The pitch hunt
Clarke, Jeanne Nora

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2017

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, 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), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
CHAPTER 5

TOP-DOWN REPAIR
OF INTERRUPTED SPEECH IN
ELECTRO-ACOUSTIC STIMULATION

CLARKE, J.,
GAUDRAIN, E.,
BASKENT, D.
Abstract

Cochlear implant (CI) users benefit differently from top-down repair of interrupted speech than normal-hearing (NH) listeners (Bhargava et al. 2014, Hear. Res. 309, 113-123). The poor pitch perception commonly reported in CI listeners might contribute to this difference, because voice pitch (F0) is a primary grouping cue that may be involved in repair mechanisms. Pitch cues can be available to bimodal CI users with low-frequency (LF) acoustic residual hearing. We thus investigated the effect of residual hearing on top-down repair mechanisms for bimodal CI users. Previous studies showed that simulating electro-acoustic stimulation (EAS) yielded better intelligibility of interrupted speech than simulating electrical stimulation alone. In another study we showed that the addition of F0 to CI simulation improved top-down repair in the spectral resolution range of CI users (Clarke et al., 2016). We thus expected to observe similar improvement in benefits with bimodal CI users wearing a hearing aid (HA) in the contralateral ear where low frequency information, rich in voice pitch cues, can be transmitted.

We tested twelve bimodal users in two hearing modes, CI only, and CI and HA together, to measure bimodal benefit. The top-down repair of speech was assessed with the phonemic restoration (PR) paradigm. The PR benefit was measured by the increase in intelligibility of interrupted sentences when the periodic silent interruptions were filled with noise. As CI users show PR benefit for different interruption parameters than NH listeners, we used different gap durations for the interruptions.

Group analysis results unexpectedly showed that adding the HA to the CI led to no bimodal benefit. Against our expectations again, no PR benefit was observed at the group level, even for conditions where it was previously shown that CI users benefited from PR. However, subjective reports showed a general preference for wearing the HA along the CI, and suggest that the additional HA provides better sound quality. Moreover, individual analysis showed that some bimodal users could benefit from the addition of the low frequency acoustic residual hearing to trigger top-down repair mechanisms. In addition, more participants could benefit from the acoustic cues delivered by their HA along their CI for top-down restoration of interrupted speech as the gaps were shortened. This suggests that every improvement in quality of bottom-up cues (in frequency and in time) may contribute with an additive effect.
to better top-down restoration, showing a powerful interaction between bottom-up cues and top-down repair mechanisms.

**Keywords:** bimodal benefit, phonemic restoration, electric and acoustic stimulation

“So close, no matter how far
Couldn’t be much more from the ear
Forever trusting what we hear
And nothing else matters”
(adapted from Metallica, 1991)
5.1. Introduction

When listening to speech in a noisy environment, some speech segments become unavailable to the listener because of the presence of competing sound sources. Top-down restoration is the brain’s capacity to reconstruct speech with missing segments. A common way to quantify top-down restoration is via the phonemic restoration (PR) paradigm (Bashford and Warren, 1987b; Başkent, 2010; Verschuure and Brocaar, 1983; Warren, 1970). The PR effect is measured as the difference in intelligibility between sentences periodically interrupted with silence intervals, or with the intervals filled with noise. When the interruptions are filled with noise (compared to silent interruptions), two things can be observed: (i) the speech is perceived as continuous behind the noise that acts as a plausible masker of the missing speech segments, (ii) intelligibility of interrupted speech increases with top-down repair mechanisms, using linguistic knowledge, expectations, and context (Bashford et al., 1992; Samuel, 1981; Verschuure and Brocaar, 1983; Wang and Humes, 2010; Warren and Sherman, 1974).

It was previously suggested that the continuity illusion perceived when the interruptions were filled with noise was a prerequisite for top-down repair of speech (Bashford et al., 1992; Başkent et al., 2009; Bregman, 1990). However, more recent studies suggested that perceived continuity and top-down repair of speech are two separate, but interacting, mechanisms (Bhargava et al., 2014; Clarke et al., 2014; Shahin et al., 2009; Shinn-Cunningham and Wang, 2008). For example, Bhargava et al. find no correlation between PR scores and perceived continuity. Clarke et al. (2014) show, in line with Bhargava et al.’s results, that PR is not impaired when voice continuity is disrupted. In this previous study, voice continuity is disrupted by manipulating voice characteristics such as the fundamental frequency (F0 — whose perceptual correlate is pitch), and the vocal tract length (VTL). Disrupted voice continuity lead participants to judge voices with VTL manipulation as from a different talker than the original voice. This indicates that perceived continuity of sentences built up with two voices from two perceived different talkers may also be disrupted. Furthermore, this lack of effect of voice manipulations on PR seems to indicate that NH listeners can rely on the linguistic context, to compensate for the manipulated voice cues, to still manage restoration of interrupted sentences. What is likely to happen in interrupted speech perception is that the silent interruptions may
introduce spurious cues, such as sudden starts and stops that might be erroneously interpreted as word boundaries and could thus impair word segmentation (Huggins, 1964; Repp et al., 1978). As a result, wrong word candidates may be activated, disturbing lexical access and thus hindering intelligibility. When the silent interruptions are filled with noise, these spurious cues may be hidden by the noise (if it is a plausible masker), resulting in the reduction of wrong word candidates being activated, and thus improving intelligibility (Bashford et al., 1992; Srinivasan and Wang, 2005; Warren and Obusek, 1971). Moreover, speech and noise are perceived as two separate streams. Good discrimination between speech and noise, along with proper linkage of the successive speech segments across time to form a coherent stream, are responsible for successful stream segregation, an underlying mechanism of top-down restoration of speech, which also contributes to perceived continuity.

Information on the voice of a talker is involved in general perceptual organization from which top-down repair of speech derives. Voice characteristics (F0 and apparent VTL) are involved in speech in noise perception notably via stream segregation (Gaudrain et al., 2007; Hartmann and Johnson, 1991; Tsuzaki et al., 2007), which is also an underlying mechanism of top-down repair of speech and perceived continuity. Voice characteristics may thus be helpful to link successive speech segments across interruptions. Voice discrimination primarily relies on voice characteristics. High pitch voices are categorized as female since women have mean F0 values in average twice as high as that of men (Titze, 1989). Besides the mean value of F0 that inform about speaker identity – gender and size – (Gaudrain et al., 2009; Ives et al., 2005), the F0 contour is the primary cue NH listeners use to recognize intonation that is also valuable information for stream segregation (Bregman, 1990), word segmentation (Cutler et al., 1997; Spitzer et al., 2007), and speech intonation perception (Peng et al., 2012). Clarke and colleagues (2014, 2016) specifically studied the effect of voice characteristics on PR with NH listeners. Although results from manipulated F0 showed no effect on PR (indicating that NH listeners may have relied on others cues, such as linguistic context, to perform PR), results from speech degraded by vocoding showed the importance of the presence of F0 cues when further spectro-temporal degradations were applied to interrupted speech.

Besides linguistic knowledge, expectations and context, top-down restoration of speech also seems to be triggered by the “right amount” of bottom-up cues (Bhargava...
et al., 2014). It has been argued that in the case of inherently degraded sound input (spectro-temporal degradations from a Cochlear Implant – CI – or vocoder), further degradation from temporal interruptions results in a bigger loss of intelligibility than for NH listeners when no such inherent degradations are present (Bhargava et al., 2016; Lacroix et al., 1979). However, Bhargava et al. (2014) showed that cochlear implant (CI) users could also benefit from top-down repair of interrupted speech, just differently than normal-hearing (NH) listeners. The poor pitch perception commonly reported in CI users might contribute to the difference in performance between CI users and NH listeners (Gaudrain et al., 2015; Heeren et al., 2012; Qin and Oxenham, 2003). Providing additional F0 information to noise-band vocoded speech improved top-down repair in the spectral resolution range of CI users (Clarke et al., 2016). This result confirms the role of bottom-up cues to trigger PR, and seem to indicate that the interaction between bottom-up and top-down mechanisms adapts to the difficulty of the listening situation. For example, when bottom-up speech cues are rich as is the case for NH listeners, the effect of a disrupted cue can be overcome by using other available cues to limit the decrease in intelligibility. However, the same disruption can become dramatic when the bottom-up speech cues are degraded themselves (such as in CIs), because the additional cues are not available or minimal, and intelligibility can break down.

