REFERENCES

Cox, R., Alexander, G. and Gray, G. (**2007**). "Personality, hearing problems, and amplification characteristics: contributions to self-report hearing aid outcomes". Ear Hear.*,* **28**, 141-162.

Jerger, J., Chmiel, R., Stach, B., and Spretnjak, M. (**1993**). "Gender affects audiometric shape in presbycusis" J. Am. Acad. Audiol., **4**, 42-49.

Gatehouse, S. (**1999**). "Glasgow hearing aid benefit profile: derivation and validation of a client-centered outcome measure for hearing aid services" J. Am. Acad. Audiol.*,* **10**, 80-103.

Gatehouse, S. and Noble, W. (**2004**). "The speech, spatial and Qualities of Hearing scale (SSQ)" Int. J. Audiol.*,* **43**, 85-99.

Glasberg, B. and Moore, B. (**1990**). "Derivation of auditory filter shapes from notched-noise data" Hear. Res., **47**, 103-138.

Hopkins, K. and Moore, B. (**2011**). "The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise" J. Acoust. Soc. Am., **129**, 334-351..

Hopkins, K., and Moore, B. (**2010**). "Development of a fast method for measuring sensitivity to temporal fine structure information at low frequencies" Int. J. Audiol.*,* **49**, 940-946.

Hopkins, K. and Moore, B. (**2007**). "Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information" J. Acoust. Soc. Am., **122**, 1055-1068.

Kishon-Rabin, L., Segal, O., and Algom, D. (**2009**). "Associations and dissociations between psychoacoustic abilities and speech perception in adolescents with severe-to-profound hearing loss" J. Speech Hear. Res.*,* **52**, 956-972.

Moore, B. (**1985**). "Frequency selectivity and temporal resolution in normal and hearing impaired listeners" Brit. J. Audiol. **19**, 189-201.

Patterson, R. (**1976**). "Auditory filter shapes derived with noise stimuli" J. Acoust. Soc. Am., **59**, 640-654.

Pichora-Fuller, K., and MacDonald, E. (**2007**). "Auditory temporal processing deficits in older listeners: From a review to a future view of Presbycusis" In Auditory signal processing in hearing-impaired listeners*.* 22nd International Symposium of Audiology and Audiological Research (ISAAR). Dau, T., Buchholz, J., Harte, J., Christiansen, T. (Eds.) 291-300.

Strelcyk, O. and Dau, T. (**2009**). "Relations between frequency selectivity, temporal fine structure processing, and speech reception in impaired hearing" J. Acoust. Soc. Am., **125**, 3328-3345.

Tyler, R., Summerfiedl, Q., Wood, E., and Fernandes, M. (**1982**). "Psychoacustic and phonetic temporal processing in normal and hearing-impaired listeners" J. Acoust. Soc. Am., **72**, 740-752.

Speech intelligibility as a function of time compression, age, word position, and signal-to-noise ratio

Anne Schlueter[1], Inga Holube[1] and Ulrike Lemke[2]

[1] *Institute of Hearing Technology and Audiology, Jade University of Applied Sciences, Oldenburg, Germany*

[2] *Phonak AG, Staefa, Switzerland*

Among other parameters, speech intelligibility depends on the rate of speech. Therefore, variation of time compression might be useful for adjusting the threshold of 50% intelligibility in speech in noise tests at fixed positive signal-to-noise ratios (SNRs). Speech rate can be modified with uniform and non-uniform algorithms. Uniform algorithms delete equally spaced segments, while non-uniform algorithms first characterize the structure of the speech and then increase the speech rate dependent on the classification. Referring to studies using fast speech, age effects have to be taken into account. To investigate fast speech in a German speech intelligibility test, sentences speeded to different time compressions were presented at different SNRs and intelligibility measurements were conducted with young and elderly normal-hearing listeners. The outcomes were used to calculate SNR-dependent discrimination functions. The results showed increasing SNRs for 50% intelligibility with increasing time compression. As expected, young listeners reached higher intelligibility than elderly listeners at equal time compressions and SNRs for 50% intelligibility were shifted to lower values. Additionally, increasing the speech rate affected word intelligibility in dependence on the words' position within the sentences. These differences in intelligibility led to shallower slopes of the discrimination functions and could possibly constrain the accuracy of the test.

