Kahneman, D. (1973). *Attention and effort*. New Jersey: Prentice-Hall.
Kiessling, J., Pichora-Fuller, K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., et al. (**2003**). "Candidature for and delivery of audiological services: special needs of older people". Int J Audiol., **42(2)**, S92-S101.
Kumar, U.A., Avinash M. C. and Meti, R.R. (**2010**). "Development of Sentences for Quick Speech-in-Noise (QuickSin) Test in Kannada". Journal of Indian Speech and Hearing Association., Vol 24(1), page 59-65
Kumar, U.A. and Thakur, A. (**2009**). "Development of sentence material for quick speech in noise test in Indian accented English". Unpublished Master Dissertation, Mangalore University, Mangalore, India.
Marian, V., Blumenfeld, H., and Kaushanskaya, M. (**2007**). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Monolinguals" J Speech Lang Hear Res., **50,** 4, 940-967.
Rakerd, B., Seitz, P., and Whearty, M. (**1996**). "Assessing the cognitive demands of speech listening for people with hearing losses". Ear Hear., **17**, 97–106.
Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., and Abrams, H. B. (**2006**). "Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing," Appl. Psycholinguist. **27**, 465– 485.
Sarampalis, A., Kalluri, S., Edwards, B., and Hafer, E. (**2009**). "Objective measures of listening effort: Effects of background noise and noise reduction". J Speech Lang Hear Res., **52**, 1230–1240.

Measuring speech-in-speech intelligibility with target location uncertainty

NIELS SØGAARD JENSEN, RENÉ BURMAND JOHANNESSON, SØREN LAUGESEN, AND RENSKJE K. HIETKAMP

Eriksholm Research Centre, Oticon A/S, Kongevejen 243, DK-3070 Snekkersten, Denmark

The most common speech-intelligibility tests do only to a limited extent reflect the situations where hearing-impaired people typically experience speech-intelligibility problems in their everyday life. One major problem is that the resulting speech reception threshold (SRT) typically is unrealistically low. In an attempt to increase the ecological validity of the Danish Dantale II speech-intelligibility test, a modified version of the test was developed. The new version includes speech masking and target location uncertainty as ways to increase the resemblance to real-life situations. In the present study, test results were obtained from 16 hearing-impaired listeners and comparison was made to results obtained in three other test conditions, all having a fixed target position but including different types of masking signals. The results showed that the introduction of speech masking as well as target location uncertainty contributed to an increase in SRT.

INTRODUCTION

A common problem experienced by hearing-impaired people is poor speech intelligibility when in a group of people where different talkers of interest frequently take turns and where other people are talking simultaneously. This problem is not reflected in the most common standardized speech-intelligibility tests, which typically are based on presentation of speech from a fixed position in front of the listener in a background of steady-state noise (e.g., the Danish Dantale II test, Wagener *et al.* 2003). The result in this type of test is often a speech reception threshold (SRT), which is very low compared to the signal-to-noise ratios where real-life problems typically are experienced. Previous research has shown that the introduction of target location uncertainty (e.g., Brungart and Simpson 2007, Kidd *et al.* 2005, and Singh *et al.* 2008) and speech masking (e.g., Francart *et al.* 2011 and Helfer and Freyman 2008) affects performance in such tests in the direction of a higher and thus more realistic SRT.

A modified version of the Dantale II test was developed, including target location uncertainty as well as speech masking. The first goal was to increase the ecological validity of the test by increasing the resemblance to everyday life (including an increase in SRT). The second (and ultimate) goal was to be able to reveal effects of

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.

hearing-aid processing schemes, which cannot be detected in the original implementation of the test. Only the first goal was investigated in the study, which is presented in the following.