The main goal of the present study was to investigate whether access to pitch can improve top-down repair of interrupted speech for actual CI users. We also investigated whether perceived continuity of interrupted speech was affected by access to pitch cues. This research question can be addressed with bimodal CI users monaurally implanted with a CI and wearing a hearing aid (HA) in the contralateral ear because they can have access to pitch cues via their low-frequency (LF) acoustic residual hearing transmitted by their HA. We expected the acoustic cues transmitted via the HA to be fused with the electric cues from the CI, contributing to a better pitch representation. The bimodal benefit is the increase in intelligibility achieved when speech is presented to both the CI and the HA, compared to the CI alone. However, Brown and Bacon (2009b) showed that a tone carrying F0 and amplitude envelope cues of the speech presented to the HA could be sufficient to achieve a bimodal benefit.

A previous study with acoustic simulations of CIs showed that simulating electro-
acoustic stimulation (EAS) yielded better intelligibility of interrupted speech than simulating electrical stimulation alone (Başkent, 2012). Similar benefit with actual bimodal CI users can be expected, as the LF cues provided by the HA would be expected to contribute to a better pitch representation, which would in turn facilitate F0 tracking across interruptions. This expectation is based on a long list of literature that showed that perception of voice characteristics may differ in hearing impairment or CI use compared to NH. First, sensory-neural hearing loss (SNHL) reduces auditory sensitivity, but also leads to suprathreshold deficiencies at level of moderate to severe hearing loss, that cannot be completely restored through amplification, and pitch perception may also be impaired (Glasberg and Moore, 1989; Moore, 1996; Plomp, 1978). Impaired pitch perception was measured in hearing impaired (HI) listeners and CI users (Gfeller et al., 2002; Sucher and McDermott, 2007; Summers and Leek, 1998), who typically show larger just noticeable difference (JND) for F0, i.e. they need a larger frequency change to perceive different complex tones compared to NH listeners. Being able to differentiate frequencies is of great help for voice discrimination, as each talker has unique voice characteristics. For example, Summers and Leek found that, unlike NH listeners with smaller JND, HI listeners with high JND could not take advantage of F0 differences between concurrent vowels to identify them. However, CI users can take advantage of their weak pitch percept for gender categorization (Fuller et al., 2014; Kovačić and Balaban, 2009). Indeed, CI users rely on F0 information (but not apparent VTL information, contrary to NH listeners) to predict the gender of a talker. For speech intonation perception task, question/statement discrimination studies have shown that, with poorer performance, cochlear implantees use F0 contours to a lesser extend than NH listeners, but also rely on the intensity of the end of the utterance that carries relevant information (Chatterjee and Peng, 2008). Chatterjee & Peng (2008) suggest that CI users can use the periodicity information in the temporal pattern of F0. Furthermore, when combined with residual hearing, CI users perform better in question/statement discrimination tasks and also have smaller F0 JND (Marx et al., 2015). This suggests that having some LF acoustic cues along the electric information improves F0 tracking needed for intonation perception. Improved F0 contours perception has also been observed for lexical segmentation in CI users with contralateral LF residual hearing (Spitzer et al., 2009). More generally, CI users that can access LF acoustic hearing in the non-implanted ear may present a bimodal benefit. That is, they perform better with their CI and contralateral HA (CI+HA) than with their CI alone (CI only), even if very little or
no intelligibility is observed with only the HA (Dorman et al., 2005; Kong et al., 2005). However, not all bimodal CI users show a bimodal benefit, and results of different studies show discrepancies; Dorman et al. (2015) showed that the bimodal benefit is best observed for sentence material (compared to isolated words) in noise (compared to quiet), when CI only performance is below 60% (to have room for improvement), and for contralateral thresholds corresponding to mild-to-moderate hearing loss (compared to severe hearing loss). Moreover, the LF residual hearing does not only convey F0 information, and better pitch tracking that facilitates sequential grouping might not be the only responsible for the bimodal benefit. Other LF phonetic cues can fall in the residual part of hearing, such as the first formant (F1) information, coarticulation (formant transitions), and voicing cues (Kong and Carlyon, 2007; Li and Loizou, 2008; Verschuur et al., 2013). These cues would favor glimpsing of the speech segments, in turn improving speech understanding, contributing to the bimodal benefit.

Although bimodal benefit was sometimes observed for speech understanding in quiet (Hamzavi et al., 2004; Shallop et al., 1992), greater benefit was usually observed for speech in noise (Armstrong et al., 1997; Zhang et al., 2010). It is not yet clear how acoustic and electric cues are integrated to improve speech perception, and how cognitive mechanisms deal with bimodal cues. The research question of the present study is to assess whether the use of a HA (providing additional pitch cues via LF residual hearing) contralateral to a CI device affects top-down repair of interrupted speech and perceived continuity. In the present study, we propose a recognition task of interrupted sentences, both with silent interruptions and when the silent gaps are filled with noise, to measure the top-down repair of speech (Experiment 1). Along with this task, perceived continuity of the same interrupted sentences (Experiment 2) is also conducted. To complement these objective measures, the short ‘Speech, Spatial and Quality of Hearing Scale’ (SSQ12) was proposed to assess the subjective outcomes of EAS (Noble et al., 2013). The phonemic restoration paradigm, i.e. difference in intelligibility between interrupted speech with filler noise and with silent, is used in Experiment 1 to measure top-down repair of speech; while the perceived continuity of the same interrupted stimuli is measured in Experiment 2. We hypothesized that access to LF would induce a bimodal benefit. If the LF information mostly contains pitch, access to LF would lead to better F0 tracking which would help sequential grouping mechanisms, and in turn lead to better continuity percept.
when the interruptions are left silent. With the filler noise interruptions, continuity percept could go both ways, as there is a trade-off between stream discrimination and sequentially linking successive speech segments. Discrimination would work better if the interleaved signals differ from each other, whereas sequential linkage would work better if the target and the masker were similar. On the other hand, if the LF information mostly contains phonetic cues, access to LF would help with lexical ambiguity, yielding better intelligibility of the remaining speech segments through glimpsing, which would boost top-down repair benefits. However, if the F0 perception is too degraded by the SNHL of the contralateral ear, addition of the HA may not bring useful extra LF information to strongly improve speech intelligibility or continuity perception.

5.2. Experiment 1: Top-down restoration of interrupted speech

5.2.1. Methods

Participants

Twelve CI users, aged 24 to 68 years (mean=57, s.d.=12) participated in this study. Details of the participants are presented in Table 5.1. All participants were bimodal CI users, i.e. implanted with a CI in one ear and wearing a HA in the contralateral ear. All but one participant were wearing their HA daily in addition to their main CI device. They all had at least one year of CI experience prior to the experiment. The use of double array electrodes was an exclusion criterion. The participants were either selected from the clinical database or replied to an online advertisement. All participants were native Dutch speakers except for one subject (who lived in the Netherlands for more than 30 years, and whose data did not deviate from more than one standard deviation from the group mean). For participants #1 to #6, we checked the patient records to make sure that they did not have deactivated electrodes in the mid- or high-frequency range. These data were not available for participants #7 to #12, who were recruited outside our clinic.

The CI is the device our participants relied primarily on for speech understanding.
Table 5.1. Details of CI participants. ‘n.a.’ indicates that the information was not available.

<table>
<thead>
<tr>
<th>Participant number</th>
<th>Sex</th>
<th>Age (years)</th>
<th>Duration of CI use (years)</th>
<th>CI Brand</th>
<th>HA Brand/Program</th>
<th>Processor</th>
<th>Strategy - Processor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>54</td>
<td>8</td>
<td>Cochlear Inc.</td>
<td>Nucleus</td>
<td>n.a.</td>
<td>n.a. - Nucleus</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>58</td>
<td>13</td>
<td>Cochlear Inc.</td>
<td>Nucleus</td>
<td>n.a. - n.a.</td>
<td>Cochlear Inc. - Freedom Widex/Naida V Sp</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>62</td>
<td>9</td>
<td>Cochlear Inc.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>MED-EL</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>60</td>
<td>5</td>
<td>Cochlear Inc.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>MED-EL</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>66</td>
<td>4</td>
<td>Cochlear Inc.</td>
<td>Nucleus</td>
<td>n.a.</td>
<td>Oticon/ChiliSP9</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>49</td>
<td>3</td>
<td>Advanced Bionics</td>
<td>HR90K HiFocus</td>
<td>n.a.</td>
<td>Oticon/AgilPro</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>60</td>
<td>10</td>
<td>Cochlear Inc.</td>
<td>Nucleus</td>
<td>n.a.</td>
<td>Otto - Rondo</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>60</td>
<td>1</td>
<td>Cochlear Inc.</td>
<td>Nucleus</td>
<td>n.a.</td>
<td>MED-EL</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>64</td>
<td>n.a.</td>
<td>Cochlear Inc.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a. - n.a.</td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>68</td>
<td>7</td>
<td>MED-EL</td>
<td>n.a.</td>
<td>n.a.</td>
<td>MED-EL</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>56</td>
<td>n.a.</td>
<td>Cochlear Inc.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>Oticon/Sumo XP</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>52</td>
<td>2</td>
<td>Cochlear Inc.</td>
<td>24RE (CS) - Nucleus</td>
<td>n.a.</td>
<td>Cochlear Inc. - Freedom</td>
</tr>
</tbody>
</table>

Note: The table provides details of participants including their sex, age, duration of CI use, CI and HA brands, and processor strategies.
Audiometric tests

First, the battery of the HA was checked. Then the HA specifications were measured (ANSI 1987), to ensure proper functionality of the HA before running the audiometric tests. The audiometric tests were conducted in a sound attenuated room, with warble tones presented at the audiometric frequencies between 125 Hz and 6 kHz. Three thresholds were measured:

i. Unaided for the residual hearing in the contralateral and ipsilateral ear of implantation (over TDH-39 headphones, separately in each ear; see Figure 5.1A and Figure 5.1B, respectively). Thresholds below 110 dB HL were plotted as not detected (N.D.). Low frequency pure tone average (LF PTA) was computed as the average of the thresholds measured at 125, 250 and 500 Hz.

ii. Aided with the CI only (free-field, binaurally; see Figure 5.2A).

iii. Aided while wearing both devices CI+HA (free-field, binaurally; see Figure 5.2B).