INTRODUCTION

In natural environments, speech often occurs together with interfering signals, which affect intelligibility. This situation is reproduced in clinical tests measuring speech intelligibility in background noise, when speech or noise is adapted in level until the threshold of 50% intelligibility (speech recognition threshold, SRT) is reached (e.g., HINT, Nilsson, 1994; OLSA, Wagener *et al*., 1999). Not only interfering signals influence intelligibility, but also speaker's speech rate. Previous studies documented, as expected, a decreasing performance with increasing speech rate (e.g., Adams and Moore, 2009). Following the example of Versfeld and Dreschler (2002), who adjusted the speech rate to reach the SRT in quiet, it could also be possible to measure the SRT by varying the speech rate at fixed signal-to-noise ratios (SNRs) in

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.

background noise. This approach would take into account two important parameters that influence intelligibility: background noise and speech rate. Furthermore, this method could be used for measuring the benefit of hearing aid algorithms like single-microphone noise reduction algorithms, because it could be used at fixed positive SNRs and could therefore test noise reduction algorithms in situations where they work beneficially.

Different algorithms can be used for compressing existing speech recordings in time, meaning an increase in speech rate without changing the pitch. Uniform time-compression algorithms use a pitch-synchronous overlap-add method for increasing the speed of speech regardless of its structure (e.g., Praat by Boetsma and Weenink, 2010). In contrast, non-uniform algorithms analyze speech material for example with regard to structure, rate, and stress (e.g., Mach1 by Covell *et al*., 1998). Consequently, they compress speech material primarily at parts that are considered less sensitive for intelligibility (e.g. pauses, vocals). Different definitions of the amount of time compression are applied in science. This article uses the time-compression factor ρ, which is the length of the time-compressed signal relative to its original (e.g., for ρ = 25% the speech was compressed to 25% of the original length). The smaller the time-compression factor, the faster the speech.

Previous studies documented further factors influencing intelligibility of time-compressed speech. For example, elderly listeners show greater difficulties in understanding time-compressed speech than young listeners (Tun, 1998; Gordon-Salant and Fitzgibbons, 2004; Schneider *et al*., 2005; Gordon-Salant and Friedman, 2011). In this study, the influence of time compression with different algorithms was investigated for young and elderly listeners using the German Oldenburg sentence test (OLSA, Wagener *et al*., 1999). Special attention was given to those time-compression factors which lead to positive SNRs. In addition, the intelligibility of words in dependence on their position within the sentences was determined.

METHOD

Participants

In total, 34 young listeners (aged under 40 years, mean age 23.3 years) participated in the experiments. A group of ten young participants took part in measurements using speech processed with the Praat algorithm, and a group of 24 young participants listened to signals compressed with the Mach1 algorithm. Additionally, nine elderly listeners (aged 60 years or older and retired, mean age 65.8 years) took part in the measurements with time-compressed speech processed with Mach1 so far. All participants had normal hearing, i.e. young listeners had a hearing level of 20 dB HL or less at all test frequencies between 0.125 and 8 kHz. For elderly listeners normal hearing was met by the same criteria, except that at test frequencies of 6 kHz and 8 kHz hearing levels of 30 dB HL and 40 dB HL, respectively, were accepted.

Stimuli

Sentences of the OLSA were presented together with a corresponding noise at different SNRs (Wagener *et al*., 1999). The sentences consisted of a name, a verb, a number, an adjective and an object. These elements were randomly chosen from lists of ten words each and built sentences with low redundancy. Noise was colored with the same long-term spectrum as the speech. The sentences were time-compressed with uniform and non-uniform algorithms to time-compression factors of ρ = 25%, 30%, 35%, 40% and 45%. The fastest speech, with a time-compression factor of ρ = 25% was presented only to the young participants, while only elderly participants listened to ρ = 45%.

