Hearing Research 350 (2017) 226–234

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

Masking release with changing fundamental frequency: Electric acoustic stimulation resembles normal hearing subjects

Alice Barbara Auinger [a], Dominik Riss [a], Rudolfs Liepins [a], Tobias Rader [b], Tilman Keck [c], Thomas Keintzel [d], Alexandra Kaider [e], Wolf-Dieter Baumgartner [a], Wolfgang Gstoettner [a], Christoph Arnoldner [a, *]

[a] Department of Otorhinolaryngology, Medical University of Vienna, Waehringer Guertel 18-20, 1090, Vienna, Austria
[b] Audiological Acoustics, Department of Otorhinolaryngology, University Hospital Frankfurt, Theodor-Stern Kai 7, 60590, Frankfurt am Main, Germany
[c] Department of Otorhinolaryngology, Krankenhaus der Elisabethinen, Elisabethinergasse 14, 8020, Graz, Austria
[d] Department of Otorhinolaryngology, Klinikum Wels-Grieskirchen, Grieskirchner Straße 42, 4600, Wels, Austria
[e] Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Spitalgasse 23, 1090, Vienna, Austria

ARTICLE INFO

Article history:
Received 1 November 2016
Received in revised form
4 March 2017
Accepted 8 May 2017
Available online 11 May 2017

Keywords:
Cochlear implant
Hearing preservation
Electric-acoustic stimulation
Release from masking

ABSTRACT

It has been shown that patients with electric acoustic stimulation (EAS) perform better in noisy environments than patients with a cochlear implant (CI). One reason for this could be the preserved access to acoustic low-frequency cues including the fundamental frequency (F0). Therefore, our primary aim was to investigate whether users of EAS experience a release from masking with increasing F0 difference between target talker and masking talker. The study comprised 29 patients and consisted of three groups of subjects: EAS users, CI users and normal-hearing listeners (NH). All CI and EAS users were implanted with a MED-EL cochlear implant and had at least 12 months of experience with the implant. Speech perception was assessed with the Oldenburg sentence test (OlSa) using one sentence from the test corpus as speech masker. The F0 in this masking sentence was shifted upwards by 4, 8, or 12 semitones. For each of these masker conditions the speech reception threshold (SRT) was assessed by adaptively varying the masker level while presenting the target sentences at a fixed level. A statistically significant improvement in speech perception was found for increasing difference in F0 between target sentence and masker sentence in EAS users ($p = 0.038$) and in NH listeners ($p = 0.003$). In CI users (classic CI or EAS users with electrical stimulation only) speech perception was independent from differences in F0 between target and masker. A release from masking with increasing difference in F0 between target and masking speech was only observed in listeners and configurations in which the low-frequency region was presented acoustically. Thus, the speech information contained in the low frequencies seems to be crucial for allowing listeners to separate multiple sources. By combining acoustic and electric information, EAS users even manage tasks as complicated as segregating the audio streams from multiple talkers. Preserving the natural code, like fine-structure cues in the low-frequency region, seems to be crucial to provide CI users with the best benefit.

1. Introduction

Over the last 25 years cochlear implantation has become the standard procedure for restoring substantial hearing in the profoundly deaf. The expansion of selection criteria for implants in patients with residual hearing started more than a decade ago. In 1999 von Ilberg et al. (1999) introduced a new treatment modality for patients with preserved low-frequency hearing and complete hearing loss at high frequencies. Since then several studies have proved the superiority of bimodal ‘electric-acoustic stimulation’ (EAS) over either modality on its own. These advantages were mainly observed in connection with increased speech recognition in noisy environments, subjective improvements in sound quality when listening to music and in other situations, where the poor frequency resolution and lack of fine spectral resolution of the electric stimulation on its own has limited performance (Arnoldner

* Corresponding author.
E-mail address: christoph.arnoldner@meduniwien.ac.at (C. Arnoldner).

http://dx.doi.org/10.1016/j.heares.2017.05.004
0378-5955/

Abbreviations

EAS	electric acoustic stimulation
F0	fundamental frequency
CI	cochlear implant
NH	normal-hearing
OlSa	Oldenburg sentence test
SRT	speech reception threshold
EM	energetic masking
IM	informational masking
ΔF0	difference in F0
SNR	signal to noise ratio
NA	not available
PSOLA	pitch-synchronous overlap-add synthesis

et al., 2009; Ilberg et al., 1999; Kiefer et al., 2005; Rader et al., 2013; Turner et al., 2008).

