for playback of test signals from the wet room in the same wet room. Lower indices were found for the other two conditions. Furthermore, a tendency becomes apparent that the externalization index is higher for playback of test signals from the reverberant room than from the dry room. The analysis indicates the assumed effect of room acoustics on binaural auralization. Further investigations with subtle graduation of room acoustic parameters, like the strength and ratio of early reflections related to the direct sound, are meaningful. Furthermore, the multimodal divergence of audio and visual context between recording and playback conditions are less investigated but seemed to play an important role in a plausible perception of a synthesized scene.

REFERENCES

Klein, F. and Werner, S. (**2011**). “Individualization of head-related transfer functions“, 37th annual convention for acoustics, DAGA, Düsseldorf.

Abou-Elleal, E. R. E. (**2003**). “Raumakustik – Interaktion visueller und auditiver Wahrnehmungen“, dissertation, RWTH Aachen.

Sass, R., Werner, S., and Siegel, A. (**2010**). “Comparison of recording methods for measurements of head-related transfer functions“. 26th, VDT Int. Convention, Tonmeistertagung, Leipzig.

Merimaa, J. and Hess, W. (**2004**). “Training of Listeners for Evaluation of Spatial Attributes of Sound“. Proc. of 117th AES Conv., preprint 6237, San Francisco.

Hammershøi, D. (**2009**). “Human localization and performance measures”. in Proceedings of ISAAR 2009: *Binaural Processing and Spatial Hearing*. 2nd International Symposium on Auditory and Audiological Research ISAAR, Denmark. Edited by J.M. Buchholz, T. Dau, J. Christensen-Dalsgaard, and T. Poulsen. ISBN: 87-990013-2-2. (The Danavox Jubilee Foundation, Copenhagen), pp. 103-111.

Møller, H., Sørensen, M. F., Jensen, C. B., and Hammershøi, D. (**1996**). “Binaural Technique: Do We Need Individual Recordings?“. J. Audio Eng. Soc, 44(6), pp. 451-469.

Møller, H., F., Hammershøi, D., Jensen, C. B., and Sørensen, M. (**1999**). “Evaluation of artificial heads in listening tests“. J. Audio Eng. Soc, 47, pp.83-100.

Hartmann, W. M. and Wittenberg, A. (**1996**). “On the externalization of sound images“. J. Acoust. Soc. Am., 99(6), pp. 3678-3688.

”HiST taleaudiometri” - A new Norwegian speech audiometry

JON ØYGARDEN

Audiology, University College of Sør-Trøndelag, 7004 Trondheim, Norway

“HiST taleaudiometri”, a new material for speech audiometry measurements, has been made available from the Audiology Programme at Sør-Trøndelag University College, the institution that teaches the Norwegian audiologists. The material will replace the old Norwegian speech audiometry material made by Quist-Hanssen in the late 1950s. The test material is constructed from four different parts – five-word sentences, three-word utterances, monosyllabic words and digit triplets. Details about the development of the five-word sentences, simulations of speech audiometry measurements and special tests included in the set are given.

INTRODUCTION

“HiST taleaudiometri” (Øygarden, 2009a and 2009b), consists of four types of speech material.

Five-word sentences

This set is developed in the tradition of Hagerman (1982) and Wagener *et al.* (1999). The sentences are constructed so that each list of 10 sentences contains exactly the same 50 words. The same syntactical structure is used for all of the sentences: (name verb numeral adjective noun). The five-word sentences in “HiST taleaudiometri” are made with a different cut and splice point which may help us achieve even more natural-sounding sentences than earlier. The material is available for different types of speech in noise tests.

Three-word utterances

These utterances are the three last words in the five-word sentences. The material has a steep slope of the performance intensity (PI) curve and is suited to measure speech recognition threshold.

Monosyllabic words

This material is a revision of the words used in the old speech audiometry. Seldom used words and words found very easy, or very difficult to recognize in a listening test were excluded from the material. The selected 160 words were mixed and repeated to make nine lists of 50 words. A Matlab mixing procedure was developed to get both the nine lists and every 10 word-group as uniform as possible.

