of IRT (similar to measures of validity and reliability on classically

designed tests), to determine its item-level psychometric properties as an assessment of listening skills in children with SNHL aged 0 to 3 years.

Test Development

Psychometric Analysis

Increasing numbers of researchers in healthcare and rehabilitation are using Item Response Theory (IRT, Lord & Novick, 1968)—instead of Classical Test Theory—for developing assessments as well as analyzing existing assessments (Edmonds & Donovan, 2011; Hula, Donovan, Kendall, Gonzalez-Rothi, 2010; Baylor, Hula, Donovan, Doyle, Kendall, & Yorkston, 2011). IRT is a modern test theory methodology in which a participant's true score reflects her *latent trait ability* rather than her numeric score on the test itself (Lord & Novick, 1968). In the case of the *IT-MAIS*, the latent trait assessed is "auditory integration," in 3 subcategories: vocal behavior, alerting to sounds, and deriving meaning from sounds. A latent trait is not directly observable, and must be inferred based on the participant's behavior (Baylor et. al., 2011). For example, auditory comprehension is a latent trait that may be assessed using accuracy on spoken yes/no questions such as, "Are you holding a doll?" Unlike Classical Test Theory, IRT analyzes item difficulty and person ability along a single interval scale, so test administrators are able to compare one participant to another and compare a single participant's performance across time. Furthermore, IRT models are especially well suited for behavioral assessments because the construct assessed is the participant's estimated latent trait ability.

Assessments developed using IRT provide different methods of administering, scoring, and analyzing assessments than those developed using Classical Test Theory. For example IRT allows test developers to make adaptive assessments, thus it is not necessary to administer all items in order to preserve validity (Baylor et al., 2011). Adaptive assessments are valuable

because they allow for more natural testing setting in which the test administrator can use her knowledge and experience to guide the assessment process (instead of adhering to rigid testing procedures). Such an assessment is especially useful when testing children who cannot or will not attend to the assessment for the amount of time necessary for all items to be administered.

Many assessments used in speech-language pathology have undergone item-level IRT analysis in recent years in order to assess the item-level properties of underlying latent trait of each assessment (e.g., *Boston Naming Test*, del Toro et al., 2011; *Functional Assessment of Communication Skills for Adults*, Donovan, Rosenbek, Ketterson & Velozo 2006; *Boston Naming Test*, Graves, Bezeau, Fogarty, & Blair, 2004; *Western Aphasia Battery*, Hula, Donovan, Kendall, & Gonzalez-Rothi, 2010; *Revised Token Test*, Hula, Doyle, McNeil, & Mikolic, 2006). In the present study, we used Rasch analysis, a 1-PL model of IRT to analyze the item-level psychometric properties of the *IT-MAIS*.

Requirements for Utilizing Rasch Analysis

The following parameters must be met to perform Rasch analysis: 1) the latent trait underlying an assessment must be unidimensional (Wright & Linacre, 1989; Edmonds & Donovan 2011; i.e. each item should assess a single latent trait, such as alerting to sound); 2) each test item must discretely assess the underlying trait (i.e. the item assesses vocal behavior only, it does not assess alerting to sounds); and 3) each item must be locally independent: previous items cannot increase the likelihood of a participant achieving the correct response on an item (Wright & Linacre, 1989; Edmonds & Donovan 2011). As a first step in the current study, we analyzed whether the *IT-MAIS* met the requirements for Rasch analysis. Our analyses indicated that the *IT-MAIS does* meet the requirements for Rasch analysis; therefore, we were

able to use Rasch analysis to measure the item-level psychometric properties (comparable to validity and reliability in Classical Test Theory) of the *IT-MAIS*.

Item Response Theory (IRT)

Prior to our description of IRT, it is important to define the unit of measure used in IRT, called the logit. On a linear scale, each logit is a constant representing the change in item difficulty (Donovan et. al 2006,Wright & Stone, 1979). Because IRT uses logits as a constant unit on a linear scale, each response category is placed on a scale that increases by a constant factor. Items ordered by increasing difficulty must increase by a constant interval to infer meaningful comparisons between participants (Edmonds & Donovan, 2011).

Because the *IT-MAIS* developers (Zimmerman-Phillips et al., 2001) only purported, but did not demonstrated that the *IT-MAIS* items were ordered in accordance with the hierarchical development of early listening skills, we used IRT to determine whether or not the *IT-MAIS* items did in fact, measure development of early auditory skills hierarchically (similarly to Erber's proposal (1982)). IRT also allows analysis of item fit, in which the participant's latent trait level is compared with the Rasch modeled latent trait. For example, a child with a high level of listening ability (auditory integration) should consistently achieve high scores on items of low difficulty. Applying IRT as described above provides a framework for developing an ecologically valid and reliable assessment. Ecological validity is important in assessment so that the participant's performance on the assessment reflects her ability in real life settings outside of the testing environment.

Although the *IT-MAIS* is widely used in the field of pediatric CIs, no one has investigated its ability to objectively measure the latent trait of "auditory integration" in infants pre- and post-CI. Furthermore, the authors (Zimmerman-Phillips, Osberger, & Robbins) do not operationally

define "auditory integration." In order to assess skills based on a theoretical construct (i.e., "auditory integration"), test developers must establish a theoretical foundation. The *IT-MAIS* authors did not define "auditory integration" or describe its theoretical foundation. Therefore, we refer to the latent trait assessed by the *IT-MAIS* as "listening development" henceforth in this paper.

If CI programs are going to continue using the *IT-MAIS* for tracking the development of listening skills in CI users, then the *IT-MAIS* must undergo thorough psychometric analysis to determine if it is a valid and reliable tool. Using an assessment with poor validity and reliability does not align with the American Speech-Hearing Association's (2013) definition of evidence-based practice and calls into question the use of the *IT-MAIS* for accurately assessing pediatric CI users because the *IT-MAIS* lacks evidence supporting its psychometric validity and reliability. Particularly, using the *IT-MAIS* as a tool track development of listening skills from birth to 3 years, pre- to post-CI.

Rasch Analysis

Rasch analysis, a 1-PL IRT model, provides a systematic process for transforming ordinal data into interval data, based on probabilistic modeling[1]. Using Rasch analysis, items are plotted on a linear continuum using 2 dimensions: participant's latent trait ability, and item difficulty (Donovan et al., 2006). Results are presented in numerous ways. For example, a Person-Item Map provides a visual representation (Figure 3, p. 49) of both item difficulty and person ability hierarchies on a single interval scale. This graphic provides useful information about the degree to which the items measure the sample's ability level. It also provides

---

[1] It is beyond the scope of this paper to provide an in-depth discussion of Rasch analysis. Interested readers should turn to Fox and Bond (2004) or any of the articles by Ben Wright found in this paper's reference section (p. 70). Here we will give the reader an overview of Rasch analysis and the role it plays in developing or evaluating assessments.

information about how well the items are distributed across the sample, and the number of levels of ability that the sample may be separated into (Baylor et al., 2011). Along with analysis of individual items, Rasch analysis provides a measure that corresponds to the validity of an assessment. Rasch analysis provides test developers and researchers with an interval scale, allowing more assessment and score comparisons than a classically designed test, which may employ an ordinal or nominal scale (Wright & Linacre, 1989; Edmonds & Donovan, 2011). An interval scale allows comparisons between participants on both overall performance and responses on individual items. Furthermore, interval scales provide clinicians the ability to report a client's progress or decline in skills pre- and post-intervention. For example if a plant grows from 1 inch to 2 inches, we can say the plant has doubled in size because, a) the inch is a fixed unit of measurement that is equal across the entire scale, and b) the inch is an interval measure that allows us to perform mathematical computations which cannot be done with nominal or ordinal scales (i.e. describing the plant as "shorter" or "taller"). In this study, we utilized Rasch item analysis with *IT-MAIS* scores taken from 23 pediatric CI users ages 11 to 36 months of age[2] taken at various time points, resulting in a total sample of 56 ratings.

Specifically, we used a polytomous scale analysis of the Rasch model (a 1-PL IRT model) to analyze the *IT-MAIS* and assess its content validity, construct validity, and reliability. The Rasch polytomous formula models the relationship between a participant's ability (i.e. trait level) and the probability of her choosing each response category (i.e. 0 - 4 Likert scale of the *IT-MAIS*) for each item. [For the polytomous formula, see p. 35]

---

[2] Rasch analysis is considered ideal for analysis for small samples (50-100). We anchored the 23 scores according to Rasch methods and by so doing, the other ratings contribute as individual ratings to estimate the model (Mallinson, 2011).

Rasch analysis assumes participants' discrimination and guessing are consistent across all items (Edmonds & Donovan, 2011). The polytomous model also permits analysis of category response curves for each item, from which one can determine whether all rating-scale categories are used or if they overlap. Overlap may indicate that respondents do not recognize the qualitative differences in the units (e.g., does "sometimes" versus does "often"; Baylor et al. 2011; Figure 1). In analyzing the *IT-MAIS,* the response curve graphs serve to easily visualize whether participants are using all units of the rating scale reliably for each item to assess a unique aspect of the underlying construct. For the purpose of our analyses, the underlying construct is listening development.
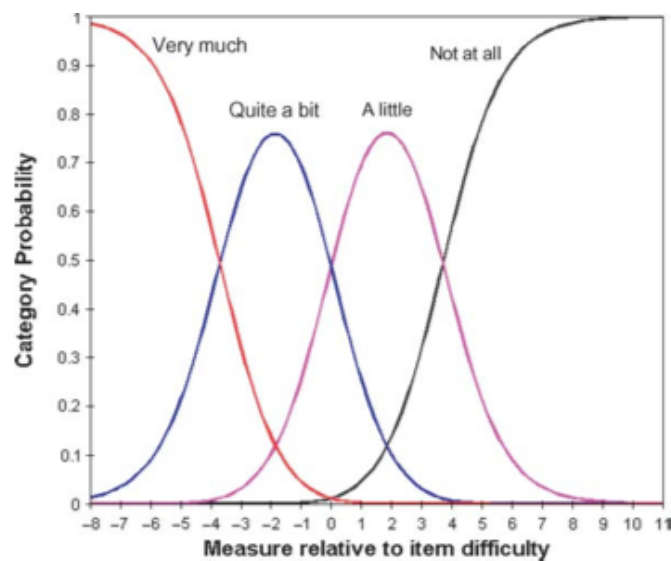
Figure 1. Category response curve taken from Baylor, et al (2011) representing an appropriate amount of overlap between categories on a Likert scale (i.e. "very much", "quite a bit", "a little", and "not at all").

Assessing Cochlear Implant Candidacy and Post-Implant Listening Development

A CI is an electronic device that requires surgical implantation of an internal receiver in the temporal bone and its attached electrode array in the cochlea. The electrode array directly stimulates the cochlea. The cochlea stimulates the auditory nerve, which sends a signal to the brain where the signal is processed and perceived as sound.

Individual differences across CI users, including age of implant, affect the person's effectiveness in utilizing the device's signal, even when the brand and model of CI device is held constant (Baudhuin, Cadieux, Firszt, Reeder, & Maxson, 2012; Neuman, Wroblewski, Hajicek, & Rubinstein, 2012). Once implanted, CIs are expected to improve speech, language, and listening skills in their recipients (Blamey & Sarant, 2000; Geers & Moog, 1994; Geers, Nicholas, & Sedey, 2003; Tomblin, et al., 1999). Tools used to evaluate pediatric CI users should ideally assess baseline skill levels pre-CI as well as measure longitudinal skill development post-CI in order to accurately track listening development and warranty appropriate, customized intervention. To quantify improvements in listening skills, it is important to assess CI users' using ecologically valid and reliable measures to ensure evidence-based practice when establishing CI candidacy and providing speech, language, and listening intervention (Edmonds & Donovan, 2011). The *IT-MAIS* was developed with such assessment and intervention in mind, but its validity and reliability are not established via psychometric analysis.

Infant-Toddler Hearing Assessment

Early cochlear implantation requires early identification of sensorineural hearing loss (SNHL). At birth, the primary tools for screening and assessing peripheral hearing function are: Auditory Brainstem Response, Auditory Steady-State Response, and Otoacoustic Emission testing. While these objective measures are indispensable for measuring peripheral hearing as well as middle and inner ear integrity, it remains the "gold standard" to assess the functional listening abilities of children (particularly CI candidates) using behavioral responses to sound (Martin & Clark, 2009). Unlike physiological and electrophysiological tests, behavioral assessments require an overt response from the child, thus assessing integration between the peripheral and central auditory systems (i.e. *functional* auditory perception and processing). Pure

tone audiometry cannot be used for many children in the first year of life because children 6-months-old and younger lack head and neck control (and sometimes, cognitive skills) necessary for Visual Reinforcement Audiometry. Alternatively, caregiver report is a common tool for assessing behavioral responses to sound in everyday listening environments. A caregiver report tool aims to assess a child's skills without directly eliciting overt responses (as is accomplished in a behavioral assessment). Thus, psychometric analysis is particularly important for parent report tools in order to be considered an ecologically valid assessment

As CI technology improves, the age at which children undergo implantation continues to decrease. In recent years, infants younger than 12 months of age received CIs (Colletti et al., 2012; Colletti et al., 2011; Hammes et al., 2002; Wright, Purcell, & Reed, 2002), despite the possible risks associated with early implantation, such as increased anesthetic risks (especially in children under 1 year old) and blood loss due to low total blood volume in young children (Waltzman & Roland, 2005).

Due to infants' cognitive and physical limitations, behavioral assessments of listening skills are not reliable. Currently, caregiver report tools serve to assess functional listening skills in CI candidates. The *IT-MAIS* (Zimmerman-Phillips, et al., 2001) is a caregiver-report tool used to assess auditory skills in infants and toddlers ages 0 to 3 years. Currently, the *IT-MAIS* is widely used to assess listening skills in young CI users, even though it lacks psychometric analysis suggesting that it is ecologically valid. (See Appendix A for the *IT-MAIS.*) Our study employed Rasch analysis to determine if the *IT-MAIS* is a valid tool for assessing infants' and toddlers' auditory skills pre-CI and tracking development of auditory skills post-CI. Specifically, we assessed item difficulty, item and person fit, as well as participants' use of each category on the Likert rating scale used by the *IT-MAIS*.

Value of Assessing Listening Development

Listening Development in Cochlear Implant Users

Again, the CI provides auditory information via direct electrical stimulation to the auditory nerve. Thus, the stimulation from a CI is significantly different than acoustic hearing (House, 1982). For instance, the CI device filters the speech signal into different, restricted frequency bands from 600-6000 Hz, whereas a healthy cochlea uses the pressure changes in the cochlear fluid to provide detailed frequency information to the auditory nerve ranging from 20-20,000 Hz (Martin & Clark, 2009). Unlike listeners with normal-hearing thresholds, CI users must actively learn to utilize the CI's restricted auditory information in order to benefit most from the CI (Tomblin et al., 2005).

Hierarchy of Listening Development

For listeners with normal hearing, Erber (1982) proposed that listening is a complex task that must be learned. He proposed learning to listen develops in 4 sequential steps: 1) sound detection, 2) sound discrimination, 3) sound identification, and 4) comprehension of sound. Among CI users, individuals differ in their rates of development and accuracy when advancing through Erber's hierarchy and learning to listen (Houston, et al., 2003; Barker, et al., 2011). Having a thorough understanding of a CI user's progress through Erber's (1892) listening hierarchy provides insight to her ability to perceive, process, and comprehend the auditory input from the electrical stimulation provided by her CI.

Again, the *IT-MAIS* addresses the following areas: vocalization behavior, alerting to sounds, and deriving meaning from sound (Zimmerman-Phillips, et al., 2001). The authors also compose the construct of "auditory integration" without operationally defining it (Zimmerman-Phillips, et al., 2001, p. 2). Pediatric CI programs use the *IT-MAIS* to track progress pre- to post-

CI. In order to track progress in listening development, items must be ordered based on an appropriate developmental listening hierarchy. To determine if *IT-MAIS* item order reflected a developmental hierarchy, we analyzed whether or not the *IT-MAIS* item order aligned with Erber's hierarchy of listening levels.

We analyzed item order because CI programs use the *IT-MAIS* to track listening development pre- to post-CI; however, the *IT-MAIS* authors do not describe a theoretical foundation on which item order was established. Establishing a hierarchy for pediatric CI users' listening development would provide insight to whether or not pediatric CI users develop listening skills in the same manner as their peers with normal hearing. Children with profound SNHL develop cognitive and language skills prior to CI stimulation; however, listening development is delayed due to minimal pure tone averages (e.g., > 90 dB) pre-CI. Children with normal hearing begin develop listening skills while concurrently developing cognitive and language skills. Due to these differences in listening development between pediatric CI users and children with normal hearing, is important to compare the trajectory of listening development in these 2 groups in order to establish a listening development hierarchy for pediatric CI users.