METHOD

Participants

The study included 16 participants (7 females, 9 males), aged 46-73 years (mean: 66 years). They all had mild-to-moderate, sensorineural, gently sloping, symmetric hearing losses, and they were all experienced hearing-aid users. During testing they wore Oticon Agil miniRITE hearing aids, fitted bilaterally two weeks prior to testing according to the prescription made by the proprietary Oticon VAC rationale. Adaptive features such as noise reduction were turned off in the hearing aids during testing. The hearing-aid microphones were fixed in an omnidirectional setting.

Test conditions

Four different test conditions were included in the study. All conditions used the Danish Dantale II sentences (Wagener *et al.* 2003) as target signal. These sentences follow the matrix-sentence format. Thus, all sentences follow the form "Name verb numeral adjective object" with ten possible alternatives for each syntactic element (e.g., "Michael had ten yellow houses"). The four conditions differed in the choice of masker signals and in the location of the target signal. All conditions were implemented in a loudspeaker set-up in an anechoic chamber (see Fig. 1). Different loudspeakers were used for target (T) and masker (M) locations, as indicated in the figure. The distance between the listener and each of the loudspeakers was 1.5 m.

Three different masker signals were included: A stationary noise, 'StatNoi' (the noise generated for the original version of the Dantale II test), a speech-modulated noise, 'ICRA4' (noise modulated as a single talker, based on track 4 of the ICRA CD; Dreschler *et al.* 2001), and recordings of running speech produced by two different Danish female talkers reading excerpts from a fairytale, '2Female'. Female talkers were selected to match the gender of the target talker. All maskers were spectrally shaped according to the target, and for the two latter (speech modulated) maskers, temporal gaps (pauses) were reduced to a maximum duration of 60 ms. In all conditions, the masker signals were presented from two loudspeakers at ±30° azimuth. In case of the StatNoi and ICRA4 noise maskers, the two signals presented from the left and the right, respectively, were identical but offset in time to obtain uncorrelated signals. Since the two different female masker-talkers read the same text, a time offset was also introduced in this case to avoid noticeable overlaps in contents.

Three conditions, one with each type of masker, used a fixed-target (FT) location directly in front of the listener (0° azimuth). The last condition used a random-target (RT) approach, where sentences were presented with equal probability from one of three locations at 0° or ±45° azimuth. Only the speech maskers were used in the RT condition. The characteristics of each test condition are summarized in Table 1.

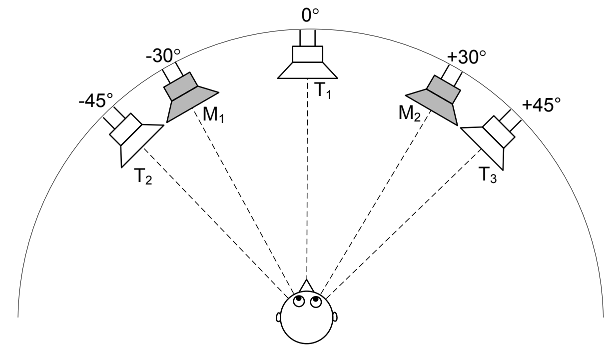

Fig. 1: Loudspeaker set-up (in anechoic chamber) used for the four test conditions. The two target loudspeakers at ±45° (T_2/T_3) are only used in the random-target condition.

Condition	**Target**	**Maskers**
FT_StatNoi	Fixed at 0° (T_1)	Unmodulated speech-shaped noise at ±30° (M_1/M_2) (Dantale II noise)
FT_ICRA4	Fixed at 0° (T_1)	Modulated speech-shaped noise at ±30° (M_1/M_2) (ICRA CD track 4, modulated as single talker, spectrally shaped as target signal, temporal gaps reduced to max. 60 ms)
FT_2Female	Fixed at 0° (T_1)	Running speech at ±30° (M_1/M_2) (Two different female speakers, spectrally shaped as target signal, temporal gaps reduced to max. 60 ms)
RT_2Female	Random at 0° (T_1) or ±45° (T_2/T_3)	Running speech at ±30° (M_1/M_2) (As above)

Table 1: Overview of the four test conditions included in the study. Target and masker positions refer to Fig. 1.