For the unaided thresholds in the implanted ear, measurements were not made for the individuals whose thresholds were lower than 100 dB HL in the patient charts.

The study was approved by the ‘Medisch Ethische Toetsingscommissie’ (Medical Ethical Review Committee) of the University Medical Center Groningen. The participants were informed about the procedure and signed a consent form prior to the experiment. Participants were paid for their participation.
Figure 5.1. Unaided thresholds for the contralateral unimplanted ear (panel A) and ipsilateral implanted ear (panel B), showing the residual hearing in each ear (circles for right ear, and crosses for left ear). N.D. refers to thresholds not detected. The color gradient indicates the LF PTA, calculated as the average thresholds at 125, 250 and 500 Hz.
Figure 5.2. Aided threshold with only CI (panel A) and with both CI and HA (panel B). Circles show binaural thresholds when the CI is on the right ear and crosses when the CI is on the left ear. The color gradient indicates the LF PTA of unaided thresholds.
Chapter 5

Stimuli

For intelligibility of interrupted sentences, we have used Dutch speech materials spoken by a male talker and digitized at 44.1 kHz sampling rate (Versfeld et al., 2000). Each sentence, consisting of four to nine words of maximum three syllables, is grammatically and syntactically correct. The corpus is divided into 39 homogeneous lists of 13 sentences. List 13 was discarded because it contains a sentence that is also present in list 21.

List 1 was used to adjust the presentation level of the stimuli. Participants were asked if the sentences were presented at a comfortable level. They were instructed not to modify their devices’ settings throughout the whole session (after they had adjusted it to an everyday life setting). Prior to testing, calibration of speech stimuli was done on uninterrupted sentences, to a fixed RMS level of 65 dB SPL. The calibration of the stimuli was performed with a Sound & Vibration Analyser (Svan 979 from Svantek) connected to a Kemar head (G.R.A.S.). If the presentation level was too soft for the participant, it was increased by 1 or 2 dB via the switch of the D/A converter.

For data collection, 24 sentence lists were used in the experiments: lists 7 to 31 (without list 13 that was discarded). Each of the 6 conditions, 3 duty cycles (DC: 50%, 62.5%, and 75%) x 2 interruption modes (IM: silent intervals and filler noise), was tested on two lists in the two hearing modes. From the remaining lists, two lists were used for baseline measurement and four lists for training of interrupted speech for the two listening modes (CI only, CI+HA).

At the beginning of each list the participants heard the same introduction sentence (“Buiten is het donker en koud” meaning “Outside it is dark and cold”) processed similarly to the sentences of the upcoming list to prepare them for the trial condition. Hearing a known sentence helped the participant get better prepared for the interruption condition of the upcoming list. Moreover, the introduction sentence was not included in the performance scores. Thus, always using the same introduction sentence allowed us to use all 13 sentences from each list, without discarding one for preparation.

To interrupt the sentences with periodic silent intervals, each sentence was modulated with a square wave with an interruption rate (IR) of 1.5 Hz and duty cycles
(DC) of 50 %, 62.5 %, and 75 %. These values were selected based on Bhargava et al. (2014) who observed a restoration effect with CI users with longer DC of 75 %. The speech-shaped noise file was provided with the sentence corpus, and was generated with white noise modulated by the long-term average spectrum of all sentences. The filler noise was produced by interrupting the speech-shaped noise with the inverse of the square wave that was used to interrupt the sentence. The filler noise was added at the SNR of 0 dB, with reference to the interrupted speech, for the conditions where the silent gaps were filled with noise bursts. A raised cosine ramp of 5 ms was applied to the onsets and offsets of the square waves to reduce spectral splatter and to prevent apparent dip in the total energy at the transition between speech and noise (Başkent et al., 2009). The sentence processing was done online in Matlab on a Macintosh computer.

**Experimental setup**

The participants were seated in an anechoic chamber during the experiment. Stimuli were sent through the S/PDIF output of an AudioFire 4 soundcard (Echo Digital Audio Corporation). They were converted to an analog signal via a DA10 D/A converter (Lavry Engineering Inc.), and played via a single loudspeaker (Tannoy) situated one meter from them. The spoken responses of the participants were recorded on a digital voice recorder (TASCAM). The participants then pressed the “Next” button on a touch screen (GPEG TFT monitor) in front of them to go to the next sentence.

**Procedure**

The experimental procedure was tested for the two hearing modes (randomly ordered): bimodal (with both devices; CI+HA) or CI only. For the CI only condition, if the participant had unaided thresholds better than 90 dB HL at the HA ear, then this ear was plugged with an Ohropax Color earplug (with more than 30 dB sound attenuation across all frequencies). Experiment 1 consisted of three parts: (1) measuring the baseline intelligibility of uninterrupted sentences, (2) a short training with interrupted sentences (with conditions not used in the experiment), (3) data collection for intelligibility of six conditions for each hearing mode. In all parts of the experiment, participants were presented one stimulus at a time, with a short beep preceding it to alert them.
Chapter 5

Every participant was tested on the same number of conditions, but the total testing times and the number of testing sessions differed across participants. The first four participants came for two two-hour sessions that were scheduled on separate days. They completed Experiment 1 followed by Experiment 2 in both sessions with different sentence lists and different random order of conditions. As it was more practical for participants to come for a single session, the procedure was adapted to fit in a single three-hour session for the last eight participants. Each condition in data collection was tested twice with two different sentence lists in a pseudo-randomized condition order. Testing time was reduced by combining the two sessions into one due to no repetition of the training and the instructions. The session included, in this order, obtaining written informed consents, conducting the audiometric tests, conducting Experiment 1, then Experiment 2, filling-in the SSQ12 questionnaire, the debriefing and occasional breaks.

1) Baseline intelligibility of uninterrupted speech

To measure the baseline speech intelligibility performance, participants listened to uninterrupted sentences. One sentence list (of 13 sentences) was used in each hearing mode (CI only and CI+HA). The baseline scores did not correlate with the age of the participant ($r^2 = 0.11, p = 0.29$) or the duration of CI use ($r^2 = 0.09, p = 0.40$).

2) Training with interrupted sentences

During training, participants listened to interrupted sentences to familiarize with the experimental paradigm. The task was to repeat what they could understand from the sentence they heard. The difference from data collection was, during training, written and auditory feedback was provided. More specifically, after the verbal response, the participants could read the complete sentence on the screen as they listened to the uninterrupted version of the sentence followed by the interrupted sentence once again (Benard and Başkent, 2013). Two sets of sentences were used that were processed with parameters that differed from the experimental conditions (i.e., IR = 1Hz and DC = 40% or IR= 2Hz and DC = 80%). One training condition was interrupted with silent intervals, the other combined with filler noise (chosen randomly for the first hearing mode, and counterbalanced for the second hearing mode).
3) Objective and subjective data collection

For data collection, participants were asked to verbally repeat, after hearing the stimulus, what they could understand from the interrupted sentence and were additionally encouraged to guess as much as possible. The data collection for the intelligibility experiment consisted of 6 conditions: 3 DCs (50%, 62.5%, and 75%) x 2 interruption modes (IM – silent intervals and filler noise), each tested twice in the two hearing modes (HM – CI only and CI+HA). Both the sentence lists and the conditions were presented pseudo-randomly. Functioning scores from the SSQ12 questionnaire were also collected. The SSQ12 questionnaire is a shorter version of the Speech, Spatial and Qualities of Hearing Scale (Gatehouse and Noble, 2004), in which the participants rate their subjective everyday hearing ability on a scale from 0 (lowest) to 10 (best ability). The full SSQ questionnaire was developed specifically for HI individuals and CI users to quantify hearing abilities in three domains (speech perception, spatial hearing, and sound quality), to complement objective measures. The full SSQ was translated from English to Dutch (available on the website: https://www.ihr.mrc.ac.uk/pages/products/ssq). We selected the Dutch translated version of the questions present in the English SSQ12 to have a Dutch version of the SSQ12. Total performance on the SSQ12 was calculated by averaging the scores of all questions. As a last task before debriefing, participants were asked to fill in the SSQ12 questionnaire two times, once as if wearing both their devices (CI+HA), then as if wearing their CI alone (CI only). Only seven participants fully filled in the questionnaire twice (CI+HA and CI only).