Proceedings of ISAAR 2011: Speech perception and auditory disorders. 3rd International Symposium on Auditory and Audiological Research. August 2011, Nyborg, Denmark. Edited by T. Dau, M. L. Jepsen, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN 87-990013-3-0. EAN 9788799001330. The Danavox Jubilee Foundation, 2012.

Digit triplets

The numerals used in the “HiST taleaudiometri” set are the numerals that were recorded by Sverre Quist-Hanssen for his speech audiometry. The numerals are organized in groups of three (digit triplets).

Table 1 shows speech recognition results for the new speech material.

	SRT, L_{eq} [dBC]	Standard deviation [dB]	SRT [dB ‘HL’]	Slope [%/dB]	SRT in noise [dB SNR]	Slope in noise [%/dB]
Digit triplets	19.2	1.4	-3.4	17		
Three-word utterances (reference)	22.6	1.0	0.0	10	-6.2	16
Five-word sentences	23.6	0.8	1.0	10	-6.0	14
Monosyllabic words	29.1	2.1	6.5	7		

Table 1: Speech recognition thresholds and slopes for the different materials in “HiST taleaudiometri” in silence and in noise.

DEVELOPMENT OF THE FIVE-WORD SENTENCES

Selection of words

The selection of Norwegian words was inspired by the words chosen for the realization of the Swedish and Danish Hagerman sentences. The chosen words had to be familiar to the groups of people who are candidates for this test as we want to measure hearing, not knowledge of words. It is desirable that even small children should know the words. The selection of names was based on the baby naming statistics for Norway presented by Statistics Norway. Five boys’ names and five girls’ names were selected among the top 100 names for each gender used in Norway during the period 1993-2002. When all the other words have been selected some of the preliminarily chosen names were substituted with others from these groups in order to improve the phonemic balance of the word list. The selected words are presented in Fig. 1.

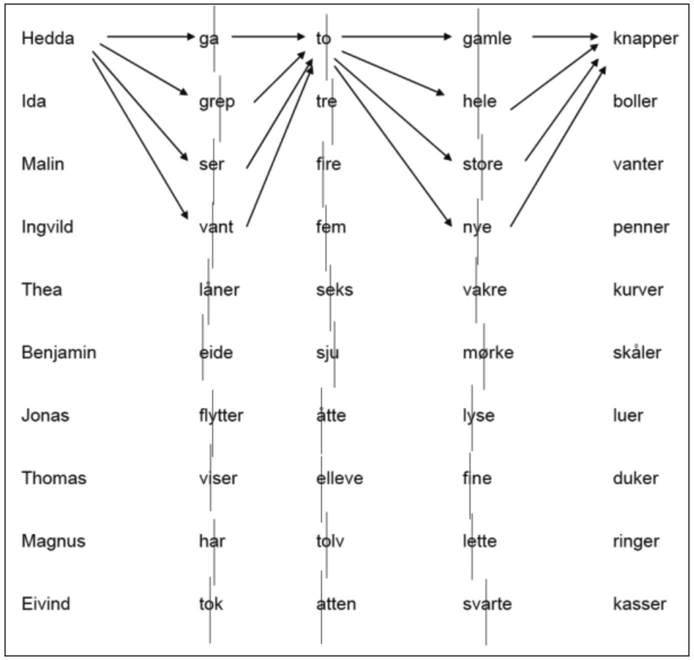


Fig. 1: The Norwegian words selected for the generation of Hagerman sentences. The arrows show how 4 of the recorded 100 sentences were read. The vertical lines show the splitting points.

Wagener developed a new method enabling them to make more natural-sounding Hagerman sentences. For each of the 10 names Wagener recorded 10 naturally spoken sentences where the name, numeral and noun were kept constant, but with a new verb and a new adjective chosen for each sentence. By using this method Wagener ended up with recordings of 100 natural sentences containing all the possible occurrences of two successive words in the given material. Wagener split each sentence between the word boundaries, except for the last two words in the sentence which could be kept together.

