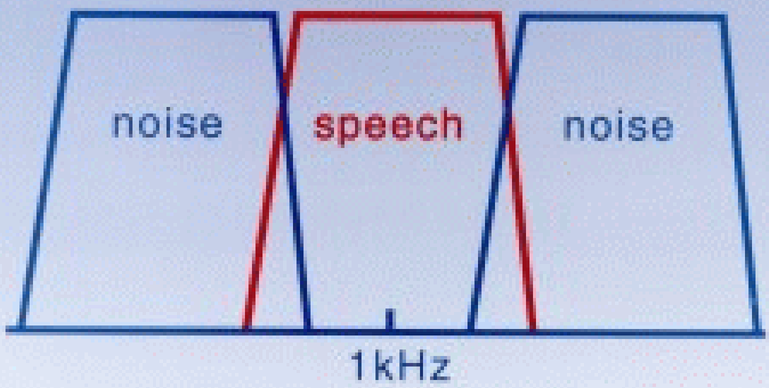
Intelligibility of narrow-band speech and its relation to auditory functions in hearing-impaired listeners

Ingrid Noordhoek



Intelligibility of narrow-band speech and its relation to auditory functions in hearing-impaired listeners

VRIJE UNIVERSITEIT

Intelligibility of narrow-band speech and its relation to auditory functions in hearing-impaired listeners

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Vrije Universiteit te Amsterdam,
op gezag van de rector magnificus
prof.dr. T. Sminia,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der geneeskunde
op woensdag 6 september 2000 om 15.45 uur
in het hoofdgebouw van de universiteit,
De Boelelaan 1105

door

Ingrid Marianne Noordhoek

geboren te Delft

Promotor: prof.dr.ir. T. Houtgast
Copromotor: dr.ir. J.M. Festen

Contents

[1]Further information on website http://www.casema.net/~astuijt/

Chapter 1

General introduction

1.1 SPEECH INTELLIGIBILITY

Hearing-impaired listeners often have problems in understanding speech, even when using a hearing aid. This produces a number of difficulties in everyday life. Hearing-impaired persons may even experience social isolation due to the problems they experience when communicating with other people. This thesis addresses the origin of reduced speech intelligibility. Section 1.1.1 describes the factors related to speech intelligibility in normal-hearing listeners. Section 1.1.2 discusses in what respect speech reception in hearing-impaired listeners differs from that of normal-hearing listeners.

1.1.1 Normal-hearing listeners

Speech communication usually takes place in the presence of ambient noise, like traffic sounds, music, or speech from competing talkers. The noise masks part of the speech, and as a result speech information is lost in the transfer from talker to listener. However, even when a large part of the speech signal is masked, intelligibility may still be acceptable for normal-hearing listeners. This shows that the speech signal contains a high level of redundancy.

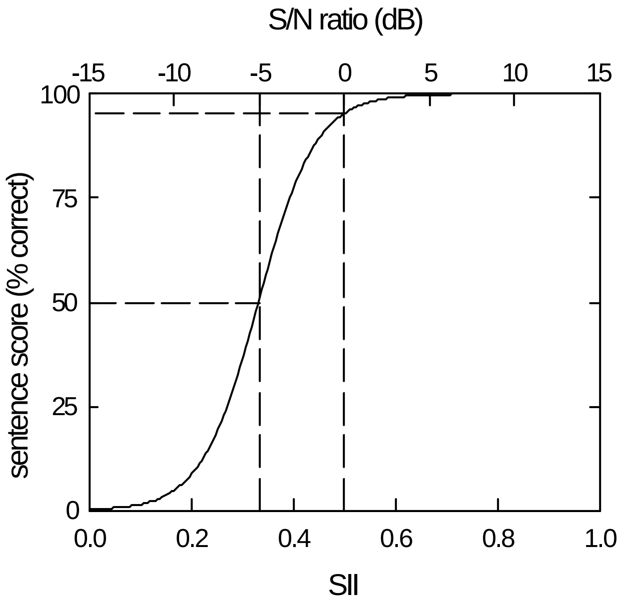


Fig. 1.1. Sentence s ng listeners, and of the signal-to-noise ratio (upper scale; for a constant S/N ratio over frequency).

The Speech Intelligibility Index (ANSI, 1997) is a physical measure designed to predict speech intelligibility for normal-hearing listeners under a variety of adverse listening conditions, such as noise masking, filtering, or an elevated hearing threshold. It is based on data obtained from many listening experiments. The SII is calculated from the speech spectrum, the noise spectrum, and the hearing threshold, and may be interpreted as the proportion of the total speech information available to the listener. When the SII is maximal (1), all speech information is available; when the SII is minimal (0), no speech information is available.

In natural speech, the instantaneous sound-pressure level fluctuates around the long-term average speech level. In the SII model, it is assumed that all speech information relevant for intelligibility is contained in a 30-dB range, symmetric around the average speech level. Thus, when for all frequencies the average speech level is 15 dB above the level of a masking noise, all speech information is available (SII=1). When the average speech signal is 15 dB below the noise level, all speech information is masked by the noise (SII=0).

By using an appropriate transfer function, the speech-intelligibility score can be predicted from the calculated SII. The form of this transfer function depends on the type of speech

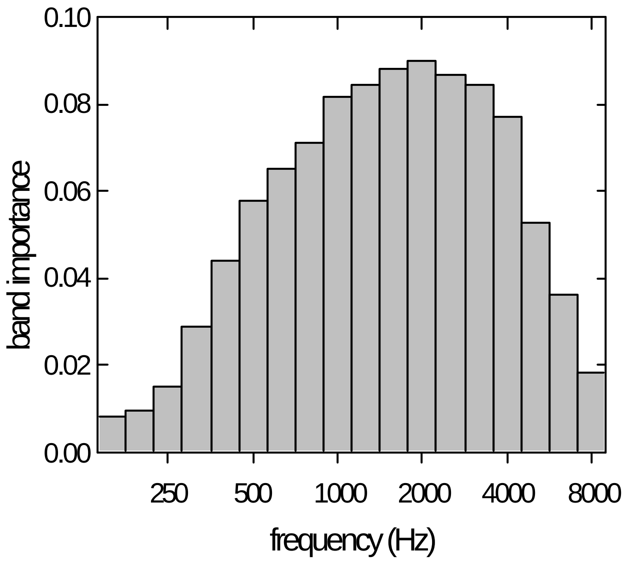


Fig dancy.

material. Figure 1.1 shows the transfer function for simple meaningful Dutch sentences. This function was measured by determining the intelligibility score for sentences as a function of signal-to-noise ratio (S/N ratio) for monaural listening, with the noise spectrum shaped according to the long-term speech spectrum (Plomp and Mimpen, 1979). The S/N ratio required for 50% sentence intelligibility is called the speech-reception threshold (SRT) in noise. For normal-hearing listeners, the SRT in noise is approximately -5 dB (i.e., the average speech level is 5 dB below the noise level). This corresponds to an SII of 0.33 (see Fig. 1.1). Thus, when 33% of the relevant speech intensity fluctuations on a decibel scale are available, 50% of the sentences are correctly understood. When speech and noise are of equal level (i.e., the S/N ratio is 0 dB), the SII is 0.5. Thus, with only 50% of the relevant intensity range available, sentence intelligibility is almost 100%.

A broad frequency region from about 125 to 8000 Hz carries the most important information for understanding speech. The information content of the speech is not distributed evenly over this frequency region. Figure 1.2 shows the band importance function, representing the contribution to intelligibility of **a**-octave frequency bands as a function of their center frequencies (Pavlovic, 1987). From this function it follows that

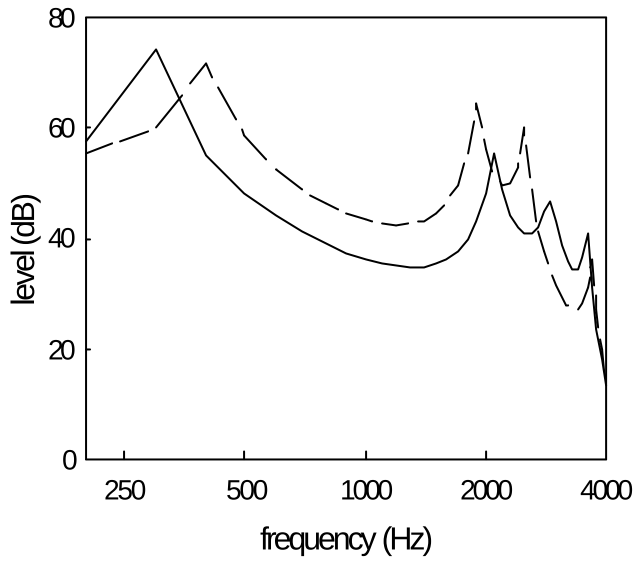


Fig. 1.3. Schematic representation of the spectral envelope of two vowels. The solid line represents the /i/, and the dashed line the /e/.

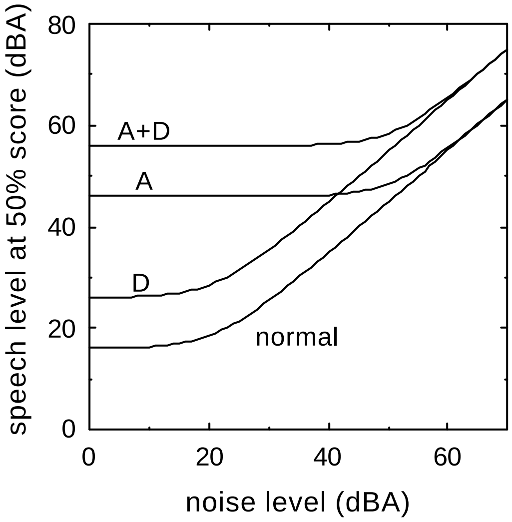
speech low-pass filtered with a cut-off frequency of 1600 Hz, as well as speech high-pass filtered with 1600 Hz, corresponds to an SII of 0.5, i.e., almost perfect intelligibility of simple sentences. So, no single part of the spectral range of speech is essential for intelligibility.

When the speech signal is audible (i.e., above the absolute threshold and not masked by noise) listeners distinguish different speech sounds on the basis of the spectral and temporal contrasts in the signal (e.g., ter Keurs, 1992; Drullman, 1994). Therefore, listeners need a sufficient auditory resolution both in frequency and in time for understanding speech. Fig 1.3 presents the envelope of the spectra of two vowels: the /i/ (solid line) and the /e/ (dashed line). This figure shows that corresponding formant peaks can lie very close to each other. This illustrates the need for a sufficient auditory spectral resolution. Since it is assumed that all normal-hearing listeners are equally able to perceive the spectral and temporal contrasts necessary for speech intelligibility, the SII model does not take a listener's auditory temporal or spectral resolution into account.

1.1.2 Hearing-impaired listeners

Hearing-impaired listeners are less sensitive for weak sounds. Therefore, for hearing-impaired listeners speech presentation levels have to be higher than for normal-hearing listeners. Amplification by a hearing aid can fully compensate a conductive hearing loss. However, amplification often does not restore sound perception to normal for listeners suffering from sensorineural hearing impairment. The reason for this is that a sensorineural hearing loss does not simply result in an *attenuation* of all sounds entering the ear, but also in a *distortion* of these sounds (Plomp, 1978, 1986).

Figure 1.4 shows the effect of attenuation (A), distortion (D), and a combination of attenuation and distortion (A+D) on the speech level required for a 50% sentence score, as a function of the noise level. The attenuation component of a sensorineural hearing loss represents the reduced sensitivity. Figure 1.4 shows that the effect of attenuation is that the listener needs higher speech levels in quiet and in conditions with low noise levels. At higher noise levels, the required speech level is not influenced by the attenuation component of the hearing loss. Therefore,



amplification using a hearing aid can compensate for attenuation. The effect of distortion is that the listener needs a higher speech to noise level, independent of the noise level itself (Fig 1.4). Therefore, amplification using a hearing aid cannot compensate for the distortion component of sensorineural hearing loss.

The distortion component of a hearing loss can be estimated by measuring the SRT in noise at high presentation levels. The amount of increase in the SRT for a hearing-impaired listeners relative to the SRT of normal-hearing listeners is related to the distortion. However, despite the high presentation level, part of the relevant 30-dB dynamic range of the speech may still fall below the hearing threshold of a hearing-impaired listener. This can happen for example when the listener has a steep frequency-dependent hearing loss. In that case, the hearing-impaired listener will have a higher-than-normal SRT in noise, because part of the speech will be inaudible. The SII model can be used to estimate the size of this so-called audibility effect, since the hearing threshold is one of the input parameters to the SII model. When a hearing-impaired listener needs the same amount of speech information as normal-hearing listeners (i.e., needs the same SII) to reach the 50% intelligibility score, this means that the origin for the higher SRT in noise is an inaudibility of a part of the speech signal. When the hearing-impaired listener needs more speech information than normal-hearing listeners (i.e., a higher SII), this indicates that the hearing-impaired listener suffers from a suprathreshold deficit that distorts the perception of speech.

A suprathreshold deficit may be either a deficit in basic auditory functions (e.g., reduced spectral or temporal resolution) or a deficit in central speech and language processing. Van Rooij and Plomp (1992) investigated the relative contribution of auditory and cognitive factors to speech perception in listeners with ages ranging from 53 to 94 years. They concluded that reduced speech perception is most likely due to deficits in auditory factors.

1.2 AIM AND EXPERIMENTAL DESIGN

The aim of this thesis is to find those auditory deficits that cause the reduced intelligibility of suprathreshold speech for hearing-impaired listeners. To investigate this issue, the relation between speech perception and basic auditory functions was investigated. Performance of hearing-impaired listeners was measured in listening experiments on speech intelligibility, and on auditory functions such as spectral resolution, temporal resolution, intensity discrimination, and frequency discrimination. Once the auditory deficits that cause reduced intelligibility can be specified, this may help designing hearing aids that improve intelligibility.

Two aspects of the present experimental approach distinguish this thesis from preceding studies that attempted to relate reduced speech intelligibility to a deterioration of specific auditory functions. First, the SII model was used to separate audibility effects from the effects of suprathreshold deficits, before relating speech intelligibility to auditory functions. Second, both speech intelligibility and auditory functions were measured in a limited frequency region, thus increasing the chance to find any clear relations.

To measure the intelligibility of speech in a limited frequency region, a novel intelligibility test (SR*B*T test) was developed. The SR*B*T test is a derivative of the well-known SRT test developed by Plomp and Mimpen (1979). SR*B*T is an acronym of speech-reception *bandwidth* threshold, and it is defined as the minimum bandwidth of speech needed to reach 50% sentence intelligibility. Complementary notched noise is added to the band-filtered speech signal, in such a way that only the frequency range within the speech passband can contribute to intelligibility.

This SR*B*T test was intended to be used with hearing-impaired listeners while the full dynamic range of the speech was above their hearing threshold. The dynamic range of most hearing-impaired listeners decreases with frequency. Therefore, the center frequency could not be chosen too high. Furthermore, the speech had to be presented in a frequency region that contributes significantly to intelligibility. Figure 1.2 shows that the frequency region that contributes most to the understanding of speech is centered around 2 kHz. Hence, as a compromise, 1 kHz was chosen as center frequency in the SR*B*T test. The frequency region around 1 kHz still contains an important portion of the speech information.

1.3 OUTLINE OF THIS THESIS

After this first introductory chapter, the thesis continues with the development of the SR*B*T test, described in chapter 2. In the SR*B*T test, the bandwidth of the speech signal is varied in an adaptive procedure. The step size to be applied in this procedure was determined in a pilot experiment. In the main experiment, the SR*B*T was measured for normal-hearing listeners at various sound-pressure levels and imposed spectral tilts. This produced a frame of reference for interpreting results of hearing-impaired listeners.

In chapter 3, the results of SR*B*T measurements with hearing-impaired listeners are presented. Additionally, thresholds using the traditional SRT test are reported for three conditions: in quiet, in noise, and while spectrally adapting speech and noise to fit the individual listeners' dynamic ranges. The origins for an elevated SRT and a broader-than-normal SR*B*T are divided into an audibility part and a suprathreshold part by applying the SII model. The suprathreshold deficits in the speech perception of each listener are quantified by the additional amount of speech information needed for an intelligibility score of 50% (as compared to normal-hearing listeners). At the end of chapter 3, the extent to which the suprathreshold deficit can be predicted from the hearing loss is investigated.

The main topic of this thesis is addressed in chapter 4, in which thresholds of hearing-impaired listeners on speech-intelligibility tests and on auditory-function tests around 1 kHz are presented. Speech intelligibility was assessed with the same tests as used in chapter 3. The measured auditory functions were detection efficiency, temporal and spectral resolution, temporal and spectral integration, and discrimination of intensity, frequency, rhythm, and spectro-temporal shape. The effect on intelligibility of increased spread of masking was investigated by modifying the SII model to include the individually measured spread of masking. Additionally, correlations between suprathreshold deficits in speech perception and auditory functions are given. The possibility of predicting suprathreshold speech perception from the measured auditory functions of the individual listeners is investigated.

In chapter 5, a general discussion is given of the most important results of this thesis. Some possibilities to improve speech intelligibility using signal processing in a hearing aid are reviewed, in the light of the suprathreshold deficits that were found detrimental for intelligibility in chapter 4. Finally, a discussion is presented on the extent to which the results on the intelligibility tests in this thesis may be considered representative for everyday speech perception.

Chapter 2

Measuring the threshold for speech reception by adaptive variation of the signal bandwidth. I. Normal-hearing listeners

An adaptive test has been developed to determine the minimum bandwidth of speech that a listener needs to reach 50% intelligibility. Measuring this speech-reception bandwidth threshold (SRBT), in addition to the more common speech-reception threshold (SRT) in noise, may be useful in investigating the factors underlying impaired suprathreshold speech perception. Speech was bandpass filtered (center frequency: 1 kHz) and complementary bandstop filtered noise was added. To obtain reference values, the SRBT was measured in 12 normal-hearing listeners at four sound-pressure levels, in combination with three overall spectral tilts. Plotting SRBT as a function of sound-pressure level, resulted in U-shaped curves. The most narrow SRBT (1.4 octave) was obtained at an A-weighted sound-pressure level of 55 dB. The required bandwidth increases with increasing level, probably due to upward spread of masking. At a lower level (40 dBA) listeners also need a broader band, because parts of the speech signal will be below threshold. The SII (speech intelligibility index) model reasonably predicts the data, although it seems to underestimate upward spread of masking.

Journal of the Acoustical Society of America 105, 2895-2902 (1999)

(http://ojps.aip.org/jasa/)

2.1 INTRODUCTION

Sensorineural hearing loss not only reduces the sensitivity for soft sounds, but often decreases the ability to understand speech presented well above the hearing threshold. The difficulties hearing-impaired listeners encounter in understanding suprathreshold speech become especially clear when speech is in some way distorted or masked. In everyday situations, ambient noise is the most frequent disturbing factor. Plomp and Mimpen (1979) developed an adaptive test for accurately measuring the speech-reception threshold (SRT) for sentences in noise. The SRT is defined as the signal-to-noise ratio required for 50% intelligibility, with the noise spectrum shaped according to the long-term speech spectrum. Hearing-impaired listeners often need higher signal-to-noise ratios than normal-hearing listeners to correctly understand 50% of the sentences. The increase of the signal-to-noise ratio required for speech understanding is denoted as hearing loss for speech in noise (Plomp, 1978).

Various efforts have been made to relate hearing loss for speech in noise to a deterioration of specific auditory functions, but they did not lead to a clear picture. Festen and Plomp (1983) investigated the relations between the pure-tone audiogram, frequency resolution, temporal resolution, and speech reception in noise for 22 sensorineural hearing-impaired listeners with moderate losses. All tests concerning frequency resolution and temporal resolution were performed at 1 kHz. Hearing loss for speech in noise was related to various measures of frequency resolution at 1 kHz (correlation coefficients from 0.49 to 0.63). No significant correlations were found between hearing loss for speech in noise and temporal resolution or the audiogram. Glasberg and Moore (1989) found a relation (R=0.56) between the SRT in noise and the audiogram for 15 subjects with moderate cochlear hearing loss, but correlations were higher with various measures of both frequency resolution and temporal resolution at 0.5, 1, and 2 kHz (correlation coefficients from 0.59 to 0.68). Van Rooij and Plomp (1990) found significant correlations between the SRT in noise and pure-tone thresholds for 72 elderly subjects. In contrast to the previous studies, frequency resolution and temporal resolution at 0.8 and 2.4 kHz could not account for an additional part of the variance. Smoorenburg (1992) showed that the pure-tone average at 2 and 4 kHz is an adequate predictor for the SRT in noise in 200 individuals with noise-induced hearing loss (R=0.72). This prediction could not be improved by taking into account two measures of frequency resolution: the critical ratio at 0.25 and 0.5 kHz for low-frequency noise (cut-off frequency of about 0.7 kHz) and the slope of its upward spread of masking.

In very general terms, one may identify two main factors underlying a listener's hearing loss for speech in noise: (1) a reduction of the effective frequency range (for instance in case of a high-frequency loss), and/or (2) a deterioration in suprathreshold sound processing. Only in the latter case, one would expect a correlation between SRT and specific auditory functions like frequency resolution and temporal resolution.

For understanding speech, a broad frequency region from about 125 to 8000 Hz is important. Therefore, correlations between the ability to understand wideband speech and auditory functions at a specific frequency may be small or absent. In order to find clear correlations between speech understanding and auditory functions at a specific frequency, it is more relevant to consider only a limited frequency range of the speech, in which the full dynamic range of the speech is above the hearing threshold.

In the present study, an adaptive test was developed that may be helpful in investigating the factors that cause reduced intelligibility of suprathreshold speech. This test determines the minimum speech bandwidth around 1 kHz required for 50% intelligibility (speech-reception *bandwidth* threshold or SR*B*T). To obtain reference SR*B*T values, the SR*B*T for normal-hearing listeners was determined.

If a hearing-impaired listener cannot reach 50% intelligibility with the same frequency region as a normal-hearing listener, then (1) this frequency region does not operate as well as for the normal-hearing listener, and (2) the hearing-impaired listener will have a wider-than-normal SR*B*T. Thus, the cause for a wider-than-normal SR*B*T must be due to changes in the auditory system in the 1-kHz frequency region. Consequently, correlations between auditory functions at 1 kHz and the SR*B*T are more obvious than correlations between auditory functions at some frequency and the wideband SRT in noise.

The SR*B*T test may be incorporated in a test battery in which, for a group of hearing-impaired listeners, several auditory functions are measured at 1 kHz. A strong correlation between a specific auditory function at 1 kHz (e.g., frequency resolution, temporal resolution, or intensity discrimination) and the SR*B*T would suggest that the deterioration of this auditory function underlies the impaired ability to understand suprathreshold speech.

The frequency region that is most important for understanding speech is centered at 2 kHz (Pavlovic, 1987). However, 1 kHz was chosen as a center frequency for the speech filter in the SR*B*T test, because this test is intended to be presented to hearing-impaired listeners with the full relevant dynamic range of the speech above their hearing threshold. The dynamic range of the speech that is relevant for intelligibility is best approximated as 30 dB (Steeneken and Houtgast, 1980). If the dynamic range (the difference between uncomfortable loudness level and hearing threshold) of a hearing-impaired listener is smaller than 30 dB in the frequency region involved in the SR*B*T test, presenting the full relevant dynamic range of the speech above hearing threshold is not possible. For many hearing-impaired listeners a small dynamic range is typically found in the high-frequency region. Therefore, a somewhat lower frequency, 1 kHz, was chosen as center frequency of the speech filter in the SR*B*T test. With this center frequency, the speech is still presented in a frequency region that contributes significantly to intelligibility. For instance, in a speech band of 1 octave, 26% of the relevant speech information is present when it is centered at 2 kHz, and 24% when it is centered at 1 kHz (Pavloic, 1987).

In the SR*B*T test, a similar adaptive procedure as with the SRT test was followed, except that now the bandwidth of the sentences was changed in the up-down procedure. Complementary bandstop-filtered noise was added to the speech to limit the frequency region contributing to intelligibility as much as possible.

The SR*B*T was measured for 12 normal-hearing listeners at four sound-pressure levels (SPLs). We also investigated the effect of imposing a spectral tilt of -6 or -12 dB/octave. This is of interest for comparing the performance of hearing-impaired listeners with a sloping audiogram to results of the normal-hearing listeners.

The SR*B*T can be expressed in SII (speech intelligibility index) values. The effectiveness of the SII model in predicting the SR*B*T of normal-hearing listeners for the various conditions was examined. If the SII model is consistent with the results of the normal-hearing listeners, the model will yield fixed intelligibility indices at the threshold. In that case, the model can be used in future experiments as a framework to predict the baseline SR*B*T of individual hearing-impaired listeners, given the speech spectrum, the noise spectrum, and the audiogram. The rationale is that the deviation of the actual performance from the SII model predictions can be used to "quantify" the degree of deterioration in suprathreshold speech processing of hearing-impaired listeners.

In the SR*B*T experiments, narrow-band speech is presented to the listeners. To investigate whether the SII at 50% intelligibility for narrow-band speech is equal to the SII at 50% intelligibility for wideband speech, the standard SRT in noise (Plomp and Mimpen, 1979) was also measured. Additionally, 50% intelligibility thresholds for two intermediate conditions were measured: the SRT for a 2½-octave band of speech, and the SR*B*T at a signal-to-noise ratio of 0 dB.

2.2 SPEECH INTELLIGIBILITY INDEX

The speech intelligibility index (SII) is a physical measure that is highly correlated with the intelligibility of speech under a variety of adverse listening conditions, such as noise masking, filtering, and reverberation (ANSI, 1997). It is a major revision of the Articulation Index (ANSI, 1969). The SII can be represented by the equation

$$SII = \sum_{i=1}^{n} I_i A_i , \qquad (2.1)$$

where n is the number of frequency bands used in the calculations. The band-importance function I_i reflects the importance of frequency band i to speech intelligibility. The band-audibility function A_i is equal to the effective proportion of the speech dynamic range within band i that contributes to intelligibility. Input variables of the SII model are the speech spectrum, the noise spectrum, and the hearing threshold.

The SII may be interpreted as a proportion of the total speech information available to the listener. When the SII is maximal (1.0), all speech information is available to the listener. The lowest value of the SII (0.0) signifies that no speech information is available.

In this study speech-reception thresholds as well as speech-reception *bandwidth* thresholds were measured. At these thresholds 50% of the sentences are intelligible for the listeners. Therefore, equal SIIs are expected at these thresholds.

2.3 METHOD

2.3.1 Materials and design

The speech material consisted of 16 lists of 13 everyday Dutch sentences (eight or nine syllables). Eight lists were pronounced by a female speaker and eight lists by a male speaker (Plomp and Mimpen, 1979; Smoorenburg, 1992). For masking, a Gaussian noise was used, shaped according to the long-term average spectrum of the sentences. The shape of the long-term average spectrum was determined separately for the two speakers. Both the speech and the noise were digitized at a sampling rate of 15,625 Hz with 16-bit resolution.

The experiment consisted of two parts. In the first part the SR*B*T was measured at various sound pressure levels and imposed spectral tilts. In the second part the combined effect of noise and bandwidth reduction was investigated.

Twelve lists of sentences were used in the first part of the experiment and four lists in the second part. The lists were presented in a fixed order. To avoid order and list effects, the presentation order of the conditions was counterbalanced over the listeners according to a Latin square, 12 x 12 for the first part and 4 x 4 for the second part of the experiment. With 12 subjects, each sequence was presented to one subject in the first part and to three subjects in the second part of the experiment.

In the SR*B*T experiments, the bandwidth of the speech signal was varied. The speech was bandpass filtered with a fixed center frequency of 1 kHz. The bandpass filters were finite impulse response filters, designed by windowing, with 256 coefficients. This allowed on-line processing of the speech.

Bandstop noise was added to the speech (Fig. 2.1) to mask speech components below and above the cut-off frequencies of the bandpass filter, and to restrict spread of excitation of the speech. Although spread of excitation in the auditory system does not provide new information, the speech could spread to a frequency region in which the integrity of the auditory system is better than in the frequency region aimed at. The bandstop noise spectrum was complementary to the speech passband. The noise was filtered off line, using finite impulse response filters with 1024 coefficients. The level and spectrum of the flanking noise were equal to the long-term average of the speech. Therefore, the SPL

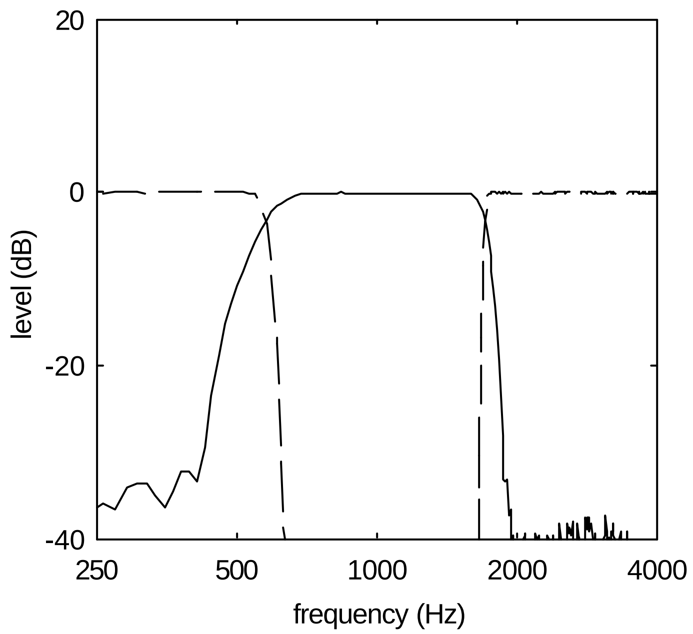


Fig. 2.1. Example of one combination of filters, used for filtering the speech (solid line) and noise (dashed lines) in the SR*B*T experiment.

of the total signal (speech plus noise) remained constant, independent of the speech bandwidth.

In the first part of the experiment the SR*B*T was measured in 12 conditions: four A-weighted SPLs (40, 55, 70, and 85 dB) each at three imposed spectral tilts (0, -6, and -12 dB/octave). The tilt was imposed on the total signal, i.e., the bandpass-filtered speech and the bandstop-filtered noise. The sound-pressure level was defined as the level of the signal before application of the spectral tilt. The spectral tilt of -6 dB/octave was obtained by low-pass filtering the signal with a cutoff frequency of 250 Hz. The filter of -12 dB/octave had a cutoff frequency of 500 Hz. After this low-pass filtering, the signal was amplified by 12 dB to obtain a fixed spectrum level at the 1- kHz center frequency of the speech band for the three spectral tilts.

In the second part of the experiment four measurements were performed at an A-weighted level of 70 dB: (1) the SR*B*T as in the first part of the experiment, (2) the SR*B*T in broadband speech-shaped noise, with a signal-to-noise ratio of 0 dB in the speech passband, (3) the SRT with a 2½-octave band of speech (center frequency: 1 kHz) in broadband speech-shaped noise, and (4) the standard SRT in noise.

2.3.2 Listeners

Twelve normal-hearing listeners, ranging in age from 19 to 34 years, participated in the experiment. The pure-tone air-conduction thresholds in the tested ear did not exceed 15 dB HL at octave frequencies from 125 to 8000 Hz.