At another visit prior to the test session, all participants were trained in all four test conditions following a program, which basically resembled the actual test program. Furthermore, half of the participants had also completed the test program with another type of hearing aid.

The test session was initiated with a training round including presentation of 30 sentences (taken from three Dantale II sentence lists) in the FT_StatNoi condition. The training was followed by tests in each of the four conditions. The order of the three FT conditions was balanced across participants, while the RT condition always was the fourth and final condition in the session due to its increased complexity as compared to the FT conditions. Each condition included a short practice run consisting of 10 sentences (except for the FT_StatNoi condition where adaptation was completed in the initial training), followed by the actual test run. The test run consisted of 30 sentences for the FT conditions and 90 sentences for the RT condition (with 30 sentences being presented from each of the three directions). The Dantale II sentence lists used within the different conditions were balanced across participants.

Adaptive test procedure

The target signal was always presented at a fixed level of 65 dB SPL, while the level of the maskers, and thus, the target-to-masker ratio (TMR) was changed adaptively before each new sentence according to the number of words repeated correctly in the previous sentence. The adaptation scheme was according to that described by Hagerman and Kinnefors (1995), however, with the adaptation target being 60% (three out of five) words correct. The start TMR was +10 dB in the initial training run, while it was determined individually (based on previous performance) in the following test rounds in order to present as many sentences as possible with TMRs close to the adaptation target.

It should be noted that in the RT condition, the adaptive procedure was followed disregarding shifts in target direction. It should also be noted that the two maskers were always presented at the same level and that a 0-dB TMR corresponded to the target and each of two maskers all having the same presentation level individually.

Calculation of results

The test results are reported as the TMR corresponding to 60% words repeated correctly (TMR_{60}), in accordance with the target used during the adaptive test. The TMR_{60} was derived from a psychometric function, which was estimated for each participant and test condition (using a maximum likelihood method as described by Brand and Kollmeier 2002). In a few cases, the estimated individual psychometric function provided a poor representation of the data points, characterized by a very shallow slope and a value of TMR_{60} being out of the range of the observed TMRs. In order to avoid such obviously erroneous data, it was decided to recalculate all individual psychometric functions with all slopes set to the median slope of the original psychometric functions across participants and conditions (since median slopes for the different conditions only differed slightly from each other). Two examples of psychometric functions, before and after applying a fixed slope, are shown in Fig. 2. The left plot in the figure shows an example where the original slope is much shallower than the fixed slope but where the change of slope still only introduces a minor change in TMR_{60} (of around 1 dB). The right plot shows an example where the original slope is very close to the median slope, and where the TMR_{60} thus is almost unaffected. It should be noted that the choice of a fixed slope means that there is a simple relationship between the reported value of TMR_{60} and the SRT (i.e., the TMR at 50% words correct). Since the median slope was around

11 %/dB, SRT equals TMR_{60} minus 0.9 dB. Fig. 2 also illustrates how TMR_{60} was calculated for a given direction in the RT condition by extracting the data points originating from presentations in that direction.