The diphone method

The five-word sentences in “HiST taleaudiometri” are generated with splitting points like that found in the diphone synthesis method. Instead of splitting between the words like Wagener, the midpoint of the first vowel in the following word was chosen as the splitting point. Fig. 2 shows a Matlab splitting tool with the recorded sentence *Jonas låner åtte vakre luer* (Jonas borrows eight pretty caps).

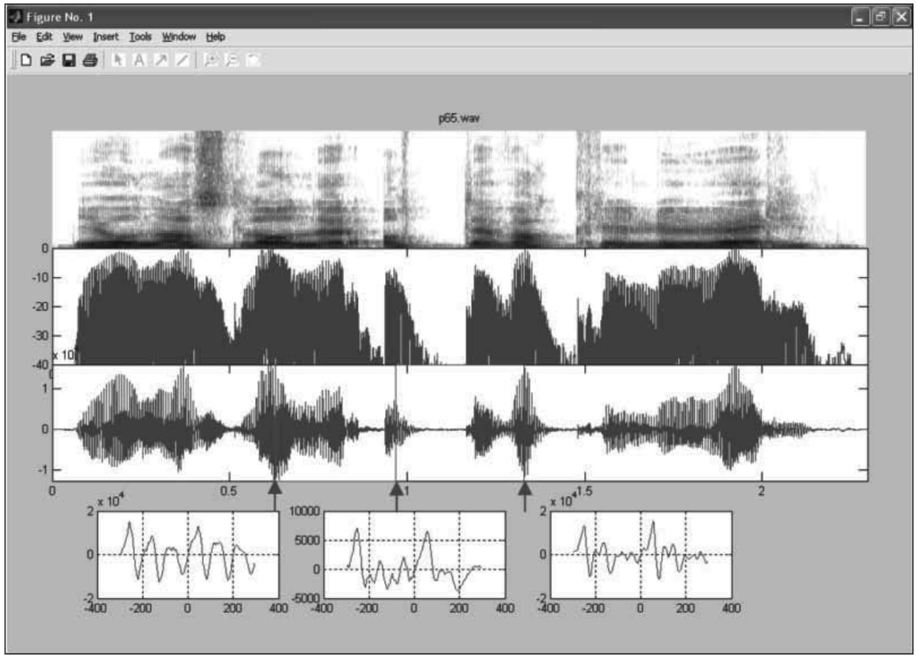


Fig. 2: The Matlab tool for generating the building blocks needed in diphone sentences. The arrows show the selected splitting points for this sentence. The three windows at the bottom show where the splitting points are automatically adjusted to the preceding zero crossing of the largest positive amplitude in one period of the vowel.

These splitting points are in a very stable phone and can easily be identified. This method gives the correct transition into the following word, which perhaps would help to get even more natural sounding sentences than earlier methods. A pair comparison test was performed comparing the naturalness of sentences generated by this method with sentences made by Wagener method, Hagerman method and natural read sentences. The sentences made by the diphone method were rated significantly more natural than sentences made by the Wagener and Hagerman method.

SIMULATION OF MEASUREMENTS

To give recommendations for measurement procedures, many different procedures were evaluated with simulations like those presented in Fig. 3. The example given here are comparisons between three-word utterances measured both with a method called quick speed method and the traditional method by measuring the PI curve. The histogram of thresholds and the numeric data show that the traditional method for this instance gives a more precise threshold with lower standard deviation. But the quick speed method were faster using only 90 items (30 three-word utterances) while the traditional method used in average 154 items.