IT-MAIS Development

The *MAIS*

The *IT-MAIS* was derived from the *Meaningful Auditory Integration Scales* (*MAIS;* Robbins, Renshaw, & Berry, 1991). Like the *IT-MAIS*, the underlying construct assessed by the *MAIS* is "auditory integration;" however, the *MAIS* authors also do not operationally define the term "auditory integration." (See Appendix B to view the *MAIS*.) The *MAIS* assesses meaningful use of sound in children with profound hearing loss aged 5 years and older. The 3 dimensions evaluated by the *MAIS* are: bonding to the device, alerting to sound, and deriving

meaning from sound (Robbins, Renshaw, & Berry, 1991). The *MAIS* consists of 10 items. The first 2 items were subsequently replaced to yield the *IT-MAIS*. The first 2 items of the *MAIS* address the child's bonding to the hearing assistive device (hearing aid, CI, or tactile aid; Robbins, et al., 1991). Alternatively, the *IT-MAIS* replaced the aforementioned items with 2 items addressing sound vocalizations, which are behavioral markers often associated with listening development in children with CIs (Connor, Craig, Raudenbush, Heavner, & Zwolan, 2006; Ertmer & Jung, 2012). The *MAIS* was developed for use with children with profound SNHL ages 5 years and older, whereas the *IT-MAIS* is used for children with profound SNHL under 3 years of age.

Like the *IT-MAIS*, the *MAIS* also lacks psychometric and empirical support. Initially, the *MAIS* was evaluated in a publication consisting of 2 different experiments (N = 50, Robbins et al., 1991). In Experiment 1, the *MAIS* was used to compare the effectiveness of different types of sensory aids (the Tactaid II, the single-channel 3M, the 22-channel Nucleus CI, and a conventional behind-the-ear hearing aid) in 50 children. Across the 4 groups of children, age of onset, age of device fitting, and years of device use were not closely matched (length of use ranged from 1.4 - 5.2 years). Closely controlled age matching is imperative when comparing experimental groups of children because development occurs rapidly and one would not expect equal performance from children age 1.4 years and those aged 3.5 years. Therefore, the CI users' age could have affected the participants' performance on the *IT-MAIS*. Thus, chronological age may have been a confounding variable (rather than device type alone affecting *IT-MAIS* scores).

Across the 50 participants, this experiment showed a significant main effect of group. Hearing aid users achieved the highest *IT-MAIS* scores, followed by both groups of CI users who received similar scores, and Tactaid II users achieved the lowest scores (Robbins, Renshaw, &

Berry, 1991). Experiment 1 also included analysis of inter-rater reliability using *MAIS* scores

from a subset of the participants (n = 28), resulting in a correlation coefficient of 0.90,

suggesting high inter-rater correlation between the 2 raters. Finally, Experiment 1 demonstrated

differences in performance on the *MAIS* based on type of device (Robbins, Renshaw, & Berry,

1991). These results demonstrated that performance on the *MAIS* showed sensitivity based on CI

device type. In order to assess the *MAIS'* validity and reliability for assessing listening

development in children ages 5 years and older, further psychometric analyses are required

Experiment 2 assessed the change of *MAIS* scores over time using a different subset of

the participants from Experiment 1 (n = 30). The parents of these children were administered the

MAIS at 6 months, and 1 year post-CI. Experiment 2 demonstrated that the children at 1 year

post-CI achieved higher scores on the *MAIS* than when they were 6 months post-CI. These

experiments provided data suggesting that overall *MAIS* scores correlated with an increase in age

and length of device use. The authors, however, did not demonstrate thorough psychometric

analysis in developing this assessment. Specifically, there was no discussion of whether the

assessment's underlying construct (listening development) is itself a valid construct to assess.

Establishing the theoretical underpinnings of an assessment (i.e. listening development) is an

essential part of test development. Without theoretical support, the validity of the test is at risk.

The researchers (Robbins, Renshaw, & Berry, 1991) also do not address how the rating scale was

developed, nor do they provide evidence that their item order was valid. The items are presented

in a hierarchy, but no explicit support was provided for how item order was determined. The *IT-*

MAIS uses 8 out of 10 items from the *MAIS*, thus the lack of psychometric analysis performed on

the *MAIS* is reflected in the validity and reliability of the *IT-MAIS*. These weaknesses raise

concern for using the *IT-MAIS* in clinical practice to assess children's listening skills and develop

therapy goals based on *IT-MAIS* results. In the present study, we performed Rasch analysis to analyze the psychometric validity and reliability of the *IT-MAIS*.

Assessing the *MAIS* via Classical Test Theory

The *Listening Progress Profile (LiP;* Archbold, 1994) is a profile measuring development of listening skills ranging from sound detection (environmental and speech), voice discrimination, to identifying one's own name (Nikolopoulos et al., 2000). Concurrent validity between the *MAIS* and the *LiP* was assessed using Classical Test Theory, reporting all correlation coefficients at significant levels ($R > 0.6$; Weichbold et al. 2004). Particularly, the data showed that the overall scores on the *MAIS* correlated with scores on the *LiP* for the 82 children with SNHL in the study, suggesting concurrent validity between assessments. The *LiP* has demonstrated sensitivity in assessing listening development in prelingually deafened CI users (Nikolopoulos, Wells, & Archbold, 2000; Nikolopoulos, O'Donoghue, Robinson, Gibbin, Archbold, & Mason, 1997); however, the *LiP* has not undergone thorough psychometric analysis. Thus, the *IT-MAIS'* concurrent validity with the *LiP* is not sufficient evidence to support psychometric validity of the *IT-MAIS*.

In summary, when developing the *IT-MAIS*, the first 2 items of the *MAIS* were replaced and the target population changed. The Classical Test Theory analysis of the *MAIS* was minimal, and no IRT analysis was performed (Robbins, Renshaw, & Berry, 1991; Nikolopoulos et al, 2000), suggesting that the *MAIS* is an assessment with unknown psychometric properties and questionable validity. These limitations of the *MAIS* suggest that the *IT-MAIS* lacks psychometric validity given that 8 out of 10 items on the *IT-MAIS* were directly taken from the *MAIS*. As previously stated, psychometric analysis is essential in developing an assessment tool in order to analyze the validity of the underlying construct and determine if the items reflect that

construct reliably. Using assessments that are reliable and valid are essential in order to provide accurate diagnoses and appropriate therapy in clinical settings (i.e. aural (re)habilitation in pediatric CI users).

IT-MAIS Development

When developing the *IT-MAIS*, the authors (Zimmerman-Phillips, et al., 2000) conducted a study assessing the overall scores of 9 total children pre- and post-CI. Children, ages 18 to 23 months, were assessed using the *IT-MAIS* during hearing aid trials pre-CI, and at 3 months post-CI. The first 2 items address vocal behavior, followed by 4 items addressing alerting to sounds, and the last 4 items address the child's ability to derive meaning from sounds. Each item in the *IT-MAIS* requires the caregiver to describe the child's behaviors using a 5-point Likert scale. The scale ranges from 0 to 4 based on frequency of the behavior (*0 / never, 1 / rarely, 2 / occasionally, 3 / frequently, 4 / always*).

In their 2000 study (Zimmerman-Phillips, Osberger, & Robbins) the majority of the children pre-CI received scores of *0 / never* on all items, suggesting the caregivers *never* witnessed any of the listening or vocalization behaviors. At 3 months post-CI, all participants demonstrated an increase in frequency of at least 7 out of 10 items (Zimmerman-Phillips et al., 2000). The authors concluded that these increases demonstrated that the *IT-MAIS* was a valuable tool to measure CI candidacy and benefit. The conclusion made by Zimmerman-Phillips and colleagues is too broad, given the weaknesses of their (2000) study. First, the small sample size and the group improvement in ratings are two important weaknesses in Zimmerman-Phillips, Osberger, and Robbins' work (2000). The small sample size (N = 9) is likely to provide results that do not appropriately represent the entire population of children 0- to 3-years-old with profound SNHL. Another limitation of the study is that the authors did not assess the validity of

the individual items; rather the authors summarized the overall changes of the group in responses between the pre- and post-CI assessments. However, these between-group differences could represent natural development of listening skills, as the post-CI group is older than the pre-CI group. Ideally, each item must be analyzed in order to determine whether the items measure unique aspects of the underlying construct (i.e. listening development) with increasing difficulty with each item, thus representing a developmental trajectory. Item analysis is important because the authors present the 10 items as if they represent increasing levels of difficulty. If item order based on difficulty (as measured by Rasch analysis) does not align with the item order in the *IT-MAIS*, it indicates that the *IT-MAIS* does not accurately indicate a child's stage in listening development. Assessing listening development is especially important for tracking progress in children post-CI as they learn to listen using electric input from the CI device. The results of Zimmerman-Phillips, Robbins, and Osberger study (2000) provided minimal evidence supporting the *IT-MAIS* efficacy, and do not address construct, concurrent or ecological validity of the *IT-MAIS*.

Ecological validity is essential in order to ensure that an assessment's results reflect a participant's abilities in everyday life. Ecological validity is particularly important for CI users because everyday settings like daycare, school or home are complex listening environments (i.e., music, conversations, television shows). Background noise is often removed from research laboratory and clinical settings; therefore, assessments performed in these settings often do not reflect listening skills in the complex listening environments found in everyday life. Attending to speech in noise is particularly difficult for CI users (Caroll & Zeng, 2007). Thus, assessments measuring listening skills in CI users should be valid outside of the laboratory's/clinic's sound booth in natural, complex listening environments. In order to appropriately assess auditory skills

in 0- to 3-year-old children the *IT-MAIS* must be ecologically valid. Currently, the *IT-MAIS* lacks

evidence supporting its validity and reliability; therefore clinical decisions (i.e. CI candidacy,

therapy goals) based on findings from the *IT-MAIS* are questionable at best.

<div align="center">

Problems with *IT-MAIS* Standardization

</div>

Classical Test Theory Analysis of the *IT-MAIS* using Chinese Infants with Normal Hearing

To date, only one study (Zheng, et al., 2009) applied psychometric analysis to the *IT-MAIS*.

In their study, Zheng and colleagues employed Classical Test Theory to determine the validity

and reliability of the Mandarin-translated *IT-MAIS* based on responses from 120 Chinese

children with Mandarin-Chinese-speaking caregivers. All children had normal hearing

thresholds. The authors used early pre-lingual auditory development as the underlying construct

upon which they sought to establish construct validity of the *IT-MAIS*. The authors compared the

participants' *IT-MAIS* scores to the developmental trajectory for normally developing infants and

toddlers of Hebrew- and Arabic-speaking parents, as reported by Kishon-Rabin (2001)[3]. Zheng

and colleagues (2009) reported that both Cronbach's α coefficient and Guttman's split-half

reliability coefficient using items 2 - 10 and items 3 - 10 of the *IT-MAIS* exceeded 0.90,

indicating strong internal consistency between the *IT-MAIS* and their unspecified measure of

auditory development. The authors concluded the items measured aspects of the same construct,

early pre-lingual auditory development (Zheng et al. 2009), thus suggesting construct validity;

however, authors created the construct "early pre-lingual auditory development" without

operationally defining it or citing evidence regarding its validity as a construct.

Weaknesses of Zheng et al. (2009)

There are a number of limitations of the Zheng et al. (2009) study. First, the English *IT-*

---

[3] Currently we only have access to the norms. The author was contacted and we requested the presentation on this
study in English, but have yet to receive a response. We cannot evaluate the study as a whole without a summary or
copy of the study translated into English.

MAIS was translated into Mandarin-Chinese for use with Chinese infants. While the authors report that translating the test from English to Mandarin Chinese was "straightforward", Mandarin Chinese and English are fundamentally different languages, particularly when examining their acoustic/auditory characteristics. For example, Mandarin Chinese relies on tonal qualities and English stress-timed language that does not rely on tonal qualities (Luo & Fu, 2004). This is particularly important because research shows that CI users receive degraded tonal information from the speech signal (relative to other aspects of the speech signal; Luo & Fu, 2004). Because the CI device does not directly encode the tonal information of the Mandarin speech signal, the performance of children with CIs exposed to Mandarin Chinese cannot be compared to those children with CIs exposed to English.

A second limitation of this study is the fact that the scores from children who reached ceiling on one or more questions were discarded. When a participant reaches a ceiling performance, it indicates she has advanced beyond the difficulty level of that item. If there is a significant ceiling effect, the item itself should be altered because the ceiling effect shows that the assessment does not provide items difficult enough to represent the entire sample. Alternatively, Zheng and colleagues (2009) did not report the ceiling effect as a weakness of the *IT-MAIS* and instead discarded the data points at the ceiling and then formed their conclusions about the *IT-MAIS*.

"Early Pre-Lingual Auditory Development" is an arbitrary construct term used consistently by the authors, but never operationally defined. This term must be operationally defined in order to validate the construct as a dependent variable and provide other researchers the ability to reproduce the methodology used by Zheng and colleagues. Designing a study so that it can be replicated is an important component of the scientific method. Replicating results

also helps to confirm a study's conclusions; however, Zheng and colleague's (2009) study cannot be replicated without an operational definition of "Early Pre-Lingual Auditory Development."

Furthermore, it is unclear how the authors assessed the children's auditory skills. Based on Erber's (1982) listening levels, the developmental hierarchy of listening skills is as follows: detection, discrimination, identification, and comprehension. We analyzed whether the hierarchy of *IT-MAIS* items adhere to the hierarchy of Erber's listening levels, which have been operationally defined over the years and are accepted in the field of aural (re)habilitation as an appropriate hierarchy of listening skills. The items on the *IT-MAIS* are presented as a hierarchy as well; therefore, we used Erber's levels to develop our hypotheses regarding *IT-MAIS* item order.

Finally, recall the children in the Zheng et al. study all had normal hearing. The *IT-MAIS* was designed to assess children ages birth to 3 years with severe to profound bilateral SNHL pre- and post-CI—not children with normal hearing. The psychometric analysis of the *IT-MAIS* needs to be performed using the population of interest in order to form conclusions on its validity and reliability as an assessment for *that* specific population. Developmental norms were established using the *IT-MAIS* for Hebrew-speaking children with normal hearing (Kishon-Rabin 2004), but not for CIs users. We analyzed pre- and post-CI *IT-MAIS* scores from 23 children who are bilaterally deaf. Using our data from these CI users, we examine whether listening skills develop based on a similar or different hierarchy in children with normal hearing than in children using CIs.

The Present Study

Given the literature reviewed within this chapter, using Rasch analysis to assess the validity and reliability of the *IT-MAIS* is the natural next step in determining the *IT-MAIS'* value

as an assessment tool for measuring auditory skills in young children who are deaf and use CIs. CI technology is improving and the age at which children undergo cochlear implantation is rapidly declining (Colletti et al. 2011). Thus, there is an urgent need for an ecologically valid tool to assess CI candidacy and post-CI benefit in children ages 0 to 3 years and to ensure evidence-based intervention for both CI candidates and users.

The *IT-MAIS* is currently used to evaluate the listening skills of children who with SNHL pre- and post-CI. Recall that the *IT-MAIS* was developed from the existing *MAIS*, which was designed for children with a different age range of deafness onset (Robbins, et al., 1991; Zimmerman et al., 2001). Neither the *MAIS* nor the *IT-MAIS* was developed using IRT, a model suited to provide item-level analysis of the assessments. Therefore, just like the *MAIS*, the *IT-MAIS* lacks confirmation of its validity and reliability as an assessment tool. So, continued use of the *IT-MAIS* without confirming its validity and reliability does not adhere to the principles of evidence-based practice. The lack of an evidence-based assessment does not allow us to make the most appropriate choices about which children are the ideal CI candidates. The decision to implant means a loss of any residual hearing that the child could ultimately utilize via hearing aids or other future technologies. In children post-CI, evidence-based assessments are essential to describe progress in listening development. Rasch analysis provides information on content and construct validity (Edmonds & Donovan 2011) and we argue it is one of the most informative psychometric analyses for an assessment with unknown item-level properties. Validity and reliability both overall and at the item level are essential to an ecologically valid assessment (Baylor et al. 2011). This project's aim was to assess the validity and reliability of the *IT-MAIS* (at the item-level and overall) using the one-parameter IRT model Rasch analysis. Our research questions were as follows:

1. Does the *IT-MAIS* data meet the assumptions for Rasch analysis?

2. Does the *IT-MAIS* demonstrate item-level psychometric properties, reflecting that the latent trait (listening development) can be measured?