4) Data Analysis

Native Dutch speaking student assistants, who were blind to the experiment purposes, scored the recorded participant responses offline following well-defined guidelines. For each sentence, the percent-correct scores were calculated as the ratio of correctly identified words to the total number of words presented. In order to correct for the small variances at extremes of the percentage scale, percent-correct scores were converted into RAU (rationalized arcsine units, Studebaker, 1985), to help fulfill the variance homogeneity assumption for the ANOVA. Each condition being tested twice, the two measured RAU scores were averaged to have one intelligibility data point per condition. These averaged intelligibility scores were used to compute the phonemic restoration scores by subtracting the scores when the interruptions
were silent from when the interruptions were filled with noise. Statistical analyses were conducted in R (Lawrence, 2015; R Core Team, 2016; Revelle, 2015; Wickham, 2009). Effect sizes are reported with generalized eta squared ($\eta^2_G$ - Bakeman, 2005).

### 5.2.2. Results

The intelligibility scores are shown in Figure 5.3. Intelligibility scores (for each DC and IM) are displayed in the upper panels, while PR benefit is displayed in the lower panel for each HM (CI only in the left panel, and CI+HA in the middle panel). Bimodal benefit (the difference between CI+HA and CI only) is displayed on the right panel. Intelligibility scores increase with increasing DC (from left to right within each panel). We conducted a 3-way repeated measures ANOVA, with the DC (50 %, 62.5 %, and 75 %), the IM (silent or filler noise) and the HM (CI+HA and CI only) as within-subject factors (Lawrence, 2015). The results are summarized in Table 5.2. We observed a significant main effect only of the duty cycle on intelligibility scores. This confirms that the duration of speech segments significantly improved the interrupted speech intelligibility on average, from 18.45 RAU at DC = 50 %, to 35.97 RAU at DC = 62.5 %, and to 52.17 RAU for DC = 75%. However, the lack of a significant effect for the IM (silent or noise) suggests that participants did not show phonemic restoration benefit (lower panels of Figure 5.3). Moreover, the lack of effect of HM suggests that the addition of the HA did not benefit the participants in interrupted speech understanding nor in phonemic restoration (right panels of Figure 5.3). The main research question of the present study was if adding a HA to the CI would produce a more robust PR effect compared to the hearing mode of CI only. To answer directly our research question, i.e. whether the bimodal benefit affects the PR benefit, we computed t-tests at each DC. The three t-tests comparing the PR benefit at each HM for the three DCs were not significant ($t(11) = 0.69, p = 0.51, t(11) = 0.27, p = 0.79, t(11) = -1.85, p = 0.092$, respectively for 50%, 62.5%, and 75% DC).
Table 5.2. Results of the 3-way RM-ANOVA on intelligibility scores.

<table>
<thead>
<tr>
<th>Within subject factors</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duty cycle (DC)</td>
<td>$F_{2,22} = 81.45, p &lt; 0.001 ***$</td>
</tr>
<tr>
<td>Interruption mode (IM)</td>
<td>$F_{1,11} = 0.44, p = 0.52$</td>
</tr>
<tr>
<td>Hearing mode (HM)</td>
<td>$F_{1,11} = 3.87, p = 0.075$</td>
</tr>
<tr>
<td>DC x IM</td>
<td>$F_{2,22} = 0.38, p = 0.69$</td>
</tr>
<tr>
<td>DC x HM</td>
<td>$F_{2,22} = 1.087, p = 0.36$</td>
</tr>
<tr>
<td>IM x HM</td>
<td>$F_{1,11} = 0.51, p = 0.49$</td>
</tr>
<tr>
<td>DC x IM x HM</td>
<td>$F_{2,22} = 1.90, p = 0.17$</td>
</tr>
</tbody>
</table>

*** Significant (p< 0.001).

**Figure 5.3.** Intelligibility (top row) of interrupted speech with silence (empty boxes) and when the interruptions are filled with noise (dark filled boxes), and PR (light filled boxes, bottom row) results for the CI only condition (left panels), the CI+HA condition (middle panels) and the bimodal benefit (right panels). Scores are displayed as boxplot ordered by increasing DC from left to right in each panel. The horizontal line indicates the median, the box indicates the 25th and 75th quartiles, and the whiskers indicate the 1.5 interquartile range (IRQ).
Individual data analysis

As group analysis only showed an effect of the DC, we looked at the results at the individual level. Indeed, results from auditory perception tasks with CI users can be subject to large individual variability (for reviews, see Blamey et al., 2013; Faulkner and Pisoni, 2013) that might cancel out at the group level. This is the case in the study by Bhargava et al. (2014), where the PR benefit (or the lack of) becomes visible among CI users only after analyzing individual data. In case individual effects cancelled each other out, we would expect participants to be ranked the same way given their performance in different listening conditions. For this purpose, as a first step, the Spearman’s rank correlation coefficients ($r_s^2$) were computed on PR benefits at the different DCs, averaged over hearing modes (see Table 5.3). The results showed strong positive correlations between DC conditions, which indicates that participants were ranked in similar order given the DC conditions. This result suggests that participants with better PR performance in one DC condition also showed better PR in the other DC conditions; and that participants with poor or no PR in one DC condition, also showed poor or no PR in other DC conditions (in line with Benard and Başkent, 2013, who showed that participants either benefited from PR or not, but PR was not affected by training). Thus, it is possible that the lack of PR effect at the group level is cancelled out due to the averaging across individuals.

Table 5.3. Results of the Spearman’s rank correlation test on PR benefit scores for each DC, averaged over the two hearing modes. The p-values are corrected for multiple comparisons with the ‘fdr’ method.

<table>
<thead>
<tr>
<th>PR benefit at DCs</th>
<th>DC = 62.5 %</th>
<th>DC = 75 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC = 50 %</td>
<td>$r_s^2 = 0.73$, $p = 0.0008$ ***</td>
<td>$r_s^2 = 0.67$, $p = 0.002$ **</td>
</tr>
<tr>
<td>DC = 62.5 %</td>
<td>-</td>
<td>$r_s^2 = 0.62$, $p = 0.004$ **</td>
</tr>
</tbody>
</table>

Additionally, we wanted to assess whether individual PR performance correlated to baseline scores (see Figure 5.4.A). The upper row of Figure 5.4. displays individual PR performance as a function of baseline scores (of uninterrupted speech) for each HM (open and filled circles for CI only and CI+HA respectively) and each DC (increasing DC across panels from left to right). Although no significant correlation was reached (correlation coefficients displayed in each panel), some participants
showed consistency in their results, especially when performing at each extreme of the present performance range. Participants #5 and #9, with better baseline speech intelligibility, showed larger PR benefit, at all DCs. Adversely, participants #2 and #7, with lower baseline speech intelligibility, did not show PR benefit. These results seem to indicate that different patterns of performance are achieved likely due to individual behaviour, and that our participants capture a large range of performance.