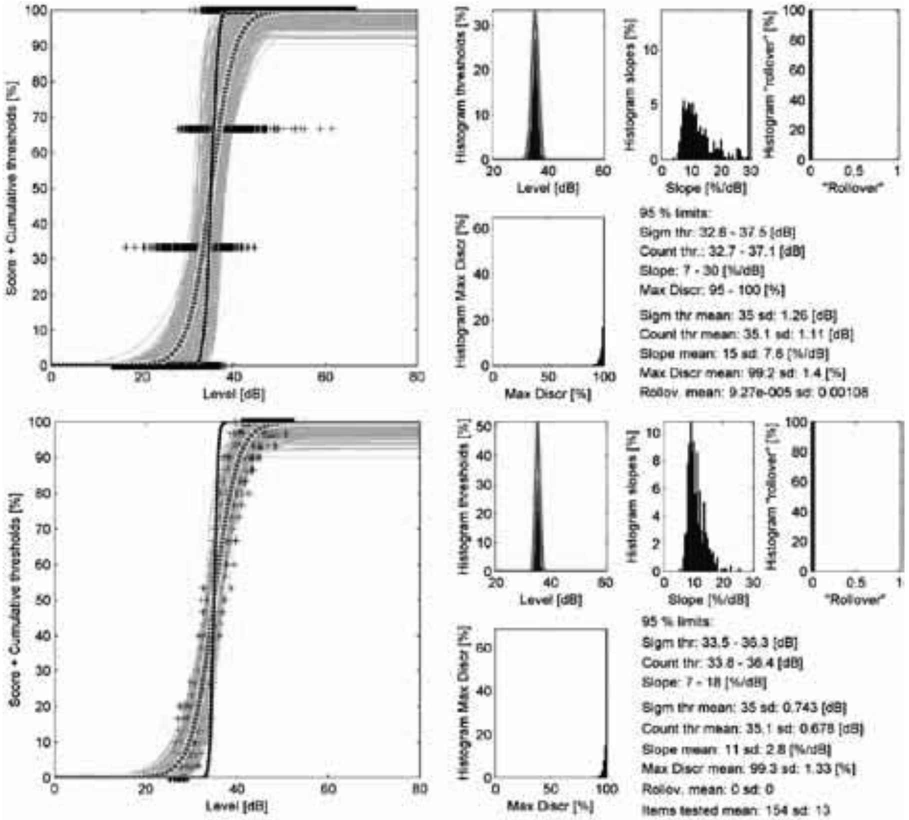


Fig. 3: Examples of simulated measurements used in developing "HiST taleaudiometri" The results of 500 simulated speech audiometry measurements with three-word utterances are presented. Top: the quick speed method. Below: the traditional method. Full explanation of the details of the figure can be found in Øygarden (2009a)

The quick speed method

This is one of the methods that are evaluated and recommended. Several of the tests in "HiST taleaudiometri" are based on this method. A recorded list of utterances/sentences with stepwise decreasing levels (steps of 2,5 dB for five-words or 1,5 dB for three-words) is presented while counting the total number of recognized words in the list. The resulting speech recognition threshold can be calculated by a formula or given from a table from the total number of recognized words.

An evaluation of the quick speed method for threshold measurements was performed by Haukland *et al.* (2010). Fig. 4 presents the results, showing approximately the same performance with the different methods.