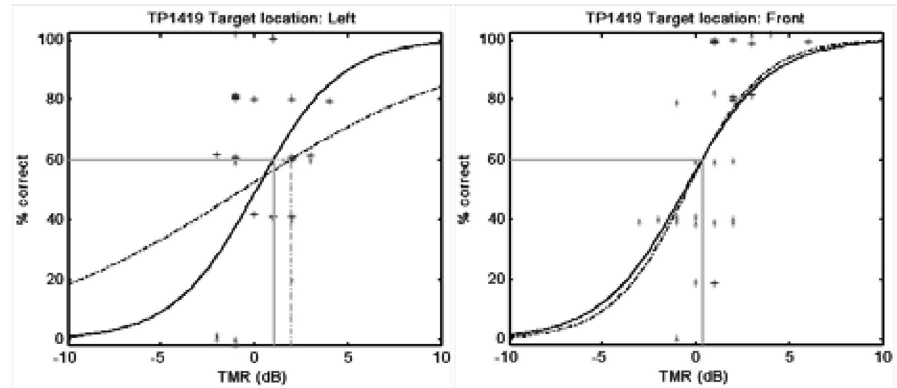

Fig. 2: Examples of data points (% words correct plotted as function of TMR) and the corresponding estimated psychometric functions from the left and front directions obtained in one run in the RT condition. The broken lines show the psychometric functions with slopes based on the actual data, while the full lines show the psychometric functions with slopes set to the median slope across participants and conditions. The result, TMR_{60}, is determined as the TMR corresponding to 60% correct. To avoid coincident data points in the plot, minor jitter along the y-axis was added.

Prior to further analysis, all TMR_{60} data were adjusted for an overall within-session training effect. The training effect was estimated from the actual data and was around 0.08 dB per 10-sentence list, which is slightly less than training effect of 0.14 dB per list reported by Wagener *et al.* (2003).

RESULTS

The mean TMR_{60} for each of the three FT conditions and for each of the three directions in the RT condition, averaged across participants, and the corresponding 95% confidence intervals are plotted in Fig. 3.

The TMR_{60} data were analysed using a repeated measures ANOVA with 'condition' as the repeated-measure factor and 'group' (dividing participants in two groups depending on whether they had been tested previously with another hearing aid) as a categorical predictor. This analysis showed a highly significant effect of condition ($F_{5,70} = 41.3$, $p < 0.000001$) but no significant effect of group ($F_{1,14} = 0.39$, $p = 0.54$) and no significant interaction ($F_{5,70} = 1.9$, $p = 0.11$). The fact that half of the participants had more experience with the test than the other half did thus not have a systematic effect on the results, indicating that the training session was sufficient to reach a stable overall performance level.

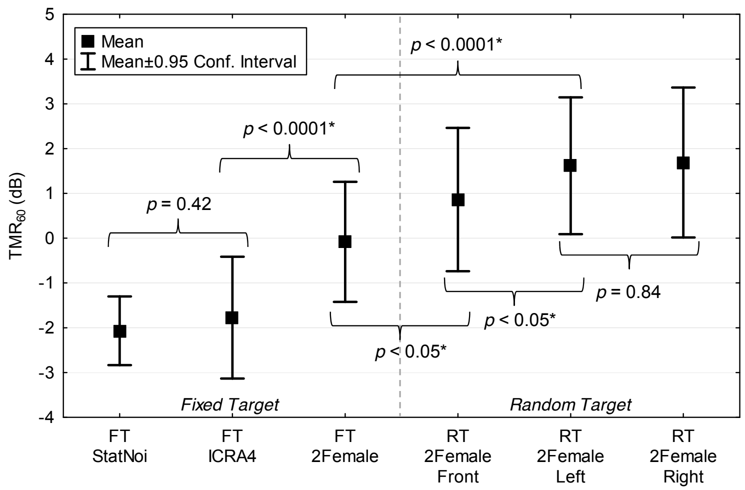

Fig. 3: Mean SRT across participants in the three FT conditions and in the three directions of the RT condition. Error bars indicate 95% confidence intervals. The figure includes selected results (level of *p*-values) from Fisher LSD post-hoc tests, assessing pairwise differences between conditions. Significant differences are indicated with asterisks.

The pairwise differences between conditions were assessed by applying Fisher LSD post-hoc testing (selected results are indicated in Fig. 3). No significant difference in mean TMR_{60} between the two FT noise conditions (StatNoi and ICRA4) was observed, whereas a significant increase (of around 1.7 dB as compared to FT_ICRA4) in TMR_{60} was observed in the FT_2Female condition. Comparing the FT_2Female condition with the RT condition showed a significant increase (of around 1 dB) in TMR_{60} in the directly comparable front direction of the RT condition, and a further (significant) increase of around 0.7 dB was observed when comparing with the left and right directions of the RT condition. There was no significant difference between TMR_{60} for the left and right directions.

