A Noninvasive, Low-cost Device to Study the Velopharyngeal Port During Speech and Some Preliminary Results

Xiaochuan Niu
Alexander B. Kain
Jan P.H. van Santen

Center for Spoken Language Understanding
OGI School of Science and Engineering
Oregon Health and Science University
Outline

- Introduction
- Device design
- Signal processing
- Experiment
- Conclusions
- Discussion
Introduction

- Speech production involves the coordination of articulators
- Appropriate opening of the velopharyngeal (VP) port
  - Normal: nasal consonants, nasal vowels, nasalized phonemes
  - Disordered: hypernasality, nasal emission
Introduction

- Motivation:
  - To monitor the status of the VP port during speech
- Difficulty:
  - The position of the VP port is “hidden”
- Aspects of consideration:
  - Invasiveness / discomfort / health risk / ease of use / expense / inherent limitations / available analysis methods
Introduction

- Review of measurements:
  - Electromyography (EMG): to measure muscle activities
  - Imaging techniques:
    - Radiography / MRI / Ultrasound (relatively static images)
    - Fiberoptic endoscopes
  - Tracking techniques:
    - Aperture tracking: Photodetection
    - Point tracking: Radiography (pellets) / Magnetometry (coils)
- Airflow:
  - Pneumotachograph mask
Outline

- Introduction
- Device design
- Signal processing
- Experiment
- Conclusions
- Discussion
Device design

- Measurement principles
  - Bernoulli's equation
  - $P_t$: pressure at the tube end
  - $P_s$: static pressure
  - $\rho$: air density
  - $V$: air velocity

\[
P_{t1} = P_s + \frac{1}{2} \rho V^2 \\
P_{t2} = P_s - \frac{1}{2} \rho V^2
\]
Device design

- Data acquisition of DC signals with a sound card:
  - Differential pressure sensor
  - Processing modules: power supply, offset and gain control, and frequency modulation (FM)
Device design
Device design

- Circuits
Outline

- Introduction
- Device design
- **Signal processing**
- Experiment
- Conclusions
- Discussion
Signal processing

- FM signal:
  \[ y(t) = K \cos(2 \pi f_c t + \Phi(t)) \]
  \[ \omega(t) = 2\pi f_c + \frac{d}{dt} \Phi(t) = 2\pi [f_c + f_d x(t)] \]

- Hilbert transform:
  \[ y^*(t) = K \sin(2 \pi f_c t + \Phi(t)) \]

- Analytical signal:
  \[ y_a(t) = y(t) + j y^*(t) = K \exp[j(2 \pi f_c t + \Phi(t))] \]
Signal processing

- Demodulation algorithm
  - Sampling & low-pass filtering (5kHz)
  - Discrete Hilbert transform to obtain the analytical signal
  - Get phase signal by canceling the carrier frequency
  - Compute the derivative of the phase signal
Signal processing

- Zero calibration
  - Zero airflow input -> Signal with the carrier frequency
  - The carrier frequency is tuned around 3kHz, but it can drift
  - A line-search algorithm is applied to a zero-input signal in order to find the optimal carrier frequency of each recording session
Outline

- Introduction
- Device design
- Signal processing
- **Experiment**
- Conclusions
- Discussion
Experiment

- **Purpose**
  - To examine whether the proposed device can provide useful information about the VP status during speech

- **Speech materials**
  - Words in carrier sentence “Say ___ please.”

<table>
<thead>
<tr>
<th></th>
<th>CVN</th>
<th>NVC</th>
<th>NVN</th>
<th>CVC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/d @ n/</td>
<td>/n @ d/</td>
<td>/n @ n/</td>
<td>/d @ d/</td>
</tr>
<tr>
<td></td>
<td>/s A n/</td>
<td>/n A s/</td>
<td>/n A n/</td>
<td>/s A s/</td>
</tr>
<tr>
<td></td>
<td>/t i: n/</td>
<td>/n i: t/</td>
<td>/n i: n/</td>
<td>/t i: d/</td>
</tr>
<tr>
<td></td>
<td>/z u n/</td>
<td>/u z/</td>
<td>/u n/</td>
<td>/z u z/</td>
</tr>
</tbody>
</table>
Experiment

- Acoustic, demodulated nasal airflow, and static airflow
Experiment

- Observations:
  - The demodulated airflow signal contain strong harmonic components during the sections of voiced speech sounds;
  - The filtered DC and low-frequency components represents the static airflow as it moves in and out of the nostril;
  - The static nasal airflow peaks occur ...
Experiment

- Phoneme influence:
Experiment

- Result summary
  - A quantitative measurement of the static nasal airflow
  - Non-speech information, such as inhalation and exhalation
  - VP information about the detailed time-course of nasal, nasalized vowel, and nasal emission events during normal speech
Outline

- Introduction
- Device design
- Signal processing
- Experiment
- Conclusions
- Discussion
Conclusions

- A small, low-cost differential pressure sensor to pick up the dynamic pressure of the airflow (the raw cost of the device is less than $100)
- Frequency modulated airflow signal recorded by a generic sound card
- Demodulation algorithm and filtering process extracting the static nasal airflow
Conclusions

- Non-invasiveness:
  - The usage of this device does neither interfere with the articulatory process during speech, nor does it cause degradation of the simultaneously recorded acoustic signal.
Discussions

- Placement sensitivity: need to design and implement appropriate mounting frame
- Measurement range: need to be more robust
- Alignment with the acoustic signal:
  - Flow velocity vs. Sound wave rate
- Q & A