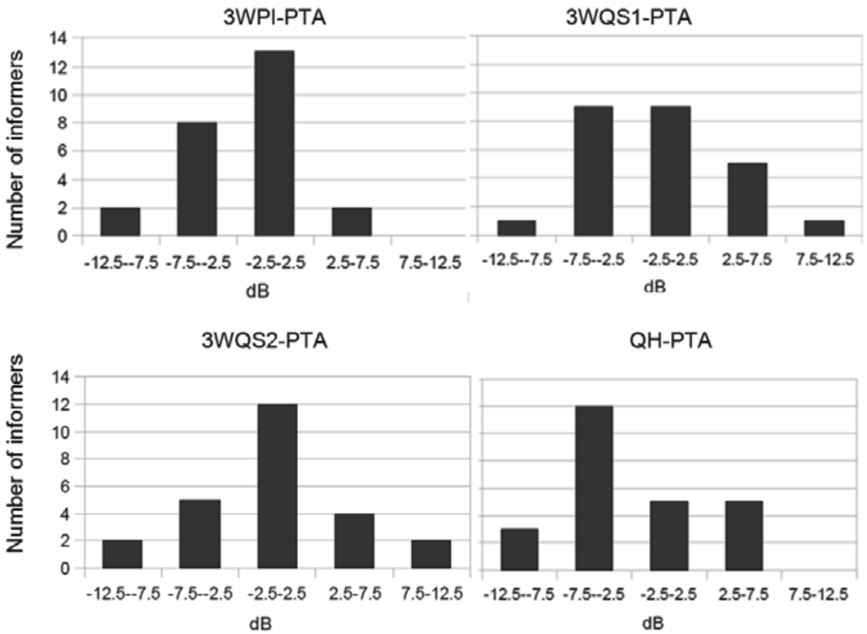
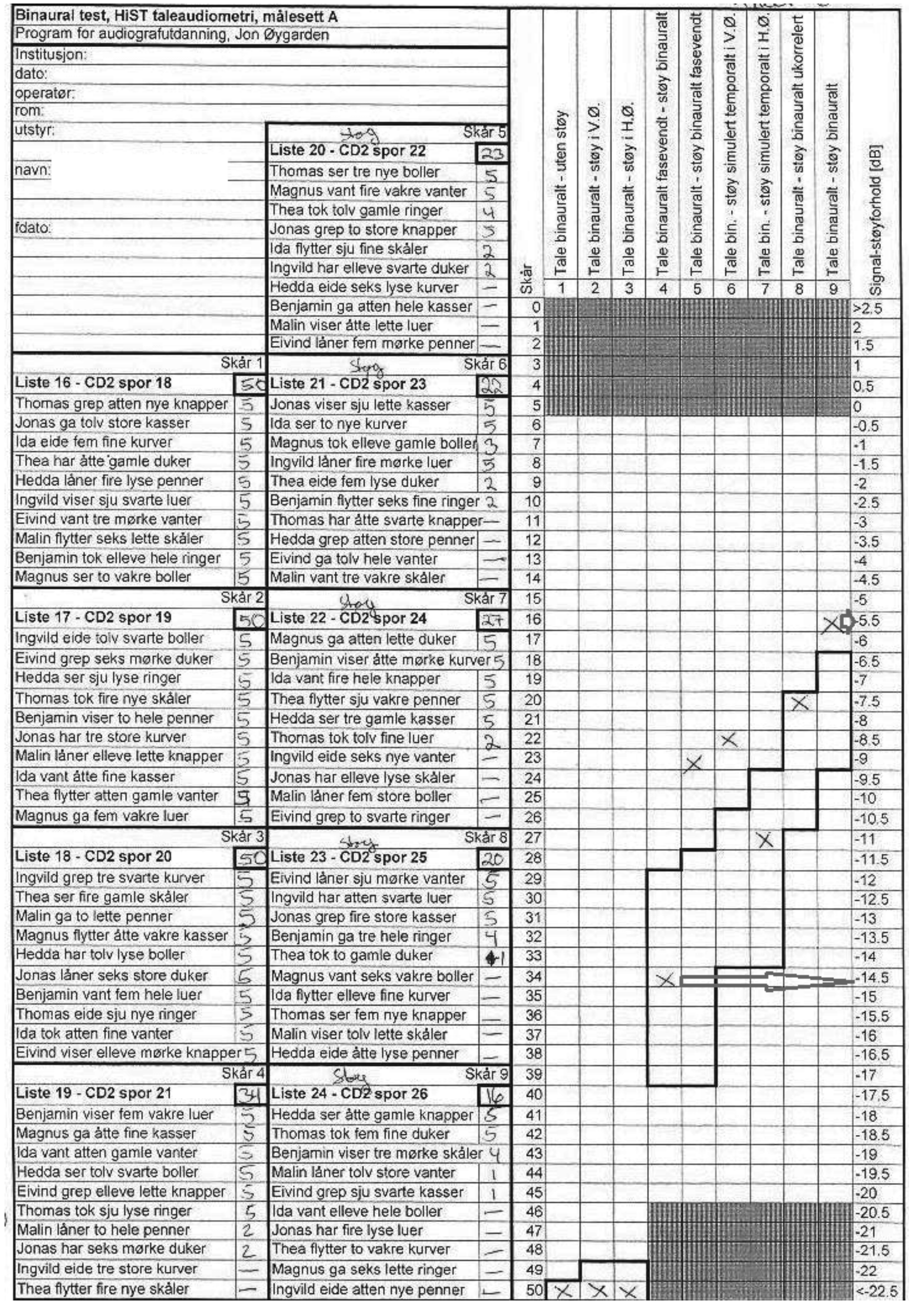


Fig. 4: Results of speech recognition threshold measured with three-word utterances both with two sets of the quick speed method (3WQS) and a PI curve fitting method (3WPI) are given below, together with results from the old Quist-Hanssen spondaic measurements (QH). All the results are given as the difference to the pure tone average, PTA.