2.3.3 Test procedure

In the SRT experiment, the level of the sentences was changed according to an adaptive procedure (Plomp and Mimpen, 1979). The first sentence of a list was presented at a signal-to-noise ratio of -8 dB. This sentence was repeated, each time at a 4-dB higher level, until the listener could correctly reproduce the sentence. The subsequent 12 sentences were then presented only once, using an up-down procedure with a step size of 2 dB. The SRT was defined as the average signal-to-noise ratio of sentences 5 to 14. Sentence 14 was not actually presented, but its signal-to-noise ratio was known from the response to sentence 13. The A-weighted level of the masking noise was fixed at 70 dB.

In the SR*B*T test, a procedure comparable with the SRT test was followed, except that now the bandwidth of the speech signal was changed adaptively. In an adaptive up-down procedure, the step size is very important. If the step size is too large, the error variance of the threshold estimate will be large. If the step size is too small, many observations are wasted in converging on the threshold (Levitt, 1971). In the SRT experiment, the slope of the intelligibility score as a function of level is about 20%/dB over the middle range (Plomp and Mimpen, 1979). This corresponds to 40% in intelligibility score per step in the adaptive procedure. To find an equivalent step size for the SR*B*T procedure, a pilot experiment was performed with 10 normal-hearing listeners in which speech intelligibility (percentage-correct score for sentences) was measured in 10 conditions of bandwidth reduction. The 10 bandwidths were equidistant on a logarithmic frequency axis. Near the 50% intelligibility score, a widening of the speech band (in Hz) by a factor 1.37 appeared to give about the same increase in intelligibility as a 2-dB increment of the signal-to-noise ratio in the SRT procedure. Hence, we chose a step size in the SR*B*T procedure corresponding to multiplication or division of the bandwidth (in Hz) by a factor 1.37.

In the SR*B*T procedure, the first sentence in a list was presented initially at a 600-Hz bandwidth. This sentence was repeated each time with the bandwidth multiplied by

$(1.37)^2$ (a double step), until the listener could correctly reproduce the sentence. The other sentences in a list were presented only once. If a sentence was repeated correctly, the bandwidth for the next sentence was divided by 1.37. If a sentence was not repeated correctly, the bandwidth for the next sentence was multiplied by 1.37. The SR*B*T was defined as the geometric mean of the bandwidth of sentences 5 to 14.

The stimuli were presented monaurally through headphones (Sony MDR-CD999) in a soundproof room. Before the experiment, a list of 13 sentences was presented, to familiarize the listeners with the procedure. During this familiarization, an SR*B*T test was performed at an A-weighted SPL of 70 dB and a spectral tilt of -6 dB/octave.

2.3.4 SII calculations

The SII was calculated following the **a**-octave band procedure of ANSI (1997). For free-field listening, the SII calculation procedure is based on the SPLs of speech and noise measured in absence of the listener at the position of the listener's head. For other listening situations, the SPLs need to be transformed to equivalent SPLs that would have been measured at the listener's position in a free field, while producing the same SPL at the eardrum of the listener as under the actual circumstances.

In this study, the SPLs of the speech and noise were measured in **a**-octave bands with the headphone (Sony MDR-CD999) on a Brüel & Kjær type 4152 artificial ear with a flat-plate coupler. These SPLs need to be transformed to equivalent free-field levels. The appropriate "artificial-ear-to-free-field transfer function" was calculated, considering that at threshold the SPL at the eardrum will be the same for both free-field and headphone measurement.

For this purpose, absolute thresholds were measured with 10 normal-hearing listeners at octave frequencies from 250 to 4000 Hz through headphones. In designing the appropriate transform, a revised free-field hearing threshold was used, because research by Killion (1978) and Berger (1981) suggests that ISO R226 (1961) is in error at low frequencies. The "artificial-ear-to-free-field transfer function" was calculated as the difference between the free-field hearing threshold (corrected by a "monaural disadvantage" of 2 dB) and the average threshold for the normal-hearing subjects listening

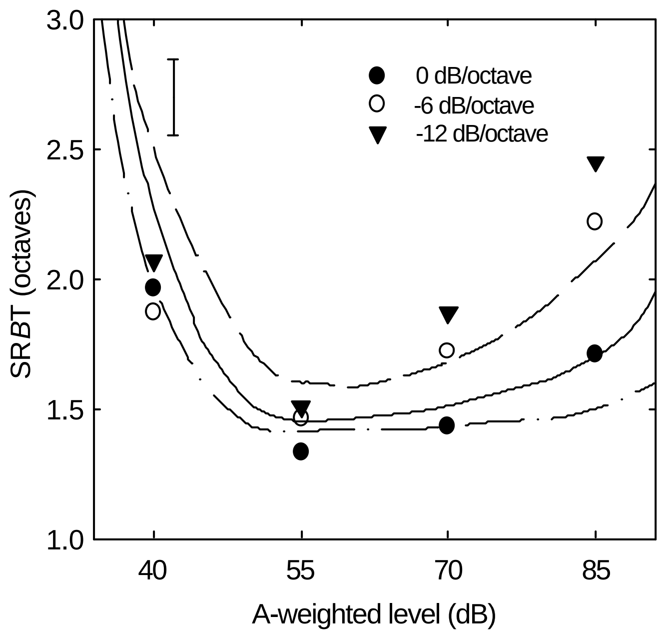


Fig. 2.2. Mean speech-reception *bandwidth* threshold for 12 normal-hearing listeners (symbols) and predictions by the SII model (curves) as a function of A-weighted sound-pressure level, with spectral tilt as a parameter. The vertical bar represents the inter-individual standard deviation. Predictions from the SII model deviate substantially from the data at 85 dBA.

through headphones. The free-field hearing threshold used in this calculation is identical to ISO (1961) at 500 Hz and above, whereas from 400 Hz downwards to 100 Hz the SPLs are between 1 and 8 dB higher.

With the "artificial-ear-to-free-field transfer function" SPLs of speech and noise, measured on the artificial ear, were transformed to equivalent free-field levels. These free-field SPLs were next converted to spectrum levels, contained within a band 1 Hz wide, as required by the SII procedure.

2.4 RESULTS

The mean SR*B*T (in octaves) for four sound-pressure levels and three spectral tilts is presented by the symbols in Fig. 2.2. The lines in this figure represent SII predictions, and will be discussed in Sec. 2.5.2. Because the inter-individual standard deviation did not significantly differ over the 12 conditions, the mean standard deviation (0.29 octave) is given only once, in the upper-left corner. The most narrow SR*B*T (1.33 octave) is obtained at 55 dBA for the spectral tilt of 0 dB/octave. Between 55 and 85 dBA, the SR*B*T increases with increasing SPL and steepening spectral tilt. The widest SR*B*T (2.45 octave) is found at 85 dBA and a tilt of -12 dB/octave. At 40 dBA the SR*B*T is on average about 0.5 octave wider than at 55 dBA.

To evaluate the effects of level and tilt, a two-factor analysis of variance for a Latin Square design with repeated measures on both factors was performed (Neter *et al.*, 1990). The effect of level was significant [$F(3,110)=52$, $p<0.001$] as was the effect of tilt [$F(2,110)=24$, $p<0.001$] and the interaction [$F(6,110)=3.8$, $p<0.01$]. Pairwise comparisons of the mean scores with the Tukey HSD test (Hays, 1988) showed that listeners need a significant wider SR*B*T than the most narrow SR*B*T (at the tilt of 0 dB/octave and 55 dBA) in eight conditions ($p<0.05$): all spectral tilts at 40 and 85 dBA, and the tilts of -6 and -12 dB/octave at 70 dBA.

In each SR*B*T measurement, 13 sentences were presented at various bandwidths around the SR*B*T. The response on each sentence (correct or incorrect) is known. From these data, the proportion of correctly repeated sentences as a function of bandwidth can be calculated. The first sentence of each list is excluded from the calculation, because this sentence was presented more than once to each listener. To calculate one psychometric function from all 1728 responses (12 sentences x 12 listeners x 12 conditions), the bandwidth of each presented sentence is converted to a relative bandwidth (*RB*) by scaling the actual bandwidth relative to the corresponding SR*B*T for that condition and listener:

$$RB = \frac{bandwidth\ (in\ Hz)}{SRBT\ (in\ Hz)} . \qquad (2.2)$$

Next, the percentage-correct score as a function of this relative bandwidth is calculated (Fig. 2.3). A logistic function is fitted to the data using the maximum-likelihood method. The obtained function can be written as

$$score = \frac{100}{RB^{-6.17}+1} . \qquad (2.3)$$

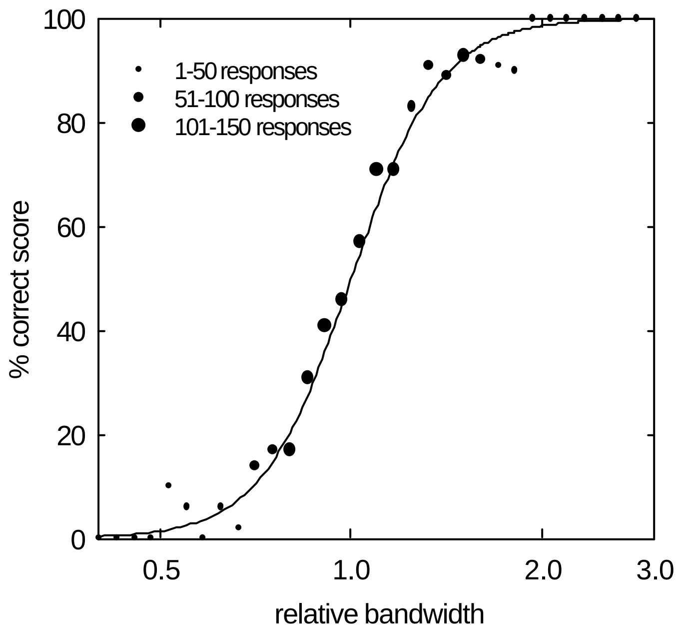


Fig. 2.3. Mean percentage-correct score for sentences as a function of the bandwidth relative to the SR*B*T, based on 1728 responses (12 sentences x 12 listeners x 12 conditions).

By definition, a relative bandwidth of 1 corresponds to 50% intelligibility. At a relative bandwidth of 0.7, only 10% of the sentences are reproduced correctly by the listeners. At a relative bandwidth of 1.4, the number of correctly repeated sentences reaches almost 90%.

Over the middle range, the slope of the psychometric function is about 40% in intelligibility score per 2-dB step in the SRT procedure. In the steepest section of Fig. 2.3, the chosen step size (multiplication of the bandwidth by 1.37) appears to result in a slightly steeper slope of about 50% in intelligibility score per step.

The means and standard deviations of the two SR*B*T and two SRT values, which were measured in the second part of the experiment are presented in Table 2.1. These conditions represent combinations of noise addition and bandwidth reduction that all lead to a 50% intelligibility. As expected, as the signal-to-noise ratio decreases listeners need a broader band of speech to understand 50% of the sentences.

The SR*B*T at the level of 70 dBA was measured in both parts of the experiment, with the same listeners. The average SR*B*T was 1.43 octave in the first part of the experiment and

Table 2.1. Four combinations of noise addition and bandwidth reduction leading to 50% intelligibility, and the corresponding modified speech intelligibility indices. Standard deviations are given between parentheses.

Condition	SRT (dB S/N ratio)	SR*B*T (octaves)	SII
No noise		1.41 (0.15)	0.319 (0.052)
2½-octave band	4.2 (2.2)		0.366 (0.054)
0 dB S/N ratio		3.39 (0.50)	0.341 (0.037)
No band-limiting	-3.8 (1.1)		0.324 (0.036)

1.41 octave in the second part. A *t* test for matched samples revealed no significant learning effect ($p<0.05$). Therefore, the standard deviation of individual SR*B*T values could be calculated. This standard deviation, which represents a reliability measure of the test, is 0.15 octave.

2.5 DISCUSSION

2.5.1 Effect of level and tilt on the SR*B*T

The SR*B*T appears to vary considerably as a function of presentation level and spectral tilt and is, in this respect, far less robust than the SRT in noise. Qualitatively, this may be understood as a result of the spectral configuration of the stimulus in the SR*B*T-experiment, i.e., a central spectral region of (physically) unmasked speech flanked by two noise bands. This configuration is vulnerable for the effect of upward spread of masking, which is expected to increase with level and negative spectral tilt. The increased SR*B*T for the 40-dBA condition probably reflects that the lower part of the relevant dynamic range of speech in the unmasked central stimulus region already falls below the hearing threshold. The stimulus configuration in the SRT test, with a low speech-to-noise ratio over the whole spectral range, is much less sensitive to these effects.

Usually, for the SRT in noise a threshold is obtained which is independent of sound-pressure level, but Smoorenburg (1992) reported that the SRT as a function of noise level increases 1.4 dB between 50 and 80 dBA for normal-hearing listeners. Probably, this was also caused by upward spread of masking. Van Dijkhuizen *et al.* (1987) investigated the effect of various spectral tilts on the SRT in noise, at an SPL of 80 dB. They concluded that spectral tilts from about -7 up to +10 dB/octave do not affect the SRT in noise. At the spectral tilt of -12 dB/octave the SRT was about 6 dB higher than without tilt. This was partially explained by upward spread of masking.

2.5.2 Predictions by the SII model

For a more quantitative approach to interpret the present data, the SII calculation model was applied, which includes the effects of upward spread of masking and speech audibility. The SII was calculated for the 12 SR*B*T values measured in the first part of the experiment. At these speech-reception *bandwidth* thresholds, 50% of the sentences are intelligible for the listeners. Consequently, equal speech intelligibility indices are expected at these thresholds.

The curves in Fig. 2.2 connect SR*B*T values with an SII of 0.33. This value corresponds to the SII for the average SR*B*T in the 70-dBA condition without spectral tilt. This SII was used as a reference for the 11 other SR*B*T values, because in the 70-dBA condition without spectral tilt all speech is well above hearing threshold and the SR*B*T is not wider due to upward spread of masking. Thus, the SII in the 70-dBA condition without spectral tilt is essentially insensitive for small errors in the input variables of the SII model (spectrum levels of speech and noise, and hearing threshold).

The observed increase in SR*B*T at 40 dBA is consistent with the predictions by the SII model. The predicted SR*B*T increases as the presentation level decreases below 55 dBA, because part of the 30-dB dynamic range of the speech falls below hearing threshold. For levels between 55 and 85 dB, both observed and SII-predicted SR*B*T increase with increasing SPL and steepening spectral tilt. The increase in the predicted SR*B*T is due to the estimated effect of upward spread of masking. At 85 dBA, the model also predicts some level distortion (i.e., the decrease of intelligibility because of a high presentation level). However, the amount of distortion is so small, that the influence on the predicted SR*B*T is negligible.

The observed SR*B*T increases faster than the predicted one between 55 and 85 dBA. This suggests that the SII model underestimates the effect of upward spread of masking of the low-frequency noise band. The exact way in which the upward spread of masking is accounted for by the model greatly influences the calculated SII for the present conditions. This holds especially for those conditions with the high SPLs and the negative spectral tilt. The SPL of the low-frequency noise was higher than the SPL of the speech in the frequency region around 1 kHz, especially in the conditions with a spectral tilt of -6 or -12 dB/octave (but also without spectral tilt), because the noise was obtained by bandstop filtering the speech-shaped noise. Furthermore, no external noise is present in the frequency region around 1 kHz, so when the upward spread of masking reaches the lower limit of the dynamic range of the speech, it already starts to affect speech intelligibility.

A possible explanation for the discrepancy between our data and the predictions by the SII model is that the slope of the masking curve in the SII model may be too steep for low-frequency maskers. In the SII model, the slope of the upward spread of masking in dB/octave (C_i) due to frequency band *i* is calculated as

$$C_i = -80 + 0.6 \cdot BL_i \,, \tag{2.4}$$

where BL_i is the SPL of frequency band *i*. This procedure is based on the **a** -octave spread of masking protocol discussed in Ludvigsen (1985), who deduced this relation from the masking curves of Zwicker (1963) with masker frequencies of 650, 1000 and 2000 Hz. In the SII model, the slope of the masking curve does not depend on masker frequency. However, below 500 Hz the relative bandwidth of the auditory filter increases with decreasing center frequency (Moore, 1986). Thus, for frequencies below 500 Hz, the slope of the masking curve may be expected to decrease with decreasing center frequency. This effect is not accounted for by the SII model.

2.5.3 Modified SII model

The present SII model cannot accurately predict the SR*B*T of the normal-hearing listeners in the various conditions of this experiment, making it unfit to predict the SR*B*T in future experiments with hearing-impaired listeners. We tried to modify the SII model in order to find a better correspondence with the SR*B*T measurements.

In the AI model (ANSI , 1969) the calculation of the upward spread of masking is frequency dependent. Therefore, the masking spectrum was recalculated using the spread-of-masking algorithm from the old ANSI S3.5 standard.

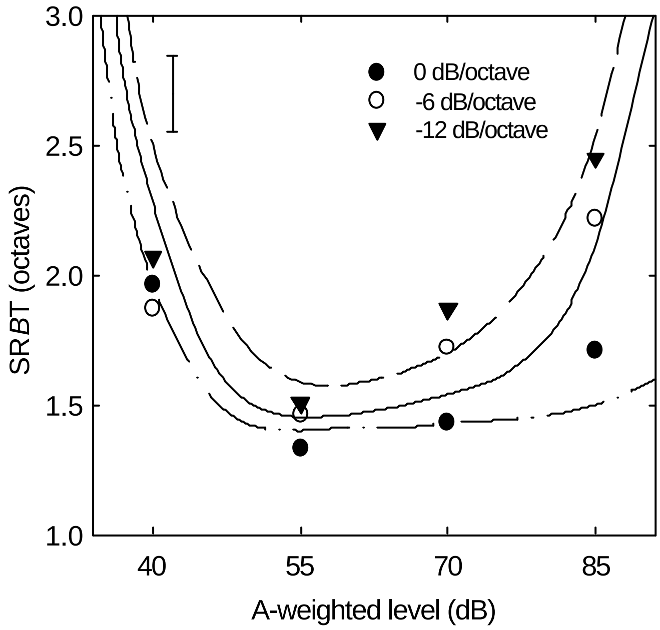


Fig. 2.4. Mean speech-reception *bandwidth* threshold for 12 normal-hearing listeners (symbols) and predictions by a modified SII model (curves) as a function of A-weighted sound-pressure level, with spectral tilt as a parameter. The vertical bar represents the inter-individual standard deviation. For all presentation levels the SII predictions are in good agreement with the data.

However, this did not lead to better predictions. At 70 dBA, the observed increase in SR*B*T when the spectral tilt becomes more negative is not predicted by the SII model with the old ANSI S3.5 spread-of-masking algorithm, while this increase is predicted by the unmodified SII model. For the negative spectral tilts at 85 dBA, the SII model with the old ANSI S3.5 spread-of-masking algorithm predicts SR*B*T values that are about the same amount too high as the predictions by the unmodified SII model are too low.

In a second attempt to obtain a better correspondence between the 12 SR*B*T measurements in the first part of the experiments and the SII model, shallower slopes of the masking curves were used for frequencies below 500 Hz. The slope of the masking curve (in dB/octave) was chosen to decrease linearly with center frequency below 500 Hz. Thus, for **a**-octave bands with masker frequencies greater or equal to 500 Hz (i.e., *i*$6), the slope of the masking curve was calculated according to Eq. (2.4) and for **a**-octave bands with center frequencies below 500 Hz (i.e., *i*<6) according to

$$C_i = [1 - (6 - i)\cdot\alpha]\cdot[-80 + 0.6\cdot BL_i] \; . \qquad (2.5)$$

The parameter α was varied on a post-hoc basis. The differences between the 12 SIIs were minimized for $\alpha = 0.08$.

Only rarely has the masking pattern of narrow noise bands below 500 Hz been measured. Maiwald (1967) measured the masking pattern of a narrow-band noise with a center frequency of 250 Hz and an SPL of 60 dB. The slope of the masking curve was about -30 dB/octave. Combination of Eq. (2.5) with $i = 3$ and $\alpha = 0.08$ results in a slope of -33 dB/octave, so the value of 0.08 seems reasonable.

With the inclusion of Eq. (2.5) in the SII model, the differences among the 12 SIIs were decreased. Figure 2.4 again shows the SR*B*T measurements (symbols), but now with the predictions by the modified SII model (lines). The equal-SII contours (SII of 0.33) in Fig. 2.4 correspond better to the SR*B*T measurements than the curves in Fig. 2.2. The 40-dBA conditions with a spectral tilt of -6 or -12 dB/octave provided the only significant differences between measured and calculated SR*B*T (*t* test, $p<0.05$). The cause may be a small deviation of the SPL or hearing threshold in the calculations from its actual value. The calculated SR*B*T increased very fast with decreasing SPL in the region around 40 dBA. Consequently, a small deviation of the SPL or hearing threshold results in a large change in SR*B*T.

The four conditions in Table 2.1 are combinations of noise addition and bandwidth reduction that all lead to a 50% intelligibility. Therefore, equal speech intelligibility indices are expected for these four conditions. The exact way in which the upward spread of masking is calculated in the SII model had a negligible effect on the SII for these four conditions. The average and standard deviation of the individual modified SII, is given in the last column of Table 2.1. An analysis of variance for a Latin Square design with repeated measures showed a significant effect of condition on the speech intelligibility index [$F(3,30)=4.55$, $p<0.05$]. Pairwise comparisons of the mean scores with the Tukey HSD test (Hays, 1988) showed that the SII in the 2½-octave-band condition was significantly higher ($p<0.05$) than in the no-noise and no-bandlimiting conditions. So far, we have no explanation for this finding.

2.5.4 Reference SR*B*T values

The SR*B*T test is meant as a tool for studying impairment of suprathreshold speech processing in hearing-impaired listeners. The aim of this experiment was to obtain reference SR*B*T values as a framework for interpreting future results with hearing-impaired listeners. It appeared that for normal-hearing listeners, four conditions in our experiment were optimal for the SR*B*T measurement (i.e., no influence of upward spread of masking or the hearing threshold): the three spectral tilts at the SPL of 55 dBA, and no spectral tilt at the SPL of 70 dBA. In these conditions, the SR*B*T is about 1.4 octave (i.e., a frequency band from 600 Hz to 1600 Hz).

Various frequency bands contribute different amounts to intelligibility. Therefore, the bandwidth required to understand 50% of the speech is frequency dependent. As mentioned in Sec. 2.1, the frequency region that is most important for understanding speech is centered at 2 kHz. Thus, if 2 kHz had been chosen as center frequency, probably a smaller SR*B*T would have been measured. According to the SII model, the importance of the frequency region from 600 to 1600 Hz, relative to the total frequency region relevant for speech intelligibility, is 33%. A frequency band centered at 2 kHz, with the same relative importance has a width of 1.2 octave (1.3 kHz - 3.0 kHz). Therefore, if the experiment had been carried out with a center frequency of 2 kHz, instead of 1 kHz, presumably an SR*B*T of about 1.2 octave, instead of 1.4 octave, would have been measured.

If the SR*B*T is measured in a condition in which there is essentially no influence of upward spread of masking or the hearing threshold, and 1 kHz is chosen as center frequency, the appropriate reference value is 1.4 octave. Reference values for the other conditions reflect the effects of upward spread of masking and/or of proportions of the speech signal falling below the hearing threshold. For these conditions, results for normal-hearing listeners are consistent with the modified SII model.

2.6 CONCLUSIONS

An adaptive test has been developed for measuring the minimum bandwidth of speech, with a center frequency of 1 kHz, required for 50% intelligibility (speech-reception *bandwidth* threshold or SR*B*T). The performance of normal-hearing listeners on the SR*B*T test was measured. Normal-hearing listeners need a speech band of about 1.4 octave (i.e., 600-1600 Hz) to understand everyday sentences at an A-weighted SPL of 55 dB. The required minimum bandwidth increases with increasing level, due to upward spread of masking. At lower SPLs, the listeners also need a broader bandwidth to understand the sentences, because parts of the speech will fall below the hearing threshold.

The SII model appears to underestimate the amount of upward spread of masking produced by the low-frequency noise in the SR*B*T measurements. Using shallower slopes of the masking curves for frequencies below 500 Hz, a better correspondence between he measurements and the SII model was obtained.

Chapter 3

Measuring the threshold for speech reception by adaptive variation of the signal bandwidth. II. Hearing-impaired listeners

*In chapter 2, an adaptive test was developed to determine the speech-reception bandwidth threshold (SRBT), i.e., the width of a speech band around 1 kHz required for a 50% intelligibility score. In this test, the band-filtered speech is presented in complementary bandstop-filtered noise. In the present study, the performance of 34 hearing-impaired listeners was measured on this SR**B**T test and on more common SRT (speech-reception threshold) tests, namely the SRT in quiet, the standard SRT in noise (standard speech spectrum), and the spectrally adapted SRT in noise (fitted to the individual's dynamic range). The aim was to investigate to what extent the performance on these tests could be explained simply from audibility, as estimated with the SII (speech intelligibility index) model, or require the assumption of suprathreshold deficits. For most listeners, an elevated SRT in quiet or an elevated standard SRT in noise could be explained on the basis of audibility. For the spectrally adapted SRT in noise, and especially for the SRBT, the data of most listeners could not be explained from audibility, suggesting that the effects of suprathreshold deficits may be present. Possibly, such a deficit is an increased downward spread of masking.*

Journal of the Acoustical Society of America 107, 1685-1696 (2000)
(http://ojps.aip.org/jasa/)

3.1 INTRODUCTION

A common complaint of listeners suffering from sensorineural hearing impairment is that they experience great difficulty in understanding speech in ambient noise. According to some studies, this difficulty is only caused by the fact that part of the speech spectrum is below the absolute threshold (Zurek and Delhorne, 1987; Lee and Humes, 1993). Other studies suggest that suprathreshold deficits (i.e., deficits that show up in a suprathreshold stimulus condition) are also involved (Glasberg and Moore, 1989; Dreschler and Plomp, 1985). In a review, Moore (1996) concluded that, for hearing losses up to about 45 dB, inaudibility of part of the speech spectrum is the dominant source of the difficulty in understanding speech, whereas for greater losses suprathreshold deficits start to play a role.

Aspects of sound perception that may be affected are spectral resolution, temporal resolution, frequency discrimination, and loudness perception. It has proved difficult to relate the reduced ability to understand speech in noise to specific suprathreshold deficits (Moore, 1996). A reason for this may be that correlations between the ability to understand broadband speech and auditory functions at a specific frequency were studied. The investigation of the factors underlying a speech-processing deficit may be simplified by restricting the research to a limited frequency region. To find clear correlations with auditory functions at a specific frequency, considering intelligibility for narrow-band speech seems more relevant.

For that purpose, the SRBT test has been developed (chapter 2). In this test, the bandwidth of speech around 1 kHz required for a 50% intelligibility score is determined (speech-reception bandwidth threshold or SRBT). The narrow-band speech is presented in complementary bandstop-filtered noise to ensure that the speech was only audible within the desired frequency band. A procedure comparable to the SRT test in noise (Plomp and Mimpen, 1979) is followed with the difference that the bandwidth of the speech signal, not the signal-to-noise ratio, is changed adaptively.

The SRBT of normal-hearing listeners is 1.4 octave under optimal conditions, i.e., when the entire speech dynamic range is above the hearing threshold, but not so loud that audibility is affected by excessive upward spread of masking. Provided that the full dynamic range of speech is above threshold, it is plausible to assume that a broader-than-normal SR*B*T points to a deterioration in sound processing in the 1-kHz frequency region. The SR*B*T test is meant as a research tool to select hearing-impaired listeners suffering from a deficit in speech processing in the 1-kHz frequency region, and may be incorporated in correlation studies of auditory functions at 1 kHz and speech perception.

Although the stimulus in the SR*B*T experiment is an artificial signal, it is not remote from everyday listening situations. In practice, it often occurs that part of the speech spectrum is masked by ambient noise with a different spectral content, like the sound from traffic, domestic equipment, or music. However, it should be clear that the SR*B*T is *not* designed as a measure for the speech communication ability of hearing-impaired listeners in real life. The human voice is probably the most common source of ambient noise. Therefore, the SRT test for broadband speech in noise, in which the noise spectrum is shaped according to the long-term average speech spectrum, is a more appropriate measure for everyday speech perception.

An alternative measure for a listener's ability to understand *narrow-band* speech is the SRT for bandpass-filtered speech in noise. However, when noise is added to the speech passband, the bandwidth of the speech must be much broader than the SR*B*T in order to allow the listeners to reach a 50% intelligibility score. For a broader speech band around 1 kHz, the correlation between intelligibility and auditory functions at 1 kHz is less obvious. Therefore, the SR*B*T test is preferred to the SRT test for bandpass-filtered speech in noise, when the experimental goal is to find correlations between a listener's ability to understand narrow-band speech and performance on psychoacoustic tests at 1 kHz.

In the present study, 34 hearing-impaired listeners and 10 normal-hearing listeners performed the new SR***B***T test and more common SRT tests. First, the SRT in quiet was measured. Next, the SRT in steady-state speech-shaped noise was measured with the noise fixed at 20 dB above each listener's SRT in quiet. Since even at these levels, it is possible that part of the speech spectrum falls below the hearing threshold, the SRT in noise was also measured with the speech and noise spectra shaped to fit the midline of the dynamic range of each individual listener. Last, the SR*B*T was determined using the same spectral shaping.

The first aim of this paper is to examine to what extent the SR*B*T of the hearing-impaired listeners differs from the SR*B*T of normal-hearing listeners. This is important, because the SR*B*T test is intended to discriminate between listeners with normal and reduced speech processing in the 1-kHz frequency region. The second aim is to investigate whether an elevated SRT or broader-than-normal SR*B*T can be explained simply within the audibility concept (losing part of the full dynamic range of the speech), or require the assumption of suprathreshold deficits. This is of significance for the audiology practice, because when the speech-understanding problem of a hearing-impaired listener is caused by a suprathreshold deficit, intelligibility cannot be restored completely by amplification of the speech signal with a hearing aid.

In this study, audibility is defined as the effective proportion of the speech dynamic range contributing to intelligibility as calculated with the speech intelligibility index, or SII (ANSI, 1997). The SII replaces the older articulation index (ANSI, 1969), and is calculated from the speech and noise spectra, and the hearing threshold. The SII model includes procedures for computing the effect on audibility of self-masking of speech, upward spread of masking, and level distortion (i.e., the decrease of speech intelligibility at high presentation levels), *for normal-hearing listeners*.

A speech-processing deficit is defined as a suprathreshold effect, not included in the SII model, that reduces intelligibility. For example, speech intelligibility can be reduced by spectral spread of masking. If the spread of masking experienced by a hearing-impaired listener in a speech intelligibility test is the same as that experienced by a normal-hearing listener at the same absolute level, this is considered an audibility effect. If the hearing-impaired listener experiences excessive spread of masking, the extent to which the spread of masking exceeds the normal spread of masking is regarded to be a speech-processing deficit.

Thus, if the performance of a hearing-impaired listener on speech intelligibility tests is consistent with the SII model, it is assumed that this listener does not suffer from a suprathreshold speech-processing deficit, and that a possible abnormal SRT or SR*B*T is due only to inaudibility of a part of the speech spectrum. If, on the other hand, performance is worse than predicted by the SII model, it is assumed that this is caused by a speech processing deficit.

3.2 METHOD

3.2.1 Materials and design

The speech material consisted of eight lists of 13 meaningful everyday Dutch sentences (eight or nine syllables), uttered by a male speaker (Smoorenburg, 1992). For masking, a Gaussian noise was used, shaped according to the long-term average spectrum of the sentences. Both the speech and the noise were digitized at a sampling frequency of 15625 Hz with 16-bit resolution.