As a next step, and to answer our research question whether the bimodal benefit had an effect on PR benefit, we further looked into individual data, because participants can show PR benefit at any DC condition. The bimodal benefit on PR, i.e. the difference between the CI+HA and CI only hearing mode conditions (displayed by arrows linking the open circles to the filled circles in Figure 5.4.A) is represented by arrows in Figure 5.4.B. We can see that more participants show bimodal benefit (arrows pointing upwards) with increasing DC (panels from left to right). To see whether a bimodal benefit on PR was present (regardless of the size of such benefit), we computed three one-sample t-tests comparing if the direction of the bimodal benefit (arrow going up, down, are flat) was greater than zero: t(11) = -1.30, p = 0.89, t(11) = -0.29, p = 0.61, t(11) = 1.82, p = 0.048 (for 50 %, 62.5 %, and 75 % DC, respectively). Only at 75 % DC is the bimodal benefit on PR significant, suggesting that more participants were able to take advantage of the additional cues from their HA to trigger top-down repair mechanisms, when they heard longer speech segments. Furthermore, it seems that the LF PTA is related to the bimodal benefit on the baseline (of uninterrupted speech). A significant positive correlation is found between the two variables ($r^2 = 0.58$, $p = 0.0041$). It indicates that the LF PTA predicted 58 % of the variance of the bimodal benefit on the baseline. Participants with lower PTA (i.e. better residual hearing), benefited from the addition of their HA along their CI for baseline performance. This suggests that the LF residual frequency plays a role in understanding uninterrupted sentences in quiet, and might be used as a predictive factor for speech understanding. However, LF PTAs were not correlated to the baseline scores, suggesting that the LF residual hearing did not predict the absolute baseline performance but the benefit gained from the addition of the HA along the CI. Such correlation was not found for intelligibility of interrupted speech or restoration mechanisms, suggesting that other factors were at play.
Further, we looked into individual data to see whether the bimodal benefit had an effect on intelligibility of interrupted speech. Figure 5.5. presents intelligibility of interrupted speech performance in a similar way as Figure 5.4. The upper row of Figure 5.5. shows the bimodal benefit on intelligibility as a function of baseline scores. The arrows linking the open circles to the filled circles in Figure 5.5.A represent the bimodal benefit, and are displayed in the lower row of Figure 5.5. It seems that more participants show bimodal benefit (arrow pointing upwards) with the lowest DC, suggesting that participants may have more room for improvement when the task is more difficult (shorter speech segments). One-sample one–tail t-tests were performed as previously done for the PR scores to assess whether a bimodal benefit on intelligibility was present. However, no significant bimodal benefit on intelligibility of interrupted speech \( (t(11) = 1.60, p = 0.069, t(11) = 1.17, p = 0.13, \) and \( t(11) = 0, p = 0.50 \) for 50 %, 62.5 %, and 75 % respectively) was observed.

**Subjective data: SSQ12**

The subjective results from the SSQ12 questionnaire are displayed in Table 5.4. The average self-perceived ratings (on a 0 to 10 scale from least to maximum hearing ability) showed a group-level bimodal benefit \( (t(6) = 3.04, p = 0.02) \). Indeed, each of the seven participant who completed the questionnaires rated their hearing abilities in the CI+HA hearing mode higher or equal to that of the CI only hearing mode. However, as most participants wore their HA 100% of waking hours, they found it difficult to evaluate their performance for the ‘CI only’ condition, thus these subjective scores should be interpreted cautiously. The aim of the SSQ12 questionnaires was to compare subjective and objective outcomes. However, the self-perceived benefit from wearing the HA along the CI was not correlated to objective measures of intelligibility of interrupted speech \( (r^2 = 0.43, p = 0.11) \) nor with PR benefit \( (r^2 = 0.15, p = 0.40) \). Thus it seems that the general preference of the bimodal users to also wear their HA along their CI, benefits the sound quality rather than their speech in noise performance.
Figure 5.4. Individual bimodal benefit on PR. A) Individual data are shown as scatterplots of PR benefit as a function of VU baseline, for 50%, 62.5% and 75% DC (panels from left to right, respectively). Open circles represent the CI only hearing mode condition labeled with the participant’s ID number, while the filled circles represent the CI+HA hearing mode condition. Circles are colored depending on the LF PTA (green for better and purple for worse LF residual hearing). Arrows are drawn for each participant linking data from the CI only condition to the CI+HA condition, representing the bimodal benefit. The regression lines for each hearing mode are presented in blue (dashed and solid line for CI only and CI+HA conditions respectively). The correlation coefficients of the regression models are given by $r^2$ for each panel, along with the ‘fdr’ corrected p-value. B) Same arrows representing the bimodal benefit as in panel A, except translated to a single ‘CI only’ origin point. The arrows point to the CI+HA condition labeled with the participant’s ID number. The bimodal benefit on PR is represented vertically, while the bimodal benefit on baseline speech intelligibility is represented horizontally.
Figure 5.5. Individual bimodal benefit on intelligibility. A) Individual data are shown as scatterplots of intelligibility of interrupted speech with silence as a function of VU baseline for 50%, 62.5% and 75% DC (panels from left to right, respectively). Open circles represent the CI only hearing mode condition labeled with the participant’s ID number, while the filled circles represent the CI+HA hearing mode condition. Circles are colored depending on the LF PTA (green for better and purple for worse LF residual hearing). Arrows are drawn for each participant linking data from the CI only condition to the CI+HA condition, representing the bimodal benefit. B) Same arrows representing the bimodal benefit as in panel A, except translated to a single ‘CI only’ origin point. The arrows point to the CI+HA condition labeled with the participant’s ID number. The bimodal benefit on intelligibility is represented vertically, while the bimodal benefit on baseline speech intelligibility is represented horizontally.
Table 5.4. Total SSQ12 subjective ratings averaged over the seven participants who completed the questionnaires for the two hearing modes (CI only and CI+HA).

<table>
<thead>
<tr>
<th></th>
<th>CI only</th>
<th></th>
<th>CI+HA</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>Mean</td>
<td>s.d.</td>
<td>Range</td>
<td>Mean</td>
</tr>
<tr>
<td>2.42 – 5.25</td>
<td>4.04</td>
<td>1.07</td>
<td>4.00  – 5.83</td>
<td>4.80</td>
</tr>
</tbody>
</table>

5.3. Experiment 2: Perceived continuity

Our main interest to measure perceived continuity in a bimodal setting was to investigate whether perceived continuity increases with the addition of the HA compared to the CI alone. This would give information on what LF cues are available to the HA, and contribute to the bimodal benefit. Indeed, if LF cues transmitted via the HA contain mainly F0 cues, we expected bimodal CI users to show stronger perceived continuity with the addition of the HA, as they would better be able to track F0 across interruptions. Furthermore, perceived continuity along with phonemic restoration measures would allow us to gain more insight on the relation between the underlying mechanisms involved in top-down repair of speech, specifically, to confirm that continuity illusion is not a prerequisite for restoration mechanisms.

5.3.1. Methods

Participants and Stimuli

The same participants took part in Experiment 2. Data from participants #1 and #2 were discarded due to a technical problem. The same sentences were presented in the same conditions and in the same order as the participants previously heard in Experiment 1.

Experimental setup

Experiment 2 was run after Experiment 1. Participants were asked to make a judgment on perceived continuity of the same sentences used in Experiment 1. For each stimulus, they were asked to judge if they heard it as continuous or interrupted (i.e. a part of the sentence missing). The participants performed this judgment task by clicking on the selected button (‘continu’ or ‘onderbroken’, meaning ‘continuous’
and ‘broken’ respectively) on the graphic interface on the touch-screen in front of them. Then, the next sentence was played. No training was provided prior to data collection as the task consisted in making a judgment (no right or wrong answer). Thus, no further training seemed necessary as participants were already familiar with the speech material and the procedure.

**Data Analysis**

The perceived continuity scores were calculated as the ratio of sentences perceived continuous to the total number of sentences presented in the set. Similar to Experiment 1, these scores were converted into RAU and averaged to have one data point per condition. Higher scores are obtained for sentences perceived as continuous, and lower scores for sentences perceived as interrupted.

### 5.3.2. Results

The perceived continuity (PC) scores are shown in Figure 5.6. PC scores (for each DC and IM) are displayed in the upper panels, while continuity benefit are displayed in the lower panel for each HM (CI only in the left panel, and CI+HA in the middle panel). Bimodal benefit (the difference between CI+HA and CI only) is displayed on the right panel. PC scores increase with increasing DC (from left to right within the panel), in line with sentences with shorter interruptions being perceived more continuous. Moreover, when the interruptions are filled with noise, the continuity percept also increased, in line with the filler noise masking the silent interruptions.

We conducted a 3-way repeated measures ANOVA, with the duty cycle (50 %, 62.5 %, and 75 %), the interruption mode (silent or filler noise), and the hearing mode (CI+HA and CI only) as within-subject factors. The results are summarized in Table 5.5. We observed significant main effects of the DC and of the IM on PC scores, as well as a significant interaction between these two variables. These results confirm that perceived continuity depends on the duration and filler of the interruptions. As gaps become shorter, interrupted speech is perceived more continuous. When the silent gaps are filled with noise, interrupted speech is perceived more continuous, especially for longer gaps. The effect of gap duration on perceived continuity is clear for each DC (see Table 5.6), whereas the effect of the filler noise improves the continuity percept at the two lower DC (longer gaps). However, the sentences are still perceived as interrupted below chance level or at chance level in the two lower duty
cycle condition (DC = 50 % and DC = 62 %, respectively). Sentences are perceived as continuous (above chance level) only at the highest DC (75 %), regardless of the interruption filler (silent or noise). However, no significant main effect was found for the hearing mode, indicating no bimodal benefit at the group level. But we cannot conclude that the addition of the HA did not provide any pitch cues to help F0 tracking across the interruptions. Further exploration of the individual data showed that very few participants perceived interrupted speech (with silent gaps) more continuous with the addition of the HA. This observation goes against the hypothesis that LF cues mostly contains pitch cues. Further analysis might be needed at the individual level to answer our research question, i.e. what LF cues are available to the HA and contribute to the bimodal benefit. However, the perceived continuity scores did not correlate with the intelligibility scores or with PR, confirming that perceived continuity is not necessary for top-down repair of interrupted speech (in line with Clarke et al., 2014).