SPECIAL TESTS

Binaural test

Fig. 5 shows a test sheet for this headphone test which has been used to measure aspects of binaural hearing. The quick speed method is used for each list with 10 five-word sentences with the speech level reduced in 2,5 dB steps. The lists consist of different binaural mixes of speech and noise: in-phase, reverse-phase or delays. The audiologist has to score number of words recognized for each sentence. The



Binaural test, HiST taleaudiometri, målesett A
Program for audiografutdanning, Jon Øygarden

Institusjon:
dato:
operatør:
rom:
utstyr:
navn:
fdato:

Skår 1

Liste 16 - CD2 spor 18	50
Thomas grep atten nye knapper	5
Jonas ga tolv store kasser	5
Ida eide fem fine kurver	5
Thea har åtte gamle duker	5
Hedda låner fire lyse penner	5
Ingvild viser sju svarte luer	5
Eivind vant tre mørke vanter	5
Malin flytter seks lette skåler	5
Benjamin tok elleve hele ringer	5
Magnus ser to vakre boller	5

Skår 2

Liste 17 - CD2 spor 19	50
Ingvild eide tolv svarte boller	5
Eivind grep seks mørke duker	5
Hedda ser sju lyse ringer	5
Thomas tok fire nye skåler	5
Benjamin viser to hele penner	5
Jonas har tre store kurver	5
Malin låner elleve lette knapper	5
Ida vant åtte fine kasser	5
Thea flytter atten gamle vanter	5
Magnus ga fem vakre luer	5

Skår 3

Liste 18 - CD2 spor 20	50
Ingvild grep tre svarte kurver	5
Thea ser fire gamle skåler	5
Malin ga to lette penner	5
Magnus flytter åtte vakre kasser	5
Hedda har tolv lyse boller	5
Jonas låner seks store duker	5
Benjamin vant fem hele luer	5
Thomas eide sju nye ringer	5
Ida tok atten fine vanter	5
Eivind viser elleve mørke knapper	5

Skår 4

Liste 19 - CD2 spor 21	34
Benjamin viser fem vakre luer	5
Magnus ga åtte fine kasser	5
Ida vant atten gamle vanter	5
Hedda ser tolv svarte boller	5
Eivind grep elleve lette knapper	5
Thomas tok sju lyse ringer	5
Malin låner to hele penner	2
Jonas har seks mørke duker	2
Ingvild eide tre store kurver	—
Thea flytter fire nye skåler	—

Støy Skår 5

Liste 20 - CD2 spor 22	23
Thomas ser tre nye boller	5
Magnus vant fire vakre vanter	5
Thea tok tolv gamle ringer	4
Jonas grep to store knapper	5
Ida flytter sju fine skåler	2
Ingvild har elleve svarte duker	2
Hedda eide seks lyse kurver	—
Benjamin ga atten hele kasser	—
Malin viser åtte lette luer	—
Eivind låner fem mørke penner	—

Støy Skår 6

Liste 21 - CD2 spor 23	22
Jonas viser sju lette kasser	5
Ida ser to nye kurver	5
Magnus tok elleve gamle boller	3
Ingvild låner fire mørke luer	5
Thea eide fem lyse duker	2
Benjamin flytter seks fine ringer	2
Thomas har åtte svarte knapper	—
Hedda grep atten store penner	—
Eivind ga tolv hele vanter	—
Malin vant tre vakre skåler	—

Støy Skår 7

Liste 22 - CD2 spor 24	27
Magnus ga atten lette duker	5
Benjamin viser åtte mørke kurver	5
Ida vant fire hele knapper	5
Thea flytter sju vakre penner	5
Hedda ser tre gamle kasser	5
Thomas tok tolv fine luer	2
Ingvild eide seks nye vanter	—
Jonas har elleve lyse skåler	—
Malin låner fem store boller	—
Eivind grep to svarte ringer	—