Signals were generated by a personal computer using TDT (Tucker-Davis Technologies) System II hardware. Speech and noise were upsampled by a factor of 2, and were each delivered through a 16-bit D/A converter (TDT DD1) at a 31250-Hz sampling frequency and low-pass filtered at 16 kHz (TDT FT5). Next, speech and noise were attenuated separately (TDT PA4), and subsequently summed (TDT SM3). The total signal was sent through a programmable filter (TDT PF1), used for frequency shaping. If necessary, the signal was passed through an amplifier.

In this study, our main interest is to compare the results of individual hearing-impaired listeners to the results of the normal-hearing listeners. Therefore, differences among listeners due to order and lists effects were avoided by presenting the lists of sentences and the intelligibility tests in a fixed order. With the eight lists of sentences, two similar blocks (test-retest) of four intelligibility tests were performed. A block consisted of three SRT tests, followed by one SR*B*T test.

The first SRT test was performed in quiet (SRTq), with the original spectrum (i.e., with the programmable filter in bypass mode). The second SRT test was performed in noise (SRTn), with the unmodified spectra for speech and noise. The level of the masking noise was 20 dB above the measured SRTq of each listener. The third SRT test was also performed in noise, with the difference that the speech and noise signals were adapted to the dynamic range of individual listeners (SRTa). The noise spectrum was shaped to fit halfway between hearing threshold and broadband uncomfortable loudness level (see Sec. 3.2.3) for frequencies between 250 and 4000 Hz. The shape of the speech spectrum was equal to the shape of the noise spectrum. The signal-to-noise ratio in the adaptive procedure was varied by adjusting the level of the speech. To obtain the desired spectra, the speech and noise signals were filtered, using a finite impulse response (FIR) filter with 160 coefficients, implemented on the TDT PF1-hardware. Below 250 Hz and above 4000 Hz, the frequency response of the filter was flat.

In the SR*B*T procedure, the bandwidth of the speech signal was varied. The speech was bandpass filtered on line with a fixed-center frequency of 1 kHz, using software filters with 256 coefficients. Complementary bandstop noise was added to the speech signal. The noise was filtered off-line, using software filters with 1024 coefficients. Both the (bandpass-filtered) speech and (bandstop-filtered) noise spectra were shaped to fit halfway the dynamic range of the individual listener, using the same FIR filter, implemented on the TDT PF1-hardware as in the preceding SRTa test.

3.2.2 Listeners

Thirty-four hearing-impaired persons were selected from the files of the University Hospital VU and served as subjects. Information about individual listeners is given in Table A.1 of Appendix A. Their age ranged from 35 to 88 years, with a mean of 64 years. Their native language was Dutch. Only listeners who could reach an intelligibility score for monosyllabic words in quiet of at least 75% were selected. Four listeners had a mixed hearing loss and 30 listeners suffered from sensorineural hearing loss in both ears. Mean, standard deviation, and range of pure-tone air-conduction thresholds at the ear under test are given in Table 3.1. For each listener, the threshold was at least 40 dB HL at one or more frequencies. Ten normal-hearing listeners, ranging in age from 18 to 33 years, participated in the experiment as a reference group.

Table 3.1. Mean pure-tone air-conduction thresholds at octave frequencies in the ear under test of the 34 hearing-impaired listeners. Also given are the standard deviation and the range of thresholds at each frequency.

Threshold	Frequency (Hz)					
(dB HL)	250	500	1000	2000	4000	8000
Mean	29	35	43	49	65	82
s.d.	19	20	17	16	17	22
Minimum	0	-5	0	5	15	15
Maximum	75	75	70	75	100	120

3.2.3 Procedure

Listeners were seated in a soundproof room. The stimuli were presented monaurally through headphones (Sony MDR-CD999). Generally, the ear with the lowest word intelligibility score in quiet was tested. Only if the risk of overhearing existed or if the listener could not reach a 75% intelligibility score in quiet with this ear, the other ear was tested.

The experiment consisted of two parts. In the first part of the experiment, the dynamic range was determined for each listener by measuring hearing threshold and uncomfortable loudness level (UCL). In the second part of the experiment, the ability of the listener to understand speech was measured with SRT tests in quiet and in noise, and with the SR*B*T test.

Dynamic range

Hearing threshold and UCL were measured with 1/3-octave bands of noise at five center frequencies: 250, 500, 1000, 2000, and 4000 Hz. Levels for intermediate 1/3-octave bands were calculated by interpolation. The maximum presentation level was 80 dB SPL in the determination of the hearing threshold and 134 dB SPL in the UCL measurement. If the hearing threshold was higher than 80 dB SPL or the UCL was higher than 134 dB SPL, the levels of 80 or 134 dB SPL were used in the calculation of the midline of the dynamic range.

For the determination of the hearing threshold, a Békésy tracking procedure was used with a step size of 1 dB. The listener was asked to push a button as long as he or she could hear a pulsating noise burst (duration: 300 ms; repetition frequency: 2.4 Hz). As long as the button was pushed, the level of the noise bursts decreased. When the button was released, the level started to increase. The threshold measurement was finished after 11 reversals. The threshold was defined as the average of all but the first reversal levels.

Uncomfortable loudness levels were determined in two steps. First, the UCL of the five 1/3-octave bands of noise was measured individually. Next, the UCL was measured with a broadband noise burst. For determination of the 1/3-octave band UCL, the listener had to push a button as soon as a pulsating noise burst with increasing level (length: 300 ms; repetition frequency: 1.7 Hz) was experienced as uncomfortably loud. The level of the noise burst increased in steps of 3 dB. When the listener pushed the button, the level of the noise burst decreased by a random amount between 21 and 31 dB. The measurement was finished after six responses. The UCL of each band was computed by averaging across the levels at which the button was pushed.

The broadband noise burst had a duration of 4 s and was spectrally shaped according to the narrow-band UCLs. To determine the UCL of the broadband noise burst, the noise burst was generated at gradually increasing levels. After each presentation the listener was asked whether the signal was experienced as uncomfortably loud. If so, this level was considered as the broadband UCL and the corresponding level in each of the five 1/3-octave bands was taken as the new UCL of that band. The broadband UCL was measured, because in the second part of the experiment a broadband masking noise was used, and since the relationship between narrow-band and broadband UCLs varies across listeners (Walker *et al.*, 1984).

Speech intelligibility

In the SRT tests, the level of the sentences was changed according to an adaptive procedure (Plomp and Mimpen, 1979). The first sentence of a list was repeated, each time at a 4-dB higher level, until the listener could correctly reproduce the sentence. The subsequent 12 sentences were presented only once, using an up-down procedure with a step size of 2 dB.

The SRTq was defined as the average A-weighted sound-pressure level of sentences 5 to 14. Sentence 14 was not actually presented, but its “would-be” level was included in the calculation of the SRT, to use the information provided by the response to sentence 13. The SRTn and SRTa were defined as the average signal-to-noise ratio (in dB) for sentences 5 to 14.

In the SR*B*T test, the first sentence in a list was presented at a 600-Hz bandwidth, much narrower than the bandwidth required for a 50% intelligibility score. This sentence was repeated each time with the bandwidth multiplied by $(1.37)^2$ (a double step), until the listener correctly reproduced the sentence. The other sentences in the list were presented only once. When a sentence was repeated correctly, the bandwidth (in Hz) for the next sentence was divided by 1.37. When the repeated sentence was incorrect, the bandwidth (in Hz) for the next sentence was multiplied by 1.37. This implies that the step size in the

SR*B*T procedure was constant on a logarithmic frequency axis. The SR*B*T was defined as the geometric mean of the bandwidth of sentences 5 to 14.
Prior to the speech intelligibility experiment, two lists of 13 sentences were presented, to familiarize the listeners with the procedure. With these lists, an SRTq test and an SRTn test were performed.

3.3 SII CALCULATIONS

As a consequence of our effort to keep the full dynamic range of the speech audible, presentation levels were different for each listener. This, in turn, may cause differences in audibility, for example because at a higher level the bandstop noise in the SR*B*T experiment produces more spectral spread of masking. The effect of level on audibility is evaluated with the speech intelligibility index (SII).
The SII for a given condition is calculated from the speech and noise spectra, and the hearing threshold (ANSI, 1997). The SII is defined by

$$SII = \sum_{i=1}^{n} I_i A_i \,, \tag{3.1}$$

where n is the number of frequency bands used in the calculations. The band-importance function I_i reflects the importance of frequency band i for speech intelligibility. The band-audibility function A_i is the effective proportion of the dynamic range of speech within band i that contributes to intelligibility. In quiet, a speech sample in band i contributes to intelligibility, when its level is higher than the level of an imaginary internal noise in the ear of the listener. This internal noise is calculated so that, if it were an external noise, it would lead to the pure-tone threshold of the listener. In noise, a speech sample in band i contributes to intelligibility when its level is higher than the level of the masking noise and the imaginary noise. Furthermore, the band audibility function A_i includes the effects on audibility of self-masking in speech, upward spread of masking and level distortion (i.e., the decrease of speech intelligibility at high presentation levels).

In the present study, the SII was calculated following the 1/3-octave band procedure of ANSI (1997), using the band-importance function for speech material of average redundancy (Pavlovic, 1987). One modification to the standard SII model was introduced. In the standard SII model, the slope for the upward spread of masking in dB/octave (C_i) due to frequency band i is given by

$$C_i = -80 + 0.6 \cdot BL_i \ , \qquad (3.2)$$

where BL_i is the level of the masker (in dB SPL) in frequency band i. In our modified SII model, shallower slopes of the masking curves were used for frequencies below 500 Hz, because calculations in chapter 2 have shown that the standard SII model underestimates the upward spread of masking produced by the low-frequency noise in the SR*B*T experiment. For 1/3-octave bands with masker frequencies below 500 Hz (i.e., $i<6$), the slope of the masking curve was calculated according to[1]

$$C_i = [1 - 0.08 \cdot (6 - i)] \cdot [-80 + 0.6 \cdot BL_i]. \qquad (3.3)$$

The level of speech and noise was measured in 1/3-octave bands with the headphone on a Brüel & Kjær type 4152 artificial ear with a flat-plate coupler. For the SII procedure, these levels were transformed to equivalent free-field levels, using the "artificial-ear-to-free-field transfer function" as derived in chapter 2. Next, the free-field levels of speech and noise were converted to spectrum levels, as is required by the SII procedure.

The spectrum level of the imaginary internal noise is

$$X = Q - R \ , \qquad (3.4)$$

where Q is the observed pure-tone threshold, measured with a psychoacoustical procedure compatible with those used for obtaining the ISO (1961) threshold, and R is the critical ratio in dB (Pavlovic, 1987). The 1/3-octave band-noise threshold is virtually identical to the pure-tone threshold (Berger, 1981; Cox and McDaniel, 1986). Therefore, the spectrum level of the internal noise was calculated as the observed 1/3-octave band-noise threshold (transformed to free-field level) minus the critical ratio in dB.

All measured SRT and SR*B*T values were subjected to SII calculations. The SII for listeners with a sensorineural hearing loss was calculated in the same way as that for normal-hearing listeners. For listeners with a mixed hearing loss, the conductive part of the hearing loss was subtracted from the level of speech, external noise, and internal noise. This means that the conductive part of the hearing loss is considered an attenuation in the transmission system, as suggested by Fletcher (1952).

3.4 RESULTS

3.4.1 Dynamic range

For the normal-hearing listeners, the difference between hearing threshold and narrow-band UCL was on average 98 dB [standard deviation (s.d.) 11 dB]. The individual narrow-band UCLs of the normal-hearing listeners had to be attenuated by 15 dB (s.d. 9 dB) on average to obtain the broadband UCL. The narrowband dynamic range of the hearing-impaired listeners ranged from about 30 dB to 100 dB at each frequency. On average, their narrowband dynamic range decreased from 70 dB at 250 Hz, to 50 dB at 4 kHz. The attenuation of the narrowband UCLs needed to obtain the broadband UCL ranged from 10 to 40 dB, with a mean of 25 dB.

3.4.2 Speech intelligibility

The thresholds of the individual listeners on the four speech-intelligibility tests are shown in Table A.2 of Appendix A, columns SRTq, SRTn, SRTa, and SR*B*T. For the normal-hearing listeners, the average SRTq was 24.0 dBA (s.d. 1.3 dBA). For the hearing-impaired listeners, the SRTq ranged from 23.5 to 84.5 dBA, with a mean of 58.8 dBA. The mean standard error (test-retest) of an individual SRTq was 2.2 dBA.

The level of the masking noise in the subsequent SRTn test was 20 dB above the individually measured SRTq. For one hearing-impaired listener, the resulting level of the masking noise was above the uncomfortable loudness level. For this listener, the level of the masking noise was set at 5 dB above the SRTq.

For the normal-hearing listeners, the average SRTn was -3.6 dB (s.d. 1.1 dB). For the hearing-impaired listeners, the SRTn ranged from -3.4 to +5.8 dB, with a mean of +0.8 dB. The average SRTa was -4.5 dB (s.d. 1.0 dB) for the normal-hearing listeners. For the hearing-impaired listeners, the SRTa ranged from -5.0 to +5.8 dB, with a mean of -1.0

dB. The mean standard error (test-retest) of an individual SRTn and SRTa was 0.96 and 0.85 dB, respectively.

The average SR*B*T of the normal-hearing listeners was 1.44 octave (s.d. 0.13 octave). The SR*B*T of the hearing-impaired listeners ranged from 1.29 to 4.02 octave, with a mean of 2.14 octave. The standard error (test-retest) of an individual SR*B*T was 0.18 octave.

3.5 DISCUSSION

3.5.1 Relationship between narrow-band and broadband UCLs

In our study, the average narrow-band UCLs (transformed to free-field levels) of the normal-hearing listeners approximately follow the equal loudness contour pattern (see also Walker *et al.*, 1984). The loudness summation formula of Stevens (1956) predicts that the narrow-band UCLs must be attenuated by 14 dB to obtain the broadband UCL. This agrees well with the average attenuation of 15 dB found for the normal-hearing listeners in our study.

Loudness summation is often reduced in hearing-impaired listeners (Moore, 1995). Therefore, an attenuation less than 14 dB would be expected for the broadband UCL of hearing-impaired listeners. However, the average attenuation for the hearing-impaired listeners was significantly higher than for the normal-hearing listeners (p<0.05). Therefore, it seems that the broadband UCL can be predicted from the narrowband UCL on the basis of loudness summation for the average normal-hearing listener, but not for the average hearing-impaired listener. For both groups of listeners, the relationship between narrow-band and broadband UCLs was highly variable, consistent with the findings of Walker *et al*. (1984).

3.5.2 SRTa versus SRTn

Figure 3.1 shows the SRT in noise for speech with an adapted spectrum (SRTa) (i.e., a spectrum halfway the dynamic range) as a function of the SRT in noise for speech with the original spectrum (SRTn) for 10 normal-hearing listeners and 34 hearing-impaired listeners. The upper limit of the one-tailed 95% confidence interval of the SRT of

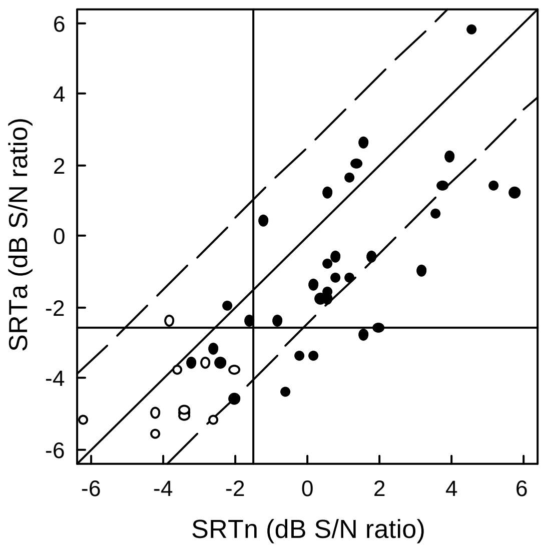


Fig. 3.1. Speech-rec rsus the standard speech-reception threshold in noise (SRTn) for normal-hearing listeners (open circles) and hearing-impaired listeners (filled circles). Solid horizontal and vertical lines represent the one-tailed 95% confidence limit for the data of the normal-hearing listeners. For data points between the dashed lines, SRTa and SRTn are not

normal-hearing listeners is chosen as the boundary between normal and elevated SRT. The separation between normal and elevated SRTn is indicated with a vertical line at -1.5 dB [i.e., -3.6 dB + $t_{0.05}(9)$*1.1 dB, where $t_{0.05}(9)$ is the critical value of the Student *t* distribution with an error probability of 0.05 and nine degrees of freedom]. Using this criterion, 7 hearing-impaired listeners had a normal SRTn, whereas 27 hearing-impaired listeners had an elevated SRTn. The boundary between a normal and an elevated SRTa lies at -2.6 dB. The SRTa was elevated for 24 of the 34 hearing-impaired listeners. The solid diagonal line in Fig. 3.1 shows where SRTn is equal to SRTa. A two-tailed *t* test, using the mean standard error (test-retest) of individual SRT values, showed that for an individual listener the difference between SRTa and SRTn is significant if it exceeds 2.5 dB (p<0.05). Thus, for the data points that lie between the dashed lines, the SRTa does not differ significantly from the SRTn. For ten hearing-impaired listeners the SRTa was significantly lower than the SRTn. In other words, for these listeners, speech

intelligibility in noise improved when the stimuli were shaped to fit in their dynamic range.
Five listeners with an elevated SRTn had a normal SRTa (see Fig. 3.1). These listeners probably performed worse than normal-hearing listeners on the SRTn test only because part of the relevant dynamic range of the speech was presented below the hearing threshold. However, for 22 of the hearing-impaired listeners, adapting the spectrum did not bring the SRT in noise back to normal. A possible explanation is that these listeners suffered from excessive spread of masking due to higher presentation levels. Another possibility is that, although many efforts were made, parts of the speech spectrum were still below absolute threshold. These possibilities were investigated with the SII model.

3.5.3 Results in relation to the SII

The speech intelligibility index is a physical measure designed to be a good predictor for speech intelligibility. It may be interpreted as the proportion of the total speech information available to the listener. For normal-hearing listeners, all conditions of equal intelligibility should result in the same SII. Thus, when the measured SRT and SR*B*T are expressed in SII values, and the SII model is consistent with the results of the normal-hearing listeners in this study, the model will yield identical SII values for the various tests. Then, an elevated SII at 50% intelligibility for a hearing-impaired listener can be considered an indication of deteriorated suprathreshold speech processing, because a listener with a higher-than-normal SII needs more speech information than normal-hearing listeners to reach the 50% intelligibility score.

Average SII

The upper two rows of Table 3.2 show the average SII values (and standard deviation) for the normal-hearing listeners (NH) and the hearing-impaired listeners (HI). For the normal-hearing listeners, the average SII for the SRTq is 0.22, and the average SII values for the SRTn, SRTa, and SR*B*T are equal: 0.31. Paired *t* tests show that the SII for the SRTq is significantly lower than the SII for the other conditions (Bonferroni test, $p<0.05$; Keren and Lewis, 1993).
Theoretically, a possible cause for the lower SII for the SRTq is an order or list effect, because the sentence lists and conditions were presented in a fixed order. However, for

Table 3.2. Average speech intelligibility index with standard deviation in parentheses for the normal-hearing listeners (NH) and the hearing-impaired listeners (HI), for four intelligibility tests. The SII was calculated according to different procedures. Modifications with respect to the standard SII procedure (ANSI, 1997) are given in the second column. 'Slopes': Shallower slopes of the masking curves are used below 500 Hz (see Sec. 3.3). 'Noise': The internal noise level is lowered by 3.6 dB (see Sec. 3.5.3). 'Desensitization': The speech desensitization factor (Pavlovic *et al.*, 1986) is included in the model (see Sec. 3.5.4). Results for the hearing-impaired listeners that differ significantly (p<0.05) from the corresponding result for the normal-hearing listeners are indicated with an asterisk.

	Modifications SII	SRTq	SRTn	SRTa	SR*BT*
NH	slopes	0.22 (0.07)	0.31 (0.04)	0.31 (0.03)	0.31 (0.03)
HI	slopes	0.26 (0.11)	0.31 (0.05)	0.36* (0.05)	0.40* (0.10)
NH	slopes, noise	0.31 (0.08)	0.32 (0.04)	0.31 (0.03)	0.31 (0.03)
HI	slopes, noise	0.30 (0.12)	0.32 (0.05)	0.37* (0.05)	0.41* (0.09)
HI	slopes, noise,	0.24* (0.08)	0.24* (0.05)	0.26* (0.05)	0.30 (0.06)

each speech-intelligibility test, the first measurement did not differ significantly from the second measurement (*t* tests for matched samples, p<0.05). This suggests that no significant order effect is present. Versfeld *et al.* (2000) did not find large list effects in the speech material of Smoorenburg (1992) used in this study: the standard deviation between lists was 0.6 dB. Therefore, it seems unlikely that the observed difference in SII in our study is a result of an order or list effect.

In quiet, the SII is very sensitive to the level of the imaginary internal noise, calculated from the hearing threshold [Eq. (3.4)]. Most likely the SII for the SRTq is too low, because the procedure for the determination of the hearing threshold in the present study using Békésy tracking resulted in thresholds that were systematically a few dB higher than would have been measured using the method of constant stimuli or the method of limits, on which the ISO (1961) threshold is based. The SII for the SRTq becomes equal to the SII in the other conditions when the SII is recalculated with the internal noise level lowered by 3.6 dB.

Rows 3 and 4 of Table 3.2 present the SII values calculated with the adapted internal noise level. Note that the adaptation of the internal noise level hardly influences the SII values for conditions other than the SRTq. For the hearing-impaired listeners, the average

SII for the SRTq and SRTn is similar to the SII for normal-hearing listeners, but for the SRTa and SR*B*T, the average SII is significantly higher (*t* test for unequal variances, $p<0.05$). Thus, for the average hearing-impaired listener, the SII corresponding to the SRTq and to the SRTn appears to be "normal" (i.e., equal to the SII of the normal-hearing listeners), suggesting the absence of the effect of suprathreshold deficits. This is consistent with the idea, mentioned in the Introduction, that the difficulty with understanding speech in background noise is only caused by part of the speech spectrum being below absolute threshold (Zurek and Delhorne, 1987; Lee and Humes, 1993). However, the mean SII for the SRTa and SR*B*T is higher. Therefore, in these conditions the effects of suprathreshold deficits *do* seem to influence speech intelligibility. This agrees with the notion that suprathreshold deficits may affect intelligibility (Glasberg and Moore, 1989; Dreschler and Plomp, 1985).

For the hearing-impaired listeners, the mean SII for the SRTa is *higher* than for the SRTn, although their mean SRTa (-1.0 dB) was 1.8 dB *lower* than their mean SRTn (+0.8 dB). This means that by adapting the spectrum, the SRT in noise did not decrease as much as predicted from audibility. The SII model would predict an equal SII for SRTn and SRTa. Thus, the SRTa can be predicted for each listener from the SII for the measured SRTn. On average, the predicted SRTa for the hearing-impaired listeners is -2.9 dB. Hence, when adapting the spectra the SII model predicts a decrease in the mean SRT in noise of 3.7 dB, instead of the observed 1.8 dB.

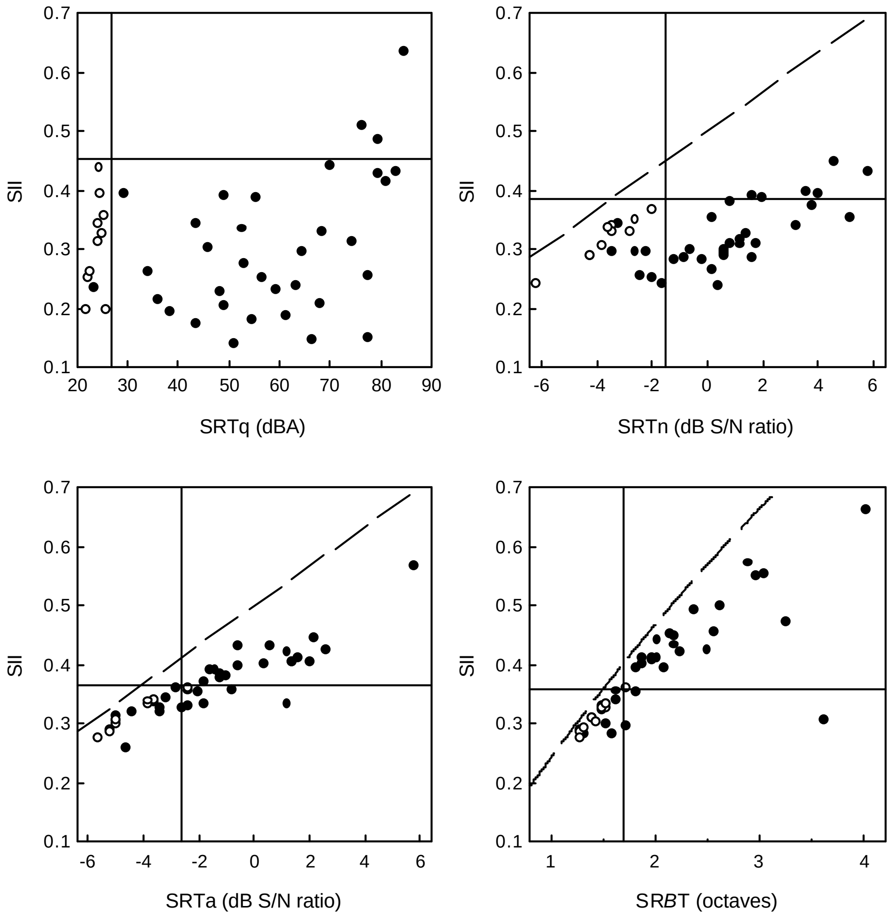
The mean SII for the SR*B*T is significantly higher than for the SRTa ($p<0.05$) for the hearing-impaired listeners. Thus, for the average hearing-impaired listener in the present study, the effect of suprathreshold deficits is most obvious for the SR*B*T condition (narrow-band speech in bandstop noise), and to a lesser extent also for SRTa (spectrally adapted broadband speech in broadband noise), but entirely absent for the standard conditions SRTn and SRTq.

One possibility that may explain why the SII varies across condition is that the suprathreshold deficit is frequency dependent. For example, in the frequency regions where the hearing loss is greater, the suprathreshold deficit may be larger [Pavlovic *et al.* (1986)]. Different frequency regions are involved in the different conditions. In the standard conditions, the listeners must use the frequency region where the original speech spectrum is above their hearing threshold. In the SRTa condition, the entire frequency

range from 250 to 4000 Hz should contribute. The suprathreshold deficit may be present in the frequency region that is below threshold in the standard conditions, but above threshold in the SRTa condition. This may explain the elevated SII for the SRTa. However, it seems improbable that a frequency-dependent deficit can explain the elevated SII for the SR*B*T, because the frequency range around 1 kHz that is important in the SR*B*T test is also important in the other tests.
Another possible cause for the dependence of the effect of suprathreshold deficits on condition can be found in the different spectra of speech and noise among conditions. For instance, when a hearing-impaired listener experiences excessive upward or downward spread of masking, the degree to which this influences speech intelligibility in a specific condition will depend on the specific spectra of speech and noise.

Individual SII

Table 3.2 shows that the interindividual spread in SII for the hearing-impaired listeners is substantially larger than the spread in SII for normal-hearing listeners. Therefore, the SII of individual listeners (after the correction of -3.6 dB of the internal noise levels) is considered. Table A.2 of Appendix A shows these individual SII values in the four conditions. In Fig. 3.2, the SII is plotted as a function of SRT or SR*B*T obtained for the individual normal-hearing listeners (open circles) and hearing-impaired listeners (filled circles). The upper limit of the one-tailed 95% confidence interval of the SII of normal-hearing listeners is chosen as the separation between normal and higher-than-normal SII (horizontal lines). The separation between normal and abnormal SRT or SR*B*T is indicated with a vertical line. The dashed lines in Fig. 2 represent the hypothetically maximum SII given the SRTn, SRTa, and SR*B*T, i.e., the SII that would have been calculated if the audibility of the speech had not been influenced by the known effects of hearing threshold, upward spread of masking, and level distortion. To calculate this hypothetically maximum SII, the full 30-dB dynamic range of the speech was assumed to be audible for all frequency bands in which the speech was not masked by noise in the same frequency band. Thus, the vertical distance from each data point to the dashed line represents the combined influence of hearing threshold, upward spread of masking and level distortion.



the maximum possible SII as a function of SRT or SR*B*T, i.e., the SII that would have been calculated if the audibility of the speech had not been influenced by the hearing threshold, upward spread of masking, and level distortion.

The horizontal and vertical lines divide the graphs of Fig. 3.2 in four quadrants. No data points fall in the upper-left quadrant. Data points in the lower-left quadrant correspond to a normal threshold. Data points in the lower-right quadrant correspond to an elevated SRT or a broader-than-normal SR*B*T, but this abnormal threshold can be explained on basis of audibility by the SII model. Finally, data points in the upper-right quadrant correspond to an abnormal threshold that cannot be explained by the SII model. The higher a data point lies above the horizontal line, the larger the speech processing deficit.

SRTq: Figure 3.2 shows that for all but three hearing-impaired listeners the SII for the SRTq is not higher than normal. The most important factor affecting the audibility of the speech in quiet is the hearing threshold. The effect of level distortion is negligible. Indeed, stepwise multiple regression shows that the SRTq (in dBA) can be predicted quite accurately ($R^2 = 0.92$) from the hearing threshold

$$SRTq = 22.2 + 0.60 \cdot HL_{0.5} + 0.32 \cdot HL_2 , \qquad (3.5)$$

where $HL_{0.5}$ and HL_2 are the hearing threshold in dB HL at 0.5 and 2 kHz, respectively.

SRTn: The results on the SRTn test also yield a rather constant SII. Only six of the listeners performed (slightly) worse than predicted on the basis of audibility. The hearing threshold is the most important factor that causes the difference between the individual SII and the dashed line. The SII calculations show that a part of the speech signal (the high frequencies) is still below the hearing threshold, even for the normal-hearing listeners. Audibility of the speech is also affected by upward spread of masking.

SRTa: In contrast to most of the elevated thresholds in noise for the *original* spectrum (SRTn), most of the elevated thresholds for an *adapted* spectrum (SRTa) cannot be explained by the SII model. For the normal-hearing listeners, the difference between the obtained SII and maximal possible SII (dashed line) for the SRTa is only caused by upward spread of masking. Upward spread of masking, some level distortion, and the hearing threshold reduce audibility for the hearing-impaired listeners. Thus, the hearing threshold still affects audibility, although speech and noise spectra were spectrally shaped to fit in the dynamic range of individual listeners from 250 to 4000 Hz. The reason is that below 250 Hz and above 4000 Hz, the frequency response of the filter was flat. For most hearing-impaired listeners, the hearing threshold at 8000 Hz was much higher than at 4000 Hz. Therefore, part of the relevant dynamic range of the speech was below the hearing threshold in the frequency region between 4000 and 8000 Hz.

SR*B*T: The SR*B*T was broader than normal for 24 of the 34 hearing-impaired listeners. For 21 hearing-impaired listeners, the broader SR*B*T cannot be accounted for by the SII model. Therefore, also in this condition, additional suprathreshold factors must have affected intelligibility. The difference between the obtained SII and the maximal possible SII (dashed line) is caused by the hearing threshold, upward spread of masking, and some level distortion. Figure 3.2 shows that for the listener with an SR*B*T of 3.6 octave, the difference between the actual SII and the dashed line is very large. The reason is that a large part of the relevant dynamic range of the speech was presented below the hearing threshold, due to the extremely narrow dynamic range of this listener.

