A comparative acoustic study on speech of glossectomy patients and normal subjects

Xinhui Zhou, Carol Y. Espy-Wilson
Department of Electrical Engineering, University of Maryland, College Park, USA

Maureen Stone
Departments of Neural and Pain Sciences and Orthodontics, University of Maryland Dental School, Baltimore, USA

INTERSPEECH 2011
Aug. 29th, 2011, Florence, Italy
Introduction

• Oral, head and neck cancer represents 3% of all cancers in the United States (about 36,000 cases in 2010), the 6th most common cancer worldwide. (American Cancer Society, 2010)

• Most of the cancer in oral cavity are related to the tongue.

• Increase in tongue cancer is 5-fold among young men and 6-fold among young women, based on tumor registry data of four European countries from 1960 to 1994 (Annertz et al., 2002)

• Tongue cancer patients often are treated by glossectomy, a surgery to remove the tumor (T1: <=2cm in the largest dimension, T2: 2-4 cm, T3: > 4 cm, T4: invades adjacent structures)
A glossectomy patient before surgery (T1 tumor)

Ref: http://www.entusa.com/oral_pictures_htm/tongue_cancer_2.htm
A glossectomy patient (T1) with primary closure
A glossectomy patient (T2) with a flap
Objectives

• To assess how the speech signal is changed as a result of the differences in speech production between normal subjects and pre- and post-glossectomy patients.

• To model the vocal tract acoustics of the glossectomy patients.

• To couple this study with the larger ongoing project -- aimed at understanding the tongue muscle mechanics and tongue motion patterns of glossectomy patients – to provide clinical guidance.
Database

- 18 normals, and 16 patients with T1 or T2 tumor on lateral part of the tongue.
Database

• Speech data
  • Vowel-consonant-vowel (VCV) words (e.g., “isi”, “ishi”…)
    • Vowels: /iy/, /ah/, /uw/, /eh/.
    • Consonants: /s/, /sh/, /g/, /l/.
    • Each word repeated at least 3 times.

• Cine-MRI image data for the tongue shape
Acoustic measurement of vowels (F1-F3 of /uw/, /eh/, /iy/, /ah/)

/ah/ in “asa”
Acoustic measurement of fricatives
(/s/ and /sh/)

• Center of gravity

\[ f_0 = \frac{\int_0^\infty f |S(f)|^2 \, df}{\int_0^\infty |S(f)|^2 \, df} \]

• Skewness

\[ \frac{\int_0^\infty (f - f_0)^3 |S(f)|^2 \, df}{(f - f_0)^2 \int_0^\infty |S(f)|^2 \, df}^{1.5} \]
F1 and F2 of vowels (/uw/, /eh/, /iy/, /ah/)

Female

Male

Normal
Patient (presurgery)
Patient (postsurgery)
(Peterson 1952)
(Hillenbrand 1995)
F1 and F3 of vowels (/uw/, /eh/, /iy/, /ah/)

(Female)

(Male)

[Diagram showing F1 and F3 frequencies for different vowels and speakers]
Center of gravity for fricatives 
/s/ and /sh/

/s/ (female)

/s/ (male)

/sh/ (female)

/sh/ (male)
Example spectra of /s/ and /sh/
Example spectrogram of /s/ and /sh/

Normal (SPH)

“A souk”

“A shell”
Example spectrogram of /s/ and /sh/  

Patient (WCS)  

“A souk”  

“A shell”
Midsagittal MR images of /s/ and /sh/<br><br>/s/ in “asouk”  /sh/ in “ashell”<br><br>Normal (SPH)<br><br>/s/ in “asouk”  /sh/ in “ashell”<br><br>Patient (WCS)
/s/ production (apical vs. laminal)

Apical

Upper apical

Apicolaminal

Laminal

(Dart 1991)
/s/ production statistics in our database

<table>
<thead>
<tr>
<th>Normal (18)</th>
<th>Apical+ Upper apical</th>
<th>Laminal+ Apicolaminal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>26</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patients (9)</th>
<th>Upper apical</th>
<th>Laminal+ Apicolaminal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6</td>
<td>11</td>
</tr>
</tbody>
</table>
Summary

• Compared four vowels (/iy/, /uw/, /eh/, and /ah/) and two fricatives (/s/ and /sh/) between normal subjects and partial glossectomy patients with T1 or T2 tumors.

• Average formants (F1- F3) for the four vowels between the normals and the patients are similar.

• Average center of gravity of fricatives for the patients is significantly lower than it is for the normals, probably due to the more posterior constrictions in patients and its resulting longer front cavity.
Thanks !