Støy Skår 8

Liste 23 - CD2 spor 25	20
Eivind låner sju mørke vanter	5
Ingvild har atten svarte luer	5
Jonas grep fire store kasser	5
Benjamin ga tre hele ringer	4
Thea tok to gamle duker	1
Magnus vant seks vakre boller	—
Ida flytter elleve fine kurver	—
Thomas ser fem nye knapper	—
Malin viser tolv lette skåler	—
Hedda eide åtte lyse penner	—

Støy Skår 9

Liste 24 - CD2 spor 26	16
Hedda ser åtte gamle knapper	5
Thomas tok fem fine duker	5
Benjamin viser tre mørke skåler	4
Malin låner tolv store vanter	1
Eivind grep sju svarte kasser	1
Ida vant elleve hele boller	—
Jonas har fire lyse luer	—
Thea flytter to vakre kurver	—
Magnus ga seks lette ringer	—
Ingvild eide atten nye penner	—

Skår	Tale binauralt - uten støy 1	Tale binauralt - støy i V.Ø. 2	Tale binauralt - støy i H.Ø. 3	Tale binauralt fasevendt - støy binauralt 4	Tale binauralt - støy binauralt fasevendt 5	Tale bin. - støy simulert temporalt i V.Ø. 6	Tale bin. - støy simulert temporalt i H.Ø. 7	Tale binauralt - støy binauralt ukorrelert 8	Tale binauralt - støy binauralt 9	Signal-støyforhold [dB]
0	▨	▨	▨	▨	▨	▨	▨	▨	▨	>2.5
1	▨	▨	▨	▨	▨	▨	▨	▨	▨	2
2	▨	▨	▨	▨	▨	▨	▨	▨	▨	1.5
3	▨	▨	▨	▨	▨	▨	▨	▨	▨	1
4	▨	▨	▨	▨	▨	▨	▨	▨	▨	0.5
5	▨	▨	▨	▨	▨	▨	▨	▨	▨	0
6										-0.5
7										-1
8										-1.5
9										-2
10										-2.5
11										-3
12										-3.5
13										-4
14										-4.5
15										-5
16									X	-5.5
17										-6
18										-6.5
19										-7
20								X		-7.5
21										-8
22						X				-8.5
23					X					-9
24										-9.5
25										-10
26										-10.5
27							X			-11
28										-11.5
29										-12
30										-12.5
31										-13
32										-13.5
33										-14
34				X						-14.5
35										-15
36										-15.5
37										-16
38										-16.5
39										-17
40										-17.5
41										-18
42										-18.5
43										-19
44										-19.5
45										-20
46				▨	▨	▨	▨	▨	▨	-20.5
47				▨	▨	▨	▨	▨	▨	-21
48				▨	▨	▨	▨	▨	▨	-21.5
49				▨	▨	▨	▨	▨	▨	-22
50	X	X	X	▨	▨	▨	▨	▨	▨	<-22.5

Fig. 5: Test sheet used in the binaural test. The framed area show normal results ± 1SD.

total score for each list is marked in the appropriate column on the right side. The measured signal-to-noise ratio can be found in the column far right. Binaural intelligibility level differences (BILD) can be computed from the response sheet. The grey arrows in Fig. 5 show that BILD between speech and noise in phase (column 9 - S0N0) and speech with binaural phase reversal and noise in phase (column 4 – SπN0) can be computed as -5.5 dB SNR minus -14.5 dB SNR = 9 dB.

Surround test

Similar test sheets are developed for this test which is also based on the quick speed method and the five-word sentences. The test material is on an audio DVD and standard home theatre systems can be used for the measurement. The sound is in Dolby Digital 5.1 surround format. The speech is in the front channel, and uncorrelated noises or distracting speech are in the four other channels. Some of the lists are made with simulated reverberation. The test can be used for evaluation of binaural performance with or without hearing aids.

