Perceptual evaluation of noise reduction in hearing aids
Brons, I.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Introduction
1.1 Introduction

During the last decades, digital hearing aids have replaced their analogue predecessors. Whereas the size of the hearing aids decreased with time, the capabilities for signal processing increased. A wide variety of signal-processing strategies has been developed for hearing aids in order to compensate for different aspects of hearing loss. Straight-forward amplification can largely compensate for deficits in the outer and middle ear (conductive hearing loss), but compensation for deficits in the cochlea (sensorineural hearing loss) requires a more sophisticated approach. Sensorineural hearing loss causes not only reduced sensitivity for soft sounds, but also a reduced dynamic range as well as a reduced spectral and temporal resolution (Moore 1996). The resulting distortions in the perceived sound make it more difficult to understand speech, especially in noisy environments. For a given level of background noise, a listener with sensorineural hearing loss needs a higher speech level than a normal-hearing listener to obtain the same performance, even if the sound is amplified (Plomp 1986). In noisy situations the hearing aid should therefore amplify speech more than background noise to improve the signal-to-noise ratio (SNR).

Improving the signal-to-noise ratio in noisy environments is one of the most difficult challenges for digital signal processing in hearing aids. For that purpose, the hearing aid should estimate from the incoming signal whether the user is in an environment with speech, noise, or both speech and noise and adjust the gain in each frequency channel accordingly. To distinguish between speech, noise, or both, the hearing aid uses signal properties that are generally different between speech and noise. Hearing aids with multiple microphone inputs can use spatial differences between speech and noise. These hearing aids are able to amplify sounds that enter the hearing aid from one direction (usually from the front) while attenuating incoming sounds from other directions. In situations where speech and noise signals enter the hearing aid from separate directions, these directional microphones effectively improve the SNR (Bentler 2005). However, speech and noise sources are not always spatially separated and even if they are, both signals can be mixed before they enter the hearing aid, for instance due to reverberations. In that case the hearing aid can no longer use spatial differences to separate between speech and noise and should use other cues to recognize speech and noise, which is more difficult because speech and noise enter the hearing aid as a mixed signal. For this purpose, hearing aids have single-microphone noise-reduction algorithms which use temporal and spectral properties to separate between speech and noise. Such algorithms are implemented both in hearing aids with one microphone input as well as in hearing aids with multiple microphone inputs where it is supplementary to directionality.
This thesis focuses on single-microphone noise reduction. The use of spatial differences between speech and noise is thus beyond the scope of this work.

Single-microphone noise-reduction algorithms generally consist of two parts. First, they classify the sound environment into speech, noise, or both. This step generally results in an estimate of the actual input SNR per frequency channel. Second, the algorithm should reduce the hearing aid gain at the right time and for the right frequencies to reduce noise, but leave the speech intact. The next sections explain the basic approaches and main properties of both parts of a noise-reduction algorithm.

1.2 Classification of the environment

The upper panels of Figure 1.1 show typical time-amplitude waveforms of speech, stationary noise and a mixture of speech and stationary noise. The figure shows that speech has a very characteristic temporal pattern, with high variation in amplitude over time corresponding to the opening and closing of the vocal tract. Noise-reduction algorithms make use of this characteristic pattern of speech for the classification of the environment. The algorithms often use the envelope of the signal for this purpose, which is plotted for the same signals in the lower panels of Figure 1.1. The envelope of speech shows large level fluctuations (lower left panel), whereas the envelope of stationary noise shows very little fluctuations (lower middle panel). Adding noise to the speech reduces the depth of the envelope fluctuations (lower right panel) compared to those of speech.

Most real-life noises have envelope characteristics that lie in between that of speech and stationary noise, with more fluctuations than stationary noise, but with lower amplitude and at other modulation frequencies than for speech (see next paragraphs).
Classification based on modulation depth

One property of the envelope signal that a noise-reduction algorithm can use for identification of speech and noise is the modulation depth (Schaub 2008). The modulation depth is the difference between the peaks and valleys of the signal envelope. As visible in Figure 1.1, modulation depth is higher for speech than for noise. Adding noise to the speech reduces the modulation depth (lower right panel in Figure 1.1) compared to that of speech alone. Therefore, the modulation depth can be used to estimate the SNR at the input of the hearing aid (input SNR). Frequency channels with high modulation depth are likely to be dominated by speech and thus to have high input SNR, whereas frequency channels with low modulation depth are dominated by noise and have low input SNR.

