Future Directions in Evaluating Frequency Compression

Michael Boretzki, Nicola Schmitt, Andrea Kegel, Harald Krueger, Julia Rehmann, Frederik Eichhorn, Katrin Meisenbacher, Juliane Raether

The Purpose of the Logatome Test

One challenge in providing auditory treatment for the hearing impaired is providing audibility in the high frequency range: can hearing instruments provide intelligibility of the fricative and plosive sounds? In the past, hearing aids have relied upon gain to make sounds audible in the high frequency range. More recently, new technologies have become available in digital hearing aids that allow for the lowering of the frequency range of high frequency sounds.

In order to measure whether changes in a hearing instrument's response relate to changes in intelligibility of high frequency phonemes, we need a speech intelligibility measure that is both specific and sensitive to the perception of high frequency phonemes. Ideally, such a test should be applicable across all degrees and configurations of hearing loss. Traditionally, sentence or word tests have been used, with scoring procedures that provide an overall speech recognition threshold or percentage correct, depending upon the test. Such tests may not be maximally sensitive to changes in intelligibility for high frequency phonemes, because scores may be influenced by top-down cognitive strategies, such as prediction of particular phonemes from the word or sentence context. One alternative is to use nonsense syllable tests to minimize the listener's ability to use context. If carefully selected, nonsense syllables may also be applicable for listeners who speak different languages. In this chapter, we present a new nonsense syllable test that has been developed specifically for the purpose of measuring intelligibility of specific high frequency phonemes. We refer to this as the Logatome Test, and the following goal was set forth in its development: to be able to measure high frequency intelligibility with different hearing aids or different settings of hearing aids. For example, we wanted the test to be sensitive to the effects of frequency compression switched on versus off, preferably in within-subject designs. In order to achieve this, the construction of the test was led by the following principles:

(1) Minimize phoneme predictability
(2) Minimize non-consonant cues
(3) Minimize ceiling and floor effects
(4) Maximize valid responses

Test Construction

The word "logatome" is a term essentially equivalent to the phrase "nonsense syllable." According to Wikipedia (2010), "A logatome is an artificial word of one or more syllables, which obeys all the phonotactic rules of a language but has no meaning. Examples of English logatomes would be the nonsense words snarp or bluck." Within our Logatome Test, all stimuli have the structure /a/-/consonant/-/a/. For example, two of the stimuli are "asa" and "asha."

The Logatome Test differs from known nonsense syllable tests in two respects. First, the test adapts the presentation level following each stimulus depending upon whether the response was correct or incorrect/unsure. This allows the test to measure the speech recognition threshold (SRT) for that particular stimulus. This test method was chosen for two reasons: (1) it maximizes specificity by providing a score for each phoneme, rather than scores averaged across a list of stimuli; and (2) the adaptive principle of the test prevents the ceiling

Address correspondence to: Michael Boretzki, Ph.D., Phonak AG, Laubisrütistrasse 28, Stäfa 8712, Switzerland, Michael.Boretzki@phonak.com

and floor effects of tests with a static presentation level. The second difference between the Logatome Test and traditional nonsense syllable tests relates to the construction of the stimuli. Specifically, all stimuli have acoustically identical initial and final /a/ taken from one base recording of a female talker. The remaining stimuli were created by excising the consonants from other recordings spoken by the same talker (e.g., "asa, asha, afa," etc.). These excised consonant portions were inserted and cross-faded between the /a/ sounds. This makes sure that subjects during the test cannot learn to use differences in level, pitch or length of the various naturally produced vowels to predict the embedded consonant. Instead, the listener must rely upon the consonant part itself and the transitions from vowel to consonant and consonant to vowel for speech sound recognition.

In summary, the Logatome Test does not measure detection or discrimination of phonemes but identification. The test conditions are controlled such that the subjects' cues for prediction of the consonant from other semantic or acoustic information are minimized. One could ask: why measure the identification of phonemes in this sophisticated manner? Wouldn't measuring phoneme detection be sufficient? We have begun by developing an adaptive intelligibility test, rather than a detection test for two reasons. First, identification of phonemes is the central function of bottom-up speech understanding, and is therefore of primary interest as an outcome measure. Second, frequency lowering techniques alter the spectrum of high frequency speech sounds, so detection thresholds alone may not be sufficient for predicting whether an audible consonant is also identified correctly.

Test Procedures

In our first prototype of the test, listeners were presented with the carrier phrase "My name is" followed by a logatome (e.g., "My name is 'asa'."). Listeners were asked to select the name from the following list of words on a touch screen: "Asa, Asha, Afa, Aka, Ata." Listeners could also indicate if the word was not understood, or repeat the stimulus if they wished. Therefore, the procedure does not force the subject to select one of the logatomes if it has not been understood. Besides the psychological advantage of not being forced to guess, there is a performance advantage of unforced-choice procedures (Kaernbach 2001) because responses near or at chance level are prevented from inappropriately steering the level adaptation.

The first prototype contained the following pool of stimuli: "aba, ada, afa, aha, aka, ala, ama, ara, asa (unvoiced), asha, awa and ata." In order to reflect the dependency of the /s/ spectrum on gender, two forms of the /s/ stimulus were created to represent male (6 kHz) and female (9 kHz) /s/.

