Epstein, M. and Florentine, M. (**2009**). "Binaural loudness summation for speech and tones presented via earphones and loudspeakers" Ear. Hear., **30**(2), 234-237.
Fastl, H. and Florentine, M. (**2011**). Loudness in Daily Environments. *Loudness*. M. Florentine (Ed.), New York, Springer.
Florentine, M. and Epstein, M. (**2006**) “Why classroom demonstrations of binaural loudness summation fail,” Proc. IHCON International Hearing Aid Research Conference.
Florentine, M. and Epstein, M. (**2010**) “Ecological loudness: Binaural loudness constancy” 20th International Congress on Acoustics. Sydney, Australia.
Hellman, R. P. (**1991**). Loudness scaling by magnitude scaling: Implications for intensity coding. Ratio Scaling of Psychological Magnitude: In Honor of the Memory of S. S. Stevens. G. A. Gescheider and S. J. Bolanowski. Hillsdale, NJ, Erlbaum**:** 215-228.
Marks, L. E. (**1978**). "Binaural summation of the loudness of pure tones" J. Acoust. Soc. Am., **64**, 107-113.
Marozeau, J., Epstein, M., Florentine, M., and Daley, B. (**2006**). "A test of the binaural equal-loudness-ratio hypothesis for tones" J. Acoust. Soc. Am., **120**(6), 3870-3877.
Marozeau, J. and Florentine, M. (**2007**). "Loudness growth in individual listeners with hearing losses: a review" J. Acoust. Soc. Am., **122**(3), EL81.
McGurk, H. and McDonald, J. (**1976**). "Hearing lips and seeing voices" Nature, **264**, 746-748.
Mohrmann, K. (**1939**). "Lautheitskonstanz im Eutfernungswechsel" Zeitschr. f. Psychologie, **145**, 145-199.
Reynolds, G. S. and Stevens, S. S. (**1960**). "Binaural summation of loudness" J. Acoust. Soc. Am., **32**, 1337-1344.
Scharf, B. and Fishken, D. (**1970**). "Binaural summation of loudness: reconsidered" J. Exp. Psych., **86**, 374-379.
Schneider, B. A. and Cohen, A. J. (**1997**). "Binaural additivity of loudness in children and adults" Percept. Psychophys., **59**, 655-664.
Smeds, K. and Leijon, A. (**2011**). Loudness and Hearing Loss. *Loudness*. M. Florentine (Ed.), New York, Springer.
Sivonen, V. P. and Ellermeier, W. (**2011**). Binaural Loudness. *Loudness*. M. Florentine (Ed.), New York, Springer.
Whilby, S., Florentine, M., Wagner, E., and Marozeau, J. (**2006**). "Monaural and binaural loudness of 5- and 200-ms tones in normal and impaired hearing" J. Acoust. Soc. Am., **119**(6), 3931-3939.
Zahorik, P. and F. L. Wightman (**2001**). "Loudness constancy with varying sound source distance" Nat. Neurosci., **4**(1), 78-83.
Zwicker, E. and Zwicker, U. T. (**1991**). "Dependence of binaural loudness summation on interaural level differences, spectral distribution, and temporal distribution" J. Acoust. Soc. Am., **89**(2), 756-764.

The role of working memory capacity and speed of lexical access in speech recognition in noise

BIRGITTA LARSBY[1,3] MATHIAS HÄLLGREN[1,2,3] AND BJÖRN LYXELL[3]

[1] *Technical Audiology, Department of Clinical and Experimental Medicine, Linköping University, SE-58185 Linköping, Sweden*

[2] *Department of Otorhinolaryngology/Section of Audiology, Linköping University Hospital, SE-58185 Linköping, Sweden*

[3] *Linnaeus Centre HEAD, The Swedish Institute for Disability Research, Linköping University, SE-58183 Linköping, Sweden*

Cognitive skills are important for speech processing and comprehension. The objective of the present investigation was to address the role of working memory capacity and speed of lexical access in speech processing in noise. This was done by measuring speech recognition thresholds (SRTs) in background noise with the Hagerman test and the Swedish HINT. Cognitive capacities were measured by a lexical-decision making test and a reading-span test. Forty hearing impaired individuals, aged 44-86 years, participated in the study. The relations between SRTs and cognitive scores were studied using correlation analyses and analysis of variance for high and low cognitive performance groups. The individuals showed significant correlations between SRTs and cognitive scores which remained also after correcting for PTA4. A higher demand on cognitive processing was found for the criteria of 80% compared to 50 % speech recognition.