Classification based on modulation spectrum

Another property of the envelope signal that can be used to identify speech and noise is the modulation spectrum (Dillon 2001; Chung 2004). The modulation spectrum provides information on how fast the signal envelope changes in level. Figure 1.2 shows the modulation spectrum for the three envelope signals in the lower part of Figure 1.1. The envelope of speech is typically dominated by modulations with frequencies below 10 Hz. In general, modulation frequencies between 3 and 6 Hz are dominant, corresponding to the rate at which syllables are produced. The envelope of stationary noise shows the same amount of modulations at all frequencies. Thus, if the modulation spectrum for a specific frequency channel is dominated by frequencies between 3 and 6 Hz it is likely that speech is present.

![Modulation spectra of speech, noise, and noise + speech](image)

*Figure 1.2:* Modulation spectra of the envelope of speech, stationary noise and the mixture of speech and noise. A higher modulation index means a higher modulation depth at that frequency.
Classification based on synchrony detection

An additional clue for speech presence is the synchrony between signal envelopes of different frequency bands (Elberling 2002; Dillon 2001). For instance, modulations at the pitch of the voice (100-400 Hz) are synchronized across frequency channels. This is illustrated in Figure 1.3, which shows the waveform of speech during a vowel in four different frequency channels. The speech envelope in each frequency channel has the same periodicity. Thus, if such synchronization across frequencies is detected, speech is likely to be present.

![Figure 1.3: Synchronization across frequency channels during speech. The upper signal shows in light gray the time-amplitude signal of one sentence, with in dark gray the fragment (part of a vowel) that is shown in the lower five signals. The 'Broadband signal' shows the fragment in its original form; the lower four signals show the fragment in four different frequency channels (each 1000 Hz wide).](image)

Classification based on speech pause detection

Another way to make use of the dynamical characteristics of speech is to use the pauses in speech to estimate the noise signal and to update this estimate over time (Loizou 2007). The most straight-forward way to do this is the use of an absolute level threshold. Signal parts where the level does not exceed this threshold are assumed to contain no speech and can therefore be considered as an update of the noise estimate, which in turn can be used to estimate the SNR. Tracking the minimum level of the incoming signal as an estimate for the noise forms the basic principle for many more sophisticated noise-estimation algorithms (Loizou 2007).
1.2.1 Reduction of the hearing-aid gain

If the properties of the incoming signal indicate the presence of noise, the noise-reduction algorithm should decide if and how it will adjust the hearing-aid gain in order to reduce the noise. The main challenge for the noise reduction is to reduce background noise while retaining the level and quality of target speech. To what extent the noise reduction succeeds in this task depends not only on the accuracy of the classification of speech and noise but also on how this information is translated into changes in hearing aid gain. The most important variables for the gain reduction will be discussed below.

**Amount of gain reduction**

The amount of gain reduction in a frequency channel is generally chosen to be proportional to the estimated SNR for that channel (Dillon 2001; Chung 2004). If the signal is clearly dominated by speech, thus at high input SNR, the gain should not be reduced in that channel, to preserve speech information. If the input SNR of a channel is very low, i.e. when noise is present but no speech is recognized in that channel, the noise-reduction algorithm will reduce gain in that channel maximally. In most hearing aids the maximum amount of gain reduction is limited, so that below a certain SNR the gain is not further reduced if the SNR decreases further (Chung 2004; Bentler and Chiou 2006). If the maximum amount of gain reduction is low, the residual background noise level will be relatively high. If the maximum amount of gain reduction is high, the residual noise will be less but that may be at the expense of speech level and sound quality. Reduction of speech level and quality can for instance occur when errors are made in the classification of the environment (speech-dominated signal parts are wrongly considered as noise-dominated and are therefore reduced in gain) (Loizou and Kim 2011). Sound quality can also be affected when the gain differs largely between neighbouring time windows or frequency channels (see next sections).

**Time constants for gain reduction**

The effectiveness of noise reduction is also determined by how fast the noise reduction reacts to changes in the environment (Bentler and Chiou 2006). Noise-reduction algorithms generally have several time constants to react properly to different changes in the input signal, so that for instance the decrease of gain as a reaction to noise can be slower than the reaction to a sudden appearance of speech. The time that an algorithm needs to reach its maximum gain reduction after noise has started varies between algorithms from several seconds to more than 30 seconds. In contrast, most algorithms react within several milliseconds to the appearance of speech by increasing the gain, in order not to loose any speech information (Chung 2004). Algorithms that change the gain very quickly can be able to reduce the noise even in the short pauses that occur
during speech. However, quick transitions in gain may be experienced as restless and can cause distortions to the speech signal, which may affect speech intelligibility. On the other hand, algorithms that adjust the gain only slowly will result in more residual noise during speech and after changes in the environment.