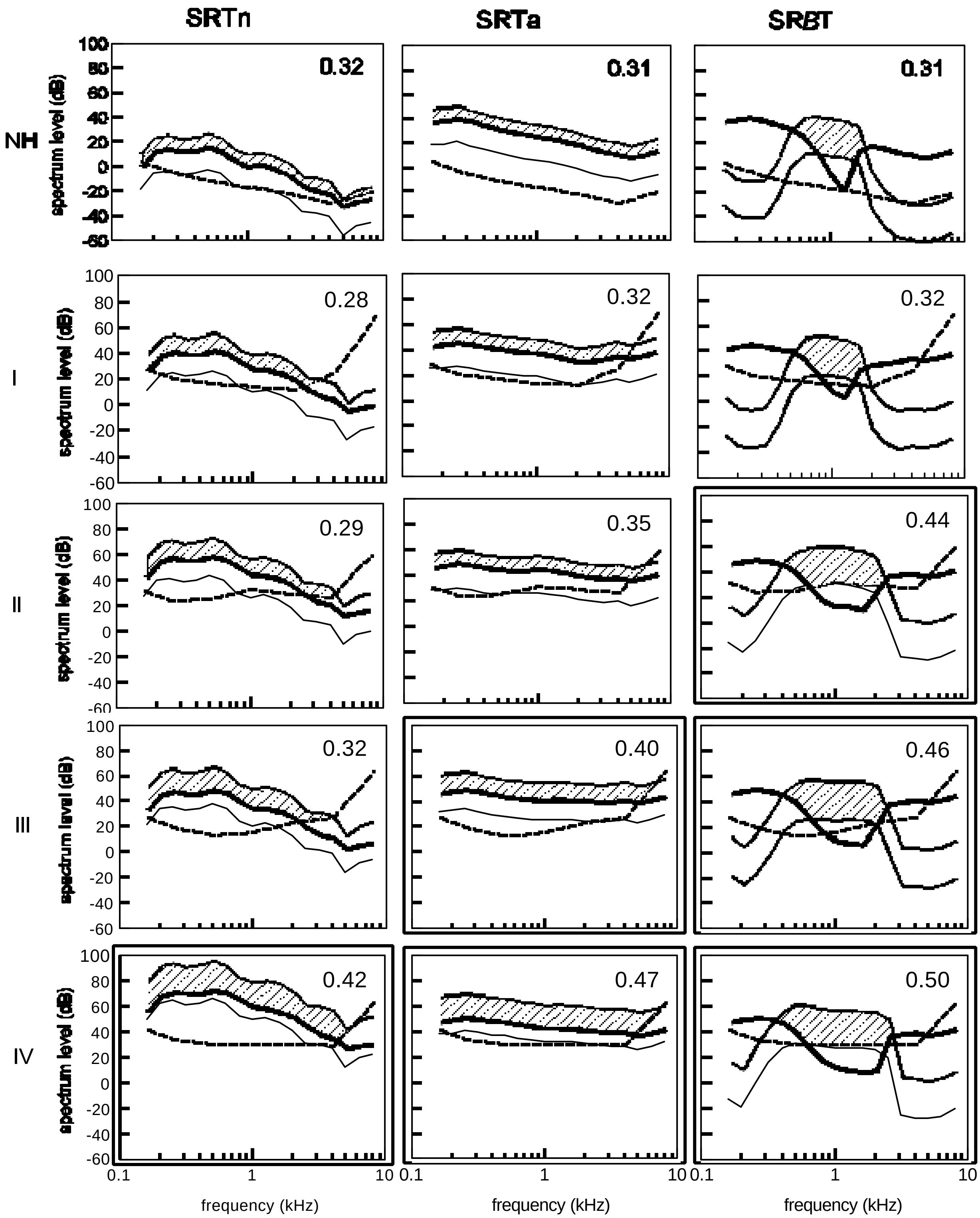
SII across tests

For a convenient discussion of the SII of the hearing-impaired listeners across tests, the listeners are categorized into groups, based on their normal or elevated SII. A measurement error is associated with each result, and this error may cause an SII to fall just on one side of a separation between normal and elevated SII. Therefore, the categorization into groups should not be viewed too absolutely, but only as a means to simplify the discussion of the individual SII across tests. The SII for the SRTq is not used for this categorization, because the SII of only three hearing-impaired listeners is distinct from the SII of normal-hearing listeners.

Considering the SII for the SRTn, SRTa, and SR*B**T***,** the SII patterns of all but three hearing-impaired listeners can be subdivided into only four groups. Figure 3.3 shows the average proportion of the speech dynamic range (hatched area) needed for a 50% intelligibility score in the three conditions, for normal-hearing listeners (NH) and the four hearing-impaired groups. The group-averaged SII is shown in the upper-right corner.

Double frames contain an elevated SII. Figure 3.3 displays no large differences in hearing threshold (dashed curves) or in the midline of the dynamic range (thick curves in the middle panels) across the four hearing-impaired groups. Thus, the differences across the groups do not seem related to major differences in hearing threshold or in the presented spectra. Each group of hearing-impaired listeners is discussed below.

Fig. 3.3. Average proportion of the speech dynamic range (hatched regions) required for a 50% intelligibility score in three conditions (SRTn, SRTa, and SR*B*T) for normal-hearing listeners (panel set NH) and four groups of hearing-impaired listeners (panel sets I through IV). The 30-dB speech dynamic range is enclosed by the two thin lines. Thick lines are the masker spectra (including upward spread of masking) for the noise. Dashed lines indicate the spectrum level of the imaginary internal noise leading to the average absolute threshold. The group-averaged SII is shown in the upper right corner. Double frames contain an elevated SII.


SRTn
SRTa
SRBT
NH
I
II
III
IV
spectrum level (dB)
100
80
60
40
20
0
-20
-40
-60
0.32
0.31
0.31
0.28
0.32
0.32
0.29
0.35
0.44
0.32
0.40
0.46
0.42
0.47
0.50
0.1
1
10
frequency (kHz)

Group I: Eleven hearing-impaired listeners have a normal SII for thresholds on all three tests. These listeners seem to process suprathreshold speech equally well as normal-hearing listeners. The four listeners with a mixed hearing loss all belong to this group. **Group II:** Four listeners have an elevated SII for the SR*B*T, but a normal SII for both SRT tests in noise. The speech-understanding problems of these listeners seem to be caused primarily by the specific spectral configuration of the stimulus in the SR*B*T experiment. A possible cause is an increased susceptibility to spread of masking.

Group III: Twelve listeners have an elevated SII for both the SRBT and the SRTa, but a normal SII for the SRTn. Like the listeners of group II, the elevated SII for the SRBT could be related to increased spread of masking. The elevated SII for the SRTa may be caused by a frequency-dependent deficit. Figure 3.3 shows that the high-frequency part of the spectra of speech and noise is lifted above the hearing threshold in the SRTa condition. It appears that, although the high-frequency part of the speech spectrum is presented to the hearing-impaired listeners of this group, they cannot use it as effectively as normal-hearing listeners. The performance of these listeners is consistent with the finding of Ching et al. (1998) and Hogan and Turner (1998) that the contribution of audibility to intelligibility is much reduced at high frequencies, where the hearing loss is severe.

Group IV: Four listeners have an elevated SII for the SRTn, the SRTa, and the SRBT**.** Two of these listeners also have an elevated SII for the SRTq. The SII for the SRTq of the other two listeners is also high: 0.41 and 0.43. Therefore, the suprathreshold deficit of these listeners does not seem to depend strongly on condition.

Overall, it seems that a normal SII for the SRTn does not imply a normal SII for the SRTa and the SR*B*T (groups II and III). However, a normal SII for the SR*B*T appears to imply a normal SII for both SRT tests in noise (group I).

3.5.4 Modifying the SII model to include suprathreshold deficits

Proficiency factor

In the literature, various attempts have been made to modify the articulation index or SII to account for the deterioration in processing of suprathreshold speech, due to the sensorineural hearing loss. Fletcher (1952) proposed using the listener-dependent proficiency factor P for this purpose. Equation (3.1) then changes to

$$SII = P \sum_{i=1}^{n} I_i A_i \ . \tag{3.6}$$

The proficiency factor depends only on the listener, and not on the listening condition. However, the SII for the listeners of groups II, III, and IV clearly depends on the listening condition (Fig. 3.3). Therefore, the proficiency factor cannot account for the performance of the listeners from these groups.

Speech desensitization factor

Pavlovic *et al.* (1986) proposed using a "speech desensitization factor" to account for the deterioration in speech processing of listeners with a sensorineural hearing impairment

$$SII = \sum_{i=1}^{n} D_i I_i A_i \ . \tag{3.7}$$

The desensitization factor D is specified as a function of the hearing loss in each frequency band. It decreases linearly from 1, for hearing losses less than 15 dB HL, to 0, for hearing losses exceeding 95 dB HL. With Eq. (3.7) the SII values are recalculated. Rows 4 and 5 of Table 3.2 show the average SII of the hearing-impaired listeners with and without the desensitization factor. An SII indicated with an asterisk differs significantly from the corresponding SII of the normal-hearing listeners (row 3). At the SR*B*T, the SII of the hearing-impaired listeners becomes equal to the SII of normal-hearing listeners, when the desensitization factor is used. However, at the SRTq, SRTn and SRTa, the SII of the hearing-impaired listeners becomes lower than the SII of normal-hearing listeners. In these three conditions, the speech desensitization factor seems to overestimate the deterioration in speech processing. It was verified that it is not possible to make the SII values of the hearing-impaired listeners equal to those of the normal-hearing listeners by choosing another linear relation between the desensitization factor D and the hearing loss. Thus, the hearing-loss dependent speech desensitization factor cannot explain our results.

Upward spread of masking

Ludvigsen (1987) modified the articulation index by including a model of auditory masking in cochlearly hearing-impaired listeners. In his model, upward spread of masking increases proportionately with the hearing loss at the frequency being masked. Increased upward spread of masking decreases the SII for the SRTn, the SRTa, and the SR*B*T. It was found to have a larger effect on the SII for the SRTn, than on the SII for the SR*B*T. Therefore, increased upward spread of masking also cannot explain why the average SII for the SR*B*T is higher than the average SII for the SRTn.

Downward spread of masking

The spectral configuration of the stimulus in the SR*B*T experiment is not only sensitive to the effect of upward spread of masking, but also for the effect of *downward* spread of masking. Increased downward spread of masking may explain the results. The noise spectrum level in the SRTn condition decreases with frequency (Fig. 3.3). Therefore, the level of downward spread of masking will be negligible compared with the level of the noise that is already physically present in the SRTn condition. Downward spread of masking will have a larger influence on the SII for the SR*B*T, because downward spread of masking from the high-frequency part of the bandstop noise will mask the (physically unmasked) bandpass-filtered speech.

For normal-hearing listeners, the slope of downward spread of masking is very steep, about 100 dB/octave (Zwicker, 1963), and independent of sound-pressure level. Abnormal downward spread of masking for hearing-impaired listeners has been observed by some authors. Glasberg and Moore (1986) measured auditory-filter shapes at three center frequencies (0.5, 1.0, and 2.0 kHz) for five listeners, and at one center frequency (1.0 kHz) for seven listeners with cochlear impairments. The slope of the high-frequency skirt of the filter, which reflects the amount of downward spread of masking, ranged from about 4 to 100 dB/octave. For four listeners, the filter shape could not even be determined at one or two center frequencies, because the filter had too little frequency selectivity. Hearing loss and high-frequency slope were not significantly correlated. Nelson (1991) obtained forward-masked psychophysical tuning curves for 21 listeners with cochlear hearing losses. Ten listeners showed abnormal downward spread of masking, when equivalent masker levels were compared. Conversely, none of these

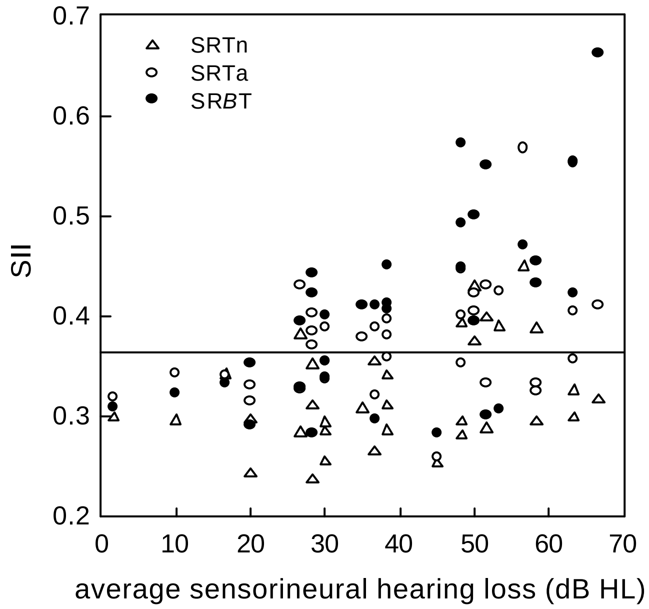


Fig. 3.4. Speech intelligibility index corresponding to the data of the hearing-impaired listeners on three intelligibility tests (SRTn, SRTa, and SR*B*T) versus their sensorineural hearing loss averaged over the frequencies 0.5, 1, and 2 kHz. The horizontal line represents the one-tailed 95% confidence limit for the SII of the normal-hearing listeners.

listeners showed abnormal upward spread of masking. The high-frequency slopes of their tuning curves ranged from 10 to 87 dB/octave. Abnormal downward spread of masking was only observed in listeners with hearing losses exceeding 40 dB HL. However, not all listeners with a hearing loss greater than 40 dB HL showed abnormal downward spread of masking.

The two above-mentioned studies show that hearing-impaired listeners may experience increased downward spread of masking. The slope of the downward spread of masking varies considerably across the listeners, and cannot be predicted from the hearing loss. Further experiments are required to investigate if, for the individual listener, increased downward spread of masking is indeed related to an elevated SII for the threshold on these intelligibility tests.

3.5.5 SII and hearing loss

As mentioned in the Introduction, Moore (1996) argued that factors other than audibility contribute to the difficulties in speech perception for listeners with a cochlear hearing loss greater than about 45 dB. To investigate the relation between SII and hearing loss for the present results, the SII for the SRTn, SRTa, and SR*B*T is plotted as a function of sensorineural hearing loss averaged over the frequencies 0.5, 1 and 2 kHz (Fig. 3.4). The horizontal line is the 95% confidence limit of the SII of the normal-hearing listeners. Figure 3.4 shows that only for average hearing losses less than 25 dB, all results can be explained on the basis of audibility by the SII model. For losses greater than 25 dB, suprathreshold factors already seem to influence speech perception for some listeners and intelligibility tests. However, high losses may still be associated with “normal” SII values.

3.6 SUMMARY AND CONCLUSIONS

The speech-understanding ability of 34 hearing-impaired listeners and a reference group of 10 normal-hearing listeners was measured with the SRT test in quiet, the standard SRT test in noise, and an SRT test in noise in which the speech and noise spectra were shaped to fit in the dynamic range of each listener. Furthermore, the threshold for speech reception was measured with the new SR*B*T test (see chapter 2). The SRBT is defined as the bandwidth of speech around 1 kHz required for a 50% intelligibility score.

All individual data were converted to SII values, with the assumption that the SII reflects the effect of audibility, and that a higher-than-normal SII value required for a 50% intelligibility score indicates a deterioration in speech processing by factors not included in the SII model. For the normal-hearing listeners, the SII model accounts for the results obtained for all four tests (SII of typically 0.31 for all SRT and SR*B*T values).

For the normal-hearing listeners, the mean SRT in quiet was 24.0 dBA. For the hearing-impaired listeners, the mean SRT in quiet was 58.8 dBA. The mean standard SRT in noise was -3.6 dB for the normal-hearing listeners and +1.0 dB for the hearing-impaired listeners. An elevated SRT in quiet or an elevated standard SRT in noise could be explained on the basis of audibility by the SII model for the great majority of hearing-impaired listeners. When the speech and noise spectra were adapted to fit in each listener's dynamic range, the SRT in noise improved on average by 0.9 dB for the normal-hearing listeners and 1.8 dB for the hearing-impaired listeners. Yet, the improvement for the hearing-impaired listeners is less than predicted on the basis of audibility by the SII model. For about half of the hearing-impaired listeners, the SRTa corresponds to significantly higher SII values.

On average, the SR*B*T of the normal-hearing listeners was 1.4 octave. The SR*B*T of the hearing-impaired listeners ranged from 1.3 to 4.0 octave. Compared with the normal-hearing reference group, 10 hearing-impaired listeners had a normal SR*B*T, whereas 24 hearing-impaired listeners had a broader-than-normal SR*B*T. Only for three hearing-impaired listeners could the broader-than-normal SR*B*T be explained on the basis of audibility by the SII model.

Eleven of the hearing-impaired listeners performed on all four tests as predicted by the SII model. The data of 23 hearing-impaired listeners on one or more of the intelligibility tests could not be explained by this model. This points to a deterioration in speech processing. The effect of such a speech processing deficit appeared to depend strongly on the test condition for hearing-impaired listeners. However, a normal SII for the SR*B*T, appears to imply a normal SII for both SRTn and SRTa. Thus, it seems that the SR*B*T test is most sensitive for speech-processing deficits. Increased downward spread of masking may be the reason for the elevated SII for the threshold on these three intelligibility tests.

Predicting the elevation of the SII from the hearing loss does not seem possible: for hearing losses greater than 25 dB, elevated SII values were found for some listeners, but for other listeners SII values were still considered normal.

NOTE

[1] It should be mentioned that the shallower slopes of the masking curves for frequencies below 500 Hz hardly influence the calculated SII in the present study (the largest SII decrement was 0.02). Equation (3.3) decreases the SII most for spectra with a negative spectral tilt and a high sound-pressure level (see chapter 2). In the present study, no spectral tilt was imposed in the SRTn test, and the imposed spectral

tilt was positive for spectra presented halfway through the dynamic range. Although the shallower slopes of the masking curves hardly influence the SII in this study, Eq. (3.3) is still used to be consistent with our previous study, described in chapter 2.

Chapter 4

Relations between intelligibility of narrow-band speech and auditory functions, both in the 1-kHz frequency region

Relations between perception of suprathreshold speech and auditory functions were examined in 24 hearing-impaired listeners and 12 normal-hearing listeners. The speech intelligibility index (SII) was used to account for audibility. The auditory functions included detection efficiency, temporal and spectral resolution, temporal and spectral integration, and discrimination of intensity, frequency, rhythm, and spectro-temporal shape. All auditory functions were measured at 1 kHz. Speech intelligibility was assessed with the speech-reception threshold (SRT) in quiet and in noise, and with the speech-reception bandwidth threshold (SRBT), previously developed for investigating ***speech perception in*** *a limited frequency region around 1 kHz. The results showed that the elevated SRT in quiet could be explained on basis of audibility. Audibility could only partly account for the elevated SRT values in noise and the deviant SRBT values, suggesting that suprathreshold deficits affected intelligibility in these conditions. SII predictions for the SRBT improved significantly by including the individually measured upward spread of masking in the SII model. Reduced spectral resolution, reduced temporal resolution, and reduced frequency discrimination appeared to be related to speech perception deficits. Loss of peripheral compression appeared to have the smallest effect on the intelligibility of suprathreshold speech.*

Journal of the Acoustical Society of America
(accepted for publication with minor modifications)

4.1 INTRODUCTION

Sensorineural hearing loss often interferes with understanding speech in a noisy environment, even when the speech is presented well above the hearing threshold. Therefore, reduced speech intelligibility cannot be just a result of the elevated hearing threshold per se. There also has to be a deterioration in suprathreshold sound processing. Various researchers studied correlations between speech perception and specific auditory functions (Tyler *et al.*, 1982; Festen and Plomp, 1983; Dreschler and Plomp, 1985; Glasberg and Moore, 1989; van Rooij and Plomp, 1990; Smoorenburg, 1992; Divenyi and Haupt, 1997a, 1997b). However, finding the specific deficits that are responsible for reduced intelligibility of suprathreshold speech has proved to be difficult. The main reason for this may be that correlations between a listener's ability to understand *wideband* speech and auditory functions in a *limited* frequency region were studied. Furthermore, auditory functions and hearing loss are usually not independent. Therefore, when examining correlations between these two factors, it is difficult to determine whether the underlying cause for reduced intelligibility lies in the deterioration of a specific auditory function, or whether it is due to part of the speech spectrum falling below the absolute threshold.

In the present study, the relation between speech perception and auditory functions was investigated while avoiding the difficulties mentioned above. The auditory functions were all measured at 1 kHz. Speech intelligibility was measured for a restricted frequency range with a center frequency of 1 kHz. For this purpose, the SRBT-test has been developed, as described in chapter 2. Thus, the auditory functions and speech perception refer to the same frequency region, which may well increase the chance to find clear correlations. The influence of audibility on speech intelligibility was evaluated with the speech intelligibility index, or SII (ANSI, 1997). After accounting for audibility, correlations between speech perception and auditory functions were examined. In this way, the effects of suprathreshold deficits were separated from the effects of audibility.

Ching *et al.* (1997) used a similar approach. They calculated the deviations of measured intelligibility scores for filtered speech from the intelligibility scores predicted by the SII model, and found that these deviations were moderately correlated with the measured spectral and temporal resolution in the same frequency region, and with the hearing threshold. Because audibility was accounted for, they concluded that the correlations with the hearing threshold may be related to suprathreshold deficits that covary with hearing loss.

In the present study, suprathreshold speech perception will not only be related to spectral and temporal resolution, but also to the results of a set of other detection and discrimination tests. The detection tests assessed spectral and temporal resolution, and spectral and temporal integration. In the discrimination tests, just-noticeable differences were measured in intensity, frequency, rhythm, and spectro-temporal shape. Many auditory functions are known to change with sound pressure level. For example, intensity discrimination improves with level, whereas spectral resolution deteriorates with level. So, differences in results for hearing-impaired and normal-hearing listeners can be caused by level effects. Such level effects can be the result of differences in either the absolute level or in the sensation level of the stimuli. The level-dependence of auditory functions was accounted for in our investigations by measuring the thresholds of the normal-hearing listeners at three presentation levels.

Speech intelligibility was assessed for narrow-band speech with the novel SR*B*T. For broadband speech, various speech-reception thresholds (SRT's) were determined, namely the SRT in quiet, the SRT in noise, and the spectrally adapted SRT in noise (in which stimuli are adjusted to fit in the individual's dynamic range). In chapter 3, the same intelligibility tests were measured using 34 hearing-impaired listeners. For most listeners, elevated SRT values in quiet and in noise could be explained on the basis of audibility. However, for the spectrally adapted SRT in noise, and especially for the SR*B*T, the data of most listeners could not be explained from audibility, suggesting that effects of suprathreshold deficits influenced intelligibility in these conditions. In the present study, an attempt is made to determine which specific suprathreshold deficits reduce intelligibility for hearing-impaired listeners.

4.2 METHOD

4.2.1 Stimuli

The speech material consisted of eight lists of 13 meaningful Dutch sentences, uttered by a female speaker (Plomp and Mimpen, 1979). For masking, a Gaussian noise that was spectrally shaped according to the long-term average spectrum of the sentences was used. Both the speech and the noise signals were digitized at a sampling frequency of 15625 Hz. Noise bands of **a**-octave were used for determination of the absolute thresholds and the levels of uncomfortable loudness.

In the measurements of the auditory functions, three types of signals were used: a tone, a click, and a Gaussian-windowed tone pulse. The tone was a sinusoid with a frequency of 1 kHz, which was gated with a 20-ms cosine squared ramp, and had a half-amplitude duration of 100 ms. The click was a filtered step-function of two octaves wide (500-2000 Hz). The Gaussian-windowed tone pulse was defined by

$$s(t) = A\sqrt{\alpha f_0}\sin(2\pi f_0 t + \frac{\pi}{4})\exp(-\pi(\alpha f_0 t)^2). \qquad (4.1)$$

The carrier frequency f_0 of the Gaussian pulse was 1 kHz and the shape factor α was 0.25. This shape factor was chosen because it defines a signal with the spectral width of a critical band (equivalent rectangular bandwidth: ¼ octave) combined with a short duration (equivalent rectangular duration: 2.8 ms). For low-level stimuli with this particular spectro-temporal shape, listeners are not able to combine stimulus information across multiple frequency channels or across "multiple looks" in time (van Schijndel *et al.*, 1999). The energy of the Gaussian pulse only depends on the amplitude factor *A*. In its generation, the Gaussian pulse was cut off at 60 dB below the top. For masking, a noise with gaussian-amplitude distribution was used with a bandwidth of four octaves (250-4000 Hz) and a duration of 400 ms. The noise masker was gated with a 20-ms cosine-squared ramp.

4.2.2 Apparatus

Stimuli were generated with a personal computer using TDT (Tucker-Davis Technologies) System II add-on hardware. In the intelligibility tests, stimuli were upsampled by a factor of two and delivered at a 31250-Hz sampling frequency. In the auditory-function tests, stimuli were delivered at a sampling frequency of 40 kHz. Stimuli were converted to analog using a 16-bit D/A converter (TDT DD1) and they were low-pass filtered at 16 kHz (TDT FT5). Signals and noise were attenuated separately (TDT PA4), and subsequently summed (TDT SM3). For frequency shaping in the intelligibility tests and in the determination of the broadband uncomfortable loudness level (see Sec. 4.2.4), the total signal was passed through a programmable filter (TDT PF1). Listeners were seated in a soundproof room. Stimuli were presented monaurally through headphones (Sony MDR-CD999). Sound-pressure levels of the stimuli were measured with the headphone placed on a Brüel & Kjær type 4152 artificial ear with a flat-plate adapter.

4.2.3 Listeners

Twenty-four sensorineurally hearing-impaired persons were selected from the files of the University Hospital VU and served as subjects. Information about individual listeners is given in Table B.1 of Appendix B. Their ages ranged from 39 to 67 years. Their native language was Dutch. Only listeners who could reach an intelligibility score for monosyllabic words in quiet of at least 75% were selected. They were tested at their better ear. The thresholds, averaged over 0.5, 1, and 2 kHz, ranged from 5 to 62 dB HL. Pure-tone air-conduction thresholds in the ear under test were at least 50 dB HL at one or more octave frequencies between 250 and 4000 Hz. Twelve normal-hearing listeners, ranging in ages from 19 to 63 years, participated in the experiment as a control group. The pure-tone air-conduction thresholds did not exceed 15 dB HL at octave frequencies from 250 Hz to 4000 Hz for their tested ears.

4.2.4 Procedure

The hearing-impaired listeners were tested in two sessions of four hours on two separate days. Session 1 consisted of six test blocks (blocks 1 to 6 described below) and session 2 of five test blocks (blocks 2 to 6 in retest) of about twenty minutes. The blocks were presented in a fixed order to avoid possible differences among listeners due to sequence effects. After each test block, the listeners had a break of the same duration. In the first test block, the listeners' thresholds and uncomfortable loudness levels were determined.

The measurements of the auditory functions were divided over four blocks (blocks 2, 3, 5, and 6). Speech reception was measured in the fourth block. The reason for presenting this speech-reception block in between the more-demanding auditory-function blocks was to give the listeners some variation in their task.
The normal-hearing listeners performed the same tests as the hearing-impaired listeners (sessions 1 and 2) and an extra third session. This extra session was included to investigate the level-dependence of the auditory functions measured in sessions 1 and 2. Auditory functions were measured at three presentation levels: 15 dB below, at the same level as, and 15 dB above the levels used in sessions 1 and 2 (determined in block 2). Session 3 consisted of nine test blocks (blocks 3, 5, and 6; each at three levels) of fifteen minutes, with breaks of fifteen minutes in between. (Block 3 was extended with the last test of block 2.) The presentation order of the levels was counterbalanced over the listeners according to a 3x3 Latin square. Each sequence of levels was presented to four normal-hearing listeners.

4.2.4.1 Threshold and UCL (block 1)

Hearing thresholds and uncomfortable loudness levels (UCL's) were measured with **a**-octave bands of noise at five center frequencies: 250, 500, 1000, 2000, and 4000 Hz. Levels for intermediate **a**-octave bands were calculated by interpolation. For the determination of the hearing threshold, a fixed-frequency Békésy procedure was used with a step size of 1 dB. After 11 reversals, the threshold measurement was finished. The threshold was defined as the average of all but the first reversal.
The UCL was determined in two steps. First, the UCL of the five **a**-octave bands of noise were measured individually. Next, the UCL was measured with a 4-second broadband noise burst, spectrally shaped according to the narrow-band UCL's. In the determination of the **a**-octave band UCL's, the listeners were presented with noise bursts that increased in level. They had to press a button when the noise burst was experienced as uncomfortably loud. After the button was pressed, the noise level was reduced by a random amount between 21 and 31 dB. The measurement was finished after six responses. The UCL of each band was computed by averaging across the levels at which the button was pushed. For details of the threshold and UCL measurement procedure, see chapter 3.

4.2.4.2 Speech intelligibility (block 4)

In this block, four intelligibility thresholds were measured: the speech-reception threshold in quiet (SRTq), the SRT in noise (SRTn), the SRT in noise with the speech and noise signals adapted to the midline of the dynamic range of individual listeners (SRTa), and the speech-reception *bandwidth* threshold (SR*B*T). In session 1 the order of the tests was SRTq, SRTn, SRTa, and SR*B*T. In session 2 the order was SR*B*T, SRTa, SRTq, and SRTn, i.e., the order was reversed, except for SRTq and SRTn. Their order was fixed, because the level of the noise in the SRTn test was taken 20 dB above the measured SRTq for each listener. In the SR*B*T procedure, the bandwidth of the speech signal was varied. Complementary shaped bandstop noise was added to the speech stimuli. Both the speech and noise stimuli were shaped to resemble the midline of the dynamic range of the individual listener, i.e., a line halfway between hearing threshold and broadband UCL (measured in block 1). A procedure comparable to the SRT test was followed, with the important difference that the bandwidth (center frequency: 1 kHz), instead of the sound-pressure level, of the speech signal is varied adaptively. For details of the intelligibility tests, see chapter 3.

Before the speech-intelligibility measurements, two lists of 13 sentences were used to familiarize the listeners with the procedure. With these lists, an SRTq test and an SRTn test were performed.