3. Does the *IT-MAIS* demonstrate adequate range of person ability and item difficulty to separate respondents into more than 2 levels?

4. Do the items on the *IT-MAIS* reflect content validity based on a priori item hierarchy rankings?

Chapter 3: Methods

We used Rasch analysis to assess the validity (construct and concurrent) and the reliability of longitudinal *IT-MAIS* data collected by the University of Iowa Children's Cochlear Implant Program.

Participants

Parents of 23 CI users aged 10 to 36 months and receiving services from the University of Iowa Children's Cochlear Implant Program completed the *IT-MAIS* during monthly visits to the center for CI candidacy assessments and post-CI care. A total of 23 children (12 male, 11 female) participated in this study. All children were born to parents with normal hearing and they were identified with bilateral SNHL within the first year of life. All caregivers reported spoken English as the primary language used at home. See Table 1 for participant demographic data.

Procedures

Upon visiting the CI center, children and caregivers participating in the comprehensive longitudinal study were administered a battery of communication assessments administered by a pediatric CI clinical researcher (other assessments include: the *Preschool Language Scales-4th Edition* (Zimmerman, Steiner, & Pond, 2002) and the *Minnesota Child Development Index* (Ireton & Thwing, 1974)). Also included in this test battery was the *IT-MAIS*. The *IT-MAIS* was administered at least twice (once pre-CI and once post-CI), and most children were assessed additional times post-CI. Pediatric CI audiologists administered and scored the *IT-MAIS*. Assessment interims ranged from between 1 month to 1 year. Assessments were scheduled for administration at 2-month intervals during the first year after CI stimulation, then at 6-month intervals until 3 years post-CI. The actual intervals between assessment dates varied due mostly to missed visits (as is a challenge when collecting longitudinal data from a clinical population).

Using these repeated measures, we analyzed a total of 56 data points (N = 23).

Table 1. Demographic data for the 23 CI users.

ID	M/F	age at HL	age at CI	CI device type	strategy	CI ear	HA use	HL etiology	mother's ed
1	M	0	11	Nucleus CI 24 RE(CA)	ACE	B	none	Waardenburg Syndrome	college
2	F	0	13	Nucleus CI 24 RE(CA)	ACE	B	B 11	CMV	post-grad
3	M	0	12	Nucleus CI 24 RE(CA)	ACE	B	B 6	hereditary (unspecified)	HS
4	M	0	16	Nucleus CI 24 RE(CA)	ACE	B	B 10	Auditory Neuropothy	college
5	M	0	14	Nucleus CI 24 RE(CA)	ACE	B	B 7	unknown	post-grad
6	M	0	14	Nucleus CI 24 RE(CA)	ACE	B	B 11	unknown	post-grad
7	F	0	12	Nucleus CI 24 RE(CA)	HiRes	B	B 5	unknown	college
8	F	0	22	Nucleus CI 24 RE(CA)	ACE	L	-	CMV	college
9	F	0	13	Nucleus CI 24 RE(CA)	ACE	B	none	CMV	post-grad
10	M	0	16	Nucleus CI 24 RE(CA)	ACE	B	B 12	unknown	-
11	M	-	13	Nucleus CI 24 RE(CA)	ACE	B	-	-	-
12	F	0	13	Nucleus CI 24 RE(CA)	ACE	B	none	unknown	college
13	M	-	12	Clarion HiRes 90K	HiRes	B	-	-	-
14	F	-	13	Nucleus CI 24 RE(CA)	ACE	B	-	-	HS+
15	F	0	9	Nucleus CI 512(CA)	ACE	B	B 6	Jervell Lange-Nielsen Syndrome	-
16	F	-	13	Nucleus CI 24 R(CA)	ACE	B	-	-	-
17	F	0	13	Nucleus CI 24 RE(CA)	ACE	B	B 5	unknown	HS+
18	M	2	13	Nucleus CI 24 RE(CA)	ACE	B	B 11	Connexin 26	-
19	M	-	14	Nucleus CI 512(CA)	ACE	B	B 7	-	-
20	F	-	13	Clarion HiRes 90K	HiRes	B	-	-	-
21	M	0	12	Nucleus CI 512(CA)	ACE	B	B 10	unknown	-
22	M	0	20	Nucleus CI 24 R(CA)	ACE	R	B 14	unknown	college
23	F	0	17	Nucleus CI 24 R(CA)	ACE	B	B 11	-	college

Note: **ID** = participant identification number; **M/F** = male or female; **age at HL** = participant's age when diagnosed with bilateral hearing loss (months); **age at CI** = participants' age at time of CI implantation (months); **CI device type** = cochlear implant device(s) used by the participant; **strategy** = processing strategy; **CI ear** = implanted ear (R = right CI, L = left CI, B = bilateral CIs); **CI use** = amount of time the participant used his/her CI relative to the time at testing (years; months); **HA use** = hearing aid use prior to implantation (R = right HA, L = left HA, B = bilateral HAs); **HL etiology** = hearing loss etiology (CMV = Cytomegalovirus, unknown = genetic testing completed but inconclusive); **mother's ed** = highest level of education self-reported from mother (HS = high school diploma, HS+ = some college classes, college = college degree, post-grad = post-graduate degree); - no information

Question 1: Does the *IT-MAIS* data meet the assumptions for Rasch analysis?

Exploratory Factor Analysis

First, we conducted an exploratory factor analysis (EFA) to test the assumption of unidimensionality of the *IT-MAIS'* 10 questions using *SPSS Factor Analysis.* Unidimensionality is an assumption of Rasch analysis and it must be present before further analyses can be performed. Factors were extracted that had a minimum eigenvalue >1.0, $\alpha = .05$.

We predicted the *IT-MAIS* items would not comprise a single construct; we believed that the 3 subcategories (vocal behavior, attending to sound, and deriving meaning from sound) assessed different factors, rather than all contributing to one factor, "listening development."

To meet the local independence assumption, we transformed the inter-item residual correlations into Fisher's *z* scores. In that form, we characterized local independence among items as $\leq 5\%$ of the non-significant pairs with correlations ≥ 2 SD from the mean (Fendrich et al., 2009; Smith, 2005).

Rasch Analysis

We chose to conduct the 1-PL Rasch analysis for polytomous rating scales using WINSTEPS 7.5 (2012). We chose Rasch analysis because it is the simplest IRT model, and it can be completed using small samples (i.e. 50 – 100, Linacre, 1994; Fox, 2007)[4]. Note that there is an ongoing sample size debate between Rasch and other IRT-model proponents (Embretson, 1996; Embretson & Reise, 2000; Hambleton, Swaminathan, & Rogers, 1991; Wright & Masters, 1982) and some measurement specialists have suggested that it is critically important to model discrimination (2-PL model) and guessing (3-PL model; Baylor, Hula, Donovan, Doyle, Kendall, & Yorkston, 2011). However, we utilized the 1-PL model in the present study. The 1-PL Rasch

---

[4] Rasch analysis is particularly useful for the present study because it provides a valid means to analyze sample sizes as small as 50. Rasch analysis allows us to anchor the scores from our small sample size (N=23) and use the remaining data to estimate the model (Mallinson, 2011).

model formula for polytomous rating scales used in this study is represented as Linacre (2002):

$$\text{Log } (P_{nik}/P_{ni(k-1)})*B_n-D_i-F_k$$

Where:

P_{nik} = the probability that person n, on encountering item I would respond (or be

observed) in category k,

$P_{ni(k-1)}$ = the probability that the response (or observation) would be in category $k-1$,

B_n = ability of person n,

D_i = difficulty of item i

F_k = a rating scale threshold defined as the location corresponding to the equal probability

of observing adjacent categories $k-1$ and k

To determine if the *IT-MAIS* met the parameters of Rasch analysis (i.e. item

unidimensionality), we analyzed individual item scores from 23 CI users pre- and post-CI. We

analyzed the structural validity of the *IT-MAIS* by assessing unidimensionality, item local

dependence, and Rasch analysis of item fit (Fendrich et al., 2007; Wolfe & Smith 2007).

Question 2: Does the *IT-MAIS* demonstrate item-level psychometric properties, thus reflecting that the latent trait (listening development) can be measured?

Item-level Psychometric Information Provided by Rasch Analysis

As stated earlier, the usefulness of any of the IRT models is that they permit a researcher

to examine the test at the level of item ability and person difficulty (and other parameters for

more complex 2- and 3-PL IRT models) rather than at the total test score level. What follows is

the item-level psychometric information used to determine how the items function, and response

reliability based on the following measures: item infit, item difficulty, person ability, and person

reliability. Our hypotheses for each analysis are presented in the following subsections.

Item InFit Statistics

In test development, it is critical that the test items are appropriate for the participants' ability levels across the full range of ability the test is designed to assess (i.e. do the items capture the trait they were developed to measure?). The item fit statistics are χ^2 values based on standardized item-person residuals (Wright & Masters, 1982). Item-person residuals are derived from unexpected responses around the person's ability. If the model and responses are a perfect match, then the fit statistic would be 1.0; however, because we are dealing with human behavior, the likelihood of perfect agreement is quite low.

There are a number of rationales found in the literature for establishing item fit criteria (Linacre, 2003; Smith, Schumacker, & Bush, 1998; Wright & Linacre, 1994). We chose a common metric of acceptable item infit set to accept infit mean square values (MnSq) < .6 and MnSq ≥ 1.4 and standardized *z*-scores (Zstd) > 2.0 (Wright & Linacre, 1994). For this study, first item fit criteria was set ((MnSq) < .6 and MnSq ≥ 1.4 and standardized *z*-scores (Zstd) > 2.0) and then misfitting items were removed one run at a time until all items exhibited adequate fit to the Rasch model.

Item Hierarchy

We were interested in examining the *IT-MAIS'* item difficulty hierarchy since the test is used to track development of listening skills. Ideally the items should measure the full range of the latent trait, without yielding significant floor or ceiling effects. We predicted that, during pre-CI assessment, children would receive scores of 0 on 8-10 of the items, thus indicating the *IT-MAIS* is an insufficient measure of their listening development because it lacks items appropriate for pre-CI assessment and/or children with hearing loss. Our hypothesis was motivated by the fact that the *IT-MAIS* was designed to assess infants' and toddlers with profound SNHL both pre-

and post-CI. However, due to profound, bilateral SNHL and 0% audibility of the speech signal, children pre-CI were expected to attain scores of *never* (i.e. 0) on all items, because the assessment measures listening skills that are not physically possible.

Item Mean/Person Mean

Comparing the item M to the person M provides an indicator of a test's internal consistency. When the 2 calibrated Ms are similar (for this study an acceptable item M/person M match was set as the actual item M/person M $\pm$ 1 SD), it indicates the test items have approximated the sample's ability. If the item difficulty and person ability match exactly results would indicate an item difficulty M = 0 and person ability M = 0 $\pm$ 1 SD. We predicted the person ability range to be much greater than the item difficulty based on our knowledge of listening development in conjunction with our judgment of the listening skills addressed by the *IT-MAIS* items.

Rating Scale Analysis

We employed rating scale analysis to assess how respondents use the units of the rating scale. Participants' responses on an ordinal rating scale can be unpredictable. In order to measure the stability of a rating scale system, Linacre (2002) established three criteria: 1) there must be at least 10 observations in each rating category, 2) the categories must advance in a step-wise fashion from lowest to highest, and 3) OUTFIT MnSq < 2. If the *IT-MAIS'* 0 - 4 rating scale met the established criteria, it would demonstrate that the units are in fact, being used in the way the developers intended (Linacre, 2002; Donovan, Velozo, & Rosenbek, 2007). We predicted the units on the rating scale would not be used predictably because the *IT-MAIS* brochure (Zimmerman-Phillips, et al., 2000) does not provide specific requirements for test administrators or clear instructions for eliciting responses from caregivers during the assessment. We predicted

that variability in *IT-MAIS* administrators and weak item verbiage would result in participants using the rating scale units unpredictably.

Person Reliability

The person reliability statistic is comparable to Cronbach's α, reported in Classical Test Theory. Cronbach's α reflects a measure of the relationship among test items (Cronbach, 1951). Thus a high Cronbach's α would suggest that items have a close relationship and should be included in the same set. An acceptable Cronbach's α is in the range of 0.8 to 1.0 (Cronbach, 1951). For this study, an acceptable person reliability statistic was set at 0.8 or above. We predicted that variability in administration (due to the same concerns for rating scale analysis, p. 38) would result in unacceptable person reliability measures ($α < 0.8$).

Question 3: Does the *IT-MAIS* demonstrate adequate range of person ability and item difficulty to separate respondents into more than 2 levels?

Person Separation

The person separation index represents an estimate of how well the instrument differentiates individuals on the measured variable (e.g., in this study naming objects and actions). A separation index > 2 is needed to attain a 0.8 reliability coefficient, which would indicate the assessment separates person ability into at least 3 statistically different levels. We predicted adequate person separation (separation index > 2), because we anticipated separation in skills demonstrated by pre-CI participants compared with post-CI participants.

Question 4: Do the items on the *IT-MAIS* reflect content validity based on *a priori* item hierarchy rankings?

Our hypothesis was that the *a priori* item hierarchy ratings by Speech-Language Pathologist masters students would establish a hierarchy with positive inter-rater correlation. Additionally, we predicted that Speech-Language Pathologist masters students would

demonstrate a hierarchy that differed from that of the *IT-MAIS*. We analyzed item order relative

to order of acquisition in listening development because CI clinical-research programs use the

IT-MAIS to measure progress from pre- to post-CI, as if the assessment were organized in

accordance with order of acquisition. The authors (Zimmerman-Phillips, Osberger, & Robbins)

do not describe theoretical foundations determining item order.

Chapter 4: Results

Question 1: Does the *IT-MAIS* data meet the assumptions for Rasch analysis?

We hypothesized that *IT-MAIS* items would *not* demonstrate unidimensionality. Specifically, we predicted that the items would address 2 different factors (vocalizations and listening skills), based on the 3 subcategories of the *IT-MAIS* items (vocal behavior, attending to sound, deriving meaning from sound). Our hypothesis was not supported. The analyses revealed that the *IT-MAIS does* exhibit unidimensionality. We measured unidimensionality using exploratory factor analysis (EFA) to answer the question: Does the *IT-MAIS* data meet the assumptions for Rasch analysis (i.e. that the construct of interest be unidimensional and that the items are not locally dependent)?

Using SPSS software (IBM, 2011), we conducted EFA using the M responses for each of the *IT-MAIS*' 10 items (Table 1). Because the items aligned in one factor, no rotations were required.

We used the Kaiser-Meyer-Olkin (KMO) test to verify adequacy of our sample. The KMO value for the present data was 0.925 ("superb" according to Field (2009)), indicating that we had an adequate sample to complete the EFA. Bartlett's test of sphericity indicated that correlation between items was sufficiently large for EFA [$\chi^2(45) = 512.005$, $p < 0.001$]. Extraction was completed for eigen values >1 with 25 iterations for convergence (Table 2).

Table 1. Descriptive statistics for scores on each item: group mean (M), standard deviation (SD), and number of data points for the individual question included in the analysis (n)

Item	M	SD	n
1. Is the child's vocal behavior affected while wearing his/her sensory aid?	2.20	1.407	56
2. Does the child produce well-formed syllables and syllable sequences that are recognized as "speech"?	2.07	1.475	56
3. Does the child spontaneously respond to his/her name in quiet with auditory cues only (no visual cues)?	2.46	1.618	56
4. Does the child spontaneously respond to his/her name in the presence of background noise with auditory cues only (no visual cues)?	1.96	1.439	56
5. Does the child spontaneously alert to environmental sounds in the home without being told or prompted to do so?	2.59	1.398	56
6. Does the child spontaneously alert to environmental sounds in new environments?	2.21	1.303	56
7. Does the child RECOGNIZE auditory signals that are part of his/her everyday routines?	2.11	1.592	56
8. Does the child demonstrate the ability to discriminate spontaneously between two speakers with auditory cues only (no visual cues)?	2.07	1.616	56
9. Does the child spontaneously know the difference between speech and non-speech stimuli with listening alone?	2.16	1.671	56
10. Does the child spontaneously associate vocal tone (anger, excitement, anxiety) with its meaning, based on hearing alone?	1.61	1.485	56

Table 2. Summary of Exploratory Factor Analysis Results for Family Relationship Measure Using Maximum Likelihood Estimation (N = 23; Note: Factor loadings over .40 appear in bold.)