Table 5.5. Results of the 3-way RM-ANOVA on perceived continuity (PC) scores.

<table>
<thead>
<tr>
<th>Within subject factors</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duty cycle (DC)</td>
<td>$F_{2,18} = 30.02, p &lt; 0.001$ ***  $\eta^2_G = 0.47$</td>
</tr>
<tr>
<td>Interruption mode (IM)</td>
<td>$F_{1,9} = 10.70, p = 0.0097$ **  $\eta^2_G = 0.14$</td>
</tr>
<tr>
<td>Hearing mode (HM)</td>
<td>$F_{1,9} = 0.051, p = 0.83$  $\eta^2_G = 0.00027$</td>
</tr>
<tr>
<td>DC x IM</td>
<td>$F_{2,18} = 3.77, p = 0.043$ *  $\eta^2_G = 0.023$</td>
</tr>
<tr>
<td>DC x HM</td>
<td>$F_{2,18} = 0.47, p = 0.63$  $\eta^2_G = 0.0012$</td>
</tr>
<tr>
<td>IM x HM</td>
<td>$F_{1,9} = 3.67, p = 0.089$  $\eta^2_G = 0.019$</td>
</tr>
<tr>
<td>DC x IM x HM</td>
<td>$F_{2,18} = 0.63, p = 0.55$  $\eta^2_G = 0.0038$</td>
</tr>
</tbody>
</table>

* Significant (p< 0.05), ** Significant (p< 0.01), *** Significant (p< 0.001).
Figure 5.6. Perceived continuity (PC) scores (top row) of interrupted speech with silence (empty boxes) and when the interruptions are filled with noise (dark filled boxes), and continuity benefit (light filled boxes, bottom row) results for the CI only condition (left panels), the CI+HA condition (middle panels) and the bimodal benefit (right panels). Scores are displayed as boxplot ordered by increasing DC form left to right in each panel. The horizontal line indicates the median, the box indicates the 25th and 75th quartiles, and the whiskers indicate the 1.5 interquartile range (IRQ).
Table 5.6. Post-hoc analysis on DC and Interruption filler variables for PC scores, pooled across hearing modes (i.e. averaged across CI only and CI+HA conditions). Results of the paired t-test ‘fdr’ adjusted for multiple comparisons are displayed along their p-values. The gray cells on the diagonal contain comparison to chance level (50 RAU).

<table>
<thead>
<tr>
<th>Interruption</th>
<th>Silent</th>
<th>Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>filler</td>
<td>DC</td>
<td>50</td>
</tr>
<tr>
<td><strong>Silent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td><strong>t = -7.99</strong></td>
<td><strong>p &lt; 0.001</strong></td>
</tr>
<tr>
<td>62.5</td>
<td>&lt;sup&gt;2&lt;/sup&gt; <strong>t = -5.01</strong></td>
<td>&lt;sup&gt;2&lt;/sup&gt; <strong>p &lt; 0.001</strong></td>
</tr>
<tr>
<td>75</td>
<td><strong>t = 2.82</strong></td>
<td><strong>p = 0.013</strong></td>
</tr>
<tr>
<td><strong>Noise</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td><strong>t = -3.07</strong></td>
<td><strong>p = 0.0095</strong></td>
</tr>
<tr>
<td>62.5</td>
<td>&lt;sup&gt;2&lt;/sup&gt; <strong>t = -0.12</strong></td>
<td>&lt;sup&gt;2&lt;/sup&gt; <strong>p = 0.91</strong> at chance level</td>
</tr>
<tr>
<td>75</td>
<td>&lt;sup&gt;2&lt;/sup&gt; <strong>t = 5.38</strong></td>
<td>&lt;sup&gt;2&lt;/sup&gt; <strong>p &lt; 0.001</strong></td>
</tr>
</tbody>
</table>

* Significant (p < 0.05), ** Significant (p < 0.01), *** Significant (p < 0.001),

C= perceived continuous (above chance level), I=perceived interrupted (below chance level)
5.4. Discussion

The main goal of the present study was to investigate whether pitch cues provided by LF residual hearing (via the HA) helped top-down repair and perceived continuity of interrupted speech in bimodal CI users. As CI users show restoration benefit for different interruption parameters than NH listeners (Bhargava et al., 2014), we used different gap duration (duty cycle) for the interruptions to change the difficulty of the task (longer gaps – smaller DC – make the task more difficult). The effect of DC is discussed first, then the effect of interruption mode, i.e. the top-down restoration. Finally, the effect of the bimodal benefit, i.e. the addition of HA along the CI, on PR and on intelligibility of interrupted speech is discussed.

Duration of the gaps: Duty cycle (DC)

Confirming our expectation, for shorter gaps, intelligibility and perceived continuity of interrupted speech both were significantly better than for longer gaps. This suggests that longer glimpses of speech facilitate intelligibility, and that F0 tracking is improved across the shorter gaps. However, DC did not affect the PR benefit. Larger DC allows for longer portions of speech to remain intact, resulting in lesser missing information. We expected CI users to show PR benefit at various DCs depending on their baseline scores, as demonstrated by Bhargava and colleagues (2014). Although getting more speech information (through shorter gaps) hinders intelligibility less, top-down repair mechanisms are not triggered adequately. This suggests that longer duration of bottom-up cues encoded in longer glimpses of speech provides better contextual information, that help improving intelligibility with both silent and noise interruptions. However, PR was not triggered as the gaps became shorter, suggesting that the filler noise could have also been a spurious cue for the bimodal CI participants of the present study.

A possible explanation is that shorter silent gaps have duration of typical closure in natural speech, and thus are more likely to be erroneously attributed to speech (Pickett and Decker, 1960; Repp, 1983). Similarly, shorter noise bursts, that share the characteristics with valid speech sounds (e.g. aperiodic bursts associated with voiceless fricatives, affricates and sibilants) and conversational sounds (e.g. nonlinguistic vocal sounds such as coughs) may be more likely to be attributed to speech, especially when the noise bursts approximate syllable duration (Verhoeven...
Furthermore, speech degraded through CI is more noise-like. Indeed, with perceptually similar spectral content, degraded speech and noise would be more difficult to segregate, and more errors can be made attributing noise bursts to be a part of speech. Additionally, perceptual similarity between degraded speech and noise might help noise to mask the absence of speech segments leading to better continuity illusion. Although in the present study, it seems that the trade-off between good segregation and masking, necessary for restoration benefit, was imbalanced. This explanation is in line with perceived continuity results from the present study that showed that bimodal CI users perceived interrupted sentences more continuous with shorter interruptions (75 % DC) regardless of the interruption mode (gaps left silent or filled with noise).

**Filling in the gaps: Interruption mode (IM)**

The main goal of the study was to investigate the effect of residual hearing on top-down repair of interrupted speech in bimodal CI users. Filling in the interruptions with noise can hide the spurious cues introduced by the silent gaps, favoring top-down restoration of speech with the filler noise interruptions. One of our hypotheses was that if the participants extract and integrate mainly phonetic cues from their LF residual hearing, phoneme identification would be facilitated, thus improving intelligibility of the speech segments glimpsed between the interruptions. Thus we expected to see an enhanced restoration benefit with the addition of HA. We also hypothesized that if participants extract and integrate mainly pitch contours from their LF residual acoustic hearing, F0 tracking would be facilitated, thus improving perceived continuity, especially when the interruptions are filled with noise.

Unexpectedly, when the gaps were left silent or filled with noise, participants performed the same, i.e. no restoration benefit was observed at the group level. However, the sentences were perceived more continuous when the gaps were filled with noise compared to when they were left silent. Even if the presence of the filler noise in the interruptions enhances perceived continuity, it did not enhance intelligibility, corroborating the independence of the two processes (Başkent et al., 2009; Bhargava et al., 2014; Clarke et al., 2014; Shahin et al., 2009).