4.2.4.3 Auditory functions (blocks 2, 3, 5, and 6)

General

The four test blocks are schematically presented in Fig. 4.1. Blocks 2, 3, and 5 each contain three detection tests. Block 6 contains four discrimination tests. Thresholds were determined using an adaptive two-interval forced-choice procedure (Levitt, 1971) with visual feedback. Each trial consisted of two intervals of 400 ms, each indicated by a light, with 400 ms of silence in between. A two-down one-up procedure was applied that estimates the 70.7 % correct point on the psychometric function. The step sizes for each test are given in the last column of Fig. 4.1. The threshold measurements were finished after 16 reversals. The threshold was defined as the mean (for block 2, and tests 3a and

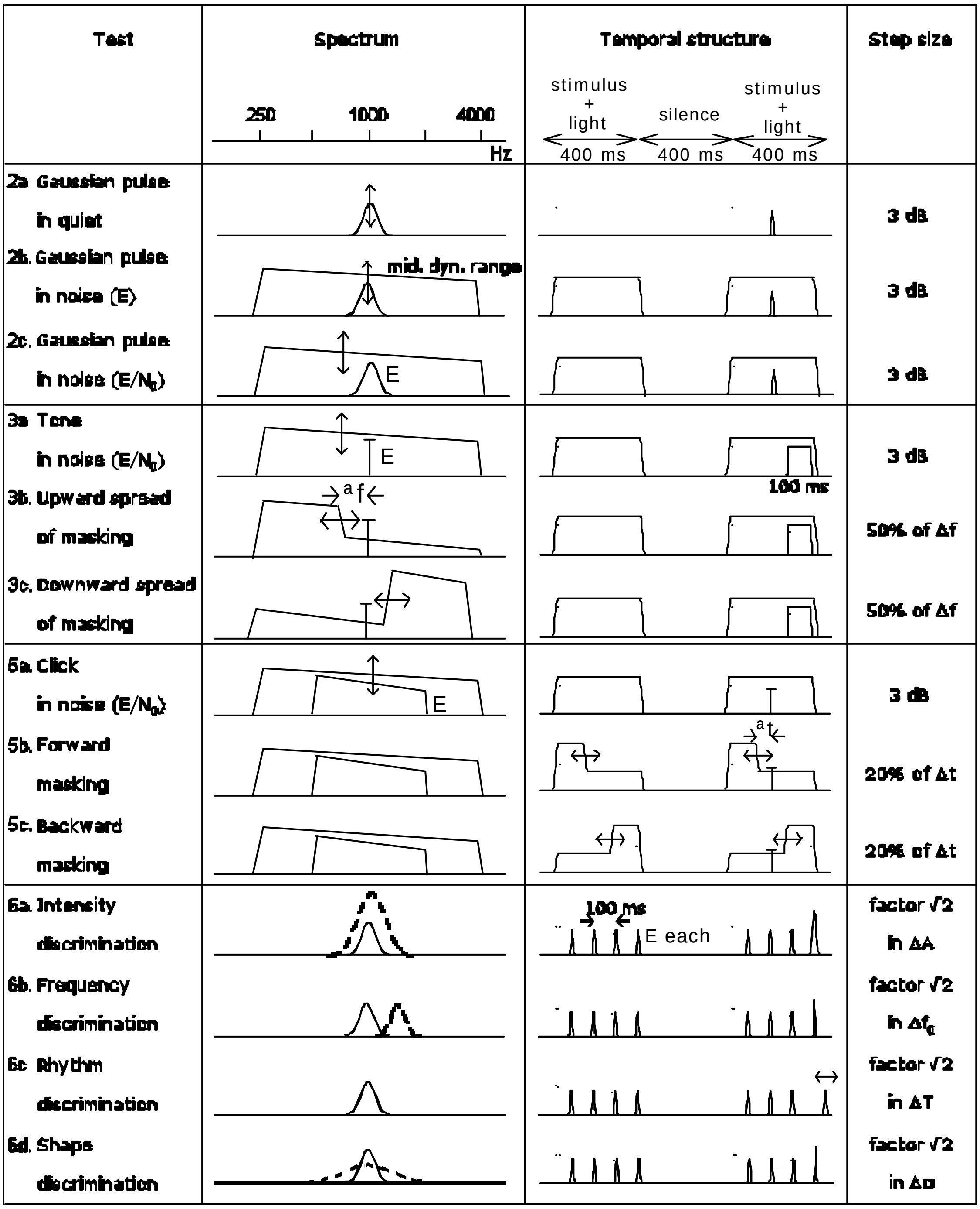


Fig. 4.1. Schematic representation of spectrum (second column) and temporal structure (third column) of the stimuli in the auditory-function tests as given in the first column. The parameter varied in the adaptive procedure is indicated with an arrow, or with a dashed line. The step size is given in the last column.

5a) or the geometric mean (for the other tests) of the last 12 reversals. A practice run with four reversals was presented before each measurement, except in session 3.

In test 2b of session 1, the energy E of the signals (tone, click, and Gaussian pulse) to be used in all subsequent tests was determined. E was chosen as the threshold energy of the Gaussian pulse in a pink noise with a bandwidth of four octaves (250-4000 Hz). The level of the pink noise was chosen so that the **a**-octave band centered at 1 kHz was presented halfway the hearing threshold and the uncomfortable loudness level of each individual listener. Consequently, the level of the pink masking noise in the subsequent detection tests was varied around the middle of the dynamic range for each listener.

Signal-detection efficiency (block 2)

First, the threshold for a Gaussian-windowed tone pulse was measured in quiet (2a). This test was included to be able to calculate the sensation level of the signals in the subsequent discrimination tests. Next, the threshold for a Gaussian pulse was measured in pink noise (2b). This test was included to determine the probe-signal energy E to be used in all subsequent tests. The threshold was defined as the energy level[1] (in dB) at the detection threshold. In these tests (2a and 2b), the level of the Gaussian pulse was varied in the adaptive procedure. Finally, the threshold for a Gaussian pulse in pink noise was determined again (2c), but now the noise level was varied in the adaptive procedure. In tests 2b and 2c, the Gaussian pulse was temporally centered in the masker. The threshold in test 2c was defined as the energy of the Gaussian pulse at detection threshold, normalized by the spectral density of the pink noise at 1 kHz, i.e., E/N_0 (in dB).

Spectral resolution (block 3)

To determine the spectral resolution without confounding it with signal-detection efficiency (Patterson *et al.*, 1982), upward spread of masking (USOM) and downward spread of masking (DSOM) were determined in two steps. In the first step the threshold for the tone in pink noise (3a) was determined. The noise level was varied in the adaptive procedure. The threshold was defined as the energy of the tone at the detection threshold, normalized by the spectral density of the noise at 1 kHz. In the second step, two measurements were performed, one to measure USOM (3b), the other to measure DSOM (3c). The USOM-masker was constructed from the noise at the threshold level of the first step of this block. The low-frequency part of the noise was amplified with 10 dB, and the high-frequency part was attenuated with 10 dB (see Fig. 4.1, 3b). To make the DSOM-masker the high-frequency part of the noise was amplified 10 dB, and the low-frequency part was attenuated 10 dB (see Fig. 4.1, 3c). In these tests, the attenuated parts of the noise are included to limit off-frequency listening. The frequency difference ª f (in Hz) between the tone and the masker cut-off was varied in the adaptive procedure. The threshold was defined as the frequency difference at the detection threshold. The USOM-masker and DSOM-masker may generate a tonal sensation, because of the steep spectral slope between the low- and high-frequency part of the noise. Therefore, the tone was presented at the end of the masker to ease detection.

Temporal resolution (block 5)

To determine the temporal resolution without confounding it with signal-detection efficiency, forward and backward masking were determined in two steps. First, the threshold for the click in pink noise (5a) was measured. The noise level was varied in the adaptive procedure. The threshold was defined as the energy of the click at the detection threshold, normalized by the spectral density of the noise. In the second step, forward masking (5b), and backward masking (5c) were measured. The forward masker was constructed from the noise at threshold level in step 5a, by amplifying the first part of the noise by 10 dB, and by attenuating the last part by 10 dB (see Fig. 4.1, 5b). For the backward masker, the last part of the noise was amplified 10 dB, and the first part was attenuated 10 dB (see Fig. 4.1, 5c). The interval ª t (in ms) between the click and transition in the masker was varied in the adaptive procedure. The threshold was defined as the interval at the detection threshold. The attenuated part of the noise limits “off-time listening” (Moore *et al*., 1988).

Discrimination (block 6)

In each measurement trial, the two intervals contained four Gaussian-windowed tone pulses repeated every 100 ms (see Fig. 4.1). In one randomly chosen interval, all four pulses were identical (reference pulses). In the other interval, the first three pulses were equal to the corresponding reference pulses and the fourth pulse deviated. The size of the difference in the fourth, odd, pulse (in intensity, frequency, latency, or shape) was varied in the adaptive procedure.

In test 6a, intensity discrimination was measured. The amplitude factor A in Eq. (4.1) was increased from A_0 to $A_0 + \Delta A$, and ΔA was varied in the adaptive procedure. The threshold was defined as the just-noticeable difference in energy (in dB), i.e., 20 $\log((A_0 + \Delta A)/A_0)$. In test 6b, frequency discrimination was measured. The carrier frequency of the odd pulse was 1000 Hz + Δf_0. The frequency difference Δf_0 (in Hz) was varied in the adaptive procedure. The threshold was defined as the just-noticeable difference in frequency between the odd and the reference pulse (in percent). In test 6c, rhythm discrimination was measured. The latency between the third reference pulse and the odd pulse was 100 ms + ΔT, and the delay ΔT was varied in the adaptive procedure. The threshold was defined as the just-noticeable increase in delay (in ms). In test 6d, shape discrimination was measured. The shape factor of the odd pulse was $\alpha_0 + \Delta\alpha$ (i.e., the duration of the fourth pulse was shorter and its bandwidth was wider). The shape-factor difference $\Delta\alpha$ was varied in the adaptive procedure. The threshold was defined as the just-noticeable difference in shape factor between the odd and the reference pulse (in percent).

Integration

The temporal and spectral integration of each listener was calculated by combining the masked thresholds for the Gaussian pulse (test 2c), the tone (test 3a), and the click (test 5a). We defined the temporal integration as the masked threshold for the 100-ms tone subtracted from the masked threshold of the 2.8-ms Gaussian pulse. Spectral integration was defined as the masked threshold for the click (two octaves wide) subtracted from the masked threshold for the Gaussian pulse (1/4 octave wide). With these definitions, maximally efficient integration corresponds to 0 dB and less efficient integration corresponds to negative numbers. The click presented in pink noise approximately fulfills the condition of equal detectability across the frequency range covered by the signal, necessary for accurately measuring spectral integration (van den Brink, 1990a).

4.2.5 SII calculations

The SII values corresponding to each SRT and SR*B*T measurement were calculated following the ANSI **a**-octave band procedure (ANSI, 1997). The SII replaces the older articulation index (ANSI, 1969). The band-importance function for speech material of average redundancy (Pavlovic, 1987) was used. For frequencies below 500 Hz, shallower slopes were used in the calculation of upward spread of masking, because previous calculations showed that, in an SR*B*T measurement, the standard SII model underestimates the upward spread of masking produced by the low-frequency noise (chapter 2).

For the SII procedure, the levels of speech and noise were transformed to equivalent free-field levels using the "artificial-ear-to-free-field transfer function" as derived in chapter 2. The spectrum level of the internal noise was calculated as the observed **a** -octave band noise threshold (transformed to free-field level) minus the critical ratio in dB.

4.3 RESULTS AND DISCUSSION

4.3.1 Results of sessions 1 and 2

The individual results of the listeners on the speech intelligibility tests and the auditory-function tests are given in appendix B, Tables B.2 and B.3. Each result is the average of the thresholds in sessions 1 and 2. Table 4.1 presents the average results for the normal-hearing listeners (column "NH") and the hearing-impaired listeners (column "HI"). The average threshold on test 2b has been calculated from just the thresholds measured in session 1, because these thresholds determined the signal energy E to be used in all subsequent tests. For eight tests, data analysis was performed on the logarithms of the thresholds, and, therefore, geometric means are presented in Table 4.1 (with the logarithms of the geometric means, and their corresponding standard deviations and individual standard errors, given in parentheses). The numbers of hearing-impaired listeners over which the averages have been calculated are given in column "number HI," because some hearing-impaired listeners could not perform one or two tests. Additionally, the thresholds of one hearing-impaired listener on the three tests of block 3 (and therefore also on temporal integration) were left out of the data analysis, because they were considered to be severe outliers[2]. The upper limit of the one-tailed 95%-confidence interval of the thresholds of the normal-hearing listeners was chosen as

Table 4.1. Average results of sessions 1 and 2 for normal-hearing listeners (NH) and hearing-impaired listeners (HI). Thresholds for the group of hearing-impaired listeners that differ significantly (p<0.05) from corresponding normal-hearing thresholds are indicated with an asterisk. “Number HI” indicates the number of hearing-impaired listeners over which the average has been calculated. “HI deviant” indicates the number of hearing-impaired listeners with a deviant threshold (i.e., a threshold beyond the one-tailed 95%-confidence limit for a normal threshold). Columns ó$_{NH}$ and ó$_{HI}$ present the standard deviation among listeners. In the last column the average individual standard error is shown. For eight tests, data analysis was performed on the logarithm of the individual thresholds. For these test, the values given in columns “NH” and “HI” represent geometric means, with the logarithm given in parentheses.

Test	dimension	NH	HI	number HI	HI deviant	ó$_{NH}$	ó$_{HI}$	individual SE
Speech reception (block 4)								
SRTq	dBA	22.7	56.0*	24	22	4.2	15.7	2.7
SRTn	dB S/N	-4.8	0.6*	23	23	0.9	2.5	0.8
SRTa	dB S/N	-5.0	-2.0*	23	17	0.7	2.3	1.0
SR*B*T	octaves	1.35	2.07*	23	19	0.14	0.52	0.15
Auditory functions								
2a. Gaussian pulse in quiet	dB	-3.4	31.8*	24	23	3.2	13.5	3.9
2b. Gaussian pulse in noise (E)	dB	42.0	53.3*	24	11	6.7	8.8	2.3
2c. Gaussian pulse in noise (E/N_0)	dB	16.7	13.6*	24	3	2.7	4.2	1.7
3a. Tone in noise (E/N_0)	dB	9.3	14.2*	23	21	1.1	2.8	1.6
3b. Upward spread of masking	Hz	142 (2.15)	213* (2.33)	23	9	(0.18)	(0.27)	(0.09)
3c. Downward spread of masking	Hz	65 (1.81)	63 (1.80)	23	4	(0.13)	(0.30)	(0.22)
5a. Click in noise (E/N_0)	dB	15.1	15.2	24	3	2.6	2.7	1.3
5b. Forward masking	ms	10.5 (1.02)	24.8* (1.39)	23	7	(0.23)	(0.30)	(0.11)
5c. Backward masking	ms	13.5 (1.13)	19.7 (1.29)	20	5	(0.19)	(0.28)	(0.07)
Temporal integration	dB	7.4	-0.8*	23	20	2.2	5.0	1.9
Spectral integration	dB	1.7	-1.6*	24	13	1.6	3.9	1.8
6a. Intensity discrimination	dB	3.9 (0.59)	2.8 (0.45)	24	1	(0.16)	(0.24)	(0.10)
6b. Frequency discrimination	%	3.1 (0.49)	6.2* (0.79)	23	10	(0.19)	(0.34)	(0.10)
6c. Rhythm discrimination	ms	25 (1.40)	26 (1.41)	24	1	(0.20)	(0.21)	(0.11)
6d. Shape discrimination	%	27 (1.43)	40 (1.60)	22	8	(0.21)	(0.26)	(0.17)

Table 4.2. Results on the auditory-function tests measured in session 3 for 12 normal-hearing listeners. Each test was performed at three levels relative to the level E in sessions 1 and 2: E-15 dB, E, and E+15 dB. Pairs of thresholds that differ significantly (p<0.05) are indicated with asterisks or daggers.

Test	dimension	E-15	E	E+15
2c. Gaussian pulse in noise (E/N_0)	dB	11.4^{r}	15.4^{r}	14.3
3a. Tone in noise (E/N_0)	dB	7.7	7.7	9.7
3b. Upward spread of masking	Hz	92$^{r\wedge}$	152^{r}	172$^{\wedge}$
3c. Downward spread of masking	Hz	57	70	50
5a. Click in noise (E/N_0)	dB	14.3	14.5	15.8
5b. Forward masking	ms	11.3	8.2	11.3
5c. Backward masking	ms	10.0	8.9	13.4
Temporal integration	dB	3.8^{r}	7.7^{r}	4.6
Spectral integration	dB	-2.8^{r}	0.9^{r}	-1.5
6a. Intensity discrimination	dB	5.5$^{r\wedge}$	3.2^{r}	2.3$^{\wedge}$
6b. Frequency discrimination	%	2.4	2.4	2.9
6c. Rhythm discrimination	ms	21	21	18
6d. Shape discrimination	%	29	22	20

the boundary between a normal and a deviant threshold (for temporal and spectral integration: lower limit). The number of hearing-impaired listeners with a deviant threshold is given in column "HI deviant." The standard deviation among listeners is given in columns "ó$_{NH}$" and "ó$_{HI}$." The individual standard error (test-retest), which represents a reliability measure of the test, is given in the last column. Average thresholds of the hearing-impaired listeners that differ significantly from those of the normal-hearing listeners are indicated with an asterisk (t test for unequal variances, p<0.05). Note that the average detection threshold for a Gaussian pulse in noise expressed as E/N_0 (test 2c) was significantly *better* for the hearing-impaired than for the normal-hearing listeners.

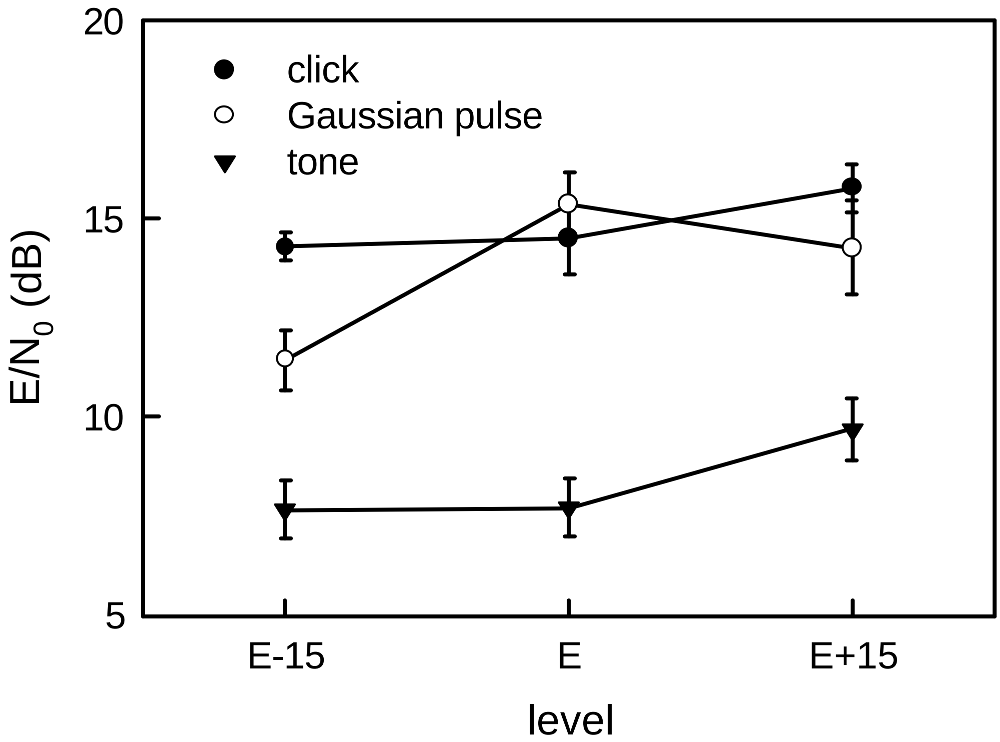


Fig. 4.2. Mean masked threshold for click, Gaussian pulse, and tone for 12 normal-hearing listeners at three levels 15 dB apart. Thresholds are expressed as the signal energy relative to the spectral density of the pink noise at 1 kHz. Vertical bars represent the standard error of the mean.

4.3.2 Level effects (session 3)

Table 4.2 presents the thresholds of the normal-hearing listeners on the auditory-function tests at three presentation levels: E-15, E, and E+15 dB. The intermediate level E was equal to the level used in sessions 1 and 2. For each test, the presence of a level effect was investigated with three Bonferroni-corrected *t* tests for matched samples ($p<0.05$ for each set of three comparisons; Keren and Lewis, 1993). Pairs of thresholds that differed significantly are indicated with asterisks or daggers. To compare the detection efficiencies for the different stimuli, the average masked thresholds for the tone, the click, and the Gaussian pulse are also presented in Fig. 4.2. Error bars show the standard error of the mean.

The signal energy E in sessions 1 and 2 differed over listeners (determined by the result of test 2b, with the noise level set at the midline of the individually measured dynamic range). For the normal-hearing listeners, the energy level E ranged from 32 dB to 52 dB. For the hearing-impaired listeners, E ranged from 36 dB to 72 dB. This range approximately overlaps the combined range of E and E+15 for the normal-hearing listeners in session 3. As can be seen in Table 4.2, no significant differences were found between the thresholds at E and at E+15 for the normal-hearing listeners. Therefore, it seems unlikely that differences in test results between the hearing-impaired and normal-hearing listeners are a consequence of the different absolute levels at which the tests were performed.
Many auditory functions are known to decline at low sensation levels. This may also cause differences in performance on the discrimination tests between normal-hearing and hearing-impaired listeners. As Table 4.2 indicates, intensity discrimination was the only discrimination ability that was worse at E-15 than at higher sensation levels for the normal-hearing listeners. However, on average intensity discrimination was better for the hearing-impaired than for the normal-hearing listener. Therefore, it appears that the lower sensation levels per se do not provide an explanation for the observed differences between the hearing-impaired and normal-hearing listeners.

4.3.3 Comparisons with the literature

Signal-detection efficiency

For the normal-hearing listeners, the detection threshold for the Gaussian pulse in noise is significantly higher at the intermediate level than at the lowest level (see Table 4.2 and Fig. 4.2). For the click and the tone, the thresholds did not significantly change with level. This is in agreement with results of Festen and Dreschler (1988) who found that the threshold for brief Gaussian-windowed tone pulses is higher at intermediate levels, and that this effect disappears both for click-like signals (duration shorter than 0.5 ms) and for signals longer than 5 ms. Van den Brink and Houtgast (1990a) did not observe a level effect in the detection of short tones, but a level-dependence of masked thresholds was also reported by Carlyon and Moore (1986), von Klitzing and Kohlrausch (1994), and Oxenham *et al.* (1997). In the last two studies, an explanation is provided in terms

of the level-dependent compressive nature of the input-output function of the basilar membrane.
At the intermediate level in our experiment, the average energy level of the Gaussian pulse was 42 dB. This corresponds to a peak amplitude of 67 dB SPL. At this sound-pressure level, the basilar membrane transfer characteristic is strongly compressive. Thus, a relatively higher sound-pressure level is needed for detection. The average masked detection threshold for the Gaussian pulse was significantly *better* for the hearing-impaired than for the normal-hearing listeners. This finding is consistent with the idea that cochlear impairment can lead to a loss of compression. Thus, because the masked threshold of the Gaussian pulse appears to be closely related to peripheral compression, it constitutes a poor measure for signal-detection efficiency.

Integration

Van den Brink and Houtgast (1990b) investigated temporal and spectral integration in masked signal detection. They found that listeners can efficiently integrate the energy of narrow-band signals (< **a** octave) over time, and the energy of short-duration signals (<30 ms) over frequency. This means that the total energy of those signals determines the masked detection threshold.
The normal-hearing listeners in our study showed even lower masked thresholds for the tone than for the Gaussian pulse, especially at the intermediate level. As mentioned above, the higher masked threshold for the Gaussian pulse at the intermediate level is probably a consequence of peripheral compression. The fact that the tone was presented at the end of the masker, while the Gaussian pulse was temporally centered in the masker, may also have caused a difference in detectability. The hearing-impaired listeners showed on average less temporal integration (as defined in this study) than the normal-hearing listeners, probably as a result from loss of peripheral compression, in agreement with data from Oxenham *et al*. (1997).
On average the hearing-impaired listeners showed less spectral integration than the normal-hearing listeners, but for both groups of listeners the spectral integration is close to 0 dB. This means that both groups of listeners can integrate the energy of the click across frequency, in agreement with the efficient spectral integration for short-duration signals reported by van den Brink and Houtgast (1990b). The difference in spectral

integration between both groups of listeners is caused by their different masked thresholds for the Gaussian pulse; their masked thresholds for the click did not differ significantly.

Spectral resolution

In Tables 4.1 and 4.2, the upward spread of masking (USOM) and downward spread of masking (DSOM) are expressed as the frequency difference (in Hz) at the detection threshold between the cut-off frequency of the masker and the frequency of the tone. The spread of masking (in Hz) can be converted to a slope (in dB/octave):

$$Slope\ (in\ dB/octave) = \frac{10\ \log(2)}{\log(1000 \pm SOM) - \log(1000)}\ , \qquad (4.2)$$

where SOM is the amount of USOM or DSOM (in Hz). USOM is subtracted from 1000, and DSOM is added to 1000. For the normal-hearing listeners, the average USOM was 152 Hz at the intermediate level (Table 4.2), corresponding to a slope of -42 dB/octave. The steepness of the slope increases with decreasing level. The slope of DSOM is about 100 dB/octave, independent of the level. These findings agree with the masking curves measured by Zwicker (1963).

Of the 23 hearing-impaired listeners, 11 listeners demonstrated increased spread of masking: either increased USOM (seven listeners), increased DSOM (two listeners), or both (two listeners). For hearing-impaired listeners, increased USOM has often been reported, and increased DSOM has occasionally been reported in the literature (for a review, see Tyler, 1986).

Temporal resolution

In the present study, forward masking and backward masking were measured for masker levels just 10 dB above the threshold level for simultaneous masking of the same probe signal. This resulted in intervals between the click and the transition in the masker of about 10 ms at the detection threshold for normal-hearing listeners. For these conditions, the thresholds for forward and backward masking did not depend significantly on the presentation level. This result differs from literature data for larger delays between masker and probe signal for which forward and backward masking decreases as a

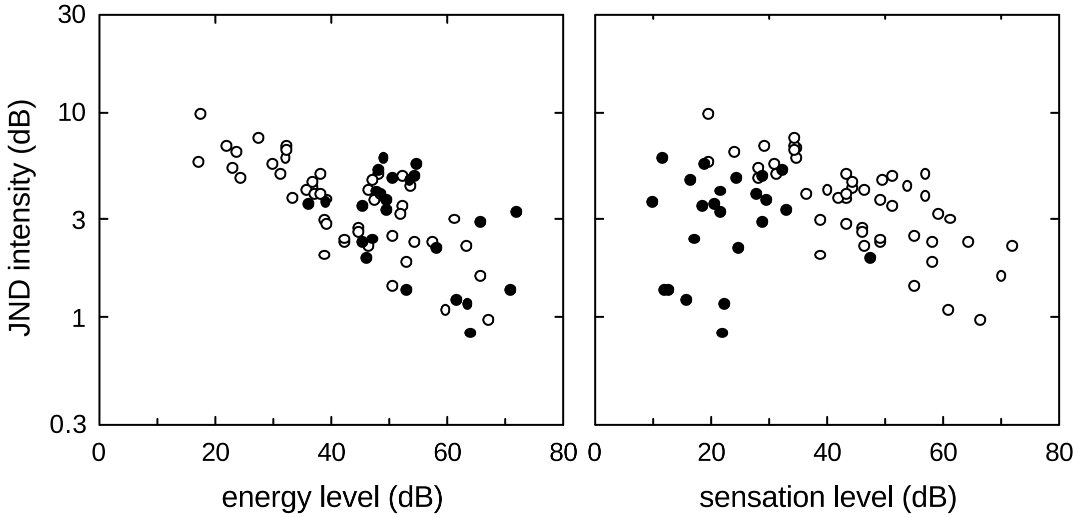


Fig. 4.3. Just-noticeable differences in intensity for Gaussian pulses versus the energy level of the pulses (left panel) and versus the sensation level of the pulses (right panel) for normal-hearing listeners (open circles) at three levels (E-15, E, and E+15) and hearing-impaired listeners (filled circles) at one level (E).

function of masker level. This level-dependence of nonsimultaneous masking becomes more pronounced with increasing masker-signal interval (Penner, 1974; Weber and Green, 1979; Oxenham and Moore, 1995). Therefore, the small intervals at threshold in the present study are probably the reason that no level effect was found in nonsimultaneous masking. On average, the hearing-impaired listeners in this study showed more forward masking than the normal-hearing listeners, in agreement with Festen and Plomp (1983), and Oxenham and Moore (1995).

Intensity discrimination

In Fig. 4.3 the just-noticeable difference (JND) in intensity is plotted against energy level (left panel) and sensation level (right panel) of the Gaussian-windowed tone pulses. The filled circles represent the thresholds of the hearing-impaired listeners averaged over sessions 1 and 2. The open circles represent the JND of the normal-hearing listeners. Four data points are shown for each normal-hearing listener: the average JND of sessions 1 and 2, and the three individual JNDs measured at the three levels in session 3.

Intensity discrimination deteriorates as stimulus duration decreases (see Florentine, 1986). Therefore, the JNDs of the normal-hearing listeners in Fig. 4.3 for the brief tone pulses used in the present study are higher than commonly reported for stimuli of longer duration. The JNDs are in the same range as those found by van Schijndel *et al.* (1999), who also used Gaussian-windowed tone pulses. Figure 4.3 shows that the JND at 1 kHz decreases with increasing signal level, consistent with Rabinowitz *et al.* (1976), and Florentine *et al.* (1987). When compared at equal sensation level, the JNDs of most hearing-impaired listeners are better than those of the normal-hearing listeners, but at equal absolute level, this difference in performance vanishes. This is consistent with Turner *et al.*, (1989), Florentine *et al.* (1993), and Schroder *et al.* (1994), although some listeners in their studies did show higher JNDs at equal absolute levels.

Frequency discrimination

As the duration of a tonal signal decreases, frequency discrimination deteriorates (Moore, 1973; Hall and Wood, 1984; Freyman and Nelson, 1986). Therefore, the JNDs of the normal-hearing listeners obtained with the 2.8-ms Gaussian pulse are poorer than is normally reported in the literature for longer-duration stimuli. On average, the hearing-impaired listeners performed poorer than the normal-hearing listeners. This is in contrast with results by Hall and Wood (1984), and Freyman and Nelson (1987), who found that performance of most of the hearing-impaired listeners was normal for brief tones, and poorer than normal for tones of longer durations.

4.3.4 Influence of audibility on speech intelligibility

The speech intelligibility index (ANSI, 1997) has been designed to predict speech intelligibility for normal-hearing listeners, and for hearing-impaired listeners without suprathreshold deficits. For normal-hearing listeners, all conditions of equal intelligibility should result in the same SII. Thus, when the measured SRT and SR*B*T of the normal-hearing listeners are expressed in SII values, the model should yield identical SII values.

For the normal-hearing listeners, the average SII values for the SRTq, SRTn, SRTa, and SR*B*T are 0.19, 0.26, 0.29, and 0.30, respectively. The SII_{SRTq} is significantly lower than the SII for the other conditions (paired *t* tests with Bonferroni correction, $p<0.05$; Keren and Lewis, 1993). This was also found in our previous study in which we assumed that

Table 4.3. Average speech intelligibility index with standard deviation in parentheses for the normal-hearing listeners (NH) and the hearing-impaired listeners (HI), for four intelligibility tests. The SII was calculated according to two different procedures. Modifications with respect to the standard SII procedure (ANSI, 1997) are given in the second column. 'Slopes': Shallower slopes of the masking curves are used below 500 Hz (see Sec. 4.2.5). 'Noise': The internal noise level is lowered by 4.6 dB (see Sec. 4.3.4). 'USOM, DSOM': The individually measured upward and downward spread of masking is included (see Sec. 4.3.6). Results for the hearing-impaired listeners that differ significantly (*t* test for unequal variances, $p<0.05$) from the corresponding results for the normal-hearing listeners are indicated with an asterisk.

	Modifications SII	SII_{SRTq}	SII_{SRTn}	SII_{SRTa}	SII_{SRBT}
NH	slopes, noise	0.29 (0.11)	0.27 (0.02)	0.30 (0.02)	0.29 (0.03)
HI	slopes, noise	0.27 (0.10)	0.30* (0.06)	0.32* (0.05)	0.39* (0.09)
NH	slopes, noise, USOM, DSOM	0.29 (0.11)	0.27 (0.04)	0.30 (0.04)	0.29 (0.04)
HI	slopes, noise, USOM, DSOM	0.27 (0.10)	0.27 (0.06)	0.30 (0.04)	0.36* (0.07)

Békésy tracking resulted in hearing thresholds that were systematically a few dB higher than the methods on which the ISO (1961) threshold is based (chapter 3).