Item	Factor Loadings
1. Is the child's vocal behavior affected while wearing his/her sensory aid?	**.629**
2. Does the child produce well-formed syllables and syllable sequences that are recognized as "speech"?	**.840**
3. Does the child spontaneously respond to his/her name in quiet with auditory cues only (no visual cues)?	**.887**
4. Does the child spontaneously respond to his/her name in the presence of background noise with auditory cues only (no visual cues)?	**.894**
5. Does the child spontaneously alert to environmental sounds in the home without being told or prompted to do so?	**.925**
6. Does the child spontaneously alert to environmental sounds in new environments?	**.875**
7. Does the child RECOGNIZE auditory signals that are part of his/her everyday routines?	**.851**
8. Does the child demonstrate the ability to discriminate spontaneously between two speakers with auditory cues only (no visual cues)?	**.869**
9. Does the child spontaneously know the difference between speech and non-speech stimuli with listening alone?	**.840**
10. Does the child spontaneously associate vocal tone (anger, excitement, anxiety) with its meaning, based on hearing alone?	**.801**
Eigenvalues	7.136
% of variance	71.358

The scree plot (Figure 2) illustrated that only one factor (listening development) accounts for all 10 items. This factor accounted for 71.36% of the variance. Thus, the *IT-MAIS* items *did* demonstrate unidimensionality and met the assumptions for Rasch analysis.
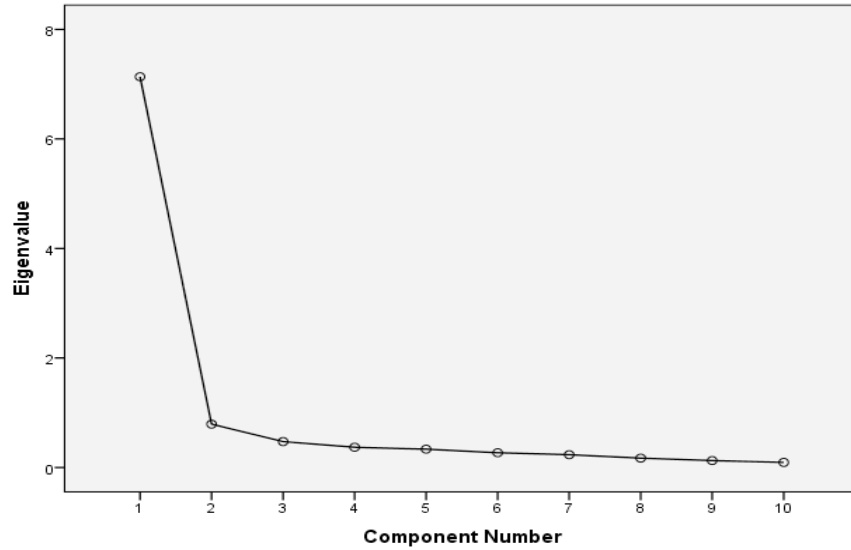

Figure 2. Scree plot demonstrating no points of inflection; thus indicating there was only one factor (listening development).

Table 3 presents correlation coefficients between each *IT-MAIS* item. Ideally, correlation coefficients should be 0.3 > 0.9. Based on these criteria, the correlation coefficients for the *IT-MAIS* items were sound.

Table 3. Correlation coefficients between all 10 items based on the factor model. (Note: See Table 2 for full-text descriptions of each item)

Correlation Matrix											
	Item #	1	2	3	4	5	6	7	8	9	10
Correlation	1	1.000	.554	.439	.471	.522	.671	.364	.513	.412	.516
	2		1.000	.679	.712	.711	.729	.701	.730	.637	.627
	3				.882	.833	.719	.736	.717	.745	.667
	4				1.000	.824	.702	.740	.736	.675	.742
	5					1.000	.807	.788	.793	.791	.674
	6						1.000	.690	.726	.719	.627
	7							1.000	.718	.704	.679
	8								1.000	.743	.625
	9									1.000	.612
	10										

We tested local independence (i.e. no correct item responses are dependent on correct responses to other items) by transforming inter-item residuals (differences between observed and expected responses) to standardized unites using Fisher's z-transformation procedure (Fendrich, 2008; Smith, 2005). Fisher's z-transformed inter-item residual correlations indicated the items demonstrated local independence based on a range of z scores from -0.173 to +0.120 (Table 4), which fell well within the established criteria for local dependence $z \geq 2.0$. This test confirmed that the data met the second assumption for performing Rasch analysis.

Table 4. Local Independence of Inter-Item Residual Correlations

IT-MAIS Item	Fisher's z transformation
2. Does the child produce well-formed syllables and syllable sequences that are recognized as "speech"?	0.026
3. Does the child spontaneously respond to his/her name in quiet with auditory cues only (no visual cues)?	-0.120
4. Does the child spontaneously respond to his/her name in the presence of background noise with auditory cues only (no visual cues)?	-0.092
5. Does the child spontaneously alert to environmental sounds in the home without being told or prompted to do so?	-0.059
6. Does the child spontaneously alert to environmental sounds in new environments?	0.121
7. Does the child RECOGNIZE auditory signals that are part of his/her everyday routines?	-0.173
8. Does the child demonstrate the ability to discriminate spontaneously between two speakers with auditory cues only (no visual cues)?	-0.034

Question 2: Does the *IT-MAIS* demonstrate item-level psychometric properties, reflecting that the latent trait (listening development) can be measured?

Item Misfit: Comparing Responses to Person Ability

We hypothesized that many of the *IT-MAIS* items would be misfitting, based on our subjective judgment that the wording used in the *IT-MAIS* items was confusing and/or vague. Consequently, we predicted that participants would not respond predictably to many of the items due to poor verbiage used in the items describing the specific behavior the items were intended

to assess. However, based on misfit criteria (MnSq < .6 and MnSq ≥ 1.4 and standardized *z*-scores (Zstd) > 2.0), item 1 [*Is the child's vocal behavior affected while wearing his/her sensory aid (hearing aid or cochlear implant)?*] and item 10 [*Does the child spontaneously associate vocal tone (anger, excitement, anxiety) with its meaning based on hearing alone?*] exceeded these criteria (Table 5).

Table 5. Item Infit Statistics based on the established infit criteria for MnSQ, and Zstd

Item	Measure	Model S.E.	Infit	
			Mnsq	Zstd
2. Does the child produce well-formed syllables and syllable sequences that are recognized as "speech"?	0.49A	0.18	1.28	1.4
3. Does the child spontaneously respond to his/her name in quiet with auditory cues only (no visual cues)?	0.52A	0.18	1.29	1.5
4. Does the child spontaneously respond to his/her name in the presence of background noise with auditory cues only (no visual cues)?	0.10A	0.19	1.43	2.1
5. Does the child spontaneously alert to environmental sounds in the home without being told or prompted to do so?	-0.26A	0.19	1.09	0.5
6. Does the child spontaneously alert to environmental sounds in new environments?	-0.45A	0.20	1.07	0.4
7. Does the child RECOGNIZE auditory signals that are part of his/her everyday routines?	0.34A	0.18	1.16	0.9
8. Does the child demonstrate the ability to discriminate spontaneously between two speakers with auditory cues only (no visual cues)?	-0.13A	0.19	1.02	0.2
9. Does the child spontaneously know the difference between speech and non-speech stimuli with listening alone?	1.23A	0.24	0.98	0.0

In order to further analyze the data using Rasch analysis, these 2 items were eliminated. When analyzing the *IT-MAIS* using only 8 items, the assessment meets the infit criteria necessary to utilize Rasch analysis. Analyzing these 8 items, the latent trait (listening development) can be measured.

Person Misfit: Responses to Items Surrounding a Participant's Ability Level

We expected that many participants would not respond predictably, and would thus exceed misfit criteria. We made this prediction because we judged the items' verbiage and

instructions to the test administrators to be confusing. For example, item 1 reads: "Is the child's

vocalizations are affected while wearing his/her sensory aid?" Item 1 then prompts the

administrator to ask the caregiver to "describe [the child's] vocalizations when you first put the

device on each day." However, "affected" is not an objectively measurable behavior and the

authors do not operationally define "affected" vocalizations (Zimmerman-Phillips et al., 2001).

Therefore, we predicted that participants would not reliably describe their children's listening

skills during the assessment. Based on our analyses, 9 participants exceeded the misfit criteria.

Misfitting persons represent caregivers who did not respond reliably to items based on their

responses to other items closest to their children's ability level.

To measure person fit, the same misfit criteria for item fit were applied (>1.4 MnSq and

>2 Zstd). The data in Table 6 presents data from participants who did not predictably respond to

the *IT-MAIS* items that were close to their established ability level.

Table 6. Misfitting persons based on the established criteria for InFit, MnSQ, and Zstd

Participant ID #	Infit MnSq	Infit Zstd
CI21	3.68	3.1
CI13	3.30	2.8
CI18	2.87	2.9
CI16	3.16	2.6
CI18	3.02	3.4
CI7	2.87	3.0
CI14	2.50	2.1
CI12	2.45	2.8
CI9	1.93	2.0

Person misfit is based on a series of iterations that Rasch analysis computes in

accordance with participants' responses to other items around their ability levels. Misfitting

persons may be deleted or retained depending on the researcher's needs in Rasch analysis. We

chose to retain misfitting persons because we do not know why these participants responded

unreliably. For example, we do not know specifics about how the items were administered (i.e., parent interview styles between administrators).

Person-Item Map

We predicted that many of the *IT-MAIS* items would measure skills at the same level of difficulty as other items. We also predicted that participants' listening skills measured pre-CI would be significantly lower than participants' skills post-CI (because children with profound SNHL have pure tone averages > 90 dB pre-CI, thus significantly restricting the participants' access to sound and subsequently limiting their listening skills). Thus, our hypothesis was that the item difficulty range would be smaller than the person ability range, indicating the items did not assess the participants' full range of listening abilities. Our analyses showed that the item difficulty range was smaller than the person ability range (Figure 3), indicating that the items did not assess the full range of participants' listening abilities.

Figure 3 shows a map of Person ability and Item difficulty where both variables are plotted on the same logit scale. Analyzing item difficulty, items 2 and 7 were both measured at about 0.5 logits. Because these items shared the same logit value, they were redundant in the latent trait (listening development) difficulty level they assessed. The remaining *IT-MAIS* items measure the latent trait (listening development) at different item difficulty levels. Rasch analysis dictates that item difficulty should reflect a range of 3 to 4 logits (typically ranging from -2 to +2 logits) for an assessment to psychometrically ideal. The item difficulty range for the *IT-MAIS* was ~1.5 logits, thus not psychometrically ideal.

Comparing person mean (M = 0.8 logits) to item mean (M = 0 logits), indicates a good match between item difficulty and person ability (person M ~ item M +/- 1). Note, in Figure 3, person ability ranges from -6 to +3.6 logits. This suggested a wide range of ability represented by

a relatively small sample. Thus, we inferred that our participants were a representative sample of the population assessed by the *IT-MAIS* (children with SNHL pre- and post-CI ages 0 to 3 years).

In determining ceiling and floor effects, if 10% of the sample is in either the floor or ceiling, it indicates that an assessment is not sensitive to the full range of person ability levels. We found no significant ceiling or floor effects for person ability (0% ceiling effect; 8.9% floor effect).

```
     person ability         item difficulty
              more      |   rare
       4                |
                      X |
                        |
            XXXX S       |
       3                |
              XXX        |
              XXX        |
                        |
            XXXX         |
                X        |
       2      XXX        |
              XXX        |
               XX        |
            XXXX         |
               XX        |
               XX        |
       1       XX |T
               M  |
               XX |
                X |S  item 2   item 7
                  |item 8
                  |item 9
       0        X |M
                X |item 4
                  |item 6
                  |item 3
                  |S
      -1          |
               XX |T  item 5
               XX S|
      -2          |
                  |
                X |
      -3          |
                X |
                T |
      -4          |
               XX |
      -5          |
               XX |
      -6 XXXXXXX |
              less      |   frequent
```
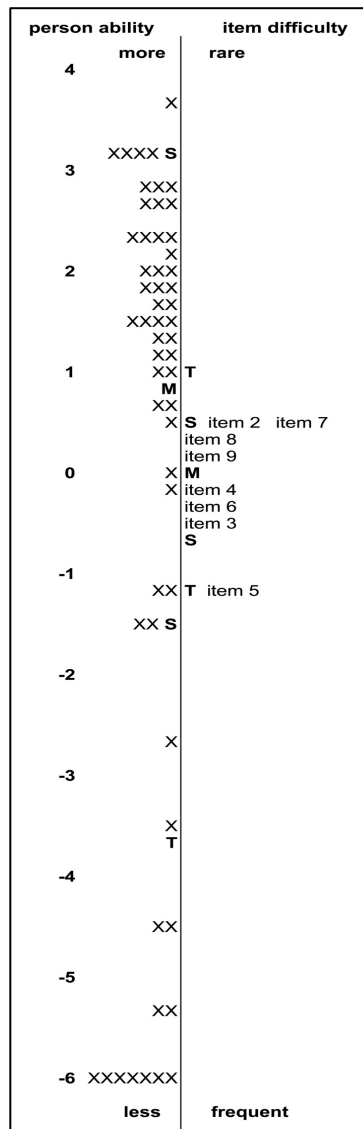
Figure 3. Map of Person ability and Item difficulty: Logit scale ranges from -6.0 to +3.6 for person ability and from -1.23 to +0.52 for item difficulty. Person ability mean is represented by the M to the left of the logit scale; item difficulty mean is represented by the M to the right of the logit scale (at 0 logits). Each "X" represents an individual child, "S" = 1 SD, "T" = 2 SD

40

Person Statistics: Measuring Person Ability

We predicted the pre-CI assessments would result in overall *IT-MAIS* scores of 0 – 3 indicating children exhibited the *IT-MAIS* skills *never, rarely*, or *occasionally*. Seven participants did not meet the minimum requirements for measuring person ability; these participants answered *0 / never* on every item, indicating that their children did not demonstrate any of the skills assessed by the items. Two participants achieved a score of *1/ rarely* on only 1 item. The data in Table 7 accounts for 11 of the 12 pre-CI assessments. The last pre-CI participant (ID: CI12) achieved an overall score of 17 on the *IT-MAIS*. This participant exceeded misfit criteria, indicating the caregiver did not report her child's behaviors reliably; therefore this score is not considered reliable.

Table 7. Participants with the lowest scores on the IT-MAIS

Participant ID #	Total Score
CI22	3
CI7	3
CI18	3
CI11	2
CI13	1
CI20	1
CI8	0
CI9	0
CI14	0
CI16	0
CI19	0

Person Reliability

We predicted a majority of participants would exceed person reliability criteria, because we hypothesized that *IT-MAIS* administration would vary across participants due to unclear administrator instructions and poor verbiage used in the *IT-MAIS* items. Person reliability was assessed using 49 data points because Rasch will not measure the 7 individuals listed who

responded *0 / never* to every item. Person reliability represents the way in which participants of a

given ability level respond to the test items. This means that a parent whose child demonstrated

high ability reliably responded with the highest rating category *(4 / always)* to items of low

difficulty. The criterion for acceptable person reliability index is 0.80 (comparable to Cronbach's

α). In the present study, the person reliability index was 0.92, which is "highly acceptable"

according to Rasch analysis.

Rating Scale Analysis

We predicted that participants would use categories *2 / occasionally* and *3/ frequently*

unpredictably. This hypothesis is based on the notion that caregivers would find it difficult to

distinguish a categorical difference between a child demonstrating a behavior 50% and 75% of

the time. The *IT-MAIS* defines ranking category *2 / occasionally* as *about 50% of the time* and

category *3 / frequently* as *at least 75% of the time* (Zimmerman-Phillips, Osberger, & Robbins,

2001). As depicted in this category response curves for each item (Figure 4), respondents did not

reliably use category *3 / frequently*. Figure 4 is an example of the response curves gathered for

all 8 items. The response curve shows the parents infrequently used category *3 / frequently* from

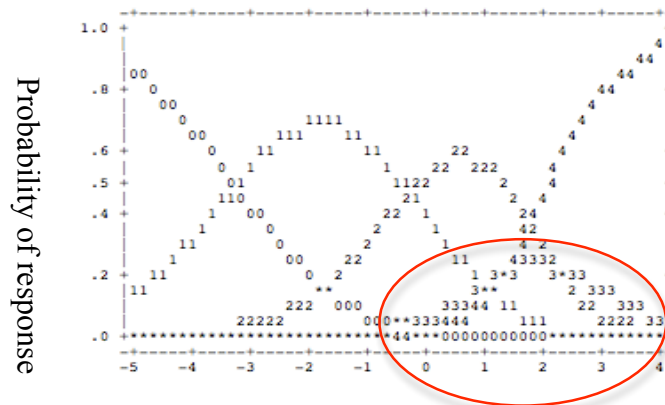the *IT-MAIS'* 0 - 4 rating scale to describe their children's listening behaviors.