**Number of frequency channels for gain reduction**

The potential benefit of noise reduction in improving the SNR is mainly due to the possibility to apply noise reduction separately for different frequency channels (Chung 2004). Because noise reduction only adjusts the overall gain for a frequency channel, the SNR within a frequency channel will not change during one time frame. Only the combination of reduced gain in some frequency channels (the noise-dominated ones) and preserved gain in others (the speech-dominated ones) makes that the instantaneous SNR of the overall signal (all frequency channels together) can improve due to noise reduction. This argues in favour of a high number of frequency channels. Within hearing aids, however, the number of channels is limited because increasing the number of channels also increases processing delay. Many state-of-the-art noise-reduction algorithms that are not (yet) implemented in hearing aids apply noise reduction in high numbers of frequency bins separately (Loizou 2007). However, applying different amounts of gain to neighbouring frequency bins introduces unwanted distortions, often referred to as musical noise (Berouti et al. 1979). It is therefore generally more favourable for the sound quality to limit the number of frequency channels, although this may not result in the most optimal improvement of SNR.

**Frequency weighting of the gain reduction**

Dividing the signal in different frequency channels allows the algorithm to assign different weights to separate frequency channels (Kuk and Paludan-Müller 2006). One can for instance decide to allow more gain reduction for the low frequencies than for the higher frequencies. Many environmental noises, for instance traffic noise, contain energy mainly in the low frequencies. For such situations listening comfort can easily be improved by reducing the gain more for lower frequencies than for higher frequencies (“high-pass filter”). This method also removes the low-frequency content of the speech signal, but if this part of the speech would otherwise be inaudible due to the noise, the filter will not reduce intelligibility. Reduction of the low-frequency noise also prevents the noise from masking high-frequency information in the speech (“upward spread of masking”) (Levitt 2001).

Several algorithms base their maximum gain reduction per frequency band on the articulation index, which describes how much the audibility of speech in each frequency band contributes to the intelligibility of the speech (Kryter 1962). Frequency bands
which are known to be important for speech intelligibility will receive less gain reduction than frequency bands with lower perceptual importance. Additionally, in order to maintain audibility of the input signal, listeners’ hearing loss at different frequencies, as well as the level of the input signal may play a role in determining the noise-reduction strength for separate frequency bands (Kuk and Paludan-Müller 2006).

1.3 Implementation in hearing aids

It may be clear from the description of the variables that there is not one optimal value for each of them. In fact, all variables involve the trade-off between reducing noise and retaining the quality of speech. Generally, most variables are preset by the manufacturer and cannot be changed by the clinician who adjusts the hearing aid settings to its user. Some hearing aids provide the possibility to adjust the maximum allowed amount of gain reduction. For the other variables, the manufacturers have the difficult task to make a well-considered choice for their settings. Because each setting has its advantages and disadvantages, and because they may interact, these choices may differ between manufacturers. Unfortunately, noise reduction in hearing aids is generally presented as a “black box” so that no information is available on the actual implementation and its underlying rationales. In order to nonetheless obtain some insight in noise-reduction implementations in hearing aids, we have made recordings of the output of some hearing aids and analyzed how the hearing aid gain differed for the same input signals if noise reduction was activated or de-activated.

Figure 1.4 shows for four different hearing aids (HA1 to HA4) how noise reduction influenced the hearing aid gain for the same input signal. As a reference condition, one “conceptual” noise reduction is added to Figure 1.4. This noise reduction received speech and noise separately and thus had knowledge on the real input SNR (which is often referred to as “ideal” noise reduction). This algorithm can obviously not be used in hearing aids, but for research purposes it provides a useful tool to investigate effects of noise-reduction settings. In Figure 1.4, this conceptual condition is added for comparison because it provides information on the real input SNR of the different time-frequency units.

The time signal at the top of the Figure 1.4 shows the input signals, consisting of stationary noise which is combined with speech after 2 seconds. Note that speech and noise are plotted separately here for clarity, but only the conceptual algorithm had the separate signals as input. The hearing aid noise-reduction systems only had the sum of both signals as input and had to estimate whether this was speech, noise, or both.
Figure 1.4: The effect of noise reduction on the hearing aid gain in four different hearing aids (HA1 to HA4). One conceptual algorithm (which received speech and noise separately and thus had knowledge on the real input SNR) was added for comparison. The time signals plotted at the top show the input noise (light gray) and speech (dark gray) signals, which were mixed before they entered the hearing aids.