When the internal noise level in the SII calculation, which represents the hearing threshold, was lowered by 4.6 dB, the differences among the four SII values were minimized. The upper two rows of Table 4.3 show the average SII values (and standard deviations) calculated with the lowered internal noise level. For the hearing-impaired listeners (row "HI"), the average SII_{SRTq} is similar to the average SII for normal-hearing listeners (row "NH"), but for the SRTn, SRTa and SR*B*T, the average SII for the hearing-impaired listeners is significantly higher than the average SII for the normal-hearing listeners (*t* test for unequal variances, $p<0.05$). An elevated SII indicates that the hearing-impaired listeners needed more speech information than the normal-hearing listeners to understand 50% of the sentences. Therefore, an elevated SII suggests that suprathreshold deficits affected intelligibility. Consistent with the results in chapter 3, the effect of suprathreshold deficits is most obvious for the SR*B*T, less obvious for the SRTa, and absent for the SRTq. In contrast with our previous results, the effect of suprathreshold deficits now is also present in the SRTn condition. In our previous study, we suggested that the dependence of the effect of suprathreshold deficits on condition may be related

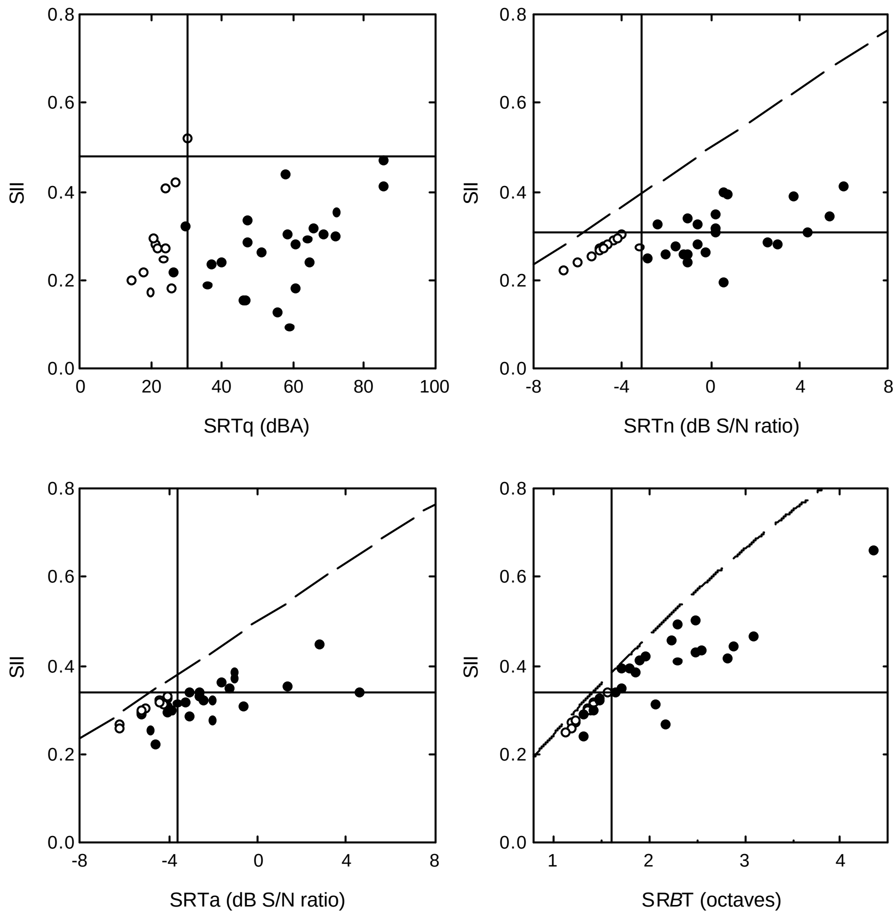


Fig. 4.4. Speech intelligibility index versus speech-reception threshold on the four intelligibility tests (SRTq, SRTn, SRTa, SR*B*T) for normal-hearing listeners (open circles) and hearing-impaired listeners (filled circles). Solid lines represent the one-tailed 95%-confidence limit for the data of the normal-hearing listeners. Dashed lines represent the maximum SII as a function of SRT or SR*B*T.

to the different spectra of speech and noise for the different conditions: when, for instance, a hearing-impaired listener suffers from excessive spread of masking, the effect on speech intelligibility will depend on the spectra of speech and noise.

Table B.2 of Appendix B shows the individual SII values for the four intelligibility tests. For each intelligibility test, the SII of individual listeners is shown in Fig. 4.4 as a function of the threshold. Data points of normal-hearing listeners are indicated with open circles and those of hearing-impaired listeners with filled circles. Dashed lines represent the maximum possible SII, i.e., the SII that would have been reached if the audibility of the speech had not been influenced by the hearing threshold, upward spread of masking and level distortion. (Since the maximum possible SII_{SRTq} is one, no dashed line has been drawn in the upper left graph). The upper limit of the one-tailed 95%-confidence interval of the SII of normal-hearing listeners is chosen as the separation between normal and higher-than-normal SII (horizontal lines). The separation between normal and deviant SRT or SR*B*T is indicated with a vertical line.

The horizontal and vertical lines divide the graphs in four quadrants. No data points fall in the upper left quadrant: a normal threshold never corresponds to an elevated SII. Data points in the lower left quadrant correspond to a normal threshold and a normal SII. Data points in the lower right quadrant correspond to an elevated SRT or a broader-than-normal SR*B*T; these result can be explained on the basis of audibility by the SII model. Data points in the upper right quadrant correspond to a deviant threshold that cannot be explained by the SII model. The higher the position of the data point above the horizontal line, the larger the speech perception deficit. The upper left graph shows that the SII model can account for all differences in SRTq between hearing-impaired listeners and normal-hearing listeners. The other three graphs show that audibility can account for some, but not all, deviant speech thresholds found for the other tests. Therefore, in these conditions, the effects of suprathreshold deficits appear to influence speech intelligibility for part of the hearing-impaired listeners.

4.3.5 Relations among the tests

Relations among auditory functions

The data of the combined groups of hearing-impaired and normal-hearing listeners form a continuum of thresholds on each test. Therefore, pooling the data of hearing-impaired and normal-hearing listeners will not bias the values of the correlation coefficients between each pair of tests. Table 4.4 shows correlation coefficients among 11 auditory functions for the combined group of all listeners. The masked thresholds for tone and click are not included, because they are not independent from (the included) temporal integration and spectral integration. Missing data were deleted casewise. Thirty complete

Table 4.4. Matrix of correlation coefficients among 11 auditory-function tests for 18 hearing-impaired listeners and 12 normal-hearing listeners. Underlined coefficients are significant at the 5% level; double-underlined coefficients are significant at the 1% level. Correlation coefficients were calculated on the average results of sessions 1 and 2, except for the coefficients in italics[3].

Test	3b USOM	3c DSOM	5b FM	5c BM	TI	SI	6a ID	6b FD	6c RD	6d SD
2c. Gaussian pulse in noise (E/N_0)	-0.21	0.05	-0.54	-0.23	*0.69*	*0.44*	0.23	0.05	-0.06	-0.05
3b. Upward spread of masking		*0.11*	0.27	0.01	*-0.38*	-0.24	-0.65	0.15	-0.17	-0.05
3c. Downward spread of masking			0.13	-0.03	*-0.12*	0.21	-0.38	0.01	-0.29	-0.21
5b. Forward masking				*0.60*	-0.73	*-0.28*	-0.23	0.43	0.30	0.52
5c. Backward masking					-0.43	*-0.02*	0.21	0.50	0.50	0.48
Temporal integration						*0.47*	0.39	-0.27	-0.01	-0.32
Spectral integration							0.14	-0.11	-0.04	-0.23
6a. Intensity discrimination								0.07	0.49	-0.01
6b. Frequency discrimination									0.29	0.53
6c. Rhythm discrimination										0.36
6d. Shape discrimination										

cases remained: 18 hearing-impaired listeners and 12 normal-hearing listeners. Of the 55 correlation coefficients, 13 are significant at the 1% level (double underlined), and five are significant at the 5% level (underlined). Within the group of hearing-impaired listeners, none of the auditory functions correlated significantly with age or absolute presentation level (not shown in the table). Correlation coefficients were calculated from the mean thresholds of test and retest, except nine correlation coefficients in italics, which were calculated from the individual results[3].

Because several auditory functions are correlated, a principal-components analysis was applied on the matrix of correlations to reduce the 11 auditory functions to a smaller set of factors. The resulting factors are linear combinations of the auditory functions. Only factors with eigenvalues greater than one were considered significant. This resulted in three factors that can account for 68% of the total variance. A varimax rotation was

Table 4.5. Factor loadings on the first three factors from a principal-components analysis on 11 auditory-function tests for the combined groups of 18 hearing-impaired listeners and 12 normal-hearing listeners. The bottom row shows the proportion of the total variance accounted for by each factor. For each test, the highest factor loading has been underlined.

Test	Factor 1	Factor 2	Factor 3
2c. Gaussian pulse in noise (E/N_0)	-0.10	-0.10	<u>0.83</u>
3b. Upward spread of masking	0.04	<u>0.63</u>	-0.35
3c. Downward spread of masking	0.07	<u>0.72</u>	0.32
5b. Forward masking	<u>0.72</u>	0.26	-0.47
5c. Backward masking	<u>0.84</u>	-0.11	-0.10
Temporal integration	0.43	0.37	<u>0.71</u>
Spectral integration	0.02	-0.07	<u>0.79</u>
6a. Intensity discrimination	0.12	<u>-0.85</u>	0.27
6b. Frequency discrimination	<u>0.78</u>	0.07	0.08
6c. Rhythm discrimination	0.56	<u>-0.57</u>	-0.04
6d. Shape discrimination	<u>0.73</u>	-0.13	-0.15
Proportion of variance explained	26%	20%	22%

performed to obtain the factors that were clearly marked by high loadings for some auditory functions. Table 4.5 shows the normalized factor loadings and the proportion of variance accounted for by each factor. Factor loadings can be interpreted as correlations between the auditory-function tests and the auditory factors. They are important for the interpretation of the auditory factors. The highest factor loading has been underlined for each test. Forward and backward masking, frequency discrimination, and shape discrimination have the highest loadings on auditory factor 1. Upward and downward spread of masking, intensity discrimination and rhythm discrimination have the highest loadings on auditory factor 2; the loadings of upward and downward spread of masking are positive, whereas the loadings of intensity and rhythm discrimination are negative, corresponding to positive and negative correlations, respectively. The auditory functions that load high on factor 3 are the masked threshold for the Gaussian pulse and

Table 4.6. Matrix of correlation coefficients between on the one hand thresholds on the four intelligibility tests and corresponding speech intelligibility indices and on the other hand 11 auditory functions, and the three factors from a principal-components analysis on these auditory functions. Correlation coefficients were calculated for the combined groups of 24 hearing-impaired listeners and 12 normal-hearing listeners, with pairwise deletion of missing data. Underlined coefficients are significant at the 5% level; double-underlined coefficients are significant at the 1% level.

	SRTq	SRTn	SRTa	SR*B*T	SII_{SRTq}	SII_{SRTn}	SII_{SRTa}	SII_{SRBT}
2c. Gaussian pulse in noise (E/N_0)	-0.44	-0.47	-0.22	-0.43	-0.03	-0.06	-0.08	-0.35
3b. USOM	0.40	0.52	0.49	0.55	0.19	0.36	0.49	0.57
3c. DSOM	0.17	-0.07	-0.06	0.33	0.23	0.36	0.22	0.46
5b. Forward masking	0.68	0.56	0.63	0.61	0.15	0.55	0.25	0.47
5c. Backward masking	0.28	0.36	0.24	0.15	0.00	0.32	0.07	0.07
Temporal integration	-0.76	-0.73	-0.66	-0.69	0.00	-0.42	-0.23	-0.57
Spectral integration	-0.36	-0.51	-0.12	-0.42	0.11	0.03	0.09	-0.37
6a. Intensity discrimination	-0.45	-0.27	-0.47	-0.44	-0.17	-0.24	-0.35	-0.45
6b. Frequency discrimination	0.42	0.49	0.59	0.43	0.08	0.66	0.38	0.43
6c. Rhythm discrimination	0.15	0.05	0.15	0.04	0.19	0.25	0.15	-0.03
6d. Shape discrimination	0.47	0.35	0.19	0.48	0.22	0.18	0.08	0.31
Factor 1	0.46	0.48	0.36	0.39	0.19	0.47	0.27	0.32
Factor 2	0.40	0.32	0.30	0.56	0.22	0.45	0.46	0.65
Factor 3	-0.54	-0.60	-0.55	-0.44	0.04	-0.02	0.01	-0.29

temporal and spectral integration. A reduction of the masked threshold for the Gaussian pulse and reduced integration are probably caused by a loss of peripheral compression (Sec. 4.3.3). Thus, auditory factor 3 is probably related to peripheral amplitude compression.

Relations between speech reception and auditory functions

Table 4.6 shows the correlation coefficients between, on the one hand, thresholds of the four intelligibility tests and corresponding speech intelligibility indices, and, on the other hand, 11 auditory functions and the three factors that resulted from a principal-components analysis on the 11 auditory functions.

Correlation coefficients were calculated for the combined groups of 24 hearing-impaired listeners and 12 normal-hearing listeners with pairwise deletion of missing data. Thirty correlations between thresholds on the four intelligibility tests (SRTq, SRTn, SRTa, and SR*B*T) and auditory functions are significant, but only 16 correlations between SII values and auditory functions are significant (double underlined: $p<0.01$; underlined: $p<0.05$). Apparently, for half of the correlations between intelligibility thresholds and auditory functions, the underlying cause is hearing threshold elevation: most auditory functions correlate significantly with hearing loss, and hearing loss correlates significantly with intelligibility thresholds (not shown).

Most significant correlations between SII and auditory functions are positive, suggesting that a suprathreshold speech perception deficit (i.e., elevated SII) may be caused by an elevated threshold on the auditory-function test in question. The SII correlates negatively with the threshold for the Gaussian pulse in noise, temporal integration, spectral integration (in case of the SII_{SRBT}), and intensity discrimination. It seems very unlikely that an elevated SII may be caused by *better* intensity discrimination. Table 4.4 shows that intensity discrimination has a significant negative correlation with both USOM and DSOM. Intensity discrimination, USOM, and DSOM did not correlate significantly with presentation level for the hearing-impaired listeners (not shown). Therefore, presentation level cannot be the underlying cause for the correlation between spread of masking and intensity discrimination. The role of spread of excitation in intensity discrimination may explain the correlation between spread of masking and intensity discrimination. The spread of excitation toward the higher and, to a lesser extent, the lower frequencies is important for intensity discrimination (Viemeister, 1972; Moore and Raab, 1974). Thus, intensity discrimination may profit from increased USOM or DSOM, as suggested by Florentine *et al.*, (1993) and Schroder *et al.* (1994). Table 4.6 shows that USOM has a significant positive correlation with SII_{SRTa}, and both USOM and DSOM have a significant positive correlation with SII_{SRBT}. Therefore, spread of masking may be the underlying cause for the negative correlation between intensity discrimination and SII_{SRTa} and SII_{SRBT}.

While auditory factor 1 only correlates significantly with SII_{SRTn}, auditory factor 2 correlates significantly with the SII values for three tests: the SRTn, SRTa, and SR*B*T. Upward and downward spread of masking, and intensity discrimination have the highest loadings on factor 2. This suggests that reduced spectral resolution is the most important cause for suprathreshold speech perception deficits. Speech intelligibility was probably affected by increased spread of masking by the noise. Furthermore, reduced spectral resolution may have interfered with the detection and discrimination of spectral features such as formants.

4.3.6 Including measured spread of masking in the SII model

The individual SII values were recalculated to investigate whether the underlying cause for the correlation between the SII and spectral resolution is the increased spread of masking by the noise. In this recalculation, the measured USOM was used instead of the SII model values for normal hearing, and the measured DSOM was used instead of no DSOM. For normal-hearing listeners, the slope of DSOM is level and frequency independent (Zwicker, 1963). Therefore, the measured slope of DSOM was used for all frequencies and levels in the SII calculations. USOM however, is level and frequency dependent. Therefore, the ratio between the measured USOM and SII-model slopes is applied as "*USOM factor*" to calculate the individual USOM for all levels and frequencies in the SII model. For hearing-impaired listeners this is probably not correct, but it is presumably closer to the real upward spread of masking than the SII model values for normal hearing.
In the standard SII model, the spectrum level for masking (Z_i) in ⅓-octave band i is calculated as

$$Z_i = 10\log\left(10^{0.1N_i} + \sum_{k=1}^{i-1} 10^{0.1[B_k + 3.32\,C_k \log(0.89\,F_i/F_k)]}\right) \qquad (4.3)$$

where N is the noise spectrum level, B is the larger of the noise spectrum level and the self-speech masking spectrum level, C is the slope of USOM, and F is the ⅓-octave band center frequency. The individually measured USOM and DSOM are included in the SII model by calculating the spectrum level for masking as:

$$Z_i = 10\log\left(10^{0.1N_i} + \sum_{k=1}^{i-1} 10^{0.1[B_k + 3.32(USOM\ factor)C_k \log(0.89F_i/F_k)]} + \sum_{k=i+1}^{18} 10^{0.1[B_k - 3.32C_D \log(0.89F_k/F_i)]}\right) \quad (4.4)$$

where C_D is the measured slope of DSOM.

The two bottom rows of Table 4.3 show the average SII values recalculated with the measured spread of masking for the normal-hearing (NH) and the hearing-impaired (HI) listeners. The SII_{SRTq} is not influenced by the modification, because no noise was present in this condition. For the normal-hearing listeners, average SII values for SRTn, SRTa, and SR*B*T do not change, but the standard deviations are larger. The average SII values for the hearing-impaired listeners are lower than without the measured spread of masking. The average SII_{SRTn} and SII_{SRTa} now are equal to those of the normal-hearing listeners. The average SII_{SRBT} is still significantly higher (p<0.05) than the SII_{SRBT} of the normal-hearing listeners, although it is closer to the normal SII_{SRBT}.

From a comparison of SII calculations with just the measured USOM and with just the measured DSOM (not shown), it can be concluded that the effect of including DSOM is negligible. Thus, the effect of including measured spread of masking shown in Table 4.3 originates primarily from including the measured USOM.

For normal-hearing listeners, the SII predictions with the measured USOM and DSOM included in the model are less accurate (larger standard deviations in Table 4.3). On the other hand, the average SII values of the hearing-impaired listeners are closer to those of the normal-hearing listeners. The change in variance in SII for the hearing-impaired and normal-hearing listeners taken together is chosen as a measure of the net effect of including the measured spread of masking. The total variance in SII_{SRTn} and SII_{SRTa} did not change compared to the original SII, but the total variance in SII_{SRBT} decreased by 46%. Therefore, it is concluded that incorporating the measured spread of masking improves the SII predictions for the SR*B*T. This may be related to the fact that USOM was measured with the stimulus at 1 kHz, and that only in the SR*B*T measurement, the speech signal was restricted to the 1-kHz frequency region.

When the measured spread of masking is included in the model, the correlation coefficient between SII_{SRBT} and USOM[4] reduces to a non-significant 0.21. This indicates that the main underlying cause for the correlation between SII_{SRBT} and USOM is that increased upward spread of masking by the noise reduces the speech range that contributes to intelligibility. With the measured spread of masking applied in the model, the correlation coefficient between SII_{SRBT} and DSOM[4] (0.35) remains significant (p<0.01). Therefore, the main underlying cause for the correlation between SII_{SRBT} and DSOM probably lies in reduced spectral resolution preserving less spectral detail of the speech signal, and thus reducing intelligibility.

4.3.7 Predicting suprathreshold speech perception from auditory functions

Multiple regression

Stepwise multiple regression was used to predict the SII [with shallower slopes of the masking curves below 500 Hz (see Sec. 4.2.5) and adjusted internal noise level (see Sec. 4.3.4)] for the four speech-intelligibility tests from the results on the auditory-function tests. Because several auditory functions are correlated, not the individual auditory functions, but the three (uncorrelated) factors extracted from the 11 auditory functions (Table 4.5), were used as predictor variables. An “auditory factor” was only included in the regression when it could account for a significant part of the variance (p<0.05). SII_{SRTq} did not correlate significantly with any of the auditory factors. For the SII values corresponding to the other intelligibility thresholds, the following regression equations were obtained:

$$SII_{SRTn} = 0.021 \cdot factor2 + 0.021 \cdot factor1 + 0.28; \qquad R^2=0.43 \qquad (4.5)$$

$$SII_{SRTa} = 0.018 \cdot factor2 + 0.31; \qquad R^2=0.21 \qquad (4.6)$$

$$SII_{SRBT} = 0.059 \cdot factor2 + 0.029 \cdot factor1 - 0.025 \cdot factor3 + 0.36 \qquad R^2=0.62 \qquad (4.7)$$

where R^2 is the variance accounted for by the regression equation. Factor 2 explains 22% of the variance in SII_{SRTn}, 21% in SII_{SRTa}, and 42% in SII_{SRBT}. Factor 1 explains an additional 21% of variance in SII_{SRTn}, and an additional 11% of variance in SII_{SRBT}. Finally, factor 3 explains an additional 9% of the variance in SII_{SRBT}.

Auditory factor 2 is related to spectral resolution (see Sec. 4.3.5). Therefore, reduced spectral resolution appears to be the most important cause of suprathreshold speech perception deficits. When factor 2 is accounted for, factor 1 can further reduce the variance in SII_{SRTn} and SII_{SRBT}. Auditory functions with a high loading on factor 1 were forward and backward masking, frequency discrimination, and shape discrimination. Auditory factor 3, probably related to peripheral compression, seems least important for suprathreshold speech perception. The negative contribution of factor 3 in Eq. (4.7) means that loss of compression results in a higher SII_{SRBT}.

The total variance accounted for by the auditory factors is largest for SII_{SRBT}. This was expected, because all auditory factors were measured at 1 kHz, i.e., the center frequency in the SR*B*T test. The auditory factors can account for a greater part of the variance in SII_{SRTn} than in SII_{SRTa}. A possible explanation is that the frequency region around 1 kHz is more important in the SRTn test than in the SRTa test. In the SRTa test, the speech signal is presented above threshold in the entire frequency range from 250 Hz to 4000 Hz. Because most hearing-impaired listeners had a high-frequency hearing loss, the high-frequency part of the speech signal was below their threshold in the SRTn test. This increases the relative importance of the 1-kHz frequency region in speech perception.

Individual approach

In this section, we will try to determine for each individual listener which deteriorated auditory function affects speech intelligibility in the 1-kHz frequency region. Figure 4.4 shows that 19 listeners had a normal SII_{SRBT} (12 normal hearing and 7 hearing impaired) and that 16 hearing-impaired listeners had an elevated SII_{SRBT}. The auditory functions in the group of 19 listeners with no speech perception deficit in the 1-kHz frequency region apparently are adequate for normal speech perception. The thresholds on the auditory-function tests can be summarized by three factors. For factors 1 and 2, the upper one-tailed 95%-confidence limit of the auditory factors in the listener group with a normal SII_{SRBT} is chosen as the separation between sound processing sufficient for speech perception, and sound processing that may cause a deficit in speech perception. For auditory factor 3, the *lower* one-tailed 95%-confidence limit is chosen as separation, because a lower score on factor 3 is probably related to the sound processing deficit "loss of compression".

Of the 16 listeners with an elevated SII, six listeners had a deviant score on factor 1, six listeners had a deviant score on factor 2, and only one listener had a deviant score on

Table 4.7. Individual SII values for the SR*B*T and thresholds on six auditory functions with a high loading on factor 1 (forward masking, backward masking, and frequency discrimination) or factor 2 (upward spread of masking, downward spread of masking, and intensity discrimination) for the 16 hearing-impaired listeners with speech-perception deficits in the 1-kHz frequency region (i.e., elevated SII_{SRBT}). The row "95% limit" gives the one-tailed 95%-confidence limit for the data of listeners with a normal SII_{SRBT} (lower limit for ID, which has a negative loading on factor 2; upper limit for other tests). Bold thresholds lie beyond this 95%-confidence limit. Empty cells mean that the listener could not perform this test. An asterisk indicates that the listener has an elevated factor score. An asterisk in parenthesis indicates that the factor scores could not be calculated, because the listener could not perform on one or two tests that load high on this factor.

	SII_{SRBT}	Factor 1				Factor 2			
		FM (ms)	BM (ms)	FD (%)	elevate factor	USOM (Hz)	DSOM (Hz)	ID (dB)	elevate factor
95% limit	0.34	30	33	8.2		331	125	2.1	
	0.35	25	19	3.0		88	61	3.0	
	0.38	25	**47**	**11.4**	~	248	**157**	4.7	
	0.39	11	9		(~)	291	28	3.7	
	0.39	26	**38**	8.0	~	138	49	4.9	
	0.41	**116**	**61**	5.1	~	314	56	2.2	
	0.41			**8.6**	(~)	118	47	2.9	
	0.41	28	25	3.5		**464**	107	2.4	~
	0.42	20	21	**8.9**	(~)	**338**	61	5.1	
	0.43	28	27	**14.2**	~	113	67	3.5	
	0.43	**31**	20	2.4		**345**	**204**	**1.2**	~
	0.44	**33**	18	7.0	~	**456**	48	**1.1**	~
	0.45	15	9	2.5		186	**161**	**0.8**	~
	0.47	**71**		**16.8**	(~)	271	98	5.9	
	0.49	**36**	**53**	**11.7**	~	**407**	89	2.3	
	0.50	19	11	5.0		**348**	**175**	3.2	~
	0.66	23	9	4.2		**483**	99	**1.3**	~

factor 3. Therefore, factor 3 seems to be least important for suprathreshold speech perception (as was also found by multiple regression in the previous section). For the 16 hearing-impaired listeners with an elevated SII_{SRBT}, the individual thresholds on six auditory functions that load high on either factor 1 or factor 2 are presented in Table 4.7. Thresholds printed in boldface lie beyond the one-tailed 95%-confidence limit for the data of listeners with a normal SII. Thresholds for rhythm and shape discrimination are not shown in the table, because on each test only one listener had a JND above the 95%-confidence limit for normal SII. Elevated scores on factor 1 or factor 2 are indicated with an asterisk. Some listeners have a slightly deviant threshold on one test, while their corresponding factor score is still normal, because they have normal thresholds on the other tests that also load high on this factor. For four listeners in Table 4.7, the factor scores could not be calculated, because they could not perform one or two auditory-function tests. This probably indicates that the auditory function that we tried to measure is very poor. Therefore, the factor on which these tests had a high loading are assumed to be elevated (asterisks in parentheses in Table 4.7). Table 4.7 shows that the speech perception deficit of individual hearing-impaired listeners seems to be related either to factor 1, or to factor 2. Only one listeners deviates on both factors, and one listener deviates on neither. For nine listeners, the speech perception deficit in the 1-kHz frequency region may be explained by reduced temporal resolution and/or reduced frequency discrimination (factor 1). For five listeners, reduced spectral resolution (factor 2) seems responsible for the speech perception deficit.

The results of the individual approach and those of the multiple regression analysis both point to factors 1 and 2 as the most important determinants of speech perception. However, the multiple-regression results identify factor 2 as the major factor, while the individual approach identifies factor 1 as the major factor. This can be partly explained from the data of the four hearing-impaired listeners with an elevated SII who could not perform all auditory-function tests. These four listeners could not be included in the multiple regression. In the individual approach, factor 1 was considered responsible for their speech perception deficit. Another difference between the two approaches that may contribute to the differences between their results is that different assumptions underlie each approach: in the individual approach we assumed that an elevated SII can be caused by one elevated factor score, independent of the other factor scores, while in multiple regression it is assumed that an elevated SII is caused by an elevation of a linear combination of the factors. A third difference is that we tried to identify the sound processing deficit for the largest possible *number* of hearing-impaired listeners in the individual approach, while multiple regression is based on explaining as much as possible of the *variance* in the SII values.

In conclusion, factor 2 (reduced spectral resolution) can explain the largest part of the variance in intelligibility of suprathreshold speech. However, to understand each individual speech perception deficit in the 1-kHz frequency range, both factor 2 (reduced spectral resolution) and factor 1 (reduced temporal resolution and/or reduced frequency discrimination) have to be taken into account.

4.4 SUMMARY AND CONCLUSIONS

Performance of 24 hearing-impaired listeners and 12 normal-hearing listeners was measured on four speech-intelligibility tests and on a set of auditory-function tests concerning detection efficiency, resolution, discrimination and integration. Intelligibility was measured in four conditions: SRT in quiet, SRT in noise (SRTn), spectrally adapted SRT in noise (SRTa), and the novel SR*B*T (chapter 2). The SR***BT*** is defined as the bandwidth of speech (center frequency: 1 kHz) in complementary notched noise required for a 50% intelligibility score.

All individual intelligibility thresholds were converted to speech intelligibility indices (SII), with the assumption that the SII includes the effect of audibility. The SRT in quiet could be explained from audibility for all listeners. However, deviant thresholds on the other intelligibility tests could not all be accounted for by audibility. A higher-than-normal SII required for a 50% intelligibility score was assumed to indicate the effect of suprathreshold deficits. Including the measured upward spread of masking in the SII model substantially improved the SII predictions for the SR*B*T, but it did not improve the SII predictions for the two SRT tests in noise. The reason for this probably lies in the fact that upward spread of masking was measured at 1 kHz, i.e., in the frequency region that is most relevant for the SR*B*T measurement.

The SR*B*T test has been developed as a research tool to identify hearing-impaired listeners suffering from a deficit in speech perception in the 1-kHz frequency region. All auditory-function tests were measured at 1 kHz. As expected, the SII_{SRBT} was most closely related to the auditory functions: 62% of the total variance in SII_{SRBT} could be explained by the thresholds on the auditory-function tests, whereas this was 43% and 21% for SII_{SRTn} and SII_{SRTa}, respectively.

The thresholds on the auditory-function tests were subjected to a principal-components analysis. This resulted in three “auditory factors”: the first factor was related to temporal resolution and frequency discrimination, the second factor was associated with spectral resolution, and the third factor was associated with peripheral compression. The auditory factor associated with spectral resolution, and the auditory factor related to temporal resolution and frequency discrimination both seemed important for suprathreshold speech perception. Reduced peripheral amplitude compression appeared least important for speech intelligibility. On an individual basis, each substantially elevated SII_{SRBT} (for 15 of the listeners) was found to be associated with a deviant value of the first and/or the second factor.

NOTES

[1]The energy level is the acoustic power times duration of the signal in decibels, defined as 10 log(E/E_0). The reference value E_0 is 10^{-12} Watt·s, which makes the numerical value of the energy level for a plane wave through a unit surface (1 m^2) during 1 second equal to the numerical value of the sound-pressure level.

[2]Severe outliers were defined as values that are more than three quartile ranges (the difference of upper and lower quartile) above the upper, or below the lower quartile.