Time-Compression Algorithms

Uniform time compression was performed with the pitch-synchronous overlap-add technique (Moulines and Charpentier, 1990) implemented in Praat (Boetsma and Weenink, 2009). The original speech signal is multiplied with a window function depending on the pitch of the signal. Segments are disregarded and remaining segments are added to the time-compressed signal. Number and position of the selected segments are dependent on the time-compression factor. Results measured with speech compressed with this processing are labeled in the following as "Praat".

The non-uniform time-compression algorithm Mach1 (Covell *et al*., 1998) estimates time-dependence of the local emphasis and the relative speaking rate. These two values are used for an estimation of the audio tension, which describes "the degree to which the local speech segments resist against changes" (Covell *et al*., 1998). Segments with a high tension are less compressible. The local target compression is adjusted with regard to the local audio tension and processed by a synchronized overlap-add technique. So Mach1 compresses pauses and silence, for example, more than it compresses speech. Results measured with speech processed by this algorithm are labeled "Mach1".

Measurement

During measurement, signals were processed by a Matlab routine and routed over a sound card (RME AD/DA-Interface ADI-2) and a headphone amplifier (Tucker Davis Technologies HB7 Headphone Driver). Participants listened to the signals presented diotically over a headphone (Sennheiser HDA 200). This measurement set-up was free-field equalized.

To familiarize themselves with the Oldenburg sentences, participants performed an SRT measurement with an adaptive adjustment of the SNR and original speech. To help participants adapt to the fast, compressed speech, the same measurement with time-compressed speech was executed and repeated when the time-compression factor was changed as for the subsequent intelligibility measurements. Feedback was given by presenting the complete sentence on a screen after the participant's answer. After the training, listeners performed the intelligibility measurement at a fixed SNR and with speech compressed to the factor used in the training. The number of correctly repeated words within a list of 30 sentences was used for the calculation of the intelligibility in percent. All measurements were conducted with a constant noise level of 65 dB SPL. Speech was presented at different levels to reach intelligibility scores between approximately 20% and 80%.

The first measurements were conducted with ten young participants using Praat. In our initial experiment, participants were invited to several sessions on different days. The presented time-compression factor was constant during each session and was selected at random. Furthermore, the SNRs were presented in random order within one session. Each SNR was measured three times. These measurements showed a training effect between the sessions but not within one session, so for the next experiments, the measurement time for each participant was shortened to one session.

During the second experiment, another group of 24 young listeners executed intelligibility measurements with Mach1. Each participant listened to speech time-compressed with two different randomly chosen factors and four or five different randomly presented SNRs.

During the third experiment, elderly participants performed the same measurements as the young listeners in the second experiment. For these measurements, time-compression factor and SNR were adjusted to achieve intelligibility scores between 20% and 80%.

RESULTS

For the presentation of the results, measured intelligibilities for different time-compression factors at different SNRs were used to model discrimination functions described by Equation 1. According to Wagener and Brand (2005), intelligibility is defined as the mean probability (p_{in}) of correctly repeated words if the sentences were presented with a defined SNR.

$$pin(SNR, SRT, slope) = 11+e4*slope*SRT-SNR \quad \text{(Eq. 1)}$$

This function was used to determine estimates of the SRT and slope at SRT for the measured intelligibility of each subject using a maximum likelihood fit. For the investigation of significant differences of the SRTs, a two-factored ANOVA with an α value of 0.5 was used.

Time Compression and SNR

The median SRT and slope for the speech time-compressed with Praat and Mach1 is documented for the young and elderly listeners in Table 1. The SRT increased and slope decreased as the time-compression factor decreased. The speech compressed with Praat led to lower SRTs and steeper slopes than the speech compressed with Mach1. This result was verified by an ANOVA, which showed significant main effects of time-compression algorithm ($p < 0.01$) and factor ($p < 0.01$) and their interaction ($p < 0.01$). Positive SRTs were reached by young subjects listening to speech compressed with Mach1 at factors of ρ = 25% and 30%.

