CHAPTER 1

General introduction
INTRODUCTION

In daily life, spoken speech is one the most important ways of communication between humans. However, speech is often interrupted by other sounds that draw our attention. Imagine arriving home after a day of work; your family awaits you, food is being prepared and children are playing. Your house will be filled with natural sounds: the laughter of children, the sounds of the cooking, notifications on your smartphone, while the wind is softly blowing through an open window, bringing the sounds of birds and a car passing in the street. In this auditory scene you are having a conversation about your day’s experiences. Your perceptual system is presented with a lot of acoustic streams that could be followed. There is a lot of competition and also masking of information. To retrieve the right message may be sometimes demanding, but as a normal-hearing listener you will be able to recognize the sounds of interest relatively well (Riecke et al. 2007). To follow a conversation in this complex acoustic environment your perceptual system makes use of a number of automatic restoration mechanisms.

Nevertheless, having a conversation in a noisy room is often reported to be difficult. The listener has to focus on the speaker without getting distracted by other conversations or by background noise. These challenging acoustic environments are even more problematic for hearing impaired listeners. The number of hearing impaired people is considerable (Stevens et al. 2013); an averaged hearing threshold across the frequencies 0.5, 1, 2 and 4 kHz of 35 dB HL or more in the better ear, has an estimated global prevalence of around 11% for above 15 years of age. Hearing impairment leads to a degraded auditory input to the auditory system for which has to be compensated. Hearing-impaired listeners often try to relieve the consequences of their hearing loss by relying on non-auditory speech cues, such as lip-reading. Even these compensation mechanisms and strategies do not always solve the problem hearing-impaired listeners face in everyday speech communication. As a result, hearing-impaired listeners report often that they tend to avoid social situations with background noise.

The hope that is driving the research presented in this thesis is that hearing-impaired listeners could learn to communicate better in complex acoustic environments, if we could provide an effective and adequate training for them.

The main aim is to understand more about the auditory and cognitive factors that are influencing, or are involved in, perceptually restoring degraded speech in such a way that the perceptual system can form it into an intelligible speech stream. This knowledge can lead to new training programs for the hearing impaired.
The reason why normal-hearing listeners are able to have a conversation in background noise is that the perceptual system uses multiple high-level cognitive mechanisms, in which many factors are involved. For example, as soon as the listeners get familiar with a speaker they benefit from the pitch and timbre of his/her voice, from the known direction of the sound and from the speech rate. Using these acoustic properties, perceptual grouping is used to find the sounds patterns pertaining to the target speech produced by the speaker of interest, which are then used to form a coherent speech stream. Segmentation is used to recognize individual words to form meaningful lexical units (Bregman 1994; Davis and Johnsrude 2007). The perceptual system makes use of highly automated restoration mechanisms in this complex acoustic environment so that the conversation can take place without much effort and can continue seamlessly, at a fast pace. These restoration mechanisms play an important role in speech perception of degraded speech, as will be shown in this thesis.

Cognitive mechanisms are involved in restoring degraded speech streams in such a way that the brain perceives it as an intelligible speech stream. Aging has been shown to affect the perception, as well as restoration, of degraded speech, even without age-related hearing loss, implying that the cognitive system is involved (Saija et al. 2013). If indeed cognitive factors play a role in the restoration of degraded speech, training hearing-impaired listeners to make better use of these cognitive mechanisms might possibly lead to better speech perception in real-life noisy environments. Therefore, the main aim of this research was to better understand how these cognitive speech restoration mechanisms work in general. The ultimate goal of the present study is to contribute to a reliable, valid, and effective training method to improve speech intelligibility in complex listening situations for hearing-impaired listeners, fitted with a cochlear implant (CI), an implantable auditory prosthesis. Even though there are a number of auditory training programs available, both in research settings and commercially, these have not been specifically designed to employ cognitive restoration mechanisms in training. For these purposes, we designed and performed a number of studies, of which the main results obtained are presented in this thesis.