Figure 4. Item 2 category response curve demonstrating respondents' use of each category unit
(0 - 4). The red circle highlights the low probability of respondents' use of category 3 /
frequently.

Recall, in order to perform a rating scale analysis, 3 requirements must be met: 1) there must be at least 10 observations in each rating category, 2) measures must advance linearly with each category, and 3) Outfit MnSq < 2 (Linacre, 2002; Donovan, 2007). The *IT-MAIS'* rating scale met all 3 of these criteria; thus, we performed a rating scale analysis. The analysis demonstrated that participants did not consistently use all 5 categories of the *IT-MAIS'* 0 - 4 rating scale. Table 6 presents each item's measures for the 3 requirements for performing rating scale analysis.

Rating scales are employed in order to attain information about a participants' degree of skill rather than a basic *yes / no* or *right / wrong* distinction (Linacre, 2002). If categories on a rating scale are not well defined and mutually exclusive, the reliability of the assessment is negatively impacted (Linacre, 2002). Irregular frequency of rating scale categories indicates that the categories are not properly calibrated in a step-wise manner (i.e. infrequent use of category 3 in the present study relative to the frequency of use for the other 4 categories; Linacre, 2002).

Question 3: **Does the *IT-MAIS* demonstrate adequate range of person ability and item difficulty to separate participants into more than 2 levels?**

We hypothesized that the *IT-MAIS* would demonstrate adequate person separation (> 2 levels), despite our prediction that many of the items would assess participants at the same difficulty levels. We also predicted distinct differences in listening abilities between children with profound SNHL pre-CI as compared with the children post-CI. Person separation was 3.41 (>2), indicating that the *IT-MAIS* separated person ability into at least 3 statistically different levels.

Table 6. Summary of category rating scale utilization criteria based on Category Rating Utilization Analysis for the 5-categoy rating scale (with misfitting persons removed).* indicates category rankings exceeding criteria for each item

Item	>10 Observations per Category	Average Measures Advance Linearly	Outfit Msq <2
2. Does the child produce well-formed syllables and syllable sequences that are recognized as "speech"?	NO	YES	NO
	0 = 13	-2.49	*-3.26
	*1 = 8	-1.15	1.72
	*2 = 8	0.81	*2.00
	3 = 16	2.17	0.58
	4 = 11	2.36	1.59
3. Does the child spontaneously respond to his/her name in quiet with auditory cues only (no visual cues)?	NO	YES	NO
	0 = 14	-3.72	*2.11
	*1 = 2	-0.64	1.49
	*2 = 5	0.21	0.21
	3 = 14	1.64	0.71
	4 = 21	2.21	*2.64
4. Does the child spontaneously respond to his/her name in the presence of background noise with auditory cues only (no visual cues)?	NO	YES	YES
	0 = 14	-3.72	1.89
	*1 = 8	0.04	1.28
	*2 = 8	1.35	0.60
	3 = 18	1.99	1.58
	*4 = 8	2.79	0.39
5. Does the child spontaneously alert to environmental sounds in the home without being told or prompted to do so?	NO	YES	YES
	*0 = 8	-5.41	0.39
	*1 = 5	-3.90	0.75
	*2 = 7	-0.30	0.76
	3 = 18	1.34	0.62
	4 = 18	2.55	0.46
6. Does the child spontaneously alert to environmental sounds in new environments?	NO	YES	NO
	0 = 10	-3.10	*4.23
	*1 = 5	-2.51	0.98
	2 = 11	0.28	0.97
	3 = 23	1.76	1.33
	*4 = 7	3.05	0.27
7. Does the child RECOGNIZE auditory signals that are part of his/her everyday routines?	NO	YES	NO
	0 = 17	-2.73	*2.74
	*1 = 3	0.70	1.38
	*2 = 6	1.07	0.46
	3 = 17	1.59	*2.46
	4 = 13	2.68	0.94
8. Does the child demonstrate the ability to discriminate spontaneously between two speakers with auditory cues only (no visual cues)?	NO	YES	NO
	0 = 16	-2.89	*2.65
	*1 = 6	-0.41	0.49
	*2 = 8	1.47	0.33
	3 = 10	2.18	1.10
	4 = 16	2.38	1.25
9. Does the child spontaneously know the difference between speech and non-speech stimuli with listening alone?	NO	YES	NO
	0 = 17	-2.63	*4.22
	*1 = 4	0.04	0.75
	*2 = 6	1.08	1.28
	3 = 11	1.78	0.71
	4 = 18	2.40	1.05

Question 4: Do the items on the *IT-MAIS* reflect content validity based on *a priori* item hierarchy rankings?

Item Statistics: Measure Order

Our hypothesis was that 4 Masters-level speech-language pathology students would rank the *IT-MAIS* items with strong inter-rater reliability. We also predicted that the item order established by the *a priori* ranking would be significantly different than the item order published in the *IT-MAIS*. We compared the *a priori* rank order with the item order established by Rasch item difficulty measures. Table 7 shows the revised item-order based on the items' level of difficulty, as determined by Rasch analysis. Table 8 reports the revised item-order based on *a priori* rankings by 4 Masters-level speech-language pathology students.

Table 7. Item Order based on difficulty (Least difficult item to most difficult item) Note: "A" = "anchored"

Item	Measure
5. Does the child spontaneously alert to environmental sounds in the home without being told or prompted to do so?	-1.23A
3. Does the child spontaneously respond to his/her name in quiet with auditory cues only (no visual cues)?	-0.45A
6. Does the child spontaneously alert to environmental sounds in new environments?	-0.26A
4. Does the child spontaneously respond to his/her name in the presence of background noise with auditory cues only (no visual cues)?	-0.13A
9. Does the child spontaneously know the difference between speech and non-speech stimuli with listening alone?	0.10A
8. Does the child demonstrate the ability to discriminate spontaneously between two speakers with auditory cues only (no visual cues)?	0.34A
2. Does the child produce well-formed syllables and syllable sequences that are recognized as "speech"?	0.49A
7. Does the child RECOGNIZE auditory signals that are part of his/her everyday routines?	0.52A

45

Table 8. Item order based on a priori rankings from 4 MA-level speech-language pathology students. The students were instructed to rank item order based on order of acquisition in listening development. Note: ***** = item ranked in the same position in both our a priori ranking and via Rasch item difficulty measures. ° = items ranked +/-1 rank position. [+] = item ranked 3 positions higher in a priori hierarchy than in item difficulty order determined by Rasch analysis.

Item
*5. Does the child spontaneously alert to environmental sounds in the home without being told or prompted to do so?
°**6.** Does the child spontaneously alert to environmental sounds in new environments?
°**3.** Does the child spontaneously respond to his/her name in quiet with auditory cues only (no visual cues)?
°**9.** Does the child spontaneously know the difference between speech and non-speech stimuli with listening alone?
°**8.** Does the child demonstrate the ability to discriminate spontaneously between two speakers with auditory cues only (no visual cues)?
°**7.** Does the child RECOGNIZE auditory signals that are part of his/her everyday routines?
[+]**4.** Does the child spontaneously respond to his/her name in the presence of background noise with auditory cues only (no visual cues)?
°**2.** Does the child produce well-formed syllables and syllable sequences that are recognized as "speech"?

To determine the validity of the 4 raters' (Masters-level speech-language pathology students) a priori item hierarchy rankings, we conducted a Spearman's Rank Order correlation to determine the relationships between the 4 raters' rank ordering of the *IT-MAIS* items. Results indicated a statistically significant strong, positive correlation between raters and the hierarchy rankings [ρ (6) = .903, $p < .01$].

Chapter 5. Discussion

Children are undergoing cochlear implantation at younger and younger ages (Colletti, et al., 2012). This decline in age and the challenges associated with accurately assessing the functional hearing of infants raises concerns regarding the tools used for CI candidacy evaluation and post-CI progress assessments. The present study focused on a parent-report tool developed with the intention to serve as a cohesive measurement of pre- and post-CI listening development—the *IT-MAIS* (Zimmerman-Phillips et al., 2001). We proposed that the *IT-MAIS* would not demonstrate ideal psychometric properties, thus calling into question clinical decisions based on *IT-MAIS* scores. Consequently, we recommend that clinical decisions such as CI candidacy and tracking listening development progress in children post-CI should not rely on *IT-MAIS* scores.

CI programs currently use the *IT-MAIS* to determine CI candidacy, establish clinical treatment goals, and track progress in listening development. However, the *IT-MAIS* lacked psychometric analysis prior to the present study.We chose to analyze the psychometric validity and reliability of the *IT-MAIS* for three main reasons: 1) an assessment used to inform clinical decisions (i.e. diagnose, establish treatment goals, and track progress) must be psychometric valid and reliable; 2) the *IT-MAIS* is widely used, but has not undergone rigorous psychometric analysis; and 3) it is particularly important that assessments used to determine CI candidacy in infants and toddlers are psychometrically sound due to the gravity of choosing cochlear implantation. Specifically, irreversible damage to inner ear hair cells and the surgical risks accompanying the decision to implant (i.e. anesthetic risks and blood loss; Waltzman & Roland, 2005) are serious factors for parents to consider prior to making a decision to implant their infants.

We chose Rasch analysis to analyze the psychometric validity and reliability of the *IT-MAIS* because it is a form of IRT that allowed us to compare item difficulty with participant ability. Rasch analysis also provided a valid means to assess data with less than 100 participants (N = 23 in the present study). Based on our analyses, the *IT-MAIS* did not demonstrate ideal psychometric properties, thus indicating the *IT-MAIS* is not a valid and reliable measure for assessing listening development in children with SNHL ages 0 to 3 years pre- and post-CI.

In the remainder of this chapter, we discuss the implications of our findings for each of the 4 research questions we proposed (p. 58-61). Additionally, we discuss the theoretical and clinical implications of our findings (p. 64) and conclude with future research directions (p. 65).

Question 1: Does the *IT-MAIS* data meet the assumptions for Rasch analysis?

Despite predicting the *IT-MAIS* would not meet the assumptions for Rasch analysis, our EFA demonstrated that the *IT-MAIS* items *were* unidimensional and locally independent. Thus, the *IT-MAIS* items met the requirements to perform Rasch analysis. Since the *IT-MAIS* data met the requirements for Rasch analysis, we were able to perform item-level analyses to assess the relationship between participants' ability and item difficulty. We were also able to thoroughly assess *IT-MAIS'* item-level psychometric properties. The results provided knowledge regarding future directions for developing a psychometrically valid and reliable assessment of listening development in pediatric CI users.

Question 2: Does the *IT-MAIS* demonstrate item-level psychometric properties, reflecting that the latent trait (listening development) can be measured?

Recall that our analyses revealed that the *IT-MAIS* did not demonstrate ideal psychometric properties, based on Rasch analysis. Our analyses of item difficulty, person ability, and person reliability revealed weaknesses of the *IT-MAIS* as an assessment for measuring listening development in children with SNHL ages 0 to 3 years pre- and post-CI. Analyses of

item difficulty showed, 2 out of 10 items exceeded the misfit criteria (>1.4 MnSq and >2 Zstd), thus they were discarded from the final analyses. Because the *IT-MAIS* only has 10 items, having 2 misfitting items raises concerns about the validity of the assessment. Our analyses also raised concern for participant reliability; specifically, 9 out of 56 data points exceeded misfit criteria. The limitations of the *IT-MAIS'* psychometric properties as measured by Rasch analysis, suggested that the *IT-MAIS* should not be used to assess listening development in children with SNHL ages 0 to 3 years pre- and post-CI. We advise that the *IT-MAIS* not be used to assess pre-CI listening skills because our analyses of person ability indicate that pre-CI children did not demonstrate *any* of the skills assessed by the *IT-MAIS* items. Specifically, the *IT-MAIS* appears to be an inappropriate tool for this population because an assessment that does not demonstrate psychometric validity should not be used clinically (i.e. to develop treatment goals and track progress over time).

Rating Scale Analysis

We analyzed participants' use of the rating scale categories to determine if participants predictably used each category ranking. Our rating scale analysis demonstrated participants used the category *3 / frequently* unreliably. This finding is particularly important in regards to the *IT-MAIS* because parents' ratings are used to evaluate their children's' listening skills (as opposed to a clinician directly eliciting behavioral responses from a child). Because the ratings and recommendations given based on those ratings are dependent solely on the caregivers' observations, caregivers should reliably select rating categories that accurately represent a child's behavior in everyday life.

One solution for improving the participants' use of the *IT-MAIS'* rating scale categories would be to alter the rating scale (e.g., reducing it to 4, instead of 5, categories) and echo the

parents' rating behaviors of the current study. However, with our relatively small sample size (N = 23) we did not have a sufficient number of observations (10 per ranking category for each item) to analyze the effects of changing the rating scale to 4 points. With a larger sample size (N ≥ 100), we could run further Rasch analysis, collapsing category *3 / frequently* into category *4 / always,* thus changing the scale into a 4-point Likert scale instead of the existing 5-point scale. This subsequent analysis would indicate whether the rating scale was used more predictably (thus, more appropriately) when the existing 5-point scale was reduced to a 4-point scale.

Question 3: **Does the *IT-MAIS* demonstrate adequate range of person ability and item difficulty to separate participants into more than 2 levels?**

In concert with our predictions and based on our data, the *IT-MAIS* separated participants into more than 2 statistically different levels. This is a strength of the *IT-MAIS* (based on our analysis) because, in order to track progress in skill development, items need to demonstrate specificity by separating participants into different ability levels. Improvement on a test measure denotes change in behavior, and significant change cannot be quantified without separating participants into different ability levels. The *IT-MAIS* achieved appropriate participant separation in accordance with the *a priori* standard of separation > 2.0 (present separation = 3.41). Therefore, we propose that the *IT-MAIS* may be a viable a starting point for the creation of a new assessment used to track listening development in children with SNHL. Specifically, researchers can utilize the participant separation demonstrated by the *IT-MAIS* as a guide for constructing new items that address the full range of person ability.

Question 4: **Do the items on the *IT-MAIS* reflect content validity based on *a priori* item hierarchy rankings?**

Item difficulty analyses indicated the order of *IT-MAIS* items was inconsistent with item order based on Rasch difficulty measures. This finding is remarkable because pediatric CI

professionals utilize the *IT-MAIS* as if the item order represents a developmental hierarchy of listening development. In other words, CI programs track listening development using the *IT-MAIS,* despite the fact that there is no theoretical evidence supporting the *IT-MAIS* items' capacity to reflect listening development. It is important to assess young CI users' listening development for 2 main reasons: 1) to inform clinicians' goals for listening therapy and 2) to establish norms for CI users' listening development (specifically order and rate of acquisition). Currently, no developmental norms exist for listening development in children with CIs. In order to accurately assess CI users' listening development, researchers must also establish listening development norms for children with SNHL post-CI. From a broad perspective, tracking listening development in CI users (using valid and reliable measures) provides information regarding CI efficacy and can shape listening intervention.

If the *IT-MAIS'* item order reflected a developmental listening hierarchy (e.g., the hierarchy established by Erber (1982)), we could use item responses from parents of children post-CI to begin developing developmental listening norms for children with CIs and determine if pediatric CI users develop listening skills following a similar pattern to children with normal hearing. Erber (1982) proposed that children with normal hearing first develop sound detection, then discrimination, identification, and finally comprehension (see p. 20 for descriptions of each listening level). However, Erber's listening hierarchy may not reflect the listening development hierarchy in CI users. CI users' listening skills *may* develop in a different order and/or at a different rate than listening skill development in children with normal hearing. Children with normal hearing develop skills (i.e. cognitive, speech, language, *and* listening) concurrently while children with profound SNHL do not develop a majority of the listening skills unless they have access to the acoustic signal via listening devices such as CIs or hearing aids. Thus, listening

51

skills are retarded in children with SNHL, however the remaining systems (e.g., vision, language, cognition, and motor) continue developing as expected. The fact that the remaining systems follow a typical path of development in children with SNHL suggests that once a child has access to the auditory signal via a CI her listening skills may develop in the same sequence as children with normal hearing.