The colours in the lower five panels show how the noise reduction influenced the gain as a function of time (horizontal axis) and frequency (vertical axis). Blue areas indicate no difference in gain between noise reduction on and off, thus no activity of the noise reduction. Red areas represent strong reduction in gain due to noise reduction. Thus, as would be expected, noise reduction was generally most active if the input signal
Introduction

consisted of noise only (the first two seconds) and less when the input signal also contained speech (starting after 2 seconds). Next sections will discuss what information can be deduced from Figure 1.4 on the variables explained before.

Amount of gain reduction

The maximum amount of gain reduction for the conceptual algorithm was set at 15 dB. All time-frequency bins where the SNR was below 0 dB, thus where noise was dominating, received maximum gain reduction. For positive SNRs, the amount of gain reduction decreased as a function of SNR, so that gain was not reduced for time-frequency units with high input SNR. The colour plot for the conceptual algorithm thus provides direct information on the real input SNR so that we can compare how well the other algorithms separated between speech and noise.

Noise reduction in the four hearing aids was set at the strongest available settings. During the first two seconds, where only noise was present and thus the algorithms should reduce gain maximally this resulted in maximum gain reduction of about 10 to 13 dB (red areas) for hearing aids HA1 to HA3, and more than 15 dB (dark red areas) for HA4 in the high frequencies.

Time constants for gain reduction

At the start of Figure 1.4, noise was already present for a longer time, so that the figure does not show how quick the algorithms adjusted the gain if the situation turned from silence to noise. However, we see that as soon as speech started (vertical line at t = 2 s), all algorithms reacted quickly with reducing noise reduction. Such a quick reaction on speech presence is important in order to retain speech information.

While speech was present, the algorithms in hearing aid HA1 and HA2 adjusted the hearing aid gain from time to time in order to reduce noise during the short pauses in the speech. Comparison with the gain-reduction pattern of the conceptual noise reduction shows that speech was most of the time recognized well by HA1 and HA2, but clearly not all speech-dominated parts were preserved. The reaction of HA3 differed between the low- and high-frequency regions. For frequencies up to 500 Hz, the gain changed only slowly. Gain in the higher frequencies reacted more quickly on changes in speech, but not by reducing gain during noise but by increasing gain somewhat during speech. HA4 adjusted its gain only slowly. After the speech started it took some seconds before gain was restored, but then gain was completely preserved as long as speech was present. So it seems that for this hearing aid preserving audibility of speech was considered more important than reducing the noise during speech.
**Number of channels for gain reduction**

The number of frequency channels within the four hearing aids can not be deduced from Figure 1.4. According to the technical specification of the hearing aids, the number of frequency channels used for noise-reduction processing varied from 8 to 20. The conceptual algorithm calculated gain for each frequency bin separately. As explained before, this had the advantage that speech and noise could be separated more accurately, but this resulted in high gain contrast between neighbouring time-frequency bins, which might be disadvantageous for the sound quality.

**Frequency weighting of the gain reduction**

Figure 1.4 reveals that the noise-reduction algorithms from the four different hearing aids have different strategies for frequency weighting. The conceptual algorithm had no rules for frequency weighting implemented. Thus if the input signal consisted of noise only, this algorithm applied the same amount of gain reduction (15 dB) over all frequencies. This was also the case for HA2. Noise reduction in HA1 was more cautious for frequencies around 1000 Hz, where it only reduced gain up to about 5 dB (green horizontal line) instead of 12 dB (red) for other frequencies. HA3 was even more careful for frequencies around 1000 and 2000 Hz. This is probably because these frequencies generally contain important speech information. In that light the behaviour of HA4 is rather unexpected: the gain during noise was preserved in the low frequencies but strongly reduced in the high-frequency regions. The rationale behind this weighting is unknown, but from recordings with other input noises we know that the frequency weighting for this hearing aid differs between noises (see for instance Chapter 3 for the gain-reduction pattern for babble noise).

Summarizing, Figure 1.4 reveals that noise reduction from different hearing aids have diverging effects on the hearing aid gain for the same speech in noise input signal. In that light, it is puzzling why there is so little public knowledge on noise-reduction implementations in hearing aids and on the motivation of manufacturers to choose for specific noise-reduction settings in their products.