CONCLUSIONS

- Five-word sentences can be generated with the diphone method to achieve a natural sounding speech
- Three-word utterances is a good alternative to spondaic words for measuring speech recognition threshold
- The quick speed method is a good alternative to measure the performance intensity (PI) curve for determining the speech recognition threshold
- Binaural tests can be performed with this material both with headphones and in free field

PRODUCT

"HiST taleaudiometri" consists of a report, two CDs and an audio DVD. It is available from the university bookstore in Trondheim:
http://butikk.tapirforlag.no/node/1303
All documentation can be found at: *http://www.hist.no/taleaudiometri*

REFERENCES

Hagerman, B. (1982). "Sentences for testing speech intelligibility in noise". Scandinavian Audiology, 11, 79-87.

Haukland, M., Kaldhol, M.R. and Meese, A. (2010). "HiST taleaudiometri — en validering av HiST hurtigtest". Bachelor thesis, HiST.

Wagener, K., Kühnel, V. and Kollmeier, B. (1999). "Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztest" (In German). Zeitschrift für die Audiologie, 38(1),4-15.

Øygarden, J. (2009a). "Norwegian Speech Audiometry", PhD thesis, NTNU, Trondheim (*http://urn.kb.se/resolve?urn=urn:nbn:no:ntnu:diva-5409*)

Øygarden, J. (2009b). "HiST taleaudiometri", Tapir akademisk forlag, Trondheim

---

cABR: A neural probe of speech-in-noise processing

SAMIRA ANDERSON[1, 2] AND NINA KRAUS[1-4]

[1, 2] *Auditory Neuroscience Laboratory http://www.brainvolts.northwestern.edu,*
[2]*Departments of Communication Sciences,* [3]*Neurobiology and Physiology,*
[4]*Otolaryngology at Northwestern University, Evanston, Il 60208, USA*

Understanding speech in noise (SIN) is a highly complex task affected by reciprocal sensory-cognitive interactions in the brain. The auditory brainstem response to complex stimuli (cABR) provides an objective index of the neural transcription of features (e.g. temporal, spectral) that are important for speech understanding. Influenced by cognitive factors such as attention and memory, measures of subcortical processing can further our knowledge of the biological mechanisms associated with deficits in SIN perception in clinical populations, such as children with learning impairments, older adults, and individuals with hearing impairment. Further, subcortical processing is modifiable through long-term experience and short-term training, making cABR a highly reliable probe for delineating the effects of training on neural speech processing.

INTRODUCTION

Understanding speech in background noise is hard for everyone, but especially so for children with language-based learning impairments, individuals with hearing loss, and older adults. Older adults frequently complain that they can hear what is said, but cannot understand the meaning, especially in background noise. They have particular difficulty understanding rapid speech, especially the rapidly-changing parts of speech contained in consonant-vowel transitions (Gordon-Salant *et al*., 2006). These difficulties occur in individuals with hearing loss and those who have audiologically normal hearing (Souza, 2007). Older adults are known to have slower neural processing. This neural slowing has been attributed to temporal jitter (Pichora-Fuller *et al*., 2007), delayed neural firing (Walton *et al*., 1998), and decreased inhibition (Caspary *et al*., 2008). We asked whether correlates of this neural slowing are found in the auditory brainstem response to complex sounds (cABR), and whether age-related deficits in subcortical spectrotemporal speech encoding are related to speech-in-noise (SIN) perception. We are currently investigating the efficacy of training for reversing age-related deficits in neural processing and SIN perception.

APPROACH

The cABR is well suited for assessing aging effects on neural processing of speech. Its fidelity to the stimulus can be seen in representation of timing features (onsets, offsets, envelope), pitch (encoding of the fundamental frequency (F_0) of the stimulus), and timbre (representation of formants above the F_0) through cycle-by-cycle neural phase-locking (Skoe and Kraus, 2010a). Given this fidelity and high reliability, the cABR has

Proceedings of ISAAR 2011: Speech perception and auditory disorders. 3rd International Symposium on Auditory and Audiological Research. August 2011, Nyborg, Denmark. Edited by T. Dau, M. L. Jepsen, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN 87-990013-3-0. EAN 9788799001330. The Danavox Jubilee Foundation, 2012.