If listening skills develop differently in CI users compared to children with normal hearing, care providers need to assess CI users' listening skills based on a developmental hierarchy that is reflective of CI users' skills—not the children with normal hearing. Because listening skills seem to develop hierarchically in all listeners, a clinician must know the order of skill emergence so that she can use intervention and target skills expected to develop next based on skills the child has mastered. Given that the *IT-MAIS* does not appear to follow a developmental hierarchy, clinicians cannot make valid inferences regarding pediatric CI users' listening development and listening intervention is likely to be reactionary and fruitless. Accurately tracking listening development is important for establishing appropriate treatment goals as well as monitoring the CI device (i.e. battery function, volume settings) to ensure every-day listening success.

Person Ability and Item Difficulty

Our analysis of person ability in comparison with item difficulty also raised concerns regarding the validity of the *IT-MAIS*. Eleven out of 12 participants who were pre-CI received *overall* scores ≤ 3. Overall pre-CI scores were as follows: 0 (n = 5), 1 (n = 2), 2 (n = 1), and 3 (n = 3), out of a possible overall score of 40. This trending floor effect indicated that the items on the *IT-MAIS* did not assess these children's listening abilities. Rather, these very low scores

indicated that the children *never* displayed the behaviors addressed by the items (according to caregiver report).

Prior to cochlear implantation, children are predicted to achieve scores of *0 / never* on most of the *IT-MAIS* items because they assess listening skills and children with SNHL have very limited listening skills due to their likely high pure tone averages (e.g., > 90 dB HL). The minimal range of sounds detectable to children with SNHL (pre-CI) brings into question the use of the *IT-MAIS* as a measure for CI candidacy. Specifically, how many items can a child achieve a score greater than *0 / never* and still be considered a candidate for CI surgery? Choosing CI surgery for a child is an important decision with irreversible effects (e.g., the remaining hair cells in the inner ear are destroyed, thus eliminating any residual hearing present before surgery and limiting the child's chance to utilize future technology). There are also a number of surgical risks that are important to consider prior to surgery. For example, anesthetic risks (i.e. increased risk of respiratory failure) are increased in children who undergo surgery at the ages of 1 year and younger. Also, blood loss (especially during drilling of the mastoid bone) must be closely monitored because of low total blood volume in young children (Waltzman & Roland, 2005). Because of the irreversible effects and risks associated pediatric CI surgery, assessments (e.g, the *IT-MAIS*) included in the CI candidacy battery must be evidence-based in order to minimize the cost-benefit ratio for each family.

Assessing Item Order

CI programs use the *IT-MAIS* to track listening development in pediatric CI users from pre- to post-CI. The *IT-MAIS* authors, however, do not suggest any theoretical underpinnings on which they based its item order. Hence, it is inappropriate to use *IT-MAIS* results to infer developmental progress in listening skills. Since the *IT-MAIS* has no supporting theoretical

construct, in our study we established an *a priori* item order based on rankings by 4 Masters-level speech-pathology students (See Table 8). Spearman's ρ indicated a positive correlation between the 4 raters' rank orders (ρ (6) = .903, *p* < .01). The strong correlation between the *a priori* item order and the item difficulty order established by Rasch supported the notion that the *IT-MAIS* would represent a developmental listening hierarchy more appropriately if the *IT-MAIS'* item order were revised to reflect the order established by Rasch analysis. If the *IT-MAIS* item order represented a trajectory of listening development, clinicians could use *IT-MAIS* scores to track progress in listening development and likely provide child-specific and developmentally appropriate intervention.

Theoretical and Clinical Implications of the Present Results

The results of the present study suggested that the theoretical foundation for "listening development" is not well established. Listening is a complex, cognitive task that researchers do not fully understand. In order to develop assessments for listening development, we need to establish a unified, biologically plausible theory of how humans listen and process spoken language. Establishing such a theoretical construct will lay the foundation for the creation of a comprehensive listening assessment with sound psychometric properties.

Clinically, our results indicate that the *IT-MAIS* did not demonstrate ideal psychometric properties; therefore, it should not be used to determine CI candidacy, establish treatment goals, or track progress in listening development. Clinical decisions based on *IT-MAIS* scores risk providing therapy that is ill-suited for the pediatric CI users and potentially hindering their progress in listening development. CI programs need a comprehensive battery of assessments (or a single comprehensive assessment) that measures cognitive, language (spoken and/or manual), *and* listening skills. However, developing a comprehensive assessment for listening development

54

requires establishing a theoretical foundation that incorporates all of the skills involved in listening development. Regrettably, we have already noted that the field is currently lacking such a theory.

Future Directions

We propose that future directions building on the present psychometric analyses of the *IT-MAIS* can follow 3 paths. The first proposed path entails revising the *IT-MAIS* with 2 main goals: 1) developing new items and re-wording the existing items in order to assess an appropriate range of listening skills in pre- and post-CI users and 2) establishing a new item order so that the order reflects a hierarchy of listening development (p. 64).

The second direction focuses on exploring listening skills to establish an operational definition for "listening development" (p. 67) and conducting more theoretically motivated research to move the field closer towards a comprehensive model of listening and spoken language processing. Specifically, we propose including the role of cognitive and communication skills in our definition and understanding of listening development. This unification of cognition and listening is important because there is a dynamic relationship between these skills. Furthermore, children with SNHL continue to develop cognitively (pre-CI) prior to developing most listening skills (due to pure tone averages > 90 dB). In contrast, children with normal hearing concurrently develop cognitive, language, and listening skills.

Finally, our third research path entails analyzing the psychometric properties of tools other than the *IT-MAIS* that CI programs use to assess listening skills in pediatric CI users (p. 68). For example, the *LittlEARS Auditory Questionnaire* (*LittleEARS;* Tsiakpini et al., 2004) is an assessment designed to track listening development in CI users who were implanted by 24 months of age. Also, the *Parents' Evaluation of Aural/Oral Performance of Children* (*PEACH;*

55

Ching & Hill, 2005a) was designed to assess listening and communication skills in children using hearing aids and/or CIs. Neither of these tools has undergone thorough psychometric analysis. However, understanding more about these tools might allow us to develop a battery of assessments for tracking listening development pre- to post- CI.

IT-MAIS **Revision**

The first path of future research we propose aims to revise the *IT-MAIS.* Using results from the item-level psychometric analyses of the present study, future researchers can re-word *IT-MAIS* items that are unclear, replace misfitting items, and develop new items in order to assess a wider range of participant ability levels. For example, new items should be developed that assess a greater range of listening comprehension skills (comprehension is the final stage of listening development according to Erber, 1982). Expanding the assessment is essential because the item difficulty/person ability map indicated that none of the existing *IT-MAIS* items assess skills demonstrated by participants with high person ability measures (i.e. many children's listening development exceeded the listening skills assessed by the *IT-MAIS* items). We predict that developing items that assess listening comprehension skills would expand the item difficulty range to assess participants in the higher person ability range. Expanding the item difficulty range to assess these participants would improve *IT-MAIS'* ability to assess listening development in pediatric CI users.

Comprehensive Approach to Listening Development

The second path of future research we propose focuses on establishing a theoretical foundation for a comprehensive definition of "listening development" in pediatric CI users. There is a dynamic relationship between listening development, cognitive development, and language development (Arlinger, Lunner, Lyxell, & Pichora-Fuller, 2009). Assessing skills in

56

each of these areas could provide a comprehensive assessment of listening development. In order to establish this comprehensive assessment, we suggest that CI programs use a battery of tools that, together, assess all of these skill sets. The *Behavior Rating Inventory of Executive Function–Preschool Version* (*BRIEF-P;* Gioia, Espy, & Isquith, 2003) was designed to measure executive functioning in children ages 2 years, 11 months old to 5 years old. The *BRIEF-P* demonstrated strong validity and reliability via psychometric analysis based on Classical Test Theory (Gioia, Isquith, Guy, & Kenworthy, 200).

In conjunction with the *BRIEF-P*, we propose using the *Preschool Language Scales, 5th Edition* (*PLS-5;* Zimmerman, Steiner, & Pond, 2011) to assess language development. The *PLS-5* demonstrates sound psychometric properties and normative data (Zimmerman et al., 2011). We suggest incorporating these two tools (*BRIEF-P* and *PLS-5*) with an assessment of listening skills (i.e. a revised *IT-MAIS*) to assess listening development. Using this battery of assessments would provide clinicians with comprehensive information for developing appropriate treatment goals and track progress (in cognitive, language, *and* listening development) pre- to post-CI.

Additionally, the field needs to conduct more research (behavioral and physiological) that can move use toward a better understanding of human listening and spoken language processing. Research that incorporates both behavioral and physiological measures is valuable for learning about how the auditory system works; specifically, how the peripheral and central auditory systems interact to *process* spoken language.

Psychometric Analysis of Additional Parent-Report Tools

The third path for future research is to perform thorough psychometric analyses of *LittlEARS* and the *PEACH.* These tools are used to assess listening development in pediatric CI users, but (like the *IT-MAIS*) they both lack *thorough* psychometric analyses. We propose

analyzing these 2 assessments because both assessments demonstrated strong psychometric properties based on the initial analyses in the literature (Bagatto, Moodie, Seewald, Bartlett, & Scollie, 2011). *LittlEARS* demonstrated adequate item difficulty range relative to participant ability (Tsiakpini et al., 2004). The *PEACH* also demonstrated good item difficulty range as well as strong agreement relative to normative data (Bagatto & Scollie, 2013). Thorough psychometric analyses of these listening assessments would provide evidence regarding whether or not CI programs should use one or both of these assessments to include in the pre- and post-CI battery. If further analyses indicate poor psychometric properties for the *LittlEARS* or the *PEACH*, we could use those results to guide revisions of the *IT-MAIS* (as we explored in the present study for the *IT-MAIS*, p. 66).

Conclusions

In this study, we analyzed the validity and reliability of the *IT-MAIS* via Rasch analysis. We chose to analyze the psychometric properties of the *IT-MAIS* because it lacked thorough psychometric analyses, yet it is widely used to assess listening skills in children with SNHL ages 0 to 3 years pre- and post-CI. Our results indicated that that *IT-MAIS* items did not demonstrate ideal psychometric properties, and the *IT-MAIS* item order did not reflect the order in which children are expected to develop listening skills. Based on these results, we suggested that the *IT-MAIS* be removed from the battery of assessments used to measure CI candidacy; additionally, the *IT-MAIS* should not be used to assess listening development post-CI until revisions are made that improve its psychometric properties.

References

Archbold, S. (1994). Monitoring progress in children at the pre-verbal stage. *Cochlear implants for young children*. London: Whurr Publishers, 197-213.

Arlinger, S., Lunner, T., Lyxell, B., & Pichora-Fuller, K., M. (2009). The emergence of cognitive hearing science. *Scandinavian Journal of Psychology*, *50*, 371-384. doi: 10.1111/j.1467-9450.2009.00753.x

Bagatto, M. P., Moodie, S. T., Seewald, R. C., Bartlett, D. J., & Scollie, S. D. (2011). A critical review of audiological outcome measures for infants and children. *Trends in amplification*, *15*, 23-33. doi: 10.1177/1084713811412056

Bagatto, M. P., & Scollie, S. D. (2013). Validation of the parents' evaluation of aural/oral performance of children (PEACH) rating scale. *Journal of the American Academy of Audiology*, *24*, 121-125.

Barker, B. A., Kenworthy, M. H., & Walker, E. A., (2011). How we do it: Employment of listening-development criteria during assessment of infants who use cochlear implants. *Cochlear Implants International, 12*, 57-59. doi 10.1179/146701010X486543.

Baudhuin, J., Cadieux, J., Firszt, J. B., Reeder, R. M., & Maxson, J. L. (2012). Optimization of programming parameters in children with the advanced bionics cochlear implant. *Journal of the American Academy of Audiology, 23*, 302-12.

Baylor, C., Hula, W., Donovan, N. J., Doyle, P. J., Kendall, D., & Yorkston, K. (2011). An introduction to item response theory and Rasch models for speech-language pathologists. *American Journal of Speech-Language Pathology, 20*, 243-259. doi:10.1044/1058-0360(2011/10-0079)

Blamey, P., & Sarant, J. (2000). Speech perception and language criteria for pediatric cochlear implant candidature. *Audiology & Neuro-Otology, 7,* 114-121. doi: 10.1159/000057659

Carroll, J., & Zeng, F. G. (2007). Fundamental frequency discrimination and speech perception in noise in cochlear implant simulation. *Hearing Research, 231,* 42-53. doi: 10.1016/j.heares.2007.05.004

Ching, T. Y., & Hill, M. (2005a). The parents' evaluation of aural/oral performance of children (PEACH) diary. Chatswood, New South Wales, Australia: Australian Hearing.

Ching, T. Y., & Hill, M. (2005b). The parents' evaluation of aural/oral performance of children (PEACH) rating scale. Chatswood, New South Wales, Australia: Australian Hearing.

Cochlear Implants in Adults and Children. NIH Consensus Statement. (1995).

http://consensus.nih.gov/1995/1995CochlearImplants100html.htm. Accessed March 31 2012.

Colletti, L., Mandala, M., & Colletti, V., (2012). Cochlear implants in children younger than 6 months. *Otolaryngology -- Head and Neck Surgery, 147,* 139-146. doi: 10.1177/0194599812441572

Colletti, L., Mandala, M., Zoccante, L., Shannon, R. & Colletti, V. (2011). Infants versus older children fitted with cochlear implants: Performance over 10 years. *International Journal of Pediatric Otorhinolaryngology, 75*, 505-509. doi: 10.1016/j.ijporl.2011.01.0

Connor, C. M., Craig, H. K., Raudenbush, S. W., Heavner, K., & Zwolan, T. A. (2006). The age at which young deaf children receive cochlear implants and their vocabulary and speech production growth: Is there an added value for early implantation? *Ear & Hearing, 27,* 628-644. doi: 1092-4388/00/4305-1185

Del Toro, C. M., Bislick, L. P., Comer, M., Velozo, C. A., Romero, S., Rothi, L.J.G., & Kendall, D.L. (2011). Development of a short form of the Boston Naming Test for individuals with aphasia. *Journal of Speech, Language, and Hearing Research*, 54, 1089-1100. doi: 10.1044/1092-4388(2010/09-0119)

Dollaghan, C. A. (2007). *The handbook for evidence-based practice in communication disorders*. (pp. 9-12). Baltimore, MD: Paul H Brookes Pub Co.

Donovan, N. J., Rosenbek, J. C., Ketterson, T. U., & Velozo, C. A. (2006). Adding meaning to measurement: Initial Rasch analysis of the ASHA FACS Social Communication Subtest. *Aphasiology*, 20, 362-373. doi: 10.1080/02687030500475184

Edmonds, L. & Donovan, N. (2011). Item level psychometrics and predictors of performance for Spanish/English bilingual speakers on an object and action naming battery. *Journal of Speech, Language, and Hearing Research, 55,* 359-381. doi: 10.1044/1092-4388

Erber, N. (1982). *Auditory Training.* (pp. 92-94). Washington DC: Alexander Graham Bell Association

Ertmer, D. J., & Jung, J. (2012). Monitoring progress in vocal development in young cochlear implant recipients: Relationships between speech samples and scores from the Conditioned Assessment of Speech Production (CASP). *American Journal of Speech-Language Pathology, 21,* 313-328. doi: 10.1044/1058-0360(2012/11-0110)

Ertmer, D., Young, N., & Nathani, S. (2007). Profiles of vocal development in young cochlear implant recipients. *Journal of Speech, Language, and Hearing Research, 50*, 393-407. doi: 10.1044/1092-4388(2007/028)

Fendrich, M., Smith, E. V., Jr., Pollack, L. M., & Mackesy-Amiti, M. E. (2009). Measuring sexual risk for HIV: A Rasch scaling approach. *Archives of Sexual Behavior*, 38, 922–935. doi: 10.1007/s10508-008-9385-2

Franz, D. C. (2002): Pediatric performance with the Med-El Combi 40+ cochlear implant system. *Annals of Otology Rhinology & Laryngology, 111*, 66–68.

Geers, A. & Moog, J. (1994). Spoken language results: Vocabulary, syntax, and communication. *Volta Review, 96,* 127-130.