### 1.4 Perceptual effects of noise reduction

Noise-reduction algorithms were at first developed with the aim of improving speech intelligibility in noisy environments (Loizou 2007). However, studies that investigated the effect of noise reduction on speech intelligibility revealed no improvement in intelligibility due to noise reduction (see for instance Nordrum et al. 2006). In fact, in some cases noise reduction even tended to reduce speech intelligibility (Hu and Loizou 2007a). The reason for this could be that, as mentioned before, noise reduction may af-
fect the speech signal and speech level due to errors in the classification and due to processing artefacts.

Even if noise reduction is able to reduce the noise level without seriously affecting the speech level and quality, this does not necessarily result in an improvement in speech intelligibility in noise. In that case, the noise that is removed by the noise reduction was not the noise that masked the speech signal, so that otherwise the auditory and cognitive system of the listener would have been able to neglect these noise parts that were now reduced by the noise reduction. Although this does not result in an objective improvement in speech intelligibility, the listener can experience more listening comfort if parts of the noise are already removed by the noise reduction. It is imaginable that this lightens the cognitive load or listening effort needed to understand speech, which in turn can lead to less fatigue due to listening in noisy environments (Sarampalis et al. 2009). Unlike speech intelligibility, which is a common objective measure for hearing aid benefit, objective measurements for cognitive load or listening effort are scarce and still in an initial phase of development and validation.

Studies evaluating hearing aid noise reduction generally use subjective measures to gain insight in the effects of noise reduction on perception (Bentler 2005). For instance, subjects can be asked to rate the sound quality within a specific condition, or to choose from different conditions the most comfortable one. In the same way, listening effort can be subjectively measured. The results of previous studies on subjective effects of hearing aid noise reduction show however no uniform results. For instance, some studies showed that listeners preferred noise reduction over no noise reduction (Boymans and Dreschler 2000; Ricketts and Hornsby 2005), whereas others found no difference in listening comfort or sound quality due to noise reduction (Alcàntara et al. 2003; Bentler et al. 2008). Also results on listening effort were inconsistent (Alcàntara et al. 2003; Bentler et al. 2008).

The studies mentioned here compared noise reduction with no noise reduction within a hearing aid, but there is little or no knowledge on how noise reduction was implemented in the specific hearing aids. As we have just seen in Figure 1.4 for a selection of four hearing aids, noise-reduction implementations can differ widely between hearing aids. It is imaginable that these differences between hearing aids will also result in different results for perceptual outcomes like listening comfort and listening effort. In this thesis, the perceptual consequences of the different noise-reduction implementations in Figure 1.4 will be investigated more systematically.
1.5 Outline of this thesis

This thesis studies the effects of single-microphone noise-reduction algorithms on perceptual outcomes (noise annoyance, speech naturalness, personal preference, speech intelligibility and listening effort).

The first part of the thesis (Chapters 2 to 5) describes the evaluation of noise reduction as implemented in different commercial hearing aids. Chapter 2 describes how we recorded the hearing aid output and removed differences in frequency response between hearing aids to allow direct comparison of noise-reduction algorithms between hearing aids. In Chapter 3 we used this method to perform perceptual measurements with normal-hearing subjects to evaluate the effects of noise reduction in linear hearing aids in terms of noise annoyance, speech naturalness, personal preference, speech intelligibility, and listening effort. Chapter 4 extends Chapter 3 in that it describes the results of comparable measurements for listeners with sensorineural hearing loss. Thus far, we evaluated noise reduction in a linear setting. However, for listeners with sensorineural hearing loss, hearing aids usually apply dynamic-range compression, which may interact with noise reduction. In Chapter 5 we therefore explored the combination of noise reduction and compression in the same commercial hearing aids as in Chapter 4. This chapter describes both acoustical measurements and perceptual measurements with hearing-impaired subjects.

The second part of the thesis (Chapters 6 and 7) describes studies using noise-reduction algorithms that were not implemented in hearing aids. Because the source code of these algorithms was available we could adjust parameters to investigate effects of different settings. In Chapter 6 we evaluated the effect of maximum attenuation strength on perceptual outcomes (noise annoyance, speech naturalness, personal preference, speech intelligibility, and listening effort) for ideal noise reduction (comparable with the “conceptual” condition in Figure 1.4 of the current chapter). Additionally, we evaluated the effect of two different (non-ideal) noise-estimation methods on the same perceptual outcomes to obtain insight in the effect of classification errors. In Chapter 7 we determined how much attenuation could be applied by a noise-reduction algorithm until subjects detected distortions caused by the noise reduction. We investigated whether this detection threshold for noise-reduction distortion differed between normal-hearing and hearing-impaired subjects and whether it was related to their individual preference for noise-reduction strength.

Chapter 8 summarizes the results of this thesis and provides some recommendations for clinical practice and for future research.