[3]The thresholds, on which the italic correlation coefficients in Table 4.4 are based, are mutually dependent within a session. For example, the noise level used in the measurements of USOM and DSOM was determined by the masked threshold for the tone in the same session. If, due to a measurement error, the masker noise level was chosen too low, this resulted in an underestimate of USOM- and DSOM-thresholds. Therefore, if the correlation coefficient between USOM and DSOM was calculated from the mean thresholds, it would have been artificially enhanced, because of error variance in their common reference (the masked threshold for the tone). However, the test of USOM and the retest of DSOM are independent, as well as the test of DSOM and the retest of USOM. Therefore, the correlation coefficient between USOM and DSOM was calculated as the average of the correlation coefficient between the

USOM test and the DSOM retest, and the correlation coefficient between the DSOM test and the USOM retest. The other italic correlations were calculated similarly.

[4]Because the SII calculated when including the measured spread of masking in the model and the measured spread of masking itself are mutually dependent, the correlation coefficient between USOM and SII was calculated as the average of the correlation coefficient between USOM test and SII retest, and the correlation coefficient between USOM retest and SII test. The same holds for the correlation coefficient between DSOM and the SII calculated when including the measured spread of masking in the model.

Chapter 5

General discussion

5.1 ORIGIN OF REDUCED SPEECH INTELLIGIBILITY

In this thesis, the origin of reduced speech intelligibility in hearing-impaired listeners is divided into (1) reduced audibility and (2) suprathreshold deficits. Section 5.1.1 describes the way in which audibility affects speech intelligibility. Section 5.1.2 reviews mechanisms by which suprathreshold deficits may affect speech intelligibility.

5.1.1 Audibility

Here, audibility is defined as the effective portion of the speech dynamic range contributing to intelligibility, as calculated using the speech intelligibility index (SII). Audibility is affected when part of the relevant dynamic range of speech falls below the hearing threshold or below a masking noise, but audibility can also be affected by auditory-masking effects *as found in normal-hearing listeners*. These auditory-masking effects include upward spread of masking, self-masking of speech, and level distortion (i.e., the decrease of speech intelligibility at high presentation levels). Thus, reduced audibility in a hearing-impaired listener includes all those factors that would affect audibility in normal-hearing listeners under the same conditions when the hearing threshold of the hearing-impaired listener is simulated by a masking noise for the normal-hearing listeners.

5.1.2 Suprathreshold deficits

The effects of possible suprathreshold deficits on intelligibility are *not* included in the SII model. Therefore, when suprathreshold deficits affect intelligibility, the SRT or SR*B*T (50% sentence intelligibility) will correspond to a higher-than-normal SII. Figure 5.1 illustrates how audibility effects are separated from the effect of a suprathreshold deficit by using the SII model for one normal-hearing listener and two hearing-impaired listeners who participated in the experiments of chapter 4. The SRT in noise of the normal-hearing listener (open circle) and of one of the hearing-impaired listeners (filled circle) correspond to a normal SII. This means that the elevated SRT in noise of this particular hearing-impaired listener was entirely due to audibility effects. The elevated SRT in noise of the other hearing-impaired listener corresponds to a higher-than-normal

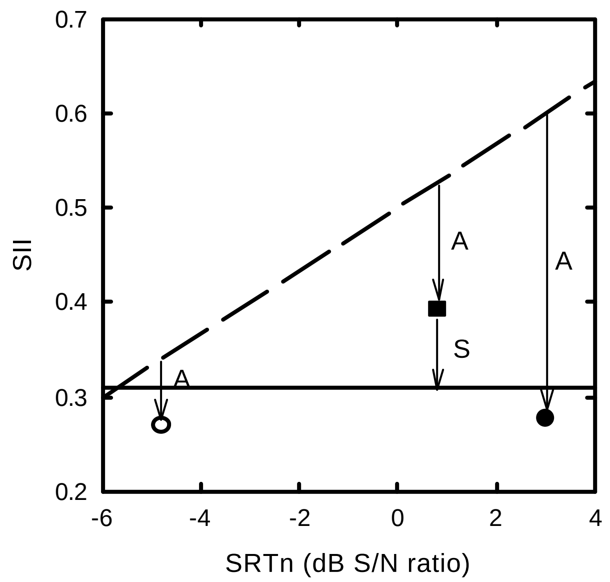


Fig. 5.1. SII versu ring-impaired listeners (filled symbols). The solid line represents the boundary between a normal and an elevated SII. The dashed line represents the growth of SII with signal-to-noise ratio if the whole dynamic range of the speech that is above the average noise level would be audible. The deviation of a data point from the dashed line, given by the distance “A,” indicates the proportion of the total amount of speech information lost due to reduced audibility, i.e., due to the hearing threshold and normal auditory masking. The elevation of a data point above the solid line, given by the distance “S,” indicates the amount of speech information lost due to a suprathreshold deficit.

SII (square). For this listener, intelligibility in the SRT test in noise was not only influenced by audibility, but also by suprathreshold deficits.

A suprathreshold deficit can be either of auditory or of cognitive origin. In chapter 4, it was found that each substantially elevated SII for the SR*B*T corresponded to an elevated auditory factor. Therefore, in these listeners, auditory factors seem sufficient to explain reduced intelligibility of suprathreshold speech. The results of chapter 4 suggest that the auditory deficits that are most detrimental to intelligibility of suprathreshold speech are reduced temporal resolution, reduced frequency discrimination, and reduced spectral resolution.

Reduced temporal resolution may cause difficulties in processing rapid changes in the speech signal, and in detecting brief silent intervals, thus making it hard to identify stop consonants. Furthermore, due to reduced temporal resolution, soft consonants may be masked by preceding or succeeding vowels that are much stronger.

Reduced frequency discrimination may interfere with speech intelligibility, because glides in formant frequency will be more difficult to detect. A formant glide is one of the cues in speech that signals the transition from one syllable to another. In addition, frequency discrimination is important for detecting changes in the pitch of a talker's voice. Pitch changes convey information about linguistic structure (Summerfield, 1987).

Reduced spectral resolution may affect intelligibility in noisy conditions due to increased susceptibility to spread of masking. In chapter 3, the SII calculations were modified with respect to the upward spread of masking. In the modified SII model upward spread of masking was taken to increase proportionally to the hearing loss, as proposed by Ludvigsen (1987). This did not improve the SII-model predictions. In chapter 4, the SII model was modified by including the actually *measured* upward spread of masking (USOM) and downward spread of masking (DSOM) for individual listeners. This did improve the SII-model predictions for the SR*B*T. This improvement originated mainly from including the measured USOM.

Also after the modification, the SII for the SR*B*T still correlates significantly with the measured DSOM. The underlying cause of this correlation may be that reduced spectral resolution not only leads to excessive auditory masking (as included in the modified SII model), but also is detrimental to speech intelligibility, because broadened auditory filters reduce the spectral contrasts in speech. When less spectral detail of the speech signal is preserved, the formant structure of vowels will be blurred.

5.2 SUPRATHRESHOLD DEFICITS AND HEARING LOSS

In chapter 3, the relation between effects on intelligibility of suprathreshold deficits and the size of the sensorineural hearing loss was investigated for 34 hearing-impaired listeners. By incorporating the results from chapter 4, the relation between the effects of suprathreshold deficits and hearing loss can be examined for a larger group of hearing-impaired listeners (see Fig. 5.2). An SII of about 0.3 is considered normal for a 50% intelligibility score for short sentences. Higher speech-intelligibility indices suggest that a suprathreshold deficit affected speech intelligibility. Figure 5.2 shows, even clearer than Fig. 3.4 in chapter 3, that the higher the average hearing loss, the more listeners

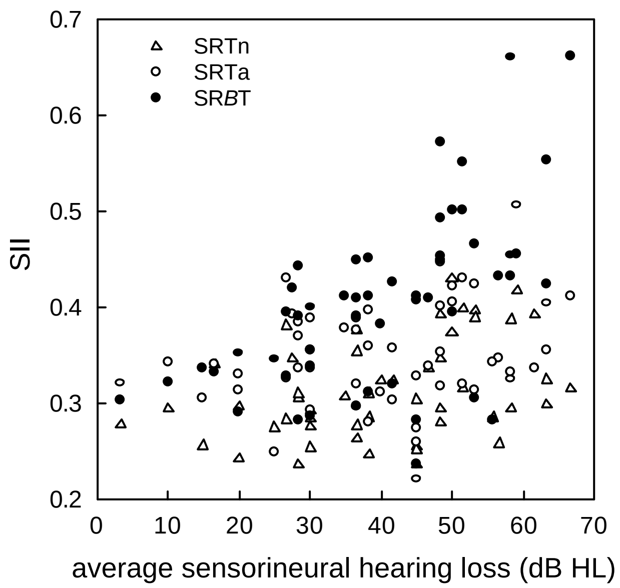


Fig. 5.2. Speech intelligibility index corresponding to the data of 53 hearing-impaired listeners on three intelligibility tests (SRTn, SRTa, and SR*B*T) versus their sensorineural hearing loss averaged over the frequencies 0.5, 1, and 2 kHz. For five listeners, who participated in two experiments (chapter 3 and 4), average results are presented.

suffer from a suprathreshold deficit (i.e., elevated SII). However, surprisingly also with a large loss normal SII values may occur. This leads to the important finding that predicting the effect of suprathreshold deficits on speech intelligibility from the hearing loss is not possible for individual listeners.

5.3 COMPENSATING FOR REDUCED INTELLIGIBILITY

5.3.1 Audibility

A hearing aid can compensate for reduced audibility due to part of the speech spectrum falling below the hearing threshold using frequency-selective amplification. In case of a conductive hearing loss, a hearing aid can fully compensate for reduced audibility by amplifying the sound by the amount it is attenuated by the hearing loss. In case of a sensorineural hearing loss, a good choice for the amplitude-frequency response of a hearing aid is one that puts the average everyday speech spectrum halfway between hearing threshold and uncomfortable loudness level (van Buuren *et al.*, 1995). This frequency shaping was applied in the SRTa test in chapters 3 and 4. In total, 32 of the 52 hearing-impaired listeners of which the SRTa was measured did not suffer from a suprathreshold deficit in this condition (i.e., they had a normal SII for the SRTa). Sixteen of those 32 listeners had a normal SRTa (see Fig. 3.2 and 4.4). This means that frequency-selective amplification could fully compensate the hearing loss in 50% of the hearing-impaired listeners without suprathreshold deficits in the SRTa condition. For the remaining 50% without a suprathreshold deficit in this condition, the applied frequency shaping could not fully compensate for the reduced audibility. This shows that in case of a sensorineural hearing loss, limitations exist to the degree to which a hearing aid can compensate for reduced audibility. One limitation is that high levels of amplification result in level distortion. Another limitation is that when the dynamic range of the hearing-impaired listener is smaller than 30 dB, it cannot contain the full relevant dynamic range of speech, because this would result in uncomfortably loud levels.

A hearing aid with fast-acting compression (syllabic compression) may compress the dynamic range of the speech into the smaller dynamic range of the listener. The compression can be either wideband or multiband. A disadvantage of wideband compression is that a strong signal in a limited frequency region suppresses weaker signals at other frequencies. A drawback of multiband compression is that it reduces the spectral contrasts in speech (Plomp, 1988). Furthermore, both types of compression reduce the temporal modulations in speech. These modulations are essential for intelligibility (Houtgast and Steeneken, 1985). Fortunately, modulation reduction by direct manipulation of the temporal envelope (like compression) is less detrimental than expected on the basis of the speech transmission index (Drullman, 1995; Noordhoek and Drullman, 1997). Compression with a factor of 2 still seems acceptable (van Buuren *et al.*, 1999).

5.3.2 Suprathreshold deficits

As already mentioned above, the frequency-shaping applied in the SRTa test is a good choice for the response of a hearing aid. Combining the results from chapter 3 and 4, it appears that in the SRTa condition 20 of the 52 hearing-impaired listeners had an elevated SII for the SRTa. Thus, almost 40% of the hearing-impaired listeners suffered from a suprathreshold deficit in this condition.

In the literature, many attempts have been described to compensate for suprathreshold deficits of hearing-impaired listeners. We will briefly discuss some attempts to compensate for the most important (according to the results of chapter 4) suprathreshold deficits: reduced temporal resolution, reduced frequency discrimination, and reduced spectral resolution.

Reduced temporal resolution results in increased susceptibility to forward and backward masking. Compression of the temporal envelope by a hearing aid reduces these effects, because it amplifies weak parts of the speech signal (consonants) relative to stronger parts (vowels). However, this positive effect may be canceled by the reduction in intelligibility caused by the degradation of temporal contrasts in speech.

As a result of *reduced frequency discrimination* glides in formant frequency or changes in the pitch of a talker's voice may be less well detected. No attempts have been found in the literature to exaggerate glides in formant frequency to compensate for reduced frequency discrimination. Grant (1987) tried to compensate for a reduced ability of profoundly hearing-impaired listeners to follow the pitch of a talker's voice. He investigated speechreading performance when information about the talker's fundamental frequency was provided as frequency-modulated sinusoids. In some conditions, the frequency variations of the sinusoid were exaggerated with respect to the talker's fundamental frequency. Speechreading with his type of representation of the fundamental frequency did not improve compared with speechreading as supplement to normal speech.

Increased susceptibility to upward spread of masking is one consequence of *reduced spectral resolution*. In conditions of seriously disturbing low-frequency noise, frequency-selective attenuation can improve speech intelligibility (van Dijkhuizen *et al*., 1991). For the degradation of spectral contrasts in speech due to broader auditory filters compensation is sought by spectral enhancement. This form of signal processing enhances the differences between the peaks and the valleys in the short-term speech spectrum. Some studies have shown modest benefits of spectral enhancement (for a review, see Moore, 1995). A fundamental limit to what can be reached by spectral enhancement is that the enhanced speech spectrum always has to pass the broadened auditory filters of the hearing-impaired listener. Another way to compensate for the degradation of the spectral contrasts in the speech signal is to reduce the masking of weak formants by adjacent intense formants by presenting adjacent formants to different ears or by moving the formants farther apart in frequency. Results show that some listeners might benefit from such formant representations (Rosen and Fourcin, 1986).

Compression and spectral enhancement have opposite effects: compression amplifies the weak parts of the speech signal, while spectral enhancement attenuates them. Therefore, it may seem surprising that compression is mentioned as a possible way to compensate for reduced temporal resolution, while spectral enhancement is mentioned as a possible way to compensate for reduced spectral resolution. The explanation for this paradox is that while the weak parts in the speech waveform (the consonants) contain important speech information, the weak parts in the speech spectrum (the valleys) do not. The important spectral information is provided by the peaks in the speech spectrum, which may become better detectable after spectral enhancement.

5.4 EVERYDAY LISTENING SITUATIONS

The speech material used in this thesis was chosen because it resembles conversational speech in real life. Therefore, the results of the intelligibility tests give insight into the problems encountered by hearing-impaired listeners in everyday speech communication. However, two aspects that are important in practice were not covered by our experiments, i.e., binaural listening and fluctuating interfering sounds.

Bronkhorst and Plomp (1992) investigated the effect of multiple speechlike maskers on speech perception in conditions simulating free-field situations. They concluded that the advantage of listening with two ears compared with best-ear listening is approximately 3 dB *for both normal-hearing and hearing-impaired listeners*. Therefore, it seems that experiments in wich the speech signal is presented to only one ear can still provide a realistic estimate of performance in binaural conditions.

In everyday situations, speech is often masked by speech from competing speakers. The temporal envelope of competing speech from a large group of people is rather constant, so steady-state noise, as used in our experiments, will be indicative of performance in such situations. However, competing speech from only one or two talkers has a strongly fluctuating envelope. For normal-hearing listeners, the SRT is substantially lower in fluctuating noise than in steady-state noise. For hearing-impaired listeners, the SRT does not substantially improve in fluctuating noise, possibly because of increased temporal masking (Festen and Plomp, 1990; Bronkhorst and Plomp, 1992; Bacon *et al.*, 1998). This means that reduced temporal resolution may be more detrimental to speech intelligibility in many practical situations than is suggested by the results of chapter 4. Therefore, investigating the relation between the SRT in fluctuating noise and temporal resolution in hearing-impaired listeners would be worthwhile.

Chapter 6

Summary

Listeners with sensorineural hearing losses often experience difficulties in understanding speech, even when the speech is presented well above their hearing thresholds. In this thesis, the origin of these difficulties was investigated by examining the relations between speech intelligibility and basic properties of the auditory system. To enhance the chances of obtaining clear correlations, the investigations were all performed in a limited frequency region around 1 kHz.

A novel intelligibility test (the SR*B*T test) to measure intelligibility of speech bandpass filtered with a fixed center frequency of 1 kHz was developed in chapter 2. In this test, the minimum speech bandwidth required for a 50% intelligibility score is determined (speech-reception *bandwidth* threshold or ***SRBT***). The narrowband speech is presented in complementary bandstop-filtered noise to ensure that the speech is only audible within the desired frequency band. The bandwidth of the speech signal is varied in an adaptive up-down procedure using a step size of a factor of 1.37 for the bandwidth (in Hz). On average, the SR*B*T of normal-hearing listeners is 1.4 octave (600-1600 Hz) under optimal conditions, i.e., when the entire speech dynamic range is above the hearing threshold, but not so loud that audibility is affected by excessive upward spread of masking.

In chapters 3 and 4, the performance of, in total, 22 normal-hearing listeners and 53 hearing-impaired listeners was measured using the SR*B*T test, as well as using more common broadband SRT (speech-reception threshold) tests, namely the SRT in quiet and the SRT in noise. For most hearing-impaired listeners, the scores on these intelligibility tests deviated from those of the normal-hearing listeners. The speech intelligibility index (SII) model was applied to separate the origin of an elevated SRT or a broader-than-normal SR*B*T into (1) reduced audibility, and (2) suprathreshold deficits. Reduced audibility includes all those factors that would affect audibility in normal-hearing listeners under the same conditions when the hearing threshold of the hearing-impaired listener is simulated by presenting a masking noise. The effect of suprathreshold deficits on speech perception is quantified by the additional amount of speech information needed for an intelligibility score of 50% (as compared to normal-hearing listeners operating in threshold-simulating noise). For short sentences, normal-hearing listeners require an SII of about 0.3 for a 50% intelligibility score. This means that normal-hearing listeners need 30% of the speech information to understand 50% of the sentences correctly. When the SII value that a hearing-impaired listener requires for a 50% intelligibility score is normal, it is assumed that a possible deviant SRT or SR*B*T is due only to inaudibility of a part of the speech spectrum. On the other hand, a higher-than-normal SII value indicates that speech intelligibility was affected by suprathreshold deficits. The results of chapters 3 and 4 show that the presence and size of the effects of suprathreshold deficits on speech perception depend on the type of intelligibility test. The SRT test in quiet shows the smallest

sensitivity to suprathreshold deficits in speech perception, while the SR*B*T test shows the largest sensitivity. Only a weak relation is observed between suprathreshold deficits and hearing loss. This shows that it is not possible to predict whether an individual listener suffers from a suprathreshold deficit from only a consideration of the size of the sensorineural hearing loss.

In chapter 4, the relations between suprathreshold speech-perception deficits and basic auditory functions were investigated. The auditory-function tests included detection efficiency, temporal resolution (i.e., forward and backward masking), spectral resolution (i.e., upward and downward spread of masking), temporal and spectral integration, and discrimination of intensity, frequency, rhythm, and spectro-temporal shape. All auditory functions were measured at or around 1 kHz. Because several of these auditory functions were correlated, the thresholds on the auditory-function tests were subjected to a principal-components analysis. This resulted in three uncorrelated "auditory factors" (i.e., linear combinations of the auditory functions). The first factor is related to temporal resolution and frequency discrimination, the second factor is associated with spectral resolution, and the third factor is associated with detection efficiency, and temporal and spectral integration.

Multiple regression was used to predict suprathreshold speech perception from the auditory factors for each of the intelligibility tests. The total variance accounted for by the auditory factors was largest for the results on the SR*B*T test (62%). This was in accordance with our expectations, because all auditory functions were measured around 1 kHz which is also the center frequency used in the SR*B*T test. The first and second auditory factors were most closely related to suprathreshold deficits in speech perception. This leads to the main conclusion of this thesis: the suprathreshold deficits that affect speech perception are (1) reduced temporal resolution, (2) reduced frequency discrimination, and (3) reduced spectral resolution.

Appendix

APPENDIX A:
INDIVIDUAL DATA FROM THE EXPERIMENTS IN CHAPTER 3

Table A.1. Hearing threshold and uncomfortable loudness level for 1/3-octave bands of noise, together with age, type of hearing loss (s = sensorineural; m = mixed), and broadband UCL attenuation (i.e., the attenuation needed to arrive at the broadband UCL with the broadband noise burst spectrally shaped according to the narrowband UCL) for 34 hearing-impaired listeners (listeners 'a' through 'H'). For listeners with a mixed hearing loss, the air-bone gap is given in parenthesis. Mean results for the normal-hearing listeners are given at the bottom of the table.

listener	age	type	hearing threshold (dB SPL)				
	(years)		0.25 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz
a	63	s	37.3	21.4	18.3	65.4	68.1
b	55	s	31.5	22.7	49.0	58.0	66.1
c	54	s	29.3	15.8	11.0	4.9	32.4
d	73	m	>80	>80 (30)	68.9 (40)	65.1 (15)	>80 (35)
e	49	s	59.9	59.4	62.8	55.2	57.8
f	65	s	45.0	43.1	46.6	59.5	60.3
g	58	s	72.7	52.5	47.7	45.1	>80
h	76	s	47.5	45.2	49.0	62.6	67.5
i	74	s	>80	62.7	57.6	48.9	55.6
j	54	m	>80	71.5 (55)	62.9 (55)	63.2 (40)	74.2 (35)
k	56	s	42.5	33.6	24.1	43.8	64.6
l	35	s	28.8	21.3	35.9	56.6	71.2
m	61	s	27.5	16.2	30.2	38.7	71.6
n	88	s	56.3	51.9	55.0	59.3	68.1
o	77	s	39.4	37.6	39.0	52.6	57.1
p	70	s	39.5	33.4	33.4	50.1	50.1
q	65	s	>80	72.2	62.5	54.4	61.9
r	45	s	55.7	58.5	50.6	53.9	67.7
s	53	m	66.3	55.5 (20)	49.8 (20)	24.7 (5)	23.7 (5)
t	85	s	56.8	53.5	61.0	69.0	73.4

UCL (dB SPL)					attenuation
0.25	0.5 kHz	1 kHz	2 kHz	4 kHz	(dB)
>134	>134	>134	>134	>134	0*
113.7	102.5	101.0	104.7	102.3	20
125.3	115.3	107.0	100.3	103.7	30
>134	127.3	>134	114.7	120.3	20
101.3	99.7	99.3	88.7	99.0	40
91.5	91.0	86.3	83.7	88.7	20
125.3	117.3	117.7	107.0	106.0	40
>134	131.3	129.3	131.0	>134	40
>134	119.2	119.0	110.0	103.3	30
>134	>134	>134	>134	>134	25
133.0	119.0	125.7	118.0	119.3	20
129.3	111.0	106.3	101.3	104.0	20
123.7	114.3	115.0	117.3	121.3	30
>134	132.0	128.3	130.2	>134	35
114.2	124.0	120.7	116.0	119.3	10
125.0	117.3	112.3	113.0	120.0	30
107.3	97.3	96.7	95.8	100.3	20
>134	>134	131.7	>134	>134	20
>134	123.8	131.0	124.3	128.0	20
126.3	117.0	113.7	114.0	121.0	10

listener	age (years)	type	hearing threshold (dB SPL) 0.25 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz
u	69	s	55.0	58.2	65.8	61.7	72.9
v	74	s	51.8	49.2	49.3	57.2	63.0
w	73	s	59.9	50.6	53.1	70.4	71.9
x	77	s	60.3	58.3	44.5	52.9	67.7
y	79	s	55.4	50.7	50.5	63.1	61.9
z	61	m	54.3	46.3 (15)	44.9 (30)	50.9 (30)	>80 (45)
A	50	s	45.5	52.1	55.2	55.8	57.1
B	65	s	31.0	18.2	25.1	42.1	40.0
C	70	s	>80	70.5	59.4	65.2	59.3
D	64	s	35.3	23.3	44.5	57.3	72.4
E	44	s	39.3	34.4	34.5	47.6	67.1
F	72	s	32.8	22.6	16.4	30.8	47.3
G	43	s	23.7	12.9	8.9	18.7	35.8
H	63	s	55.8	41.9	39.3	62.8	>80
mean normal	26		26.2	14.0	8.7	9.0	7.1

* For the first hearing-impaired listener, the extended procedure for determining the broadband UCL was not followed. For this listener the adapted spectra in the SRTa and SR*B*T test were positioned halfway between threshold and *narrowband* UCL. (This listener did not experience the resulting presentation level as uncomfortably loud.)

UCL (dB SPL)					attenuation
0.25 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz	(dB)
130.2	115.0	110.0	105.7	111.3	20
116.0	117.7	104.3	108.5	109.3	30
125.0	119.0	115.0	129.0	125.0	35
120.7	118.2	117.0	108.3	123.3	30
104.0	103.3	102.0	100.7	100.7	20
106.3	101.3	101.3	96.2	103.3	30
114.0	105.0	101.7	99.3	116.3	15
>134	125.7	113.7	111.7	107.7	20
110.3	103.0	104.7	107.0	103.7	20
126.8	117.0	111.7	117.3	>134	40
>134	128.0	119.3	125.3	128.3	30
102.7	95.0	88.7	86.8	88.0	30
106.0	100.7	95.0	95.4	100.7	10
97.6	91.8	101.3	92.0	112.0	30
121.1	112.4	108.4	107.5	104.8	15

Table A.2. Individual thresholds and corresponding Speech Intelligibility Indices of 34 hearing-impaired listeners (listeners 'a' through 'H') and 10 normal-hearing listeners (listeners '1' through '10') on four speech-intelligibility tests: SRT in quiet (SRTq), SRT in noise (SRTn), SRT in noise for speech with an adapted spectrum (SRTa), and the SR*B*T. Each threshold is the average of two measurements. The SII is calculated with the internal noise level lowered by 3.6 dB (see Sec. 3.5.3). The hearing-impaired listeners are categorized into four groups based on the SII corresponding to their SRTn, SRTa, and SR*B*T (Group I: normal SII for all tests; Group II: elevated SII for only the SR*B*T; Group III: normal SII for only the SRTn; Group IV: elevated SII for all tests).

listener	SRTq		SRTn		SRTa		SR*B*T		group
	(dBA)	SII	(dB)	SII	(dB)	SII	(oct.)	SII	
a	49.0	0.390	5.2	0.351	1.4	0.403	2.25	0.422	III
b	48.5	0.228	1.6	0.285	-2.8	0.359	2.14	0.451	II
c	29.5	0.394	-2.6	0.295	-3.2	0.342	1.49	0.322	I
d	77.5	0.150	-2.4	0.254	-3.6	0.336	1.62	0.339	I
e	68.0	0.207	1.6	0.389	2.6	0.424	3.63	0.306	
f	51.0	0.139	0.2	0.264	-3.4	0.320	1.72	0.297	I
g	70.0	0.440	3.8	0.374	1.4	0.405	2.63	0.500	III
h	59.5	0.229	1.8	0.310	-0.6	0.396	1.97	0.412	III
i	84.5	0.635	1.2	0.316	1.6	0.411	4.02	0.661	III
j	77.5	0.255	-3.4	0.296	-5.0	0.314	1.29	0.290	I
k	46.0	0.301	0.8	0.310	-1.2	0.384	1.32	0.283	
l	38.5	0.193	0.4	0.237	-1.8	0.370	2.02	0.442	III
m	36.0	0.215	-0.2	0.283	-3.4	0.327	1.49	0.329	I
n	79.5	0.486	4.0	0.392	2.2	0.446	2.90	0.572	IV
o	49.0	0.204	1.2	0.307	-1.2	0.378	2.02	0.411	III
p	52.5	0.334	0.2	0.354	-1.4	0.389	1.87	0.410	III
q	81.0	0.413	4.6	0.449	5.8	0.567	3.25	0.471	IV
r	61.5	0.186	-2.0	0.252	-4.6	0.259	1.58	0.283	I
s	55.5	0.387	-3.2	0.341	-3.6	0.340	1.53	0.333	I
t	74.5	0.313	1.4	0.325	2.0	0.404	3.04	0.553	III
u	66.5	0.147	0.6	0.294	-1.8	0.332	2.56	0.454	II
v	76.5	0.510	3.6	0.398	0.6	0.430	2.97	0.550	IV

listener	SRTq		SRTn		SRTa		SR*B*T		group
w	79.5	0.428	5.8	0.429	1.2	0.422	2.08	0.395	IV
	(dBA)	SII	(dB)	SII	(dB)	SII	(oct.)	SII	
x	64.5	0.295	-1.2	0.280	0.4	0.401	2.37	0.493	III
y	68.5	0.330	2.0	0.387	-2.6	0.325	2.19	0.433	
z	54.5	0.179	-1.6	0.242	-2.4	0.330	1.81	0.352	I
A	63.5	0.237	-2.2	0.294	-2.0	0.353	2.19	0.448	II
B	43.5	0.341	0.8	0.381	-0.6	0.430	1.81	0.395	III
C	83.0	0.429	0.6	0.298	-0.8	0.356	2.50	0.423	II
D	53.0	0.274	3.2	0.340	-1.0	0.381	1.97	0.406	III
E	43.5	0.173	0.6	0.293	-1.6	0.389	1.87	0.400	III
F	34.0	0.260	-0.8	0.284	-2.4	0.355	1.62	0.355	I
G	23.5	0.234	-0.6	0.298	-4.4	0.319	1.40	0.308	I
H	56.5	0.251	0.6	0.287	1.2	0.332	1.53	0.300	I
1	21.8	0.198	-4.2	0.289	-5.6	0.274	1.40	0.309	
2	24.4	0.437	-3.4	0.338	-5.0	0.304	1.40	0.308	
3	22.2	0.250	-2.0	0.367	-3.8	0.332	1.53	0.326	
4	25.9	0.197	-3.8	0.306	-2.4	0.361	1.72	0.360	
5	22.5	0.260	-2.6	0.349	-5.2	0.288	1.49	0.327	
6	24.1	0.343	-6.2	0.241	-5.2	0.285	1.29	0.284	
7	24.9	0.326	-2.8	0.330	-3.6	0.339	1.44	0.301	
8	24.7	0.394	-3.4	0.330	-5.0	0.299	1.29	0.275	
9	25.5	0.356	-3.6	0.337	-3.8	0.338	1.53	0.332	
10	24.1	0.312	-4.2	0.288	-5.0	0.305	1.32	0.293	

APPENDIX B:
INDIVIDUAL DATA FROM THE EXPERIMENTS IN CHAPTER 4

Table B.1. Hearing threshold and uncomfortable loudness level for 1/3-octave bands of noise, together with age, and broadband UCL attenuation (i.e., the attenuation needed to arrive at the broadband UCL with the broadband noise burst spectrally shaped according to the narrowband UCL) for the individual hearing-impaired listeners. Mean results for the normal-hearing listeners are given at the bottom of the table.