INTRODUCTION

Hearing speech yields not only the sensation of incoming auditory information, but also involves processing and interpretation of speech in the context of knowledge and previous experiences. Information regarding how speech is categorized and related, for example contextually, phonologically, lexically, syntactically, and semantically, is stored in a long-term memory. Speech processing includes elements of both top-down and bottom-up processing at all stages of information-processing (cf., Davis and Johnsrude, 2007).

The absolute hearing threshold is often a major factor in explaining the variation in speech recognition scores for hearing-impaired listeners in quiet as well as in noise (Plomp, 1986). Remaining variations may be caused by inter-subject variations in peripheral supra-threshold distortion (e.g. Plomp, 1986; Glasberg and Moore, 1989), and from variations in central auditory and cognitive processes (Humes, 2005; Pichora-Fuller, 2003).

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.

A number of previous studies have shown the importance of cognitive processing in auditory tasks (Humes, 2005) as well as in various kind of speech processing (e.g. Hällgren *et al.*, 2001; Lunner, 2003; Lyxell *et al.*, 2003; Pichora-Fuller, 2003; Larsby *et al.*, 2005; Foo *et al.*, 2007). Distortion or limitation of an incoming speech stimulus or limitations caused by hearing loss slow down the processing of speech since it become more dependent on cognitive processing. Processing of spoken information is partly dependent on the characteristics of three cognitive abilities; capacious working memory, fast lexical processing of information and phonological skills (Lyxell *et al.*, 2003).

In several studies of cognitive functions and speech recognition in noise different types of methods have been used, varying in terms of type of sentence materials and scoring criteria. Seldom have the methods been compared and evaluated. The Hagerman sentences (Hagerman, 1982), is a closed-set sentence material, which have an identical structure and are combined from a limited set of words, making them highly redundant on a syntactic level, but not on a semantic level. The Swedish HINT (Hällgren *et al.*, 2006), is an open-set material, with short meaningful everyday sentences with redundancy also on a semantic level. In the Hagerman test the standard scoring criteria is to find the signal-to-noise ratio where 50% of the words are recognized. For HINT the scoring criteria is 50% correct sentences, corresponding to a higher percentage of recognized words. For the Hagerman sentences, Lunner and Sundewall-Thorén (2007) compared 50% with 80% correctly recognized words and the relationship to working memory capacity and showed that scores at 80% speech-recognition level were more sensitive to cognition, than those at 50% level. Larsby *et al.* (2008) compared speech recognition thresholds both for 50% and 80 % word recognition with the Hagerman sentences in unmodulated and modulated noises. Different response patterns for the 50% and 80% response criteria were observed indicating that other mechanisms, such as cognitive abilities and/or central auditory functions, also affected the word recognition performance. The aim of the present study was to study the effect of sentence material and scoring criteria in speech recognition tasks in noise and their relationships to the two cognitive skills, working memory capacity and speed of lexical access.

METHODS

Auditory tests

Speech recognition in noise was measured using the Swedish HINT sentences (Hällgren *et al.*, 2006) and the Hagerman sentences (Hagerman, 1982). In the HINT test, the noise level was varied in 2 dB steps in an adaptive method, to reach either 50% correctly repeated sentences or 50% correctly repeated keywords. In the Hagerman test, the noise signal was adjusted adaptively in an interleaved method to estimate threshold for 50% and 80% correctly repeated words. In both tests, the speech was presented at a fixed level of 70 dB SPL, and the noises were spectrally shaped according to the long-term average of the speech material of the corresponding set.

Cognitive tests

Lexical-decision test. The task is to judge whether a combination of three letters is a real word or a non-word by pressing predefined response buttons for "yes" and for "no". One hundred items were used. Both accuracy and speed of performance are assessed (Hällgren *et al.*, 2001). The outcome measure for speed of lexical access throughout this study is the quotient between accuracy and response time.