PHONEMIC RESTORATION

By studying the restoration capacities of the auditory system of normal hearing listeners we gained more knowledge of how auditory perception of interrupted speech works in general. These studies are interesting per se for understanding more about the auditory perception of hearing-impaired listeners. In order to find (partial) answers to the main aim of this thesis, a well-controlled auditory scene, or rather, a well controlled and well-understood auditory stimulus is employed. This stimulus is periodically interrupted sentences with and without filler noise in the silent intervals, and induces restoration. Figure 1.1 shows this stimulus schematically; silent intervals interrupt the spoken sentence in such way that only fragments of the original sentence are audible.
Research on interrupted speech goes back to the fifties. Miller & Licklider (1950) described that an interrupted word can be perceived as continuous through noise. In their paper they described that if speech is interrupted, it is similar to seeing a landscape through a picket fence.

“It is much like seeing a landscape through a picket fence - the pickets interrupt the view at regular intervals, but the landscape is perceived as continuing behind the pickets” Miller & Licklider (1950)

This reflects the ability of the listener to perceive “glimpses” of speech between the silent interruptions or between the loud noise fillers of the manipulated speech signals. Miller and Licklider showed that words interrupted by silent intervals were less intelligible than uninterrupted speech and that interrupting with non-stationary maskers lead to better perception. They used a wide range of interruption rates in their study. Warren (1970) was the first to show that listeners believed they heard in a sentence a phoneme, which was only masked by an extraneous sound, while in reality the phoneme did not exist at all. The stimulus was a recording of the following sentence,

“The state governors met with their respective legislatures convening in the capital city”.

The first “s” in the word “legislatures” and the phonemes adjoining were replaced by a recorded cough. Participants were asked to encircle the exact position of the cough on the sentence written on paper. No participant identified the exact position of the cough and half of the participants circled beyond the word “legislatures”. This special type of auditory induction is called phonemic restoration, and is evoked when a portion of speech is perceived as present, while it is missing in reality, and replaced by noise (Warren and Sherman 1974). Phonemic restoration was also studied with periodic temporal interruptions filled with silence or with filler noise. For example, Powers & Wilcox (1977) examined interrupted speech with several interruption rates, and found that interrupted sentences were significantly more intelligible when interruptions were filled with noise (Fig.1.2) than when left as silent (Fig.1.1).
Subsequent studies have shown the positive influence of context and linguistic rules. Restoration seems to be easier with speech segments that are linked by context and linguistic rules (Sivonen et al. 2006b). Context influences possibly the prior knowledge needed to restore the interrupted speech (Shahin and Miller 2009). Linguistic skills, vocabulary, and verbal comprehension were indirectly suggested to be of importance on the perception of interrupted speech (Bashford and Warren 1987; Bashford et al. 1992; Bronkhorst et al. 1993).

Another factor influencing the benefit from restoration mechanisms is hearing impairment. Hearing-impairment changes the quality of the bottom-up speech signals, which makes it more difficult for the highly cognitive restoration mechanisms to do their work. Başkent, Eiler, & Edwards (2010) demonstrated that mildly hearing impaired listeners experience a similar restoration benefit from filler noise as NH listeners. However, negligible or no restoration benefit was observed in moderate hearing impairment. Front-end signal processing of hearing aids may have a negative affect on the sound signals presented to the hearing-impaired listener as well (Başkent et al. 2009). Nelson & Jin (2004) reported that severely hearing-impaired listeners using a CI experience even more difficulties in understanding interrupted speech. The spectral resolution is degraded in CI speech processing and the temporal fine structure is not fully transmitted (Nelson and Jin 2004; Chatterjee et al. 2010). CI users often face challenges in speech perception in complex acoustic listening situations, possibly due to limited help from the benefit of high-level restoration mechanisms. A possible cause is that the signal does not contain the necessary speech cues to induce cognitive restoration mechanisms.