listener	age	hearing threshold (dB SPL)					UCL (dB SPL)					att.
	(years	0.25 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz	0.25 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz	(dB)
a	65	64.8	57.9	48.2	54.8	55.1	109.5	105.7	96.2	98.5	96.8	30
b	67	45.0	34.7	33.9	41.4	56.1	>120	>120	117.7	116.3	>120	40
c	62	31.0	22.8	38.6	72.8	>80	>120	112.0	106.8	103.5	104.2	20
d	65	32.3	31.8	44.8	52.8	73.4	114.0	106.3	101.0	107.5	112.8	30
e	58	76.0	70.2	57.4	65.6	74.5	109.5	100.5	84.7	90.5	88.5	*35*
f	63	59.9	51.8	48.4	50.7	49.1	>120	113.3	112.8	115.0	117.0	30
g	53	53.5	56.0	52.1	63.8	72.1	105.7	94.3	102.3	99.0	92.0	30
h	57	45.9	58.6	59.5	57.7	48.4	78.2	82.0	89.5	82.0	75.5	40
i	52	48.5	58.7	56.5	58.7	54.9	>120	>120	119.0	117.0	>120	20
j	62	25.9	22.8	30.9	39.1	64.0	>120	>120	116.7	116.8	108.7	15
k	58	60.6	56.7	52.4	61.8	39.5	108.7	104.8	99.3	107.2	98.8	35
l	39	27.4	28.7	29.6	66.3	50.0	>120	118.2	113.0	109.2	100.0	20
m	54	47.7	50.8	62.9	66.6	69.3	>120	>120	114.2	117.7	109.7	35
n	39	42.9	38.3	49.0	45.3	24.8	97.7	95.3	91.7	93.5	85.3	20
o	60	45.8	36.7	31.9	33.1	59.7	96.0	90.2	88.0	84.8	87.8	30
p	67	31.2	20.0	8.8	25.2	49.4	>120	119.8	106.0	107.7	114.2	40
q	45	27.7	20.4	12.5	31.3	47.8	>120	114.2	100.3	101.0	105.5	20
r	65	33.0	23.3	20.9	71.1	71.7	>120	>120	>120	>120	>120	20
s	64	60.3	49.3	47.2	74.9	>80	93.3	96.7	93.0	96.3	115.7	35
t	65	65.3	65.4	52.1	61.8	49.4	>120	104.5	93.8	98.5	80.2	30
u	67	79.9	73.7	64.0	66.8	67.5	119.0	111.7	113.3	113.3	113.3	30

listener	age	hearing threshold (dB SPL)					UCL (dB SPL)					att.
	(years	0.25 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz	0.25 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz	(dB)
v	64	60.8	54.9	60.8	58.1	64.9	>120	113.3	>120	>120	113.3	25
w	41	51.6	38.3	45.8	50.3	46.5	101.5	91.3	85.3	87.8	85.2	30
x	64	38.5	38.7	43.6	75.5	>80	119.0	108.5	101.8	111.3	113.3	30
mean normal	27	27.8	18.6	8.1	7.6	6.3	118.9	115.3	111.7	108.9	105.0	22

Table B.2. Individual thresholds and corresponding Speech Intelligibility Indices of 24 hearing-impaired listeners (listeners ‘a’ through ‘x’) and 12 normal-hearing listeners (listeners ‘1’ through ‘12’) on four speech-intelligibility tests : SRT in quiet (SRTq), SRT in noise (SRTn), SRT in noise for speech with an adapted spectrum (SRTa), and the SR*B*T. Each threshold is the average of two measurements. An empty cells means that the listener could not perform this test. The SII is calculated with the internal noise level lowered by 4.6 dB (see Sec. 4.3.4).

listener	SRTq		SRTn		SRTa		SR*B*T	
	(dBA)	SII	(dB)	SII	(dB)	SII	(oct.)	SII
a	66.0	0.316	-0.6	0.324	-1.6	0.358	2.50	0.426
b	47.5	0.283	0.2	0.306	-3.0	0.337	1.72	0.391
c	36.0	0.185	0.6	0.192	-0.6	0.304	1.49	0.319
d	51.5	0.258	6.0	0.410	-1.0	0.370	2.31	0.489
e	85.5	0.468					4.35	0.660
f	61.0	0.276	-1.0	0.337	-2.6	0.339	1.92	0.410
g	56.0	0.123	-1.0	0.237	-4.6	0.221	1.32	0.236
h	69.0	0.302	0.8	0.393	4.6	0.336		
i	65.0	0.237	0.2	0.347	-2.4	0.318	2.25	0.453
j	37.0	0.231	-0.6	0.276	-4.0	0.292	1.32	0.287
k	61.0	0.178	0.6	0.397	-3.2	0.314	3.11	0.465
l	40.0	0.237	3.0	0.277	-3.8	0.296	1.81	0.391
m	72.0	0.296	-0.2	0.258	-1.2	0.347	2.56	0.433
n	46.0	0.153	-2.4	0.323	-3.60	0.312	1.87	0.382
o	47.5	0.334	-1.6	0.275	-4.8	0.249	1.72	0.346
p	30.0	0.317	-2.0	0.256	-4.0	0.306	1.67	0.336
q	26.5	0.215	-1.0	0.257	-4.0	0.323	1.44	0.298
r	58.0	0.435	5.4	0.343	-1.0	0.383	1.97	0.418
s	64.5	0.289	2.6	0.282	1.4	0.351	2.19	0.266
t	59.0	0.090	-2.8	0.246	-3.0	0.281	2.08	0.311
u	86.0	0.411	3.8	0.387	2.8	0.444	2.90	0.440
v	72.5	0.351	0.2	0.316	-2.0	0.320	2.50	0.500
w	47.0	0.151	-1.2	0.255	-2.6	0.328	2.83	0.412
x	58.5	0.300	4.4	0.304	-2.0	0.274	2.31	0.407

listener	SRTq		SRTn		SRTa		SR*B*T	
	(dBA)	SII	(dB)	SII	(dB)	SII	(oct.)	SII
1	23.5	0.244	-5.4	0.249	-5.2	0.287	1.40	0.295
2	30.5	0.516	-4.8	0.275	-5.0	0.301	1.49	0.324
3	21.5	0.278	-4.4	0.285	-5.2	0.293	1.44	0.314
4	21.0	0.292	-5.0	0.268	-5.2	0.291	1.44	0.312
5	20.0	0.171	-6.6	0.219	-4.2	0.311	1.25	0.267
6	14.5	0.196	-4.0	0.300	-4.4	0.318	1.36	0.299
7	24.0	0.406	-4.2	0.290	-5.2	0.297	1.21	0.271
8	26.0	0.179	-3.2	0.271	-6.2	0.263	1.58	0.335
9	22.0	0.270	-5.0	0.264	-4.4	0.321	1.21	0.255
10	24.0	0.270	-4.6	0.277	-4.4	0.313	1.25	0.274
11	27.0	0.417	-4.8	0.270	-4.0	0.329	1.14	0.245
12	18.0	0.215	-6.0	0.237	-6.2	0.255	1.36	0.297

Table B.3. Individual thresholds on the auditory-function tests as described in Sec. 4.2.4.3 of 24 hearing-impaired listeners (listeners ‘a’ through ‘x’) and 12 normal-hearing listeners (listeners ‘1’ through ‘12’). Each threshold is the average result of two measurements, except the threshold on test 2b which is based on one measurement only. An empty cells means that the listener could not perform this test.

listener	2a	2b	2c	3a	3b	3c	5a	5b	5c			6a	6b	6c	6d
	GPq	GPn	GPn	Tone	USOM	DSOM	Click	FM	BM	TI	SI	ID	FD	RD	SD
	(dB)	(dB)	(dB)	(dB)	(Hz)	(Hz)	(dB)	(ms)	(ms)	(dB)	(dB)	(dB)	(%)	(ms)	(%)
a	27.1	45.7	16.9	15.7	113	67	13.6	28.0	27.3	1.3	3.4	3.5	14.2	23	62
b	25.8	54.7	17.4	13.2	138	49	17.9	25.9	37.5	4.3	-0.5	4.9	8.0	27	67
c	26.2	48.0	8.6	12.4	172	18	15.7	13.6	11.6	-3.9	-7.1	4.1	5.2	27	18
d	48.9	45.7	10.7	13.3	407	89	12.5	36.1	53.0	-2.6	-1.9	2.3	11.7	35	44
e	41.2	53.1	9.6	15.2	483	99	14.0	22.8	8.6	-5.6	-4.5	1.3	4.2	25	69
f	37.0	66.0	16.1	14.9	118	47	20.1			1.2	-4.0	2.9	8.6	49	31
g	35.8	54.8	13.0	12.5	80	18	13.9	18.7	39.8	0.4	-0.9	5.5	3.7	42	67
h	58.3	66.2	22.4				14.6	117.3			7.9	1.3	65.5	30	24
i	42.2	64.2	9.8	13.5	186	161	12.4	15.4	8.7	-3.7	-2.6	0.8	2.5	10	13
j	16.7	49.7	10.7	11.2	303	58	12.9	13.9	12.0	-0.5	-2.3	3.3	2.0	16	11
k	37.7	54.2	13.4	17.5	271	98	14.6	70.8		-4.1	-1.2	5.9	16.8	66	
l	20.1	49.7	11.2	12.7	291	28	16.8	11.1	8.9	-1.5	-5.6	3.7		11	31
m	46.2	61.9	11.5	13.2	345	204	11.3	31.1	20.3	-1.7	0.2	1.2	2.4	19	40
n	37.1	53.7	14.7	18.4	248	157	16.3	24.7	47.2	-3.8	-1.7	4.7	11.4	35	43
o	15.5	36.3	11.9	12.0	88	61	12.3	24.6	19.0	-0.2	-0.4	3.5	3.0	27	57
p	20.9	43.7	23.7	12.2	106	52	21.5	12.3		11.5	2.2	3.9	11.1	29	42
q	-1.4	46.2	22.4	11.0	217	74	17.6	6.6	7.4	11.4	4.9	1.9	3.5	12	28
r	15.9	48.2	11.9	12.1	338	61	14.2	20.1	20.7	-0.2	-2.3	5.1	8.9	47	
s	29.3	44.2	8.9	13.0	241	29	14.2	24.9	18.1	-4.1	-5.3	3.6	5.2	31	66
t	26.3	50.7	13.3	14.5	54	29	16.8	24.4	24.3	1.8	-3.5	4.7	3.5	35	77
u	41.4	63.7	12.9	17.9	456	48	13.2	33.3	18.2	-5.0	-0.3	1.1	7.0	30	90
v	50.7	72.2	13.2	12.0	348	175	16.1	18.8	11.3	1.2	-2.9	3.2	5.0	16	27
w	30.3	47.4	12.3	15.8	464	107	12.4	27.8	25.2	-3.6	-0.2	2.4	3.5	15	28
x	33.4	58.2	11.0	22.5	314	56	20.1	116.0	60.6	-11.0	-10.8	2.2	5.1	24	81

listener	2a	2b	2c	3a	3b	3c	5a	5b	5c			6a	6b	6c	6d
	GPq	GPn	GPn	Tone	USOM	DSOM	Click	FM	BM	TI	SI	ID	FD	RD	SD
	(dB)	(dB)	(dB)	(dB)	(Hz)	(Hz)	(dB)	(ms)	(ms)	(dB)	(dB)	(dB)	(%)	(ms)	(%)
1	-1.2	44.9	19.4	10.7	182	70	18.2	10.6	9.3	8.7	1.2	2.7	3.3	20	21
2	0.1	46.5	19.5	9.4	163	69	16.6	16.7	18.5	10.1	2.9	4.2	5.7	53	85
3	-2.5	32.3	13.5	8.5	97	84	15.0	7.9	10.2	5.0	-1.5	6.6	2.2	32	22
4	-5.1	38.3	14.9	9.8	106	44	13.4	10.2	17.1	5.1	1.5	5.0	2.0	17	31
5	-4.3	50.9	18.8	8.8	281	57	17.2	5.9	7.9	10.0	1.6	2.5	1.6	26	19
6	-8.7	48.4	21.3	10.7	208	49	17.8	5.1	8.9	10.6	3.5	3.8	5.9	22	20
7	-4.0	39.4	15.2	10.4	153	68	13.8	6.6	16.0	4.8	1.4	3.7	2.3	20	25
8	0.9	52.4	19.3	10.3	106	88	17.9	9.6	13.0	9.0	1.4	4.8	6.0	14	55
9	-1.9	32.6	14.9	7.7	117	60	13.8	13.3	17.5	7.3	1.2	6.8	3.4	51	20
10	-0.1	38.9	13.3	8.8	68	42	12.7	10.1	8.3	4.5	0.6	3.0	2.9	14	31
11	-7.4	37.0	14.4	7.3	129	123	9.3	38.7	34.3	7.2	5.1	4.2	3.2	40	15
12	-6.9	42.6	16.8	9.6	226	58	15.4	12.0	16.5	7.3	1.4	2.3	2.2	21	26

References

ANSI (**1969**). ANSI S3.5-1969, "American national standard methods for the calculation of the articulation index" (American National Standards Institute, New York).

ANSI (**1997**). ANSI S3.5-1997, "American national standard methods for calculation of the speech intelligibility index" (American National Standards Institute, New York).

Bacon, S. P., Opie, J. M., Montoya, D. Y. (**1998**). "The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds," J. Speech Language Hear. Res. **41**, 549-563.

Berger, E. H. (**1981**). "Re-examination of the low-frequency (50-1000 Hz) normal threshold of hearing in free and diffuse sound fields," J. Acoust. Soc. Am. **70**, 1635-1645.

Bronkhorst, A. W. and Plomp, R. (**1992**). "Effect of multiple speechlike maskers on binaural recognition in normal and impaired hearing," J. Acoust. Soc. Am. **92**, 3132-3139.

Carlyon, R. P., and Moore, B. C. J. (**1986**). "Detection of tones in noise and the "severe departure" from Weber's law," J. Acoust. Soc. Am. **79**, 461-464.

Ching, T., Dillon, H., and Byrne, D. (**1997**). "Prediction of speech recognition from audibility and psychoacoustic abilities of hearing-impaired listeners," in *Modeling Sensorineural Hearing Loss*, edited by W. Jestaedt (Erlbaum, Hillsdale, NJ).

Ching, T. Y. C., Dillon, H., and Byrne, D. (**1998**). "Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification," J. Acoust. Soc. Am. **103**, 1128-1140.

Cox, R. M., and McDaniel, D. M. (**1986**). "Reference equivalent threshold levels for pure tones and 1/3-oct noise bands: Insert earphone and TDH-49 earphone," J. Acoust. Soc. Am. **79**, 443-446.

Divenyi, P. L., and Haupt, K. M. (**1997a**). "Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. II. Correlation Analysis," Ear Hear. **18**, 100-113.

Divenyi, P. L., and Haupt, K. M. (**1997b**). Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. III. Factor representation," Ear Hear. **18**, 189-201.

Dreschler, W. A., and Plomp, R. (**1985**). "Relations between psychophysical data and speech perception for hearing-impaired subjects. II," J. Acoust. Soc. Am. **78**, 1261-1270.

Drullman, R. (**1994**). "Intelligibility of temporally degraded speech," Ph. D. dissertation, Vrije Universiteit, Amsterdam.

Drullman, R. (**1995**). "Temporal envelope and fine structure cues for speech intelligibility," J. Acoust. Soc. Am. **97**, 585-592.

Festen, J. M., 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**, 1725-1736.

Festen, J. M., and Dreschler, W. A. (**1988**). "Irregularities in the masked threshold of brief tones and filtered clicks," in *Basic Issues in Hearing - Proceedings of the 8th International Symposium on Hearing*, edited by H. Duifhuis, J. W. Horts, and H. P. Wit (Academic, London), pp 295-301.

Festen, J. M. and Plomp, R. (**1983**). "Relations between auditory functions in impaired hearing," J. Acoust. Soc. Am. **73**, 652-662.

Fletcher, H. (**1952**). "The perception of speech sounds by deafened persons," J. Acoust. Soc. Am. **24**, 490-497.

Florentine, M. (**1986**). "Level discrimination of tones as a function of duration," J. Acoust. Soc. Am. **79**, 792-798.

Florentine, M., Buus, S., and Mason, C. R. (**1987**). "Level discrimination as a function of level for tones from 0.25 to 16 kHz," J. Acoust. Soc. Am. **81**, 1528-1541.

Florentine, M., Reed, C. M., Rabinowitz, L. D., Durlach, N. I., and Buus, S. (**1993**). "Intensity Perception. XIV. Intensity discrimination in listeners with sensorineural hearing loss," J. Acoust. Soc. Am. **94**, 2575-2586.

Freyman, R. L., and Nelson, D. A. (**1986**). "Frequency discrimination as a function of tonal duration and excitation-pattern slopes in normal and hearing-impaired listeners," J. Acoust. Soc. Am. **79**, 1034-1044.

Freyman, R. L., and Nelson, D. A. (**1987**). "Frequency discrimination of short- versus long-duration tones by normal and hearing-impaired listeners," J. Speech Hear. Res. **30**, 28-36.

Glasberg, B. R., and Moore, B. C. (**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.

Glasberg, B. R., and Moore, B. C. (**1986**). "Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments," J. Acoust. Soc. Am. **79**, 1020-1033.

Glasberg, B. R., and Moore, B. C. (**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.

Grant, K. W. (**1987**). "Encoding voice pitch for profoundly hearing-impaired listeners," J. Acoust. Soc. Am. **82**, 423-432.

Hall, J. W., and Wood, E. J. (**1984**). "Stimulus duration and frequency discrimination for normal-hearing and hearing-impaired listeners," J. Speech Hear. Res. **27**, 252-256.

Hays, W. L. (**1988**). *Statistics* (Holt, Rinehart and Winston, New York), 4th ed.

Hogan, C. A., and Turner, C. W. (**1998**). "High-frequency audibility: Benefits for hearing-impaired listeners," J. Acoust. Soc. Am. **104**, 432-441.

Houtgast, T., and Steeneken, H. J. M. (**1985**). " A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria," J. Acoust. Soc. Am. **77**, 1069-1077.

International Organization for Standardization (**1961**). ISO R226-1961, "Normal equal-loudness contours for pure tones and normal threshold of hearing under free field listening conditions" (available from American National Standards Institute, New York).

Keren, G., and Lewis, C. (**1993**). *A Handbook for Data Analysis in the Behavioral Sciences* ***(Hillsdale, New Jersey).***

Killion, M. C. (**1978**). "Revised estimate of minimum audible pressure: Where is the "missing 6 dB"?," J. Acoust. Soc. Am. **63**, 1501-1508.

Lee, L. W., and Humes, L. E. (**1993**). "Evaluating a speech-reception threshold model for hearing-impaired listeners," J. Acoust. Soc. Am. **93**, 2879-2885.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467-477.

Ludvigsen, C. (**1985**). "Relations among some psychoacoustic parameters in normal and cochlearly impaired listeners," J. Acoust. Soc. Am. **78**, 1271-1280.

Ludvigsen, C. (**1987**). "Prediction of speech intelligibility for normal-hearing and cochlearly hearing-impaired listeners," J. Acoust. Soc. Am. **82**, 1162-1171.

Maiwald, D. (**1967**). "Beziehungen zwischen Schallspektrum, Mithörschwelle und der Erregung des Gehörs," Acustica **18**, 69-80.

Moore, B. C. J. (**1973**). "Frequency difference limens for short-duration tones," J. Acoust. Soc. Am. **54**, 610-619.

Moore, B. C. J. (**1986**). *Frequency Selectivity in Hearing* (Academic Press, London).

Moore, B. C. J. (**1995**). *Perceptual Consequences of Cochlear Damage* (Oxford University Press, Oxford).

Moore, B. C. J. (**1996**). "Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids," Ear Hear. **17**, 133-161.

Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (**1988**). "The shape of the ear's temporal window," J. Acoust. Soc. Am. **83**, 1102-1116.

Moore, B. C. J., and Raab, D. H. (**1974**). "Pure-tone intensity discrimination: some experiments relating to the "near-miss" to Weber's law," J. Acoust. Soc. Am. **55**, 1049-1054.

Nelson, D. A. (**1991**). "High-level psychophysical tuning curves: Forward masking in normal-hearing and hearing-impaired listeners," J. Speech Hear. Res. **34**, 1233-1249.

Neter, J., Wasserman W., and Kutner M. H. (**1990**). *Applied Linear Statistical Models* (IRWIN, Homewood, IL; Boston, MA), 3rd ed.

Noordhoek, I. M., and Drullman, R. (**1997**). "Effect of reducing temporal intensity modulations on sentence intelligibility," J. Acoust. Soc. Am. **101**, 498-502.

Noordhoek, I. M., Houtgast, T., and Festen, J. M. (**1999**). "Measuring the threshold for speech reception by adaptive variation of the signal bandwidth. I. Normal-hearing listeners," J. Acoust. Soc. Am. **105**, 2895-2902.

Noordhoek, I. M., Houtgast, T., and Festen, J. M. (**2000**) "Measuring the threshold for speech reception by adaptive variation of the signal bandwidth. II. Hearing-impaired listeners," J. Acoust. Soc. Am. **107**, 1685-1696.

Noordhoek, I. M., Houtgast, T., and Festen, J. M. "Relations between intelligibility of narrowband speech and auditory functions, both in the 1-kHz frequency region," J. Acoust. Soc. Am. (accepted for publication with minor modifications).

Oxenham, A. J., and Moore, B. C. J. (**1995**). "Additivity of masking in normally hearing and hearing-impaired subjects," J. Acoust. Soc. Am. **98**, 1921-1934.

Oxenham, A. J., Moore, B. C. J., and Vickers, D.A. (**1997**). "Short-term temporal integration: Evidence for the influence of peripheral compression," J. Acoust. Soc. Am. **101**, 3676-3687.

Patterson, R. D., Nimmo-Smith, D. L., Weber, D. L., and Milroy, R. (**1982**). "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold," J. Acoust. Soc. Am. **72**, 1788-1803.

Pavlovic, C.V. (**1987**). "Derivation of primary parameters and procedures for use in speech intelligibility predictions," J. Acoust. Soc. Am. **82**, 413-422.

Pavlovic, C. V., Studebaker, G. A., and Sherbecoe, R. L. (**1986**). "An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals," J. Acoust. Soc. Am. **80**, 50-57.

Penner, M. J. (**1974**). "Effect of masker duration and masker level on forward and backward masking," J. Acoust. Soc. Am. **56**, 179-182.

Plomp, R. (**1978**). "Auditory handicap of hearing impairment and the limited benefit of hearing aids," J. Acoust. Soc. Am. **63**, 533-549.

Plomp, R. (**1986**). "A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired," J. Speech Hear. Res. **29**, 146-154.

Plomp, R. (**1988**). "The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function,"J. Acoust. Soc. Am. **83**, 2322-2327.

Plomp, R., and Mimpen, A. M. (**1979**). "Improving the reliability of testing the speech reception threshold for sentences," Audiology **18**, 43-52.

Rabinowitz, W. M., Lim, J. S., Braida, L. D., and Durlach, N. I. (**1976**). "Intensity Perception. VI. Summary of recent data on deviations from Weber's law for 1000-Hz tone pulses, J. Acoust. Soc. Am. **59**, 1506-1509.

Rosen, S., and Fourcin, A. (**1986**). "Frequency selectivity and the perception of speech," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic Press, London), pp 373-487.

Schroder, A. C., Viemeister N. F., and Nelson, D. A. (**1994**). "Intensity discrimination in normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **96**, 2683-2693.

Smoorenburg, G. F. (**1992**). "Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram," J. Acoust. Soc. Am. **91**, 421-437.

Steeneken, H. J. M., and Houtgast, T. (**1980**). "A physical method for measuring speech-transmission quality," J. Acoust. Soc. Am. **67**, 318-326.

Stevens, S. S. (**1956**). "Calculation of the loudness of complex noise," J. Acoust. Soc. Am. **28**, 807-832.

Summerfield, Q. (**1987**). "Speech perception in normal and impaired hearing," British Medical Bulletin **43**, 909-925.

ter Keurs, M. (**1992**). Intelligibility of spectrally smeared speech," Ph. D. dissertation, Vrije Universiteit, Amsterdam.

Turner, C. W., Zwislocki J. J., Filion, P.R. (**1989**). "Intensity discrimination determined with two paradigms in normal and hearing-impaired subjects," J. Acoust. Soc. Am. **86**, 109-115.

Tyler, R.S. (**1986**). "Frequency resolution in hearing-impaired listeners," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic Press, London), pp 309-371.

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

van Buuren, R. A., Festen, J. M., and Houtgast, T. (**1995**). "Evaluation of a wide range of amplitude-frequency responses for the hearing-impaired," J. Speech Hear. Res. **38,** 211-221.

van Buuren, R. A., Festen, J. M., and Houtgast, T. (**1999**). "Compression and expansion of the temporal envelope: Evaluation of speech intelligibility and sound quality,"J. Acoust. Soc. Am. **105**, 2903-2913.

van den Brink, W. A. C., and Houtgast, T. (**1990a**). "Efficient across-frequency integration in short-signal detection," J. Acoust. Soc. Am. **87**, 284-291.

van den Brink, W. A. C., and Houtgast, T. (**1990b**). "Spectro-temporal integration in signal detection," J. Acoust. Soc. Am. **88**, 1703-1711.

van Dijkhuizen, J. N., Anema, P. C., and Plomp, R. (**1987**). "The effect of varying the slope of the amplitude-frequency response on the masked speech-reception threshold of sentences," J. Acoust. Am. **81**, 465-469.

van Dijkhuizen, J. N., Festen, J. M., and Plomp, R. (**1991**). "The effect of frequency-selective attenuation on the speech-reception threshold of sentences in conditions of low-frequency noise," J. Acoust. Soc. Am. **90**, 885-894.

van Rooij, J. C. G. M., and Plomp, R. (**1990**). "Auditive and cognitive factors in speech perception by elderly listeners. II: Multivariate analyses," J. Acoust. Soc. Am. **88**, 2611-2624.

van Rooij, J. C. G. M., and Plomp, R. (**1992**). "Auditive and cognitive factors in speech perception by elderly listeners. III. Additional data and final discussion," J. Acoust. Soc. Am. **91**, 1028-1033.

van Schijndel, N. H., Houtgast, T., Festen. J. M. (**1999**). "Intensity discrimination of Gaussian-windowed tones: Indications for the shape of the auditory frequency-time window," J. Acoust. Soc. Am. **105**, 3425-3435.

Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (**2000**) "Method for the selection of sentence materials for efficient measurement of the speech reception threshold," J. Acoust. Soc. Am. **107**, 1671-1684.

Viemeister, V. F. (**1972**). "Intensity discrimination of pulsed sinusoids: The effects of filtered noise," J. Acoust. Soc. Am. **51**, 1265-1269.

von Klitzing, R., and Kohlrausch, A. (**1994**). "Effect of masker level on overshoot in running- and frozen-noise maskers,"J. Acoust. Soc. Am. **95**, 2192-2201.

Walker, G., Dillon, H., Byrne, D., and Christen, R. (**1984**). "The use of loudness discomfort levels for selecting the maximum output of hearing aids," Aust. J. Audiol. **6**, 23-32.

Weber, D. L., and Green, D. M. (**1979**). "Suppression effects in backward and forward masking," J. Acoust. Soc. Am. **65**, 1258-1267.

Zurek, P. M., and Delhorne, L. A. (**1987**). "Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment," J. Acoust. Soc. Am. **82**, 1548-1559.

Zwicker, E. (**1963**). "Über die Lautheit von ungedrosselten und gedrosselten Schallen," Acustica **13**, 194-211.

Dankwoord

Het onderzoek dat in dit proefschrift beschreven wordt, heeft plaatsgevonden aan de Vrije Universiteit te Amsterdam in de periode van 1 mei 1995 tot 1 september 1999. Het was een mooie tijd! Dat is te danken aan het interessante onderwerp waaraan ik mocht werken, maar zeker ook aan mijn collega's.

Mijn promotor, prof.dr.ir. Tammo Houtgast, wil ik bedanken voor het initiëren van mijn promotieproject en voor zijn inspirerende begeleiding bij de invulling ervan. Mijn copromotor, dr.ir. Joost Festen, wil ik bedanken omdat zijn deur altijd (in ieder geval figuurlijk, maar meestal ook letterlijk) open stond voor vragen, en omdat hij kromme zinnen in mijn manuscripten zo mooi recht wist te buigen.

De leden van de leescommissie, prof.dr.ir. F.A. Bilsen, prof.dr.ir. W.A. Dreschler, prof.dr. H.F. Mahieu, prof.dr. G.F. Smoorenburg, dr. J. Verschuure en dr. J. Wouters, ben ik zeer erkentelijk voor de tijd die ze hebben willen besteden aan de beoordeling van dit proefschrift.

Hans van Beek was de reddende engel als mijn computer "raar deed." Hans, bedankt voor je (hulp bij het) programmeerwerk en voor de leuke mailtjes. Terwijl ik lekker op dezelfde plek mocht blijven zitten, heb ik toch vijf verschillende kamergenoten gehad. Dat waren achtereenvolgens Ronald van Buuren, Judith Kessens, Laura Daalder, Niek Versfeld en Dick Buitelaar. Bedankt voor de plezierige werksfeer! Niek, ook bedankt voor je goede voorbeeld wat betreft efficiënt onderzoek doen en je aansporingen aan mij om op schema te blijven. Finn Dubbelboer, Johannes Lijzenga en Nicolle van Schijndel wil ik bedanken voor hun gezelligheid tijdens de lunch en de koffiepauze. Soms hadden we zulke geanimeerde gesprekken dat het nodig was de deur van de koffieruimte te sluiten om niemand te storen. Johannes wil ik daarnaast ook bedanken voor het verbeteren van mijn Engels. Nicolle, jouw aanwezigheid gaf niet alleen een significante, maar ook een grote bijdrage aan de plezierigheid van mijn tijd bij de VU. Verder konden we tijdens het gezamenlijk afgelegde deel van ons woon-werk verkeer goed over onze onderzoeksprojecten (die steeds meer raakvlakken gingen vertonen) discussiëren en over andere perikelen uit het AIO-bestaan.

Mijn vrienden en familie wil ik bedanken voor hun interesse in mijn promotieonderzoek. Papa en mama, bedankt voor alles. Arjen, bedankt voor je liefde (en het inrichten van mijn "promo-hok").

Tenslotte bedank ik alle proefpersonen voor hun luisterbijdrage!

Curriculum vitae

Ingrid Noordhoek werd geboren in Delft op 19 mei 1971. Vier jaar later verhuisde zij naar Hardinxveld-Giessendam, waar ze de lagere school bezocht. In 1989 ontving ze haar VWO diploma aan de Willem de Zwijger Scholengemeenschap in Papendrecht en begon ze met de studie Technische Natuurkunde aan de Technische Universiteit Delft. Tijdens haar studie liep ze twee stages: bij de afdeling Oogheelkunde van het Academisch Ziekenhuis Nijmegen onderzocht ze, onder leiding van dr.ir. J.M. Thijssen, de snelheid en de verzwakking van ultrageluid in oogweefsel en bij het Hahn-Meitner-Institut in Berlijn deed ze onder leiding van dr. H. Rossner onderzoek naar kristalstructuren door interpretatie van photo-absorptie spectra. Ze voerde haar afstudeerproject uit onder leiding van prof.dr.ir. F.A. Bilsen bij de vakgroep Akoestiek, sectie Akoestische Perceptie. Het onderwerp was de verbetering van de richtingsgevoeligheid van een hoortoestel gebaseerd op array-technologie. Na haar afstuderen in 1995 verliet ze Delft, maar niet het vakgebied Akoestiek. Ze vond een promotieplaats bij de afdeling Audiologie van de vakgroep Keel- Neus- en Oorheelkunde, faculteit Geneeskunde, Vrije Universiteit Amsterdam. Hier onderzocht ze de oorzaken die ten grondslag liggen aan de problemen die slechthorenden ondervinden bij het verstaan van spraak. Dit proefschrift is hiervan het resultaat. Sinds september 1999 werkt Ingrid Noordhoek bij TNO-TPD in Delft op de afdeling Buitenluchtakoestiek van de divisie Geluid.