ULTRAFEST III

The poor man’s 3D:
Reconstructing 3D tongue surfaces from multiple sagittal or coronal planes

Heather Flowers (1), Tim Bressmann (1), Brent Carmichael (2), Chiang-Le Heng (1)

(1) University of Toronto, Department of Speech-Language Pathology
(2) University of Toronto, Institute for Biomechanics and Biomedical Engineering

Introduction

• The problem: The tongue is a non-rigid 3-dimensional hydrostat – but B-mode ultrasound can only capture 2D planes
• Our goal is to develop a clinically feasible method of dynamic 3D ultrasound imaging

Previous research: 3D static imaging

• Stone & Lundberg (1996): reconstruction of 3D surfaces from multiple coronal and sagittal scans
• Watkin (1999): Use of a commercial 3D software for 3D reconstruction of static speech sounds
• Bressmann et al. (2005; in press): Reconstruction of 3D surfaces in normal speakers and partial glossectomies

‘4D’ ultrasound

• Problems:
  • Proprietary software
  • Focus on visualization, rather than quantitative 3D measurement
  • Copious post-processing
  • Costs out of range for most speech labs

Previous research: 3D dynamic imaging

• Wein et al. (1988; 1993): Pseudo-3D reconstructions of tongue movement from multiple parasagittal scans; evaluation of 3D tongue ‘movement’ in a patient with apraxia of speech
• Yang & Stone (2002): Reconstruction of tongue movement from 8 para-sagittal and coronal scans; time-alignment based on a Dynamic Programming algorithm

Goals for clinical research

Decrease:

• time for data acquisition (minimum number of repetitions per scan)
• invasiveness (accurate scan angle without forced head positioning)
The Present Study

- Exploratory data collection from normal speakers and a partial glossectomiee
- Static surface reconstruction
- Dynamic surface reconstruction
- Comparison of scanning positions and number of scans

Method

- Tools for Data Acquisition
  1. CHASE II
     Scan angles set to variable degrees with lever system
  2. ScreenBlast 3.0 /Sony Media Soundforge 7.0
     Align ultrasound frames according to the acoustic information
  3. UltraCATS
     Semi-automatic tracing and measurement of tongue contours

Screen Blast

Sound Forge

Ultra-CATS

Method

3D static reconstruction

- A sustained production of each of 10 sounds (acquisition time ca 5-7 sec)
  - Phonemes: /a/, /i/, /u/, /s/, /sh/, /n/, /l/, /r/, /ng/, /th /
  - CHASE lever is moved to preset positions throughout the production of a sustained sound
Data acquisition (Sustained speech sounds)

Results for Sustained /a/

Results for Sustained /i/

Results for Sustained ‘sh’

Results for Sustained ‘ng’

Method: Dynamic 3D surface reconstruction
- Three to five repetitions of each sentence at different scan angles using 3 sagittal (15 degrees apart) or 4 coronal planes (10 degrees apart)
- Speech tempo and sentence stress were paced with a digital metronome set at 120 beats per minute. The subject was instructed to speak at half tempo
- Visualization of the acoustic signal (oscillogram) in order to select tokens of comparable length
- Lever was repositioned at a new angle prior to each set of repetitions:
  i) golly, ii) It ran a lot, and iii) It rang a lot (Yang & Stone, 2002)
Data acquisition (Multiplanar paced 2D)

Results: 3D based on sagittal scans
- Plausible 3D surfaces
- Minor height differences in the posterior lateral tongue
- Asymmetry related to:
  - scan angles?
  - inter-token variability?

"golly" reconstructed from sagittal scans
"It ran a lot" reconstructed from sagittal scans

Inter-token variability (right p35)
Inter-token variability (right 0)
Inter-token variability (right a35)

Partial Glossectomee

Partial Glossectomee
Results: “It ran a lot”

Results: 3D based on coronal scans
- Symmetrical results with 7 data points extracted in four coronal planes
  - Right 15° to left 15°

3D surfaces
(Yang and Stone, 2002)

“It ran a lot” reconstructed from coronal scans
“It rang a lot” reconstructed from coronal scans

“It ran/ rang a lot”

Limitations
- Deviations in head position and the relative transducer position may skew results
- Could possibly be improved by:
  - taking additional repetitions per scan angle (especially for speech disorders)
  - averaging data from multiple repetitions

Conclusions
- Timing with a digital metronome is effective for those participants who can pace their speech
- Data acquisition does not take very long, nor does the data processing and analysis
- Spatial alignment during the acquisition process needs to be sorted out, especially for sagittal planes

Future Directions
- Compare different methods to validate data (e.g. 3D static from Freescan compared to 3D static reconstructions using CHASE II)
- Evaluate test-retest reliability
- Make additional dynamic 3D tongue surface reconstructions for normal speakers
- Analyze disordered speech (initially with partial glossectomy)

Questions?

Voice and Resonance Lab webpage:
http://www.slp.utoronto.ca/People/Labs/TimLab/index.htm

Contact: heather.flowers@utoronto.ca