Reading-span test. The subject's task is to comprehend three-word sentences and to recall either the first or the final words of a presented sequence of sentences in the correct serial order. The sentences are presented in a word-by-word fashion. Half of the sentences are meaningful and the others are absurd. The individuals' task is to respond "yes" (for a normal sentence) or "no" (for an absurd sentence). After a sequence of three to five sentences the test leader indicates that the subject should start to recall either the first or the final word for each presented sentence. The average ratio of correctly reported words is used as outcome measure and estimate of the working memory capacity.

Individuals

Forty listeners participated in the study, ranging from 46 to 86 years (average = 66.2, SD = 8.7). All individuals had a bilateral mild to moderate sensorineural hearing impairment. The average audiogram of the individuals is shown in figure 1.

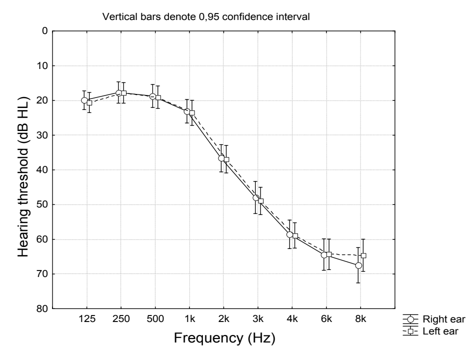

Fig. 1: Average hearing-threshold levels for the 40 HI listeners.

Procedures

The subjects were seated in a sound-attenuated room. The auditory stimuli of the SRT tests were presented over a loudspeaker at a distance of one meter in front of the subject. No hearing devices were used during testing.

Statistics

Pearsson correlations were used to study relations between SRT scores and cognitive measures. ANOVA was carried out to further analyze differences in SRT scores between cognitively high vs. low scoring persons.

RESULTS

In the result section descriptive statistics from the different tests are presented in Table 1. Correlations between SRTs and cognitive measures and the correlations between different cognitive measures are then given followed by the analysis of SRTs for cognitive low and high performance groups, respectively.

		Avg.	SD	Units
Age		66.2	8.7	Years
PTA4		32.5	7.3	dB
SRT measures				
HINT	keyw. score	1.0	2.7	dB SNR
HINT	sent. score	3.5	3.1	dB SNR
Hagerman	50% words	-1.8	3.7	dB SNR
Hagerman	80% words	5.4	6.1	dB SNR
Cognitive measures				
Lexical decision	Acc/Rstm	106	22	%/sec.
Reading span	Acc	34	11	%

Table 1: Mean values and SDs for the 40 hearing-impaired listeners. For Lexical Score, and Reading-span Score, higher numbers indicate better scores.

Correlations between SRTs and cognitive measures and between working memory capacity and speed of lexical access measures: There were strong correlations between the SRTs with the lexical test and the reading-span test. In addition there were many significant correlations between the pure-tone average thresholds (PTA4) and both the SRTs and the cognitive tests. In order to control for the confounding factor of peripheral hearing, partial correlations corrected for the PTA4 were calculated. The correlation-coefficients between the SRTs and the cognitive tests decreased with this correction, but there were still significant correlations between the SRTs of the HINT test (sentence scoring) and Hagerman test (80% word recognition) with the lexical score and reading-span score, see table 2.

	HINT keyw. score	HINT sent. score	Hagerman 50% words	Hagerman 80% words
Age	0.11	0.14	0.23	0.22
Lexical decision	-0.26	-0.37*	-0.23	-0.44**
Reading span	-0.31	-0.33*	-0.07	-0.32*

Table 2: Partial Pearson correlations, corrected for PTA4 for 40 HI listeners. The symbols "*" and "**" indicate significance at the 5% and 1% level, respectively.

Analysis of SRTs for cognitively low and high performance groups: For each cognitive test the individuals were divided into three cognitive sub-groups based on their results in the lexical decision test and in the reading span test, respectively. For each of the speech recognition tests, HINT and Hagerman, two-way ANOVAs were performed with SRT as dependent variable and cognitive group (high, low) as

between-group factor, and scoring procedure (level1: HINT 50% keyword, Hagerman 50% word; level2: HINT 50% sentence, Hagerman 80% word) as within-subject factor.