The lack of PR benefit at the group level is difficult to explain as we used the same listening material with similar interruption parameters, almost the same inclusion
criteria for the participants (all except the NVA scores that were limited to above 70% in Bhargava’s study), and almost similar performance of baseline scores were achieved as Bhargava and colleagues (2014). The NVA scores were shown to be not correlated to PR benefit ($r^2 = 0.0051, p = 0.88$) in the present study, and thus it does not seem that our larger range of inclusion for NVA scores explains the difference of results observed between the two studies. Moreover, in the present study, the lack of PR does not seem to be due to failure to discriminate the speech from the noise at a group level, which could occur with limited number of electrodes in a CI, where the transmitted signal would not have enough details to make the difference between speech and noise. However, we could not identify specific factors explaining the differences in performance of the CI users in the present study from those in Bhargava et al (2014). Note that phonemic restoration is a noisy measure and subject to large individual differences across and within participants (as observed in the present study, and in line with Benard et al., 2014; Benard and Başkent, 2013, 2014; Bhargava et al., 2014). This might have affected the group-averaged results in the present study. Bhargava (2016) suggested that a large number of conditions should be measured with CI users in order to globally capture restoration benefit. In the present study, we only tested one SNR (0 dB) for the filler noise, whereas Bhargava and colleagues (2014) tested four different SNRs (5 dB, 0 dB, -5 dB, and -10 dB). It is possible that our set of conditions did not cover a range of conditions large enough to capture PR at the group level. A possible explanation for the participants who did not show PR is that the filler noise might have been more of a distractor for these CI users, whereas it usually facilitates comprehension – compared to silent gaps – for NH listeners.

Furthermore, to account for the lack of restoration benefit at the group level, even at the easiest condition of the present study (DC=75%), we analyzed the individual data. Some participants showed a restoration benefit, which was correlated to their baseline score (VU baseline) for each duty cycle. These results slightly differ from those of Bhargava and colleagues (2014). Participants in Bhargava et al. (2014) had a greater restoration benefit when they had a better baseline intelligibility for the more difficult condition (DC=50%), whereas, they showed a restoration benefit regardless of their baseline intelligibility, for the easier condition (DC=75%). Figure 5.7 shows the PR benefit as a function of VU baseline scores from the conditions that were common between Bhargava et al. (2014) (open circles) and the present
That is, for DCs of 50% (on the left panel) and 75% (on the right panel), for CI only hearing mode (our CI+HA condition was not included), and for SNR of 0 dB (other SNRs from Bhargava et al.’s study were not included). When selecting this subset of data, correlations are not significant, as it is with the full set of data reported in Bhargava et al. (2014). This corroborates the effect of individual variability on PR that can occur for different conditions for CI users. However, the results with the subset show the same trends as the full set results. And we can see that in the present study, participants’ baseline spread on a wider range, towards weaker baseline intelligibility than in the study from Bhargava et al. (2014), where participants were selected to be star performers. This may explain why baseline intelligibility and restoration benefit are also correlated at our easier condition (DC=75%). However, when selecting our participants from the same baseline range as Bhargava et al. (2014), we still do not observe similar correlation patterns as in their study. The baseline intelligibility predicted 50% of variance of the PR benefit (with similar correlation values of about $r^2=0.50$) in the present study as well as in

Figure 5.7. Scatter plot of PR benefit shown as a function of baseline scores for 50 % DC (left panel) and 75 % DC (right panel). Each individual data is labeled with the participant’s number. The regression lines are represented for both data (the dashed line for open circles and the solid line for filled dots). The correlation coefficients of each regression model are given by $r^2$ for each panel, along with the ‘fdr’ corrected p-value.
Bhargava et al. (2014). This indicates that other parameters might come into play, such as inter-personal differences in linguistic knowledge, acceptance of noise, selective attention, age-related effects, etc.

**Bimodal benefit on PR**

In the present study, no PR benefit was observed at the group level with the addition of the HA. It was previously shown in HI listeners that the more severe the hearing loss, the more impaired PR benefit was (Başkent et al., 2010) and the poorer understanding of interrupted speech was (Başkent et al., 2010; Jin and Nelson, 2010). Maybe, in the present study, the participants’ hearing loss in the non-implanted ear was too severe to observe any performance benefit. However, individual analysis (displayed on panel B of Figure 5.4) showed that some participants benefited from the addition of the HA for top-down restoration of interrupted speech regardless of their residual hearing (i.e. upward arrows in different colors displaying different LF PTA). This bimodal benefit on PR was also subject to individual variability. First, some participants showed a consistent negative restoration effect, i.e. were apparently disturbed from the addition of noise in the silent gaps when listening with only their CI. But with the addition of the HA, their intelligibility performance with interruptions filled with noise equated that of silent interruptions, suggesting that the filler noise was less of a distractor when listening with both their CI and HA. It seems that for these participants LF acoustic cues helped better discard the noise bursts as a masker. Second, other participants benefited from the additional LF acoustic cues to trigger or improve PR. On a positive note, more participants were able to take advantage of the additional cues from their HA to trigger top-down repair mechanisms, when they heard longer speech segments (i.e. at higher DC). Thus it seems that bimodal users benefit best from the additive effect of the better quality cues, both in frequency and in time.

**Bimodal benefit on intelligibility and perceived continuity of interrupted speech**

As stated before, the main goal of the study was to investigate the effect of residual hearing on top-down restoration and perceived continuity of interrupted speech in bimodal CI users. We hypothesized that the LF residual hearing provided by the HA would improve pitch perception, allowing a better intelligibility of the successive
speech segments. Thus we expected speech perception to improve with the addition of the HA. Unexpectedly, we did not observe a bimodal benefit for the interrupted speech perception task at the group level. The bimodal benefit was also not observed for the perceived continuity task at the group level. The present results do not fully support the grouping theory (improved voicing representation to help segmentation and improve sentence recognition), as perceived continuity was stronger when the silent gaps were filled with noise, but did not improve with the addition of the HA. This suggests that the better sequential linking of successive speech segments across interruptions may have relied on other cues than tracking of F0 contours.

The present results show large individual variability, as generally observed for CI users (Dunn et al., 2005; Heo et al., 2013; Mok et al., 2006). The factors that contribute to such large variability in CI outcome have been identified as duration of deafness, duration of CI use, aided/unaided residual thresholds, worse/better ear implanted, age-related cognitive factors (Blamey et al., 2013), but it is not clear which of these factors are involved in the present study. For example, Heo and colleagues (2013) observed large individual variability in bimodal benefits across participants and type of speech materials used, but could not identify a single factor to explain these differences. However, some individual factors are rejected as influencing bimodal benefit, such as the duration of device use (CI and CI+HA) that does not seem to affect bimodal benefit (Dunn et al., 2005). Thus, it is difficult to generalize our findings to the whole EAS population. Here, the discussion will be based on our small sample of participants and speculations will be limited accordingly. Comparison with simulation studies are also made with caution, as such studies are capitalizing on the potential bimodal benefit, while results from actual bimodal patients could differ due to additional factors that can affect performance but not always included in acoustic simulations of CIs (Başkent et al., 2016; Bhargava, 2016). Indeed, in simulation studies, NH listeners are presented with a fixed degree of spectrotemporal degradation, and individual differences inherent to the hearing loss of the CI users are not always captured.

In the present study, all participants had severe hearing loss in the contralateral ear. Only 5 participants out of 12 had low-frequency pure tone averages (PTA), computed as the average unaided thresholds at 125, 250, and 500 Hz, below 62 dB HL (represented in green in the color gradient). Dorman et al. (2015) suggested
that low PTAs might explain the lack of bimodal benefit. In the present study, LF PTAs were correlated with intelligibility of uninterrupted speech (baseline scores), with participants with little residual hearing showing no bimodal benefit for uninterrupted speech. However, no such correlation was found for intelligibility of interrupted speech or phonemic restoration. Moreover, Dunn and colleagues (2005) speculated that the differences in bimodal benefits across participants for speech perception may be due to the amount of residual hearing and speech information amplified by each participant’s HA (but residual hearing was not assessed). Mok and colleagues (2006) showed that the individual differences in bimodal benefit scores between participants could be accounted for by differences in aided thresholds in the non-implanted ear. This finding was also observed by Jang and colleagues (2014) who specified that the aided thresholds are more effective in predicting localization ability than speech perception performance. The aided thresholds negatively correlated with speech perception in quiet: participants with more favorable aided hearing levels (above 50 dB HL) performed better than those with poorer thresholds (below 50 dB HL), and no such correlation was found for the unaided thresholds. Unfortunately, in the present study, aided thresholds in the non-implanted ear were not measured during audiometric tests. But, in line with Jang and colleagues (2014), we found no correlation between the unaided thresholds in the non-implanted ear and bimodal benefit on intelligibility scores.