Despite all efforts, everyday communication environments, where multiple concurrent speakers are involved, still present a major challenge for hearing impaired people. Even normal-hearing people often have difficulties following conversations in background noise. This is because the original speech signal the subject tries to follow (target signal) is obscured by two different types of masking. *Energetic masking* (EM) occurs when the spectral energy of both the target and masking signals overlaps so that portions of one or both of the speech signals are rendered inaudible at the periphery. Two signals akin to each other regarding their spectral energy cause similar excitation patterns along the basilar membrane (Carhart et al., 1969). *Informational masking* (IM) is supposed to be a central auditory phenomenon and results when both signal and masker are audible, but the listener is unable to distinguish between the two elements due to the listener's inability to ascertain which element belongs to the target and which to the masker signal (Durlach et al., 2003a, 2003b). All listeners are affected by EM and IM, but due to their poor resolution hearing impaired listeners are more affected. Brungart et al. (2001) studied the interactions between target and masking talker in normal hearing listeners. The coordinate response measure (CRM) (Bolia et al., 2000), which was originally used for military purposes, served as their test material and consists of sentences with the common structure “Ready <call sign>, go to <color> <number> now”. Two phrases from the test corpus were presented simultaneously. Listeners had to repeat the color and number following the <call sign> “Baron”. Masker sentences were randomly selected with different call signs, colors and numbers. Results indicated that IM rather than EM dominates speech perception performance, which was lowest when the masking speech was taken from the same male talker. Performance increased when a different talker of the same gender was used as masker and increased even more for female masking talkers. Aside from a variety of cues like spatial separation, interaural time differences, F1 formants and co-articulation cues that contribute to resolving complex hearing tasks (Bronkhorst, 2015; Brown and Bacon, 2009; Dillon et al., 2015; Kong and Carlyon, 2007; Sheffield and Zeng, 2011), the fundamental frequency (F0) in the low-frequency range seems to be crucial to speech segregation in challenging listening situations; the more distinct the difference in F0 between two sources, the better the speech perception and a so-called release from masking occurs (Brungart, 2001; Carroll et al., 2011; Cullington and Zeng, 2009; Zhang et al., 2010).

In the current study, we focused on F0 as a major cue for speech discrimination and evaluated whether the superior performance of EAS patients in background speech can in fact be attributed to the better perception of F0 differences (ΔF0) in this group of patients. The influence of F0 on speech understanding has been widely investigated. Wilson et al. (2005) stated that one reason why EAS is superior to both electric and acoustic stimulation alone is, that the so-called fine structure – the very detailed frequency information the brain uses to enhance the hearing experience – is presented without modification in the low-frequency range. This fine structure information includes F0s and the first couple of harmonics along with at least some indication of first-formant frequencies for speech.

In contrast, cochlear implant users are particularly adversely affected by competing noises due to their limited spectral resolution and their reliance on envelope-based speech coding strategies. In quiet, this limited spectral resolution can be sufficient for speech understanding, whereas in background noise it is responsible for a definite deterioration in performance in CI patients (Brown et al., 2016; Cullington and Zeng, 2008; Nelson et al., 2003; Turner et al., 2008). In actual CI users the influence of ΔF0 on increasing speech performance is less clear: some authors found no effect of F0 on intelligibility (Mulhern and Cullington, 2014; Qin and Oxenham, 2003; Stickney et al., 2007, 2004; Summers and Leek, 1998) or only minor effects (Chatterjee and Peng, 2008; Visram et al., 2012). How F0 contributes to enhanced speech performance is still discussed controversially and little is known about how EAS users could benefit from ΔF0 between target and masker signals. In contrast to performance in continuous noise, better segregation and glimpsing cues have been discussed for enhanced speech intelligibility in multi-talker situations by means of F0 information (Li and Loizou, 2008; Nelson et al., 2003; Turner et al., 2004).

We therefore aimed at investigating the influence of ΔF0 between target and masker talker in actual EAS users, as this has so far only been investigated in bimodal users (CI ipsilateral and hearing aid contralateral) or simulation studies. Pyschny et al. (2011) tested the influence of bimodal stimulation upon speech recognition in the presence of a single competing talker. In their study, the Oldenburg sentence test (OlSa) served as test material. Masker sentences were manipulated regarding their F0 and formants. Across all target to masker conditions, improved target-masker separation with bimodal fitting could not be found, but for all three listening conditions (only CI ipsilateral, only hearing aid contralateral, CI and hearing aid together) a large ΔF0 enhanced speech performance (especially in the bimodal condition when F0 was raised by 80 Hz). One has to bear in mind that they included only good CI performers, who achieved speech understanding of at least 90% in quiet with the OlSa material. This could be the reason why their CI subjects also benefited from ΔF0, though to a lesser extent. Similar results were observed by Shpak et al. (2014). The authors found that increasing ΔF0 between target and competitor speech did result in higher performance scores in both the bimodal condition and the CI alone condition. Interestingly, they stated that the lower the unaided and aided thresholds at low frequencies are, the greater the bimodal benefit is. Conclusions from the latter two studies should be treated with caution, as they assessed performance in a bimodal (CI ipsilateral and hearing aid contralateral) condition and no “real” EAS condition. Still, if ΔF0 in bimodal stimulation can help to improve speech performance, the question remains how this affects the performance of real EAS listeners, which we aimed to shed light on in this study.

To assess the influence of ΔF0 on speech intelligibility, we chose the OlSa testing material. With its low predictability, it seemed to be an appropriate speech test for the purpose of our study. One sentence served as a masker; the F0 was shifted upwards by 4, 8 and 12 semitones. The masker sentence was presented