Geers, A. E., Nicholos, J. G., & Sedey, A. L. (2003). Language skills of children with early cochlear implantation. *Ear & Hearing, 24,* 46S-58S.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Test review behavior rating inventory of executive function. *Child Neuropsychology*, *6*(3), 235-238. doi: 10.1080/09297041003679344

Graves, R. E., Bezeau, S. C., Fogarty, J., & Blair, R. (2004). *Boston Naming Test* short forms: A comparison of previous forms with new item response theory based forms. *Journal of Clinical and Experimental Neuropsychology*, *26*, 891–902. doi: 10.1080/13803390490510716

Hammes, D. M., Novak, M. A., Rotz, L. A., Willis, M., Edmondson, D. M., & Thomas, J. F. (2002). Early identification and cochlear implantation: Critical factors for spoken language development. *Annals of Otology, Rhinology, and Laryngology–Supplement, 189,* 74–78.

Hayes, D. & Northern, J. L. (1997). *Infants and Hearing.* United States of America: Singular Publishing Group.

Houston D. M., Pisoni D. B., Kirk K. I., Ying E.A., & Miyamoto R. T. (2003). Speech perception skills of deaf infants following cochlear implantation: A first report. *International Journal of Pediatric Otorhinolaryngology, 67,* 479–495. doi: 10.1016/S0165-5876(03)00005-3

Hula, W., Donovan, N. J., Kendall, D. L., & Gonzalez- Rothi, L. J. (2010). Item response theory analysis of the Western Aphasia Battery. *Aphasiology*, *24*, 1326–1341. doi: 10.1080/02687030903422502

Hula, W., Doyle, P.J., McNeil, M. R., & Mikolic, J. M. (2006). Rasch modeling of *Revised Token Test* performance: Validity and sensitivity to change. *Journal of Speech, Language, and Hearing Research*, *49*, 27–46. doi:10.1044/1092-4388(2006/003)

IBM Corporation (2011). IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corporation

Ireton, H., & Thwing, E. (1974). *Minnesota Child Development Inventory*. Minneapolis, MN:

University of Minnesota.

Kishon-Rabin L., Taitelbaum R, Elichai O, Maimon D, Debyiat D, & Chazan N. (2001) Developmental aspects of the *IT-MAIS* in normal-hearing babies. *Israeli Journal of Speech and Hearing, 23,* 12-22.

Lazaridis, E., Therres, M. & Marsh, R. R. (2010). How is the children's implant profile used in the cochlear implant candidacy process? *International Journal of Pediatric Otorhinolaryngology, 74*, 412-415. doi: 10.1016/j.ijporl.2010.01.022

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*, 85-106.

Lord, F. M., & Novick, M. R. (1968) *Statistical theories of mental test scores*. Reading MA: Addison-Welsley Publishing Company

Mallinson, T., (2011). Rasch Analysis of Repeated Measures. *Rasch Measurement Transactions, 251,* 1317.

Martin, F. N. & Clark, J. G. (2009). *Introduction to Audiology.* Boston MA: Pearson Education

Neuman, A. C., Wroblewski, M., Hajicek, J., & Rubinstein, A. (2012). Measuring speech recognition in children with cochlear implants in a virtual classroom. *Journal of Speech, Language and Hearing Research*, *55*, 532. doi:10.1044/1092-4388(2011/11-0058)

Nikolopoulos, T. P., O'Donoghue, G. M., Robinson, K. L., Gibbin, K. P., Archbold, S. M., & Mason, S. M. (1997). Multichannel cochlear implantation in post-meningitic and congenitally deaf children. *American Journal of Otology, 18,* s147-s148.

Nikolopoulos, T. P., Wells, P. & Archbold, S. M. (2000). Using Listening Progress Profile (LiP) to assess early functional auditory performance in young implanted children. *Deafness Education, 2*, 142-151. doi: 10.1002/dei.85

Nikolopoulos, T., Dyar, D., Archbold, S., & O'Donoghue, G. (2004). Development of spoken language grammar following cochlear implantation in prelingually deaf children. *Archives of Otolaryngology-Head & Neck Surgery, 130*, 629-633. doi: 10.1001/archotol.130.5.629

Osberger, M. J. & Zimmerman-Phillips, M. J., & Koch, D. B. (2002). Cochlear implant candidacy and performance trends in children. *Annals of Otolology, Rhinology & Laryngology, 111*, 62-65.

Robbins, A. M., Koch, D. M., Osberger, M. J., Zimmerman-Phillips S., Kishon-Rabin, L. (2004). Effect of age at cochlear implantation on auditory skill development in infants and toddlers. *Archives of Otolaryngology-Head & Neck Surgery, 130*, 570-574. doi: 10.1001/archotol.130.5.570

Robbins, A. M., Renshaw, J. J., & Berry, S. W. (1991). Evaluating meaningful auditory integration in profoundly hearing-impaired children. *American Journal of Otology, 12*, 144-150.

Roberts, P. M. (2008). Issues in assessment and treatment for bilingual and culturally diverse patients. In R. Chapey (Ed.), *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders5th Edition* (pp. 245-275). Philadelphia: Lippincott Williams and Wilkins.

Smith Jr, E. V. (2005). Effect of item redundancy on Rasch item and person estimates. *Journal of Applied Measurement*, *6*, 147.

Tomblin, B., Barker, B. A., & Hubbs, S. (2007). Developmental constraints on language development in children with cochlear implants. *Journal of Audiology, 46*, 512-523. doi: 10.1080/14992020701383043

Tomblin, B., Barker, B. A., Spencer, L. J., Zhang, X., & Gantz, B. J. (2005). The effect of age at cochlear implant at initial stimulation on expressive language growth in infants and toddlers. *Journal of Speech, Language, and Hearing Research, 48*, 853-867. doi: 1092-4388/05/4804-0853

Tomblin, J. B., Spencer, L., Flock, S., Tyler, R., & Gantz, B. (1999). A comparison of language achievement in children with cochlear implants and children using hearing aids. *Journal of Speech, Language, and Hearing Research, 42,* 497-511.

Tsiakpini, L., Weichbold, V., Kuehn-Inacker, H., Coninx, F., D'Haese, P., & Almadin, S. (2004). *LittlEARS Auditory Questionnaire*. Innsbruck: MED-EL.

Valencia, D. M., Rimell, F. L., Friedman, B. J., Oblander, M. R., & Helmbrecht, J. (2008). Cochlear implantation in infants less than 12 months of age. *International Journal of Pediatric Otorhinolaryngology, 72*, 767-773. doi: 10.1016/j.ijporl.2008.02.009

Weichbold, V., Anderson, I. & D'haese, P. (2004). Validation of three adaptations of the *Meaningful Auditory Integration Scale* (MAIS) to German, English, and Polish. *International Journal of Audiology, 43*, 156-161. doi:10.1080/14992020400050021

Weichbold, V., Tsiakpini, L., Coninx, F., & D'Haese, P. (2005). Development of a parent questionnaire for assessment of auditory behaviour of infants up to two years of age. *Laryngology-Rhinology-Otologie*, *84*, 328. doi: 10.1055/s-2004-826232

Wolfe, E. W., & Smith, E. V. (2007). Instrument development tools and activities for measure validation using Rasch models: Part II-Validation activities. In E.V. Smith Jr. & R. Smith (Eds.), *Rasch measurement: Advanced and specialized applications* (pp. 243-290). Maple Grove, MN: JAM Press.

Wright, B. D., & Linacre, J. M. (1989). Observations are always ordinal; measurements, however, must be interval. *Archives of Physical Medicine and Rehabilitation*, *70,* 857–860.

Wright, M., Purcell, A., & Reed, V. A. (2002). Cochlear implant and infants: Expectations and outcomes. *Annals of Otology, Rhinology, and Laryngology- Supplemental, 189,* 131-137.

Zheng, Y., Wang, K., Meng, J., Meng, Z., Xu, K. & Tao, Y. (2009). A normative study of early prelingual auditory development *Audiology Neurotology, 14*, 214-222. doi: 10.1159/000189264

Zimmerman-Phillips, S., Osberger, M.J., & Robbins, A.M. (May, 1997). *Assessment of auditory skills in children two years of age or younger*. Paper presented at the 5th International Cochlear Implant Conference in New York, NY.

Zimmerman-Phillips, S., Robbins, A. M., & Osberger, M. J. (2000). Assessing cochlear implant benefit in very young children. *The Annals of Otology, Rhinology & Laryngology. Supplement, 109*, 42-43.

Zimmerman-Phillips, S., Osberger, M. J., & Robbins, A. M. (2001). *Infant-Toddler Meaningful Auditory Integration Scale*. Sylmar, CA: Advanced Bionics Corporation.

Appendix A.
Infant-Toddler Meaningful Auditory Integration Scale (Zimmerman-Phillips, 2001)

Item 1: Is the child's vocal behavior affected while wearing his/her sensory aid (hearing aid or cochlear implant)?

The benefits of auditory input are often apparent first in the speech production skills of very young children. The frequency and quality of vocalizations may change when the device is put on, turned off, or not working properly.

1. Ask the parent: Describe _____'s vocalizations when you first put his/her device on each day. Have the parent explain how and if the child's vocalizations change when the sensory aid is first turned on and auditory input is experienced at the start of each day.

2. Ask the parent: If you forget to put the device on ____, or if the device is not working properly, do you and/or others notice that ____'s vocalizations are different in any way (e.g., quality, frequency of occurrence)?

3. Or ask: Does the child test the device by vocalizing when the device is first turned on?

_____ 0 = Never
No difference in the child's vocalizations with the device turned on versus the device turned off.

_____ 1 = Rarely
Slight increase in the frequency of the child's vocalizations (approximately 25%) with the device on (or similar decrease with the device off).

_____ 2 = Occasionally
Child vocalizes throughout the day, and there are increases in vocalizations (approximately 50%) with the device turned on (or similar decrease with the device turned off).

_____ 3 = Frequently
Child vocalizes throughout the day, and there are noticeable increases in vocalizations (approximately 75%) with the device on (or a similar decrease with the device off). Parents may report that individuals outside the home notice a change in the frequency of child's vocalizations with or without the device.

_____ 4 = Always
Child's vocalizations increase 100% with the device on compared to the frequency of occurrence with the device turned off.

Parent Report:

Item 2: Does the child produce well-formed syllables and syllable-sequences that are recognized as speech?

This type of utterance is characteristic of the speech of developing infants. The utterances contain speech sounds and syllables that are recognized as speech by the parents. Parents often assert the baby is "talking."

1. Ask the parent: Does _____ "talk" to you or to objects?

2. Ask the parent: As_____ plays alone, what kinds of sounds do you hear when the device is on?

3. Ask the parent: Does _____ say sounds and words used in nursery rhymes or playing with toys? (e.g., hop hop, moo, baaa, choo choo, mmmmm).

4. Ask the parent: For specific examples of the types of utterances the child produces, as well as the frequency with which they are produced.

_____ 0 = Never
Child never produces speech-like utterances, child only produces undifferentiated vocalizations, or the parents cannot give any examples.

_____ 1 = Rarely
Child produces speech-like utterances once in awhile (approximately 25% of the time), but only when provided with a model (spontaneous imitation).

_____ 2 = Occasionally
Child produces speech-like utterances 50% of the time when provided with a model (spontaneous imitation).

_____ 3 = Frequently
Child produces these utterances approximately 75% of the time; parents can give many examples. Child produces the syllable sequences spontaneously, but with a limited phonetic repertoire. The child can clearly and reliably imitate sequences with a model (spontaneous imitation).

_____ 4 = Always
Child produces syllable-sequences consistently and on a spontaneous basis (i.e., without a model). The utterances consist of a varied repertoire of sounds.

Parent Report:

Item 3: Does the child spontaneously respond to his/her name in quiet with auditory cues only (i.e., no visual cues) when not expecting to hear it?

Infants and toddlers demonstrate a variety of behaviors in response to sound. Examples of such responses in very young children may be: momentary cessation of an activity (e.g. stops moving, playing, sucking, crying), searching for the sound source (e.g., looks up or around after hearing their name), widening or blinking their eyes.

1. Ask the parent: If you called _____'s name from behind his/her back in a quiet room with no visual cues, what percentage of the time would s/he respond the first time that you called his/her name? Many young children commonly demonstrate an off-response when auditory stimulation stops; any repeatable behavior is considered a response, provided the child demonstrates the behavior consistently.

2. Ask the parent: For specific examples of the types of responses that the parent observes, especially to assign the highest ratings.

_____ 0 = Never
Child never responds to his/her name, or the parents cannot give any examples.

_____ 1 = Rarely
Child responds to his/her name only about 25% of the time on the first trial, or only with multiple repetitions.

_____ 2 = Occasionally
Child responds to his/her name about 50% of the time on the first trial, or does it consistently but only after the parent repeats the name more than once.

_____ 3 = Frequently
Child responds to his/her name at least 75% of the time on the first trial.

_____ 4 = Always
Child responds to his/her name reliably and consistently on the first trial.

Parent Report:

Item 4: Does the child spontaneously respond to his/her name in the presence of background noise with auditory cues only (i.e., no visual cues)?

1. Ask the parent: If you called _____'s name from behind his/her back with no visual cues in a noisy room (e.g., people talking, children playing, the TV on), what percentage of time would s/he respond to you the first time that you called his/her name? Use the response criteria specified in Question 3 to score the parent's observations.

Remember that in general, the younger the child, the more subtle the responses observed. Rather than overt responses to stimuli such as searching for the source of the sound, a cessation in activity or a freezing behavior is commonly observed. As long as the behavior is observed consistently, it is considered a response.

2. Ask the parent: For specific examples of the types of responses that the parent observes.

_____ 0 = Never
Child never responds to his/her name in noise, or the parents cannot give any examples.

_____ 1 = Rarely
Child responds to his/her name in noise about 25% of the time on the first trial, or only with multiple repetitions.

_____ 2 = Occasionally
Child responds to his/her name in noise about 50% of the time on the first trial, or does it consistently but only after the parents repeat the name more than once.

_____ 3 = Frequently
Child responds to his/her name in noise at least 75% of the time on the first trial.

_____ 4 = Always
Child responds to his/her name in noise reliably and consistently on the first trial.

Parent Report:

Item 5: Does the child spontaneously alert to environmental sounds (dog, toys) in the home without being told or prompted to do so?

1. Ask the parent: Tell me about the kinds of environmental sounds to which _____ responds at home and in familiar situations (e.g., grocery store, restaurant, playground) without prompting.

Give me examples. Question parents to be sure the child is responding via audition, without visual cues.

2. Ask the parent: To provide specific examples, such as alerting to the telephone, TV, dog barking, smoke alarm, toys that make sounds (e.g., music boxes, music mobiles see-and-say toys, horns honking, dishwasher, microwave bell).

The child must alert spontaneously to the sound without prompting from the parent. Recall that very young children demonstrate various responses to sound including momentary cessation of activity, searching for the sound source, widening and/or blinking their eyes. Young children often respond when a sound ceases, rather than at the onset.

Any repeatable behavior is considered a response provided it is demonstrated consistently.

_____ 0 = Never
Child never demonstrates the behavior, the parents cannot give any examples, or child responds only after a prompt.

_____ 1 = Rarely
Child responds about 25% of the time to different sounds. Parents can give only one or two examples, or give several examples of sounds that the child responds to on an inconsistent basis.

_____ 2 = Occasionally
Child responds about 50% of the time to more than two environmental sounds. If there are a number of sounds that regularly occur to which the child does not alert (even if he consistently responds to two sounds such as the phone and the doorbell), assign a score no higher than Occasionally.

_____ 3 = Frequently
Child consistently responds to many environmental sounds at least 75% of the time.

_____ 4 = Always
Child basically responds to all environmental sounds reliably and consistently.

Parent Report:

Item 6: Does the child spontaneously alert to environmental sounds in new environments?

1. Ask the parent: Does _____ show curiosity (verbal or nonverbal) about sounds when in unfamiliar settings (e.g., such as in someone else's home, unfamiliar store, or a restaurant without being prompted?). Examples include clanging dishes in a restaurant, bells dinging in a department store, PA system in public buildings, baby crying in another room, smoke alarm, an unfamiliar toy at a playmate's home. A younger child may provide nonverbal indications that s/he has heard a new sound with eye-widening, a frown or a smile, searching for the source of the new sound, imitation of the new sound (such as when playing with a new toy), starting to cry after a loud or unusual sound, or looking to a parent for information. The response behaviors may be demonstrated when the sound is first detected or when it ceases.