		Young		Elderly	
	ρ [%]	**SRT [dB]**	**Slope [%/dB]**	**SRT [dB]**	**Slope [%/dB]**
Praat	**25**	-2.6 (1.5)	5.4 (0.3)	-	-
	30	-3.4 (1.8)	8.8 (1.5)	-	-
	35	-4.9 (1.5)	10.8 (1.6)	-	-
	40	-6.0 (1.5)	11.7 (3.3)	-	-
	45	-	-	-	-
Mach1	**25**	4.8 (5.6)	2.9 (0.8)	-	-
	30	0.0 (2.3)	5.8 (0.7)	8.0 (5.5)	3.3 (0.4)
	35	-1.6 (3.0)	6.7 (1.1)	2.0 (2.9)	5.3 (0.5)
	40	-2.3 (1.9)	8.0 (1.6)	1.5 (3.9)	5.6 (2.0)
	45	-	-	-0.4 (0.9)	5.0 (2.4)

Table 1: Median SRTs, slopes, and interquartile range of the discrimination functions calculated for each individual subject

Age Effect

Table 1 compares the intelligibility of speech compressed with Mach1 measured with young and elderly listeners. Young participants reached lower SRTs and therefore higher performance than elderly listeners. An ANOVA confirmed these results and showed a significant effect of the main factors age ($p < 0.01$) and time-compression factor ($p < 0.01$) and their interaction ($p = 0.016$). Elderly participants achieved positive SRTs at ρ = 30%, 35% and 40%.

Word Position

Figure 1 summarizes the intelligibility of words depending on their position within a sentence for the young participants. The figure shows results measured with ρ = 25% and 35% and a median overall intelligibility of about 50%. Therefore, the results measured for Praat (see Figure 1a) at an SNR of -3 dB for ρ = 25% and -5 dB for ρ = 35% are shown. Intelligibilities measured with Mach1 are depicted in Figure 1b; here, speech was presented at SNRs of 5 dB and -1 dB for ρ = 25% and 35%, respectively. Although the overall results show a nearly equal median intelligibility, the individual words' intelligibility differed. Names and objects showed the highest intelligibility, while verbs, numbers and adjectives were less intelligible. The differences between names or objects and adjectives, for example, were higher for Mach1 than for Praat and tended to increase with decreasing time-compression factor.

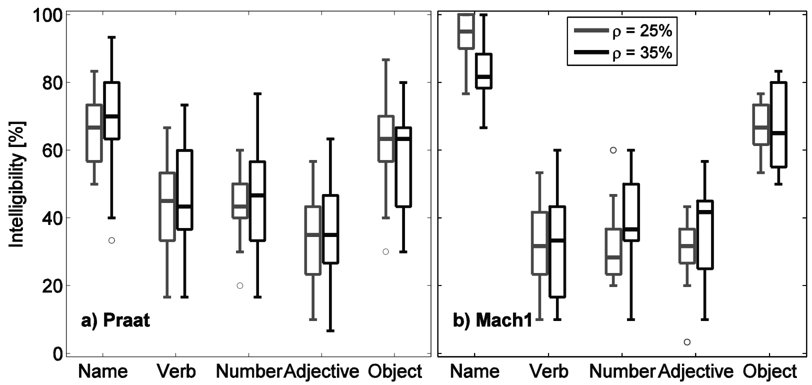

Fig. 1: Intelligibility of words depending on their position in the sentence. The plot depicts results of young participants at nearly equal median overall intelligibility of 50% measured with speech time-compressed to ρ = 25% and 35% using a) Praat and b) Mach1.

DISCUSSION

Time Compression and SNR

The median SRTs of our experiments showed a decreasing performance with decreasing time-compression factor and therefore increasing speech rate. This expected effect has been documented in previous studies (Versfeld and Dreschler, 2002; Adams and Moore, 2009).