DISCUSSION

The results show no significant difference between results in unmodulated (StatNoi) and modulated (ICRA4) noise. This is in line with other studies (e.g., Festen and Plomp 1990), which showed no or limited benefit of 'listening in the dips' for hearing-impaired listeners, whereas observations made on normal-hearing listeners showed a benefit in performance (i.e., a decrease in SRT) in modulated noise as compared to unmodulated noise.

The significant increase in TMR_{60} introduced by the speech maskers (2Female) as compared to the speech-modulated noise (ICRA4) can be explained by the informational masking introduced by the (comprehendible) speech maskers. Similar

effects have been observed in other studies (e.g., Francart *et al.* 2011 and Helfer and Freyman 2008). However a direct comparison of the magnitude of the effect observed in this study (around 1.7 dB) to that obtained in other studies is complicated by the fact that the effect depends highly on the exact choice of target as well as masker signals.

A significant difference of around 1 dB was observed between the values of TMR_{60} obtained from the front direction in the FT and RT conditions, respectively. This result is in line with results from other studies (e.g., Brungart and Simpson 2007, Kidd *et al.* 2005, and Singh *et al.* 2008), which showed that *a priori* knowledge about the location of the target (i.e., where to direct attention) contributes to better speech-intelligibility performance as compared to conditions where the location of the target is uncertain.

A small but significant increase in TMR_{60} of around 0.7 dB was observed for the left and right target directions as compared to the front direction in the RT condition. In a discussion of this result, it should be noted that the participants were allowed to turn their heads during the measurement, but in practice, this option was used very rarely. Most participants faced the front loudspeaker during the entire test, also when the target was presented from the side. Thus, the difference may be explained by participants directing more attention to the front direction than the two side directions. Another explaining factor could be that the benefit of access to binaural (time and level) cues in the front direction superseded a (monaural) benefit of the 'baffle effect' (a level increase of around 3 dB at the ear closest to the target), which was introduced when the target was moved from the front to the side.

After completing the measurements in all conditions many participants spontaneously commented that the RT condition reminded them much more of the problems they encountered during their daily life. These statements do, together with observed increase in SRT, indicate that the goal of increasing the ecological validity of the test actually was reached.

The data presented above do not provide evidence of whether the second goal of developing the modified test (i.e., an ability to reveal effects of hearing-aid features, which cannot be detected in the original Dantale II test) was reached. However, inspection of the individual data showed response patterns, which clearly deviated from the mean pattern shown in Fig. 3, e.g., in terms of the SRT differences between the front and left/right directions in the RT condition. This indicates that the RT condition may offer possibilities to differentiate between listeners, which are not offered by the FT conditions.

All in all, the results suggest further work with the test. Besides assessing the test's ability to detect differences between hearing-aid features, more work is needed to assess the test-retest reliability in the RT condition. Furthermore, the test could be further modified, e.g., by using other target or masker signals or by modifying the loudspeaker set-up.

CONCLUSIONS

The mean TMR_{60} (SRT) data obtained in the study showed no difference between performance in unmodulated and speech-modulated noise, indicating no benefit of dip-listening for this group of people with mild-to-moderate hearing losses.

The introduction of real speech as masking signal led to a significant increase in SRT, indicating an additional informational-masking effect. Adding target location uncertainty by randomly presenting the target from three different directions resulted in a significant increase in SRT of around 1 dB in the front direction.

The increase in SRT due to the introduction of speech masking and target location uncertainty does – together with subjective comments made by participants – indicate that the goal of increasing the ecological validity of the Dantale II test was reached.

ACKNOWLEDGEMENTS

We would like to thank our colleagues Christian Stender Simonsen and Henrik Lodberg Olsen from Oticon A/S, who provided us with the core of the test software and the recordings of the female talkers, which were used as maskers.