_____ 0 = Never
Child never demonstrates the behavior or the parents cannot give any examples.

_____ 1 = Rarely
Child demonstrates the behavior but does so only about 25% of the time; parents can give only one or two examples of this behavior.

_____ 2 = Occasionally
Child demonstrates the behavior numerous times (about 50%) of the time, and parents can give a number of different examples.

_____ 3 = Frequently
Child demonstrates the behavior about 75% of the time, parents can give many different examples, and responses are a common occurrence.

_____ 4 = Always
Very few new sounds occur without the child showing a response or curiosity about them.

Parent Report:

Item 7: Does the child spontaneously RECOGNIZE auditory signals that are part of his/her everyday routines?

1. Ask the parent: Does _____ regularly recognize or respond appropriately to auditory signals at daycare, preschool, or in the home with no visual cues or other prompts? Examples of this may be looking for a familiar toy that the child hears but cannot see, looking at the microwave when it goes off or the telephone when it rings, looking at the door when the dog is outside barking, wanting to come in the house, looking at the door when hearing the garage door opening, putting hands over his/her eyes if you stand behind the child and verbally initiate an interactive play game such as "Peek-a-boo." Other games include "Pat-a-cake" or "So Big."

_____ 0 = Never
Child never demonstrates the behavior, or the parents cannot give any examples.

_____ 1 = Rarely
Parents can give one or two examples of the behavior. Child responds to these signals 25% of the time. If there are a number of sounds that occur regularly to which the child does not alert, assign a score no higher than Occasionally.

_____ 2 = Occasionally
Parents can provide more than two examples. Child responds to these signals about 50% of the time.

_____ 3 = Frequently
Parents can give many examples. Child demonstrates consistent response to these signals at least 75% of the time.

_____ 4 = Always
Child clearly has mastered this skill and routinely responds to auditory signals that are part of everyday routines. There are very few sounds that the child does not recognize within the daily routine.

Parent Report:

Item 8: Does the child demonstrate the ability to discriminate spontaneously between two speakers with auditory cues only (i.e., no visual cues)?

Examples of this behavior include discriminating between the voices of mother or father and that of a sibling, or discriminating between the voices of mother and father. Examples of this behavior may be attending/responding to the parent who spoke when only auditory cues are present.

1. Ask the parent: Can _____ tell the difference between two voices, like Mom or brother/sister, just by listening to them?

2. At a more difficult level ask: If _____ is playing with two siblings and one sibling spoke, would _____ look in the direction of the appropriate brother/sister?

_____ 0 = Never
Child never demonstrates the behavior, or the parents cannot give any examples.

_____ 1 = Rarely
Child can discriminate between two very different voices (adult/child) about 25% of the time. Ask parents to provide examples.

_____ 2 = Occasionally
Child can discriminate between two very different voices (adult/child) about 50% of the time. Ask parents to provide examples.

_____ 3 = Frequently
Child discriminates between two very different voices (adult/child) 75% of the time; sometimes discriminates between two similar voices (e.g., voices of two children). Ask parents to provide examples.

_____ 4 = Always
Child always discriminates between two very different voices; very often discriminates between two similar voices.

Parent Report:

Item 9: Does the child spontaneously know the difference between speech and non-speech stimuli with listening alone?

The purpose of this question is to evaluate whether the child has categorical perception between speech and non-speech stimuli. We address this by inquiring about instances where the child may confuse these two stimuli, or show that s/he is not confused. For example, if a child has an established response to certain stimuli (e.g. rocking in response to music), does s/he ever exhibit this behavior in response to speech stimuli?

1. Ask the parent: Does _____ recognize speech as a category of sounds that are different from non-speech sounds? For example, if you are in a room with your child and you called to him/her, would s/he look for you or for a favorite toy?

2. Ask the parent: Does _____ ever search for a family member's voice versus looking for a familiar toy?

_____ 0 = Never
Child does not know the difference between speech versus non-speech stimuli, or parents cannot give any examples.

_____ 1 = Rarely
Child demonstrates speech/non-speech distinction about 25% of the time; parents can give only one or two examples. Child often confuses speech and non-speech stimuli.

_____ 2 = Occasionally
Child demonstrates speech/non-speech distinction at least 50% of the time; parents can give a number of different examples.

_____ 3 = Frequently
Child demonstrates speech/non-speech distinction at least 75% of the time; parents can give many different examples.

_____ 4 = Always
Child consistently and reliably demonstrates the behavior; child makes essentially no errors in discriminating speech from non-speech stimuli.

Parent Report:

Item 10: Does the child spontaneously associate vocal tone (anger, excitement, anxiety) with its meaning based on hearing alone?

In the very young child, does the child recognize changes in emotion conveyed by voice associated with the use of motherese or chiild-directed speech? Examples of this include laugh or coo in response to large fluctuations in the intonation or changes in voice and upset when scolding or told firmly no-no, even with no substantial increase in the loudness of the voice.

1. Ask the parent: By listening only, can _____ tell the emotion conveyed in someone's voice such as an angry voice, an excited voice, etc.? (e.g., Mother yells, and child startles and cries in response or child laughs or smiles in response to changes in intonation and prosody in parents' voices without seeing their faces).

_____ 0 = Never
Child does not demonstrate the behavior, parents cannot give any examples, and child has no opportunity to show the behavior.

_____ 1 = Rarely
Child demonstrates the behavior about 25% of the time. Ask parents to provide examples.

_____ 2 = Occasionally
Child demonstrates the behavior about 50% of the time. Ask parents to provide examples.

_____ 3 = Frequently
Child demonstrates the behavior about 75% of the time. Ask parents to provide examples.

_____ 4 = Always
Child consistently and appropriately responds to a range of vocal tones. Parents can provide numerous examples.

Parent Report:

MEANINGFUL AUDITORY INTEGRATION SCALE (MAIS)
Amy M. Robbins
Indiana University School of Medicine
Indianapolis, IN 46202

NAME_____ DATE_____

INTERVAL_____

CONDITION(device)_____

EXAMINER_____

INFORMANT_____

Item 1: Score item 1a if the child is younger than age 5 and item 1b if the child is older than age 5.

1a. Does the child wear the device all waking hours WITHOUT resistance?

Ask the parent, "What is your routine for putting on _____'s device each day?" Have the parent explain how long
the child wears the device and determine if the child wears it all waking hours **WITHOUT** resistance or for only restricted periods of time. Ask. "If one day you didn't put the device on _____ would _____show any indication that s/he missed wearing it (such as pulling or pointing to his/her ear, going over to where the device is kept when not in use, looking upset or quizzical, etc.)" An additional query would be, "Does you child give any nonverbal indication that s/he is upset when the device is removed (such as crying or fussing)?".

_____0=Never: If parent seldom puts the device on the child because the child resists wearing it.

_____1=Rarely: If the child wears the device for only short periods of time but resists wearing it.

_____2=Occasionally: If child wears device for only short periods of time but without resistance.

_____3=Frequently: If the child wears the device all waking hours without resistance.

_____4=Always: If the child wears the device all waking hours and provides some indication if the parent forgets to put it on one day and/or some indication that s/he is upset or misses the device when it is not on.

PARENT REPORT:

1b. Does the child ask to have his or her device put on, or put it on him/herself WITHOUT being told?

Ask "What is _____'s routine for putting on his/her device each day?" Have parent explain if it is the parent or the child who takes responsibility for it. Ask, "If one day, you didn't put the device on _____ and didn't mention it, would_____ask to wear it and be upset by not having it?" An additional query would be, "Does your child basically wear it according to routine (such as all day at school and one hour at night) or does s/he want it on all waking hours?" (for example, s/he puts it on at night even after his/her bath)? The latter would indicate a child who is more boded and dependent on his her device than the former.

_____0=Never: If the child resists wearing it.

_____1=Rarely: If the parent says child wears it without resistance, but would never ask for it.

_____2=Occasionally: If child might inquire about it and is content to wear it with a set time routine.

_____3=Frequently: If the child wears the device all waking hours without resistance.

_____4=Always: Only if child wears it all waking hours and it's part of his body (like glasses would be).

PARENT REPORT:

Item 2: Does the child report and/or appear upset if his/her device is non-functioning for any reason?

Ask parent to give examples of what the child has done (verbally or nonverbally) when the device was not working.

Ask also, "Have you ever checked _____'s device and found it was not working (or headpiece had fallen off), but s/he had not noticed or had not told you?" In the case of the younger child, ask "Have you ever checked _____'s device and found it wasn't working but s/he had not provided any nonverbal indication (such as crying, reaching for the headpiece, etc.) that it was not working?"

_____0=Never: If child has no awareness of the device working or not.

_____1=Rarely: If parent says child might only notice a malfunctioning device (using verbal or nonverbal indication) once in a while.

_____2=Occasionally: If parents can give some examples of when the child would recognize a malfunctioning device (or if headpiece has fallen off) more than 50% of the time and may be beginning to distinguish some device problems from others.

_____3=Frequently: If parent gives examples and/or child can often distinguish different types on malfunction (e.g. bad cord vs. weak batteries).

_____4=Always: If child would never go without immediately detecting and reporting a problem with his/her unit and can easily identify what the problem is.

PARENT REPORT:

Item 3: Does the child spontaneously respond to his name in quiet when called auditorially-only (with no visual cues)?

Ask, "If you called _____'s name from behind his back in a quiet room with no visual cues, what percentage of the time would he respond the first time you called?"

_____0=Never: If the child never does.

_____1=Rarely: If he has done it only once or twice or only with multiple repetitions.

_____2=Occasionally: If he does it about 50% of the time on the first trial or does it consistently but only when parent repeats his name more than once.

_____3=Frequently: If he does it at least 75% of the time on the first try.

_____4=Always: If he does this reliably and consistently, responding every time just as a hearing child would. Ask for examples.

PARENT REPORT:

Item 4: Does the child spontaneously respond to his name in the presence of background noise when called auditorially only with no visual cues?

Ask, "If you called_____'s name from behind his back with no visual cues in a noisy room, with people talking and the TV on, what percentage of time would he turn around and respond to you the first time you called"?

_____0=Never: If the child never does.

_____1=Rarely: If the child has done it only once or twice or only with multiple repetitions.

_____2=Occasionally: If he does it about 50% of the time on the first trial or does it consistently but, only when the parent repeats his name more than once.

_____3=Frequently: If he does it at least 75% of the time on the first try.

_____4=Always: If he does this reliably and consistently , responding every time just as a normal hearing child would.

Ask for examples.

PARENT REPORT:

Item 5: Does the child spontaneously alert to environmental sounds (doorbell, telephone) in the home without being told or prompted to do so?

Ask, "Tell me about the kinds of environmental sounds_____responds to at home and give me examples".

Question parents to be sure the child is responding auditorially only with no visual cues. Examples could be asking about the telephone, doorbell, dog barking, water running, smoke alarm, toilet flushing, engines revving, horns honking, microwave bell, washer changing cycles, thunder, etc. Examples must be child alerting spontaneously and not prompted by parent.

_____0=Never: If parent can give no examples or if child responds only after a prompt.

_____1=Rarely: If parent can give only one or two examples, or give several examples where the child's responses are inconsistent.

_____2=Occasionally: If child responds about 50% of the time to more than two environmental sounds.

_____3=Frequently: If child consistently responds to many environmental sounds at least 75% of the time.

_____4=Always: If child basically responds to environmental sounds the way a hearing child would. If there are a number of sounds which regularly occur to which the child does not alert (even if he consistently responds to two sounds such as the phone and the doorbell) he would score no higher than Occasionally.

PARENT REPORT:

Item 6: Does the child alert to auditory signals spontaneously when in new environments?

Ask, "Does your child show curiosity (verbally or nonverbally) about new sounds when in unfamiliar settings, such as in someone else's home or a restaurant by asking, "What was that sound?" or "I hear something?" A younger child may provide nonverbal indications that s/he has heard a new sound with eye widening, looking quizzical, searching for the source of the new sound, imitation of the new sound (such as when playing with a new toy). Examples parents have reported are children asking about clanging dishes in a restaurant, bells dinging in a department store, PA systems in public buildings, unseen baby crying in another room.

_____0=Never: If parents can give no examples.

_____1=Rarely: If parents can give only one or two examples.

_____2=Occasionally: If child has done this numerous times and parents can give examples.

_____3=Frequently: If parents can give numerous examples and this is a common occurrence.

_____4=Always: If very few sounds occur without the child asking about them (or, in the case of the younger child, showing curiosity nonverbally).

PARENT REPORT:

Item 7: Does the child spontaneously RECOGNIZE auditory signals that are part of his/her school or home routine?

Ask, "Does _____ regularly recognize or respond appropriately to auditory signals in his/her classroom (e.g., school bell, PA system, fire alarm) or in the home (e.g., running to the window to see which family member is home when s/he hears the garage door opening; going to the table when the bell of the microwave goes off, signaling that the food is cooked and it is time to eat) with no visual cues or other prompts?"

_____0=Never: If s/he never does it.

_____1=Rarely: If there are one or two instances.

_____2=Occasionally: If s/he responds to these signals about 50% of the time.

_____3=Frequently: If many examples are given and the child does it 75% of the time.

_____4=Always: If s/he has clearly mastered this skill and does it all of the time.

PARENT REPORT:

Item 8: Does the child show the ability to discriminate spontaneously between two speakers, using audition alone (such as knowing mother's vs. father's voice, or parents' vs. sibling's voice)?

Ask, "Can _____ tell the difference between two voices, like Mom or Dad's (or Susie's or John's) just by listening to them?"

_____0=Never: If parent can give no examples of the child discriminating between two speakers.

_____1=Rarely: If one or two examples are given.

_____2=Occasionally: If several examples are given and the child does this at least 50% of the time.

_____3=Frequently: If many examples are given and the child does this 75% of the time.

_____4=Always: If always done and the child shows no errors in doing this.

PARENT REPORT:

Item 9: Does the child spontaneously know the difference between speech and nonspeech stimuli with listening alone?

Ask, :"Does _____ recognize speech as a category of sounds that are different from nonspeech sounds? For example, if you were standing behind your child and a noise occurred, would s/he ever say, "What was that noise?"
In the case of the younger children, ask, "Would _____ever run into the next room to search for a family member's voice versus looking out the window for a dog or fire truck.?"

_____0=Never: If parent can give no examples of the child discriminating speech from nonspeech.

_____1=Rarely: If one or two examples are given.

_____2=Occasionally: If several examples are given and the child does this at least 50% of the time.

_____3=Frequently: If many examples are given and the child does this 75% of the time.

_____4=Always: If always done and the child shows no errors in doing this.

PARENT REPORT:

Item 10: Does the child spontaneously associate vocal tone (anger, excitement, anxiety) with its meaning based on hearing alone?

Ask, "By listening only, can _____ tell the emotion conveyed in someone's voice such as angry voice, and excited voice, etc.?" (e.g., Father yells at child to "hurry up" through the bathroom door and the child responds, "Why are you mad? and yells back at him. In the case of the younger child, the child starts to cry because of the angry sound in his/her voice). Another example is if the parent is reading a new book to a young child while s/he is sitting on the parent's lap and cannot see their parent's face, (e.g., Mom says "the boy yelled "Let's go!" and the child says "The boy is happy to go to the park").

_____0=Never: If the parent can give no examples or if the child has never had the opportunity to do this.

_____1=Rarely: If the child does it 25% of the time.

_____2=Occasionally: If the child does it about 50% of the time.

_____3=Frequently: If s/he does it 75% of the time.

_____4=Always: If s/he consistently can identify more than one emotion in the listening alone condition.

PARENT REPORT:

Total points correct:_____/40

Appendix C.
Statement Regarding Institutional Review Board Approval

In the present study, we retrospectively analyzed data collected at the University of Iowa Pediatric Cochlear Implant Center. Thus, this study did not require Institutional Review Board (IRB) approval. The relevant IRB approval information for the data used in the present analysis is as follows:

IRB ID #: 201205706

Title: Children's CI Project VI

Vita

Anne Schubert was raised in Novi, Michigan. She graduated from the University of Michigan, Ann Arbor with a bachelor's degree in Linguistics. Following graduation, she attended Louisiana State University to pursue her Master's degree in Communication Disorders. She worked in Dr. Brittan Barker's Spoken Language Processing Laboratory on a Louisiana Board of Regents research grant exploring the effects of talker-specific information on toddler's listening. Anne completed her thesis under Dr. Brittan Barker in partial fulfillment of the requirements for a Master of Arts degree. Upon graduating, Anne hopes to work as a speech pathologist in an acute care setting.