Evaluation of the Prototype Logatome Test

In an initial study, Meisenbacher (2008) tested the measurement principle with a sample of normal hearing subjects with respect to required effort, duration and reliability, and applicability in quiet and in white noise. With a small group of subjects with mild hearing losses it was also determined that the majority of the stimuli differentiated between mild hearing loss and normal hearing (i.e., that Logatome Test scores for most stimuli differed significantly between these groups).

A further investigation evaluated whether the Logatome Test could be used with listeners who speak different languages. In a preliminary internal study English, German and Thai native speakers with normal hearing were tested using the stimuli "aba, ada, afa, aga, aha, aka, ala, ama, ana, apa, asa (6 kHz), asa (9 kHz), asha, and ata." Whereas the SRTs and measurement errors were very similar for the English and German group, the Thai group showed remarkably elevated thresholds and measurement errors for some of the tested phonemes: /k/ (which is not a part of the Thai language), and also /s/ and /sh/.

A further study (Boretzki and Kegel 2009) focused on determining whether the test was sensitive to changes in hearing instrument function in adult listeners with mild hearing losses. Two sets of stimuli were used. Set One consisted of the stimuli "asa (6 kHz), asa (9 kHz) and ada." Set Two consisted of "asa (6 kHz), asa (9 kHz), ada, afa, aka, asha and ata." The average hearing thresholds of the subjects were 15, 15, 21, 32, 42, 56, 51 dB HL at the frequencies .25, .5, 1, 2, 4, 6 and 8 kHz. The hearing aid model was Exelia Art, and open fitting was used. Below 3.5 kHz, the instruments were set such that they did not provide any insertion gain. Above 3.5 kHz, a target insertion gain of 25 dB was used, with modifications to ensure stable gain and minimal internal noise. Listeners were tested with frequency compression on or off. The SRTs of nearly all consonants were improved for aided (without frequency compression) versus unaided hearing. Enabling frequency compression showed a further significant advantage for the 9 kHz /s/ and a less than significant improvement for the 6 kHz /s/. This was true in both stimulus sets.

A further study showing the sensitivity and specificity of the Logatome Test in children has been conducted by Wolfe and colleagues (2010) and is reported elsewhere in this volume.

Development of a Clinical Version of the Logatome Test

The studies cited above revealed one weakness of the first prototype of the Logatome Test: the reliability was not yet acceptable. Individual test-retest deviations of measured SRTs were as much as 10 dB. Therefore we explored several methodological factors with the aim to reduce this measurement error. Inspection of sources of measurement variance led to the following observation: during presentations near SRT, repetitions of identical stimuli may be heard as non-identical. For example, the sequence "asa – asa – asa" may be perceived as "ada – afa – asa." The reasons for this phenomenon may include variation in internal noise, (e.g., breathing) or that human phoneme decoding does not yield identical results if perceptual input is poor. A remarkable reduction of the measurement error was achieved by presenting the logatomes in repetition rather than only once. We eliminated the carrier phrase, and presented each logatome three times. The listener was instructed to select a logatome if all three presentations were the same. If they were not the same, the listener indicates this instead, and the test adapts to use a higher presentation level. Catch trials (i.e., trials in which the three presentations are intentionally different) are used throughout the adaptive sequence, in order to ensure that listeners remain vigilant in their monitoring of whether the three were identical or not.

As expected, this modification produced slightly increased SRT values, but reduced the test-retest deviations to approximately 4 dB. Test-retest error was reduced especially for those phonemes that showed large measurement errors with the single presentation approach.

Summary and Future Directions

In summary, our goal was to develop a language-neutral intelligibility test that is sensitive and specific to high frequency phoneme intelligibility. The procedure, called the Logatome Test, is a nonsense syllable test with adaptive measurement of the SRT per phoneme. Fricative phonemes spoken by a female talker are used in the test in order to challenge high frequency hearing loss. In order to remove spurious cues of vowel and duration, all non-consonant cues were removed by careful acoustic editing. The stimuli are presented more than once to ensure that the participant clearly recognizes the stimulus. For this reason, the measured threshold represents phoneme intelligibility rather than phoneme detection. This improves the consistency and reliability of the measurement. The Logatome Test has been shown to be sensitive to the effects of frequency compression in adults with mild and moderate hearing losses, and in children with moderate to moderately severe hearing loss (Wolfe et al. 2010). Further evaluations and developments are in progress and offer the promise of a clinically usable version of the test across languages and ages.

References

Boretzki, M., and Kegel, A. 2009. The benefit of nonlinear frequency compression for people with mild hearing loss. *AudiologyOnline*. Retrieved 11/23/2009 from: http://www.audiologyonline.com/articles/article_detail.asp?wc=1&article_id=2317.

Kaernbach, C., 2001. Adaptive threshold estimation with unforced-choice tasks. *Perception and Psychophysics* 63(8): 1377–1388.

Meisenbacher, K. D., 2008. Entwicklung und Evaluation eines adaptiven Logatomtests zur Ermittlung der Konsonantenverständlichkeit (Development and evaluation of an adaptive logatome test for measuring the intelligibility of consonants). Diploma thesis, Fachhochschule Oldenburg (University of Applied Science, Oldenburg, Germany).

Wikipedia.(2010). *Logatome.* Retrieved Nov. 2, 2010, from: http://en.wikipedia.org/wiki/Logatome.

Wolfe, J., John, A., Schafer, E., Nyffeler, M., Boretzki, M., and Caraway, T. (2010). Evaluation of non-linear frequency compression for school-age children with moderate to moderately-severe hearing loss. *Journal of the American Academy of Audiology* 21(10): fast track article.