Gordon-Salant and Fitzgibbons (2001) documented that processing of fast speech is attributed more to a modification of consonants than vowels or pauses. This finding is supported by our results, which showed a lower performance measured with Mach1 compared to Praat. Mach1 added artifacts to the speech signal. Some short consonants at the ends of words were deleted and the names were perceived relatively longer in comparison to the other words in the sentence. In contrast, Praat maintained the speech rhythm and all phonemes. The processed speech sounded more natural. Hence, the speech artificially compressed by Mach1 was less intelligible than speech compressed with Praat. Nevertheless, comparison of the performance measured with Praat and Mach1 was difficult. The method used for the measurements with Praat was different to that used for the measurements with Mach1. The change of the method was necessary because a training effect occurred during repeated measurements on different days and intelligibility increased from session to session.

Positive SNRs at SRT are essential for the processing of single-microphone noise reduction algorithms with maximum benefit (Schlueter, 2007). These were reached by the young participants listening to sentences time-compressed with Mach1 at the

very low time-compression factor of ρ = 25% and extremely fast speech. But higher time-compression factors could be used for results at positive SNRs if an intelligibility of 80% instead of 50% was selected for the intended threshold.

Age Effect

Higher SRTs were measured with elderly participants than with young listeners. This means that young listeners were better able to understand time-compressed speech than elderly listeners were. Furthermore, the estimated slopes of the discrimination functions were steeper for young participants than for elderly participants. These findings reproduced the results measured by Versfeld und Dreschler (2002). The reason for this difference in performance with age is widely discussed. Age-related decline has been thought to influence the performance of elderly listeners. Slowing of processing (Tun, 1998; Gordon-Salant and Fitzgibbons, 2004), decline in processing resources (Gordon-Salant and Fitzgibbons, 2001) and auditory decline (Schneider *et al*., 2005) are possible reasons for the effect. Recent findings by Gordon-Salant and Friedman (2011) indicated that the decline in recognition of rapid speech in noise is not only dependent on the aging process. Greater attention to auditory information may reduce the expected age-related decline in auditory temporal processing.

Elderly listeners reached positive SNRs at less compressed speech with a factor of ρ = 40% or lower. Therefore, speech did not have to be compressed as severely for elderly people to attain positive SNRs as for the young participants. This should be taken into account in investigations of speech in noise tests with time-compressed speech, because elderly people are the main user group of hearing aids with noise reduction algorithms. Nevertheless, these results are preliminary and should be confirmed with a larger number of elderly participants.

Word Position

Wagener and Brand (2005) investigated the dependence of SRTs on word position within a sentence for the original OLSA. Only small differences in intelligibility were documented. The difference between the SRT of names and the other words was only -0.6 dB. In contrast, the current study showed a higher intelligibility for names and objects compared to verbs, numbers and adjectives. This effect was enhanced by the Mach1 algorithm and by a higher time-compression factor. Under the assumption of a limited cognitive capacity of listeners for speech comprehension, primacy and recency effects of memory are primarily focused on the name and the object of an OLSA sentence. Additionally, increasing SNRs might not lead to equal increases in intelligibility of each word in a sentence especially when the maximum intelligibility of the name and object is already achieved. This difference in intelligibility of the words might influence the steepness of the discrimination functions. The shallower slope of the discrimination functions might constrain the accuracy of the test.

ACKNOWLEGEMENTS

We would like to thank Micha Lundbeck for his support on data collection and Simon Tucker (Department of Computer Science, University of Sheffield) for providing the Mach1 algorithm.

REFERENCES

Adams, E. M. and Moore, R. E. (**2009**). "Effects of Speech Rate, Background Noise, and Simulated Hearing Loss on Speech Rate Judgment and Speech Intelligibility in Young Listeners" J. Am. Acad. Audiol., **20**, 28-39.

Boetsma, P. and Weenink, D. (**2009**). "Praat – doing phonetics by computer (version 5.1.15)" www.praat.org (last viewed 08/02/2011).

Covell, M., Withgott, M. and Slaney, M. (**1998**). "Mach1: Nonuniform Time-Scale Modificaton of Speech" in Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signalprocessing, Seattle, **1**, 349-352.

Gordon-Salant, S. and Fitzgibbons, P. J. (**2001**). "Sources of Age-Related Recognition Difficulty for Time-Compressed Speech" J. Speech, Lang., and Hear. Res., **44**, 709-719.