Other factors than the residual hearing level may influence the bimodal benefit. Rather than audibility in the non-implanted ear (NIE), that has been studied more extensively as it was thought to be of main influence on the bimodal benefit, the nature and amount of the LF cues present in the residual hearing has been the object of a debate in recent studies. On one hand, Zhang and colleagues (2010) argue that F0 accounts for the majority of the bimodal benefit, as it is observed even when the residual acoustic hearing is simulated by low-passed (LP) filtered speech at 125 Hz. The information from F0 improves voicing representation when speech is presented in quiet, which reduces the word candidates in the lexicon, improving speech perception. When speech is presented in noise, the better voicing representation would give access to LF acoustic landmarks that indicate syllable structure and word boundaries, which can help segmentation, improving sentence recognition (grouping theory). Similarly, LF cues, even when unintelligible when presented alone in the contralateral ear with normal hearing, improve speech in
noise perception when combined with the CI ear, indicating that additional LF cues are used to separate two talkers (Cullington and Zeng, 2010). On the other hand, other EAS simulations have shown that other cues than F0 contours are used for the bimodal benefit in noise. Kong and Carlyon (2007) contend that LF phonetic cues such as F1 (below 500 Hz), coarticulation cues, LF consonant cues, and voicing cues (presence/absence of F0), contribute to the advantage of adding LP filtered speech to vocoded speech through better glimpsing of the speech cues in the presence of a fluctuating masker (glimpsing theory). Indeed, Li and Loizou (2008) showed that the large bimodal benefit observed when the masker is a competing talker disappeared when the masker is a steady-state noise. In steady-state noise, there is no fluctuation in the temporal envelope, which prevents glimpsing of the target as is possible in the dips of a competing voice (when there is a drop in energy). Of course, when comparing conditions with and without F0 information, both F0 contours and voicing cues are usually affected, and have not been investigated independently to our knowledge. Furthermore, regarding F0 contours, CI users with weak F0 representation adapt to rely on other cues of stress and intonation, such as amplitude and duration (Hegarty and Faulkner, 2013). If bimodal participants also primarily rely on those other cues, that are transmitted in the CI, addition of the HA along the CI may reduce the bimodal benefit. However, in the present study, some participants did not show a bimodal benefit. It seems improbable that the additional cues from the LF residual hearing (via the HA) became redundant with the amplitude and duration cues transmitted in the CI, that the whole bimodal benefit would be counteracted (although it might be reduced).

Before the EAS became widespread, it was speculatively assumed that potential perceptual incompatibilities between electrical stimulation by CI and acoustic stimulation by HA may arise, such as pitch, dynamic range, and loudness differences. However, the bimodal benefit is now well established, and pitch differences between the two ears were investigated by Green and colleagues (2014) who showed that the bimodal benefit is not negatively affected by an overlap of frequencies between the CI and the HA ears. If present, a pitch mismatch does not impede bimodal benefit. It is unclear though how loudness mismatch between the two device affects the bimodal benefit and whether a bimodal fitting for loudness balance is favorable to the bimodal benefit (Blamey et al., 2000).
Despite the lack of bimodal benefit at the group level in the present study, the participants anecdotally reported preferring wearing their HA along their CI for daily listening situations, as supported by the significant difference in subjective functioning scores (SSQ12) between the two hearing modes (CI only and CI+HA). This result is in line with Ching and colleagues (Ching et al., 2004) who found that almost half their participants reported better functioning in everyday life despite no bimodal benefit in objective measures was observed. In more details, subjective outcomes of EAS via the full SSQ questionnaire showed that ratings in the sound quality domain may be related to the ability to process and integrate the electric and acoustic signals, whereas ratings in speech-related domain were not associated with objective performance (Heo et al., 2013). The present experiment did not capture a bimodal benefit on objective measures from the addition of the HA, nonetheless, the lack of improvements in interrupted speech intelligibility does not exclude other potential benefits, such as a better sound quality representation and/or reduced listening effort when they are wearing their HA. Furthermore, it has been argued that the LF information provided by the HA allows better glimpsing and better simultaneous integration of the glimpses over frequencies, with the LF coming from the HA and the HF from CI (Li and Loizou, 2008). Better ability to glimpse the target contributes to EAS benefit for speech in noise, via the LF SNR advantage in voiced segments. The case of interrupted speech, where the target and the masker periodically alternate, is a particular case of glimpsing where the most favorable SNR occur on the ON speech segments and worse SNR on OFF speech segments. Our participants reported preferring wearing their HA along their CI compared to only their CI, and showed a bimodal benefit on the subjective measure (SSQ12), but did not show bimodal benefit on objective measures (intelligibility and restoration of interrupted speech) at the group level. This suggests that the HA may provide new information (that is glimpsed) that fails to be integrated with CI information to elicit the percept of a fused sound from the electric and acoustic information. Although this question whether the electric and acoustic signals are fused into a single sound was not assessed in the present study, one participant anecdotally reported hearing the sounds coming from the implanted ear after the aided ear. Overall, the possible fusion failure might explain their better subjective sound quality but absence of intelligibility improvement with the addition of the HA.

Another factor affecting the bimodal benefit is the type of listening material used and
the experimental task. A greater bimodal benefit was found for sentence recognition with a competing talker (of opposite sex) than other maskers or in quiet (e.g. Heo et al., 2013; Li and Loizou, 2008). There are only few studies, to our knowledge, on bimodal benefit with interrupted speech. First, the work of Başkent and colleagues (2010, 2012) showed a small bimodal benefit when LP filtered speech (<500 Hz) was added to high spectral resolution CI simulation (16 and 32 channels), whereas the bimodal benefit was larger for lower spectral resolution (4 and 8 channels). These results suggest that additional LF bottom-up cues are of greater help to link successive speech segments when spectral degradations are more severe. Second, very recent studies with simulated bimodal hearing showed that bimodal benefit reduces when speech is interrupted with silent gaps compared to uninterrupted speech (Kong et al., 2015; Oh et al., 2016), suggesting that additional LF bottom-up cues are of lesser help for temporal interruptions than for more severe spectral resolution degradations. Oh and colleagues further suggest that the bimodal benefit rely on interactions between bottom-up cues and top-down repair mechanisms, in line with results from actual CI users from Bhargava et al. (2014).

Binaural hearing advantage for bilateral CIs over unilateral CI has been found for spatial hearing, segregation, naturalness, and listening effort subjective ratings (Noble et al., 2008). When contralateral residual hearing is available, then bimodal hearing is recommended and encouraged before mentioning a second CI. Stimulation of the NIE via sound amplification (with a HA) would avoid further deterioration due to auditory deprivation. The addition of an HA along a CI showed better sound quality and better quality of life ratings compared to unilateral CI use (Farinetti et al., 2014), as well as improved speech intelligibility in noise (Armstrong et al., 1997; Zhang et al., 2010). However, large individual differences were observed across participants (Dunn et al., 2005; Heo et al., 2013; Mok et al., 2006), and Dunn and colleagues suggested that a bimodal fitting of the CI and the HA could benefit bimodal users instead of independently fitting each device. To try to improve the bimodal benefit, new strategies and fitting methods are developed (Francart et al., 2015; Morera et al., 2012; Perreau et al., 2013; Ullauri et al., 2007; Veugen et al., 2016). In conclusion, some of the participants who did not show bimodal benefit for interrupted speech in the present study might benefit from an optimal fitting and synchronization of the two devices (CI and HA) together.
Acknowledgments

The authors would like to thank the participants, as well as the student assistants Julia Verbist, Wilke Bosma, Charlotte Deblecourt and Emma Agneray who assisted in collecting the data and transcribed participants’ responses. Many thanks to clinical staff, especially Cochlear Implant Northern Netherlands (CINN), Rolien Free, Bert Maat and Emile de Klein, for their support. The authors also acknowledge William Noble (University of New England, Australia) and Stuart Gatehouse (MRC Institute of Hearing Research, Scotland) for developing The Speech, Spatial and Qualities of Hearing Scale, as well as Liesbeth Royackers (ExpORL, K.U. Leuven, Belgium) for the Dutch translation, and Sophia Kramer (VU MC, Amsterdam, The Netherlands), Wouter Dreschler (AMC, Amsterdam, The Netherlands), Hans Verschuure (ErasmusMC, Rotterdam, The Netherlands), William Damman (AZ St. Jan, Brugge, Belgium), Astrid van Wieringen (ExpORL, K.U.Leuven, Belgium), and Heleen Luts (ExpORL, K.U.Leuven, Belgium) for the evaluation of the translation. This study was supported by a VIDI grant from the Netherlands Organization for Scientific Research, NWO, awarded to the last author (grant no. 016.096.397; from Netherlands Organization for Health Research and Development, ZonMw). Further support came from a Rosalind Franklin Fellowship from University of Groningen, University Medical Center Groningen, and funds from Heinsius Houbolt Foundation. The study is part of the ‘Healthy Aging and Communication’ research program of the Otorhinolaryngology Department of University Medical Center Groningen.