On a combination of amplitude and frequency modulation used for processing speech signals in cochlear implants

Svetlin Antonov¹ and Snejana Pleshkova-Bekiariska²

Abstract – A research, which combines the measurement of both amplitude and frequency modulation of speech signals and their processing in the processing unit of the cochlear implant, is being proposed. Numeric simulation is used as the basis for a comparison between the usage of the aforementioned combination of both modulations and the usage of only amplitude modulation. Using the proposed algorithm, a comparison between the original and processed signals is drawn.

Keywords – Cochlear implants, amplitude and frequency modulation, speech processing.

I. INTRODUCTION

Acoustic characteristics in speech signals allow listeners to derive not only the meaning of the speech but also the speaker’s identity and emotion. Previous studies using either naturally produced whispered speech [1] or artificially synthesized speech [2] [3] have isolated and identified several important acoustic cues for speech recognition. For example, computers relying on primarily spectral cues and human cochlear-implant listeners relying on primarily temporal cues can achieve a high level of speech recognition in a quiet environment [4] [5].

The goal of this study is to verify the relative contributions of spectral and temporal cues to speech recognition in realistic listening situations. A speech signal produced by a male talker is chosen for this purpose. We propose a combination of slowly varying amplitude modulation (AM) and frequency modulation (FM) from a number of frequency bands in speech signals and testing their relative contributions to speech recognition in acoustic and electric hearing. Different from previous studies using relatively fast FM to track formant changes in speech production [8] [11] or fine structure in speech acoustics [9] [10], the “slow” FM used here tracks gradual changes around a fixed frequency in the subband. We evaluate the AM-only, AM plus FM, and the original unprocessed speech signal to compare these 3 situations, and to extract the MSE and the distortion.

II. METHODS

We conducted an experiment to test this hypothesis about

¹Svetlin Antonov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: santonov@yahoo.com
²Snejana Pleshkova-Bekiariska is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: snejpl@tu-sofia.bg.
IV. DISCUSSION

Because the FM cue is derived from phase, the present study argues strongly for the importance of phase information in realistic listening situations. We note that for at least two decades phase has been suggested to play a critical role in human perception [17], yet it has received little attention in the auditory field.

The most direct and immediate implication is to improve signal processing in auditory prostheses. Currently, cochlear implants typically have 12–22 physical electrodes, but a much smaller number of functional channels as measured by speech performance in a quiet environment [18]. The results of our research strongly suggest that frequency modulation in addition to amplitude modulation should be extracted and encoded to improve cochlear implant performance. Recent perceptual tests have shown that cochlear implant subjects are capable of detecting these slowly varying frequency modulations by electric stimulation [19].

ACKNOWLEDGEMENT

Research in the subject has recently commenced through a NIS TU - Sofia funded project, № 121 ПИ 0063-07.

REFERENCES