REFERENCES

Brand T. and Kollmeier B. (**2002**). "Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests" J. Acoust. Soc. Am., **111**, 2801-10.

Brungart, D.S. and Simpson, B. D. (**2007**) "Cocktail party listening in a dynamic multitalker environment" Percept. Psychophys., **69**, 79-91.

Dreschler, W. A., Verschuure, H., Ludvigsen, C., and Westermann S. (**2001**): "ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment" Audiology, **40**, 148-57.

Festen J. and Plomp R. (**1990**). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing" J. Acoust. Soc. Am., **88**, 1725-36.

Francart, T., van Wieringen, A., and Wouters, J. (**2011**). "Comparison of fluctuating maskers for speech recognition tests" Int. J. Audiol., **50**, 2-13.

Hagerman B. and Kinnefors C. (**1995**). "Efficient adaptive methods for measuring speech reception threshold in quiet and in noise" Scand. Audiol., **24**, 71-77.

Helfer, K. S. and Freyman, R. L. (**2008**). "Aging and Speech-on-Speech Masking" Ear Hear., **29**, 87-98.

Kidd, G., Jr., Arbogast T. L., Mason, C. R., and Gallun F. J. (**2005**). "The advantage of knowing where to listen" J. Acoust. Soc. Am., **118**, 3804-15.

Singh, G., Pichora-Fuller, M. K., and Schneider, B. A. (**2008**). "The effect of age on auditory spatial attention in conditions of real and simulated spatial separation" J. Acoust. Soc. Am., **124**, 1294-1305.

Wagener, K., Josvassen, J. L., and Ardenkjaer, R. (**2003**). "Design, optimization and evaluation of a Danish sentence test in noise" Int. J. Audiol., **42**, 10-17.

Confusion of Danish consonants in white noise

MORTEN L. JEPSEN[1] AND TORSTEN DAU[1]

[1] *Centre for Applied Hearing Research, Technical University of Denmark, DK-2800 Lyngby, Denmark*

Several classical studies have investigated how the perception of consonants is affected by interfering noise. The present study is aimed at developing a consonant confusion test, considering fifteen frequently used Danish consonants presented in initial position in combination with the vowels [æ], [i] and [u] and spoken by ten speakers. Consonant recognition was measured in the presence of a white-noise masker at a number of signal-to-noise ratios. The presented data were obtained in eight normal-hearing listeners. The results were analyzed in terms of recognition rates and formation of confusion groups. It was hypothesized that the confusion groups are similar to those found in English, but that consonants that are common in Danish would fall into specific groups. The results showed that there iss a large asymmetry in the confusions, and that the confusion groups are not as clear as in other languages. Additionally, this study introduces the "wheel of confusions" (WoC) which is a new graphical representation of consonant confusion data. The results will be used as reference data in on-going studies on consonant perception in listeners with hearing impairment and will address the individuality of consonant perception with and without hearing instruments.

INTRODUCTION

Understanding speech is one of the primary functions of human hearing. The processes underlying speech perception are sophisticated and comprise many stages that are challenging to investigate in isolation. Often, the recognition of whole sentences are considered, however, it may be of utmost importance to understand perception of smaller fragments of speech, e.g. on a phonemic level. In understanding naturally spoken sentences with meaningful context, many of the speech sounds do not necessarily have to be heard. It is therefore important for effortless speech understanding that the individual speech segments are perceived correctly, particularly for people with hearing impairment. Several studies on perception of spoken consonants (e.g., Miller and Nicely, 1955) have unveiled many interesting aspects, such as the categorization of consonants in voicing, place and manner of articulation. In recent years, studies of consonant recognition and confusions have gained popularity and have led to new insights on the relation between different spoken consonants, and new tools have been developed to analyse confusion data and the acoustic bases of consonants in relation to perception (e.g., Régnier and Allen, 2008). Recent data measured in listeners with hearing

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.