When the individuals were divided into groups based on their results on the lexical test, there was a main effect of scoring procedure, both in the HINT, (p<0.001) and in the Hagerman test (p<0.001). In the HINT test the SRTs required for 50% correctly scored keywords (level1) were significantly lower than 50% correctly scored sentences (level2). In the Hagerman test, the thresholds were lower for 50% (level1) than for 80% (level2) word recognition. There was no main effect of cognitive group, neither in HINT nor in Hagerman. However, there was an interaction between scoring procedure and cognitive group, in HINT (p=0.02) as well as in Hagerman (p=0.05), see figure 2A. There were much larger differences between the cognitive groups for level2, than for level1.

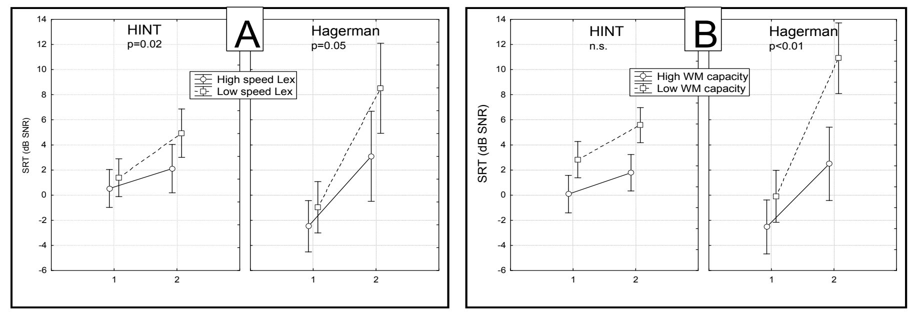

Fig. 2: Mean SRTs for level1 and level2 scoring in the HINT test and the Hagerman test for the group with high (□) and low (o) scores in the lexical test, fig. 2A, and the reading span test, fig. 2B. Vertical bars denote 0.95 confidence intervals.

When the individuals were divided into groups based on their results on the reading-span test, there was a main effect of scoring procedure, both in the HINT, (p<0.001) and in the Hagerman test (p<0.001). In both HINT and Hagerman the SRTs were lower at level1 compared to level2. There was a main effect of cognitive group, in HINT (p=0.001) and in Hagerman (p=0.001). In general the high working memory capacity group performed better (lower SRTs). There was also a significant interaction between scoring procedure and cognitive group, in Hagerman (p=0.001), but not in HINT, see figure 2B.

DISCUSSION

The peripheral hearing function is usually measured by pure-tone audiometry. Hearing professionals often observe that two persons with identical audiograms may have varying ability to make use of amplification or to deal with adverse listening conditions. In addition to audibility, supra-threshold discrimination abilities, age, and cognitive abilities contribute to the total variance for SRTs in noise (Glasberg and Moore, 1989; Akeroyd, 2008; Houtgast and Festen, 2008). The present study focuses on how two different cognitive abilities important for speech understanding, working memory and speed of lexical access, affect speech recognition at two different response criteria and with two different speech materials.

Correlations between cognitive tests and speech-in-noise tests: Several SRT results correlate significantly with the results of the cognitive lexical-decision and reading-span tests (table 2). For the individuals in the present study, peripheral hearing is a dominant factor; the average pure-tone thresholds correlate with test scores and may be a confounding factor in the correlations between SRTs and cognitive scores. However, after correction for pure-tone hearing thresholds, strong correlations between cognitive scores and SRTs can still be seen between the lexical score and the reading-span score on the one hand and the HINT with sentence scoring and the Hagerman test with the 80% word-recognition criterion on the other hand (table 2). For the 50% word/keyword recognition criteria (level1), there were no significant correlations between SRTs and cognitive scores after PTA4 correction was applied.

One explanation for the different degrees of correlation between the cognitive scores and the SRTs for level1 and level2 is that, in the more difficult listening situations (level1), the most critical factor is the lack of audibility. In contrast, in easier listening situations (level2), the larger amount of auditory information makes it possible to compensate for disturbed auditory information using explicit processing, leading to increased understanding (see also Larsby *et al.*, 2008). The observation of high demands on cognitive processing at level2 has also been verified by Foo *et al.* (2007), in a reading span task, using both the Hagerman test and the Swedish HINT.

