

2016 PROGRAM GUIDE

A photograph of the Chicago skyline at dusk, featuring the Willis Tower and other skyscrapers. The foreground shows a beach with waves crashing onto the shore, and the sky is a mix of blue and orange.

COSM COMBINED
OTOLARYNGOLOGY
SPRING MEETINGS

HYATT REGENCY CHICAGO
CHICAGO, ILLINOIS

MAY
18-22
2016

American Academy of Facial Plastic and Reconstructive Surgery
American Broncho-Esophagological Association
American Laryngological Association
American Neurotology Society
American Otological Society
American Rhinologic Society
American Society of Pediatric Otolaryngology
The Triological Society

Acoustic analysis and computerized reconstruction of speech in Laryngectomised Individuals

Dr J Allen, FRACS ORL HNS

Dr H Sharifzadeh, PhD

Professor I McLoughlin,

Professor H Sarrafzadeh,

Dr I Ardekani, PhD

Introduction

In laryngectomised individuals, rehabilitation options include esophageal speech, tracheoesophageal puncture (TEP), and electrolarynx devices enabling patients to communicate. These do not generate natural sounding speech; at best they are monotonous or robotised. We present a non-surgical, non-invasive alternative which combines whisper signal analysis with direct pitch insertion and formant enhancement, to 'reconstruct' missing speech. This pilot study assesses the acoustic features of laryngectomees speech and required enhancement for natural speech regeneration.

Method

Audio recordings were made of five laryngectomised patients in a sound proof room from three randomisations of a balanced word list. Formant contours of speech samples in /hVd/ structure were analysed through combined segmentation extraction methods. Formant characteristics were used to establish a vowel formant space for laryngectomy speech. Data was translated by spectral enhancement and pitch insertion algorithm into phonated words. A subjective listening test using mean opinion scores (MOS) was conducted for five normal hearing listeners proficient in English. Subjects were randomly presented with reconstructed words and each subject scored word quality over a five-point scale.

Results

All audio recordings were able to be analysed. Vowel formant information was adequately produced in all samples. Frequency of the first two formants were measured and were comparable to corresponding formant frequencies in phonated speech. Mean opinion scores demonstrated excellent recognition of words and vowels. Reconstructed words were recognisable to naive listeners.

Conclusions

Computational speech regeneration demonstrates promising results in reproducing acoustic features of laryngectomee speech. This approach generates a more natural voice than alternative methods, using a non-invasive, portable external device.