Normal-hearing (NH) listeners presented with acoustic simulations of CIs face similar difficulties, suggesting that the spectrotemporally reduced speech signal, as it can occur in a real CI, may be contributing to the lack of beneficial phonemic restoration. CIs divide the sound into frequency channels and eventually stimulate the auditory nerve fibers via electrodes on the array implanted in the cochlea. The number of electrodes is limited; up to 24 electrodes are available, depending on the manufacturer of the CI, however, the actual functional spectral channels have been shown to be even smaller, possibly due to channel interactions (Friesen et al. 2001). This small number of independent stimulation channels stimulated simultaneously leads to spectral degradation of sound. This degradation has a detrimental effect on the restoration benefit of filler noise in interrupted speech (Başkent 2012), when tested with acoustic simulations of CIs. However, a recent study by Bhargava, Gaudrain, & Başkent (2014) found more promising results in CI users; a restoration benefit was observed when the speech segments were made longer.

![Figure 1.2](image1.png)

**Figure 1.2:** Schematic representation of interrupted speech with filler noise. The speech stream is periodically interrupted and the silent intervals are filled with bursts of speech shaped noise.
than the interruptions. Further, a restoration benefit was observed with a few CI users, who were also star participants. This suggests that there is a possibility that CI users can achieve benefit from restoration; however, it currently does not work for them as effectively as it does for normal-hearing individuals.

**Definitions**

*Phonemic restoration:* the brain fills in for missing parts of speech thus forming a continuous percept of the speech stream

*Phonemic restoration benefit of filler noise:* the increase in intelligibility of periodically interrupted speech when the silent intervals are filled with noise bursts

**OUTLINE OF THE THESIS**

Chapter 1 *Introduction* is a general introduction to restoration mechanisms of degraded speech. A short overview is given, describing phonemic restoration and the influences of hearing impairment on speech communication and on the effectiveness of restoration mechanisms.

**PART I**

In the first part of the thesis we report studies with NH listeners who were tested with interrupted speech without further spectral degradations. Before the studies presented in this thesis were performed it was under debate if interrupted speech without filler noise was an ecologically valid stimulus, since it does not occur in natural circumstances. Thus, it was unknown if the phonemic restoration benefit of filler noise was a real perceptual phenomenon, or simply a side effect of the unfamiliarity with this stimulus. If the effect is real, it would be persistent after intensive training with interrupted speech without filler noise, which would remove the unfamiliarity factor. Furthermore, it was implied in literature that linguistic and cognitive factors would influence the restoration capacities of people, however, it was unknown exactly which of these factors would be important for speech restoration.

In Chapter 2 *Perceptual Learning of Interrupted Speech* we explored the effect of training on the perception of interrupted speech with silent intervals and with filler noise. Furthermore, we studied how the restoration benefit of the filler noise would evolve during training with feedback (auditory feedback/sentences presented on monitor). An indication of the listening-effort involved is given.
In Chapter 3 ‘Individual differences in top-down restoration of interrupted speech: Links to linguistic and cognitive abilities’ we investigated whether linguistic skills, such as receptive vocabulary, and cognitive skills, such as overall intelligence, might account for the individual differences in the use of top-down restoration mechanisms of trained participants.

PART II

In the second part of the thesis we took a step towards the ultimate goal, namely, to be able to teach CI users to make better use of high-level repair mechanisms and thus to improve speech perception in complex listening situations. We tested and trained NH listeners with interrupted speech, that was degraded as it may happen in CI speech processing (spectrotemporally degraded sentences). Further, auditory and visual factors influencing the phonemic restoration benefit of filler noise, with or without such degradations on the speech signal, are studied.

In Chapter 4 ‘Perceptual learning of temporally interrupted spectrally degraded speech’ we studied whether the perception of CI simulations of interrupted sentences can improve despite poor baseline intelligibility performances. Then, we studied if listeners can derive a restoration benefit of the filler noise through training with feedback.

In Chapter 5 ‘The effect of visual cues on top-down restoration of temporally interrupted speech, with and without further degradations’ the influence of visual speech cues, in the form of an accompanying video of the speaker, on the enhancement of the intelligibility and phonemic restoration benefit of filler noise is investigated. This is performed with interrupted speech with and without spectrotemporal degradations, with and without additional visual cues.

In Chapter 6 ‘General conclusions’ is a discussion of the main outcomes of the studies in this thesis and of their implications and future perspectives for hearing impaired listeners and cochlear implant users.
PART I
TEMPORALLY DEGRADED SPEECH