



# Stabilization techniques for ultrasound imaging of speech articulations



Lisa Davidson and Paul De Decker, Department of Linguistics, New York University

## 1. Introduction

### Aim of ultrasound imaging in speech research

To use information about tongue shape and place and amount of constriction to inform both phonetic and phonological hypotheses.

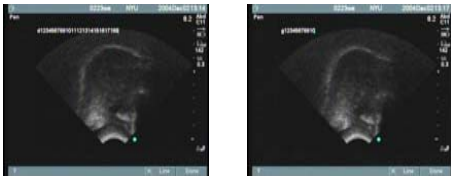
### A challenge for ultrasound in speech research

- To ensure that differences among images result from changes in tongue shape, not head or transducer movement (Stone and Davis, 1995; Gick, Bird and Wilson 2005).
- Ideally, to find an inexpensive, transportable stabilization system to carry out fieldwork (e.g. Miller-Ockhuizen, Namaseb, and Iskarous, 2004)

### Goal of this Study

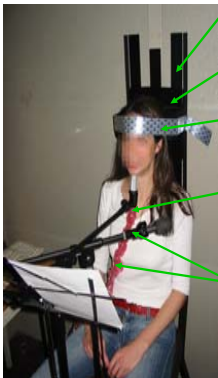
To validate the head stabilization system in the New York University Phonetics/Phonology Laboratory.

E.g., are these two images the same or different?



## 2. Stabilization Equipment

- Speakers are seated on a chair in a large soundproof booth.
- Moldable head stabilizer (intended for elderly patients with low head and neck tone) is attached to the wall with velcro.



- Head is further stabilized with a velcro head restraint.
- Transducer is held rigid with microphone stand.
- Synchronized ultrasound output and audio captured directly to computer with Canopus ADVC-1394 capture card and Adobe Premiere.



<http://www.comfortcompany.com>

## 3. Validation Procedure

### Methodology

- One measurement dot with a 6mm diameter was placed between the speaker's eyes, and another was placed on the front of the transducer. This primarily captures translation along the x (right-left) and y (up-down) axes.
- Participants were seated in the stabilization set up and a Panasonic Palmcorder IQ VHS-C video camera recorded the speaker's face and transducer. The video output was streamed directly to Adobe Premiere.



Placement of dots on Speaker 1

### Speakers 1 and 2

- Stimuli were divided into 5 blocks of 18 sentences. Each block was printed on one 8.5x11" sheet of paper. The sheets were placed on a music stand in front of the speaker at eye level. Speakers read at their own pace (approximately 1 minute per page). Once the speaker finished reading one sheet, the experimenter changed the page for the speaker.

### Speakers 3 and 4

- Speakers read the blocks of sentences off a laptop computer set up at approximately eye level. Each sentence appeared for 3 seconds, and then the following sentence appeared in exactly the same place on the screen.
  - This modification addresses whether reading from top to bottom causes greater head movement than displaying sentences in the same location.

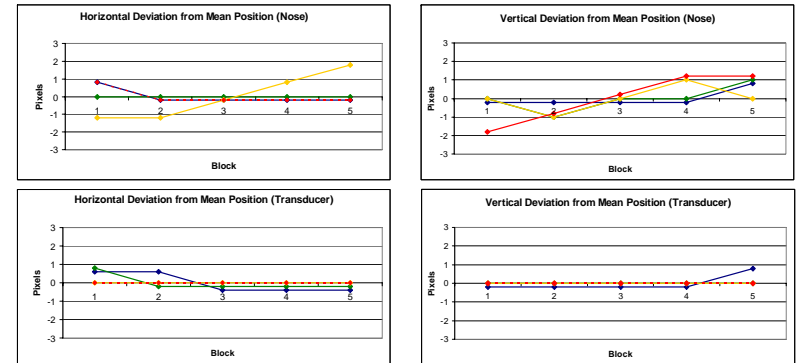
For each of the 5 blocks, every 10<sup>th</sup> video frame was selected for measurement.

### Movement-Tracking Procedure

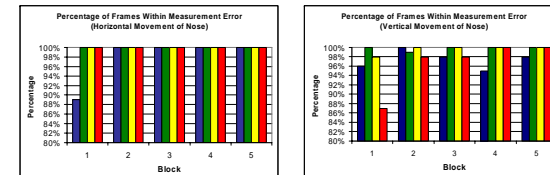
- A semi-automated technique implemented in Matlab tracked the movement of the dot.
- Contrast between dots and background was enhanced with histogram equalization for each color channel (red, green, and blue).
- For each dot, a region of interest (ROI) was defined such that the dot remained completely inside the region for all frames in the sequence.
- For each ROI, the intensities of the three color channels were weighted and summed to give a composite image. The weightings were determined manually from the first image in the sequence.
- To determine the center of dot, the composite image was convolved with an idealized circular dot of the predetermined diameter. The pixel in the convolution result with the maximum value was taken to be the center.

## 4. Results

### Deviation from Mean Position Across 5 Blocks



### Percentage of frames in which center of dot is within 1mm of the mean



### Measurement Information

- Ultrasound measurement error = 1mm
- Speaker 1: 1mm = 2 pixels
- Speakers 2, 3, 4: 1mm = 3 pixels

## 5. Discussion

- Overall, there is very little head movement beyond the measurement error of 1mm along either the vertical or horizontal dimension.
  - Maximum amount of movement of nose: 11%. Most movement occurs in Block 1, when speaker is still adjusting. This suggests that either the first block should be removed, or sufficient practice given.
  - There is more vertical movement than horizontal movement (possibly due to movement of forehead skin, not whole head).
- No transducer frames move beyond the measurement error.
- Reading down a sheet of paper does not cause greater movement than reading off of a laptop screen.

**Conclusion:** The moldable head stabilizer attached to a wall in combination with microphone stands is an effective, inexpensive method for head stabilization both in the lab and in the field.

Special thanks to David Goldberg (Weill Cornell Medical College) and Serafina Shishkova (NYU) for their help with this study.