Analysis of variance: The results from the ANOVAs support the theory that cognitive capacity is more essential in level2 than in level1. When the individuals with the highest and the lowest score on the lexical test were compared, significant interactions between the cognitive group and scoring level (figure 2A) emerge, the difference between the two cognitive groups is more pronounced at level2 than at level1.

The results from the ANOVA, where individuals were divided into cognitive groups according to results on the reading-span test, show main effects of cognitive group both for the HINT and the Hagerman test (figure 2B). Furthermore, there was a significant interaction between cognitive group and scoring level in the Hagerman test, but not in the HINT test. Lunner and Sundewall-Thorén (2007), studying speech recognition in noise with the Hagerman test and relating it to cognitive function, as measured with a visual letter monitoring task, found that scores at 80% speech-recognition level were more sensitive to cognition, than those at 50% level. The different demands on cognitive processing at 50% and 80% word recognition level in the Hagerman test have in the present study been verified also with the reading-span and the lexical test results as dividers of cognitive capacity. For the first time, a higher degree of cognitive demand at level2 compared to level1 was also shown for the HINT test with the lexical test as divider between the cognitive groups.

Different speech materials: In the analyses of the results in the HINT test with reading-span score as divider there is no interaction, but a main effect of cognitive group (figure 2B). As there is a difference in peripheral hearing between the cognitive groups, the “high” cognitive group has better hearing; the main effect can be explained by peripheral hearing. However, peripheral hearing cannot explain the interaction effects. It can thus be concluded that the level2 scoring compared to level1 is more sensitive to cognitive capacity as measured by reading-span, letter monitoring (Lunner and Sundewall-Thorén, 2007) and lexical decision making.

Communication requires speech recognition, interpretation of information, and decision making. In difficult listening situations, we can use context and visual cues from lip-reading to improve communication. Compared to the complexity of communication, the task to repeat sentences in the SRT measurements is rather straightforward and less demanding. Both the Hagerman and the Swedish HINT test were included in the present investigation. The purpose was to study the effect of cognitive functions in speech recognition, both for a closed and open set sentence material, assessing both general and material specific components. In the HINT test, whole sentences need to be retrieved, whereas in the Hagerman test recognition is primarily on a word level. The Hagerman sentences (Hagerman, 1982) have an identical structure and are combined from a limited set of words, making them grammatically correct and highly redundant on a syntactic level. However, the sentences are low-redundant on a semantic level. The Swedish HINT sentences (Hällgren *et al.*, 2006) are short meaningful everyday sentences with redundancy both on syntactic and semantic levels.

For both the Hagerman and the HINT test there are significant remaining correlations, after correction for peripheral hearing, between speech recognition thresholds and cognitive capacity at level2 (table 2). This support the idea that a common underlying mechanism which is not related to speech material per se plays a role. This statement is also supported by the ANOVA results. Working memory capacity is highly critical for 80% word recognition in the Hagerman test. The Hagerman sentences have less semantic redundancy than the HINT sentences which force the listener to recognize, store and recall several single units rather than a whole meaningful sentence.

ACKNOWLEDGMENTS

This research was supported by grants from The Swedish Council for Working life and Social Research (FAS) and the European Union FP6, Project 004171 HearCom. The authors wish to thank Gunilla Wänström for the help with the data collection.

REFERENCES

Akeroyd, M.A. (**2008**). “Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults” Int J Audiol, **47, Suppl 2**, S53-71.