Gordon-Salant, S. and Fitzgibbons, P. J. (**2004**). "Effects of stimulus and noise rate variability on speech perception by younger and older adults" J. Acoust. Soc. Am., **115**, 1808-1817.

Gordon-Salant, S. and Friedman, S. A. (**2011**). "Recognition of Rapid Speech by Blind and Sighted Older Adults" J. Speech, Lang., and Hear. Res, **54**, 622-631.

Moulines, E. and Charpentier, F. (**1990**). "Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones" Speech Comm., **9**, 453–467.

Nilsson, M., Soli S. D. and Sullivan, J. A. (**1994**). "Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise" J. Acoust. Soc. Am., **95**, 1085-1099.

Schlueter, A. (2007) "Perzeptive Beurteilung von Sprache im Störgeräusch", Master's thesis, Oldenburg University.

Schneider, B. A., Daneman, M. and Murphy, D. R. (**2005**). "Speech Comprehension Difficulties in Older Adults: Cognitive Slowing or Age-Related Changes in Hearing?" Psychology and Aging, **20**, 261-271.

Tun, P. A. (**1998**). "Fast noisy speech: age differences in processing rapid speech with background noise" Psychology and Aging, **13**, 424-434.

Versfeld, N. J and Dreschler, W. A. (**2002**). "The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners" J. Acoust. Soc. Am., **111**, 401-8.

Wagener, K. and Brand T. (**2005**). "Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: Influence of measurement procedure and masking parameters" Int. J. Audiol., **44**, 144-156.

Wagener, K., Kühnel, V. and Kollmeier, B. (**1999**). "Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests" Z. Audiol., **38**, 4-15.

Tinnitus, hyperacusis and their relation to hearing loss in professional symphony orchestra musicians

JESPER HVASS SCHMIDT[1,2,3], ELLEN RABEN PEDERSEN [1,3], HELENE PAARUP [1,3], TURE ANDERSEN[2,3], JAKOB CHRISTENSEN-DALSGAARD[4], TORBEN POULSEN[5], JESPER BÆLUM[1,3]

[1] *Dept. of Occupational and Environmental Medicine, Odense University Hospital, DK-5000 Odense C, Denmark*

[2] *Dept. of Audiology Odense University Hospital, DK-5000 Odense C, Denmark*

[3] *Institute of Clinical Research, University of Southern Denmark, DK-5230 Odense M, Denmark*

[4] *Institute of Biology, Centre for Sound Communication, University of Southern Denmark, DK-5230 Odense M, Denmark*

[5] *Centre for Applied Hearing Research, Department of Electrical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark*

Background: Musicians are exposed to loud sounds which can lead to hearing loss and hearing associated symptoms such as tinnitus and hyperacusis. Tinnitus and hyperacusis may be particular predominant in a population of musicians since musicians especially pay attention to audiologic symptoms. However, tinnitus and hyperacusis may or may not be associated with hearing loss. Purpose: To investigate the association between subjective hearing symptoms and objective hearing thresholds. Methods: Questionnaire data from 351 musicians from five symphony orchestras were used to estimate the frequency of subjective hearing loss, tinnitus and hyperacusis. Data from user operated Two Alternative Forced Choice Audiometry were available from 223 musicians and 199 of these included questionnaire data as well. Results: Subjective hearing loss was significantly ($p<0.001$) related to the hearing thresholds. Tinnitus was not related to the hearing thresholds in musicians. Subjects with hyperacusis were shown to have better hearing thresholds compared to musicians without hyperacusis. This was significant for the left ear after correction for age and gender ($p<0.02$). Conclusions: Auditory symptoms such as tinnitus and hyperacusis were not related to a reduced sensitivity in musicians. Hyperacusis was shown to be associated with a more sensitive hearing in musicians.

INTRODUCTION

Tinnitus can be described as an auditory phantom sensation where the experience is ringing in the ears when no external sound is present. Tinnitus is a common disorder in the population with an increasing prevalence with increasing age affecting

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.