Davis, M.H. and Johnsrude, I.S. (**2007**). "Hearing speech sounds, Top-down influences on the interface between audition and speech perception" Hear Res **229**, 132-147.
Foo, C., Rudner, M.., Ronnberg, J. and Lunner, T. (**2007**). "Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity" J Am Acad Audiol, **18**, 618-31.
Glasberg, B. and Moore, B. C. J. (**1989**). "Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech" Scand Audiol, **Suppl 32**, 1-25.
Hagerman, B. (**1982**). "Sentences for testing speech intelligibility in noise" Scand Audiol, **11**, 79-87.
Houtgast, T. and Festen, J.M. (**2008**). "On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise" Int J Audiol, **47**, 287-95.
Humes, L. (**2005**). "Do 'auditory processing' tests measure auditory processing in the elderly?" Ear and Hearing, **26**, 109-19.
Hällgren, M., Larsby, B. and Arlinger, S. (**2006**). "A Swedish version of the Hearing In Noise Test (HINT) for measurement of speech recognition" Int J Audiol, **45**, 227-237.
Hällgren, M., Larsby, B., Lyxell, B. and Arlinger, S. (**2001**). "Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired subjects" J Am Acad Audiol, **12**, 357-370.
Larsby, B., Hällgren, M., Lyxell, B. and Arlinger, S. (**2005**). "Cognitive performance and perceived effort in speech processing tasks: Effects of different noise backgrounds in normal-hearing and hearing-impaired subjects" Int J Audiol, **44**, 131-143.
Larsby, B., Hällgren, M. and Lyxell, B. (**2008**). "The interference of different background noises on speech processing" Int J Audiol, **47, Suppl 2**, S83-90.
Lunner, T. (**2003**). "Cognitive function in relation to hearing aid use" Int J Audiol, 42: S49-S58.
Lunner, T. and Sundewall-Thorén, E. (**2007**). "Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid" J Am Acad Audiol, **18**, 539–552.
Lyxell, B., Andersson, U., Borg, E. and, Ohlsson, I.-S. (**2003**). "Working-memory capacity and phonological processing in deafened adults and individuals with severe hearing impairment" Int J Audiol, **42, Suppl 1**, S86-S89.
Pichora-Fuller, M.K. (**2003**). "Processing speed and timing in aging adults: psychoacoustics, speech perception, and comprehension" Int J Audiol, **42, Suppl 1**, S59-67.
Plomp, R. (**1986**). "A signal-to-noise ratio model for the speech-perception threshold of the hearing impaired" J Speech Hear Res, **29**, 146-154.

Side-effects of binaural tone vocoding on recognising target speech presented against spatially separated speech maskers

MARTIN R. ANDERSEN, MICHAEL S. KRISTENSEN, TOBIAS NEHER AND THOMAS LUNNER

Eriksholm Research Centre, Oticon A/S, DK-3070 Snekkersten, Denmark

Previous experiments have indicated that monaural Temporal Fine Structure (mTFS) information aids Speech Reception. In these experiments mTFS was either kept or substituted using a tone-vocoder. Results showed that hearing-impaired subjects were not able to utilise mTFS information to the same degree as normal-hearing subjects. A first step towards a more ecological experiment would be to exploit the tone-vocoder paradigm in a simulated spatial setup, and measure binaural TFS (bTFS) benefit. However, by the introduction of a binaural tone-vocoder, a concern arose that artificial ITD cues pointing to a direction determined by the phase difference between the carriers of the two channels, would be introduced in addition to the intended removal of the original Interaural Time Difference (ITD) by vocoding. This experiment investigated this concern, by measuring speech reception for target speech presented against spatially separated speech maskers. 21 young normal hearing, 10 elderly normal hearing and 11 elderly hearing impaired subjects were tested in a fixed spatial condition with either the artificial ITD pointing forward (0º azimuth) or ±50º.

INTRODUCTION

Recent studies (Hopkins *et al*. 2008; Lunner *et al.*, in press) have shown that temporal fine structure (TFS) is important for speech intelligibility in complex situations with concurrent masking talkers. In these studies the benefit from TFS was measured in terms of Speech Reception threshold (SRT) changes using speech-on-speech tests. The speech signals were presented monaurally over headphones, using tone vocoding to remove the original monaural TFS (mTFS) cues.

A first step towards a more ecological experiment would be to exploit the tone-vocoder paradigm in a simulated spatial setup, and to measure binaural TFS (bTFS). However, by the introduction of a binaural tone-vocoder, a concern arose that artificial ITD (AITD) cues pointing to a direction determined by the phase difference between the carriers of the two channels, would be introduced in addition to the intended removal of the original Interaural Time Difference (ITD) by vocoding.

In an earlier pilot study, this concern was tested on 8 normal-hearing (NH) subjects (Andersen *et al*., 2010), where a condition with the AITD pointing forward (0°

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.