5pSC20: EM sensor measurements of glottal structure versus time

1st Pan-American/Iberian Meeting on Acoustics
Dec. 1-6, 2002
Cancun, Mexico

John F. Holzrichter*, Lawrence C. Ng, and Gerald J. Burke
Lawrence Livermore National Laboratory, Livermore, CA
***Also UC Davis, Dept. of Applied Physics**
James B. Kobler and John J. Rosowski
Massachusetts Eye and Ear Infirmary, Boston, MA

Work supported by U.S. DOE, Darpa, and NSF
see web site at http://speech.llnl.gov

I. The Good News is:

- GHz Electro-Magnetic waves easily penetrate human tissue:

- They reflect from all of the Speech articulator tissues

- EM radar-like sensors are important because they...
  - provide real time information, unaffected by Acoustic noise,
  - use very low power EM waves, < 1mW ,
  - can be very low in cost, < $5 each, in quantity
  - are very small.

II. The Not so Good News is:

- EM waves reflect, refract, scatter, and are partly absorbed by all dielectric and conductivity interfaces in their path

- Interferometric sensor signals may be ambiguous regarding the origins of their reflected signals due to
  - Longitudinal location ambiguity, and
  - Product of (Target Area x Movement) can be ambiguous
  - Sensor movement relative to the targeted tissues

- Transverse resolution is typically 2-3 cm, but longitudinal resolution is a few microns

III. Good News: EM sensors robustly sense Voiced speech signals

- Direct measurements of vocal fold movements--with and without contact: < 1mm to 1 cm movements

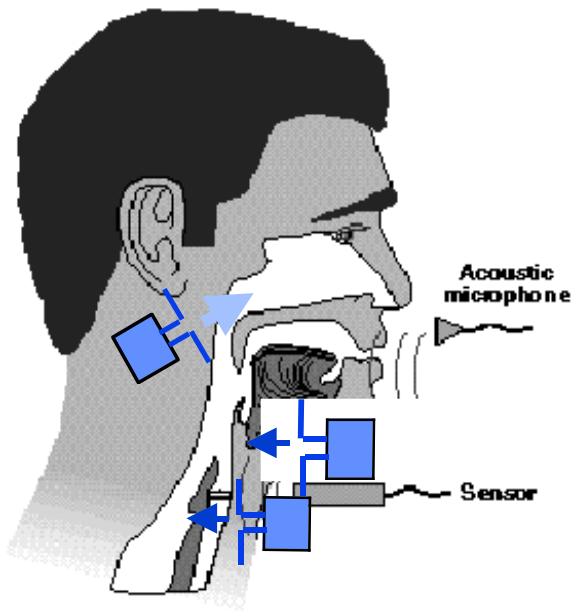
- Pressure induced trachea wall movements (primarily the anterior wall):
  10 -20 micron movements.

- Pressure induced vocal resonator wall and surface motions--e.g., cheek, tongue, lips, pharynx, sinus surfaces:
  e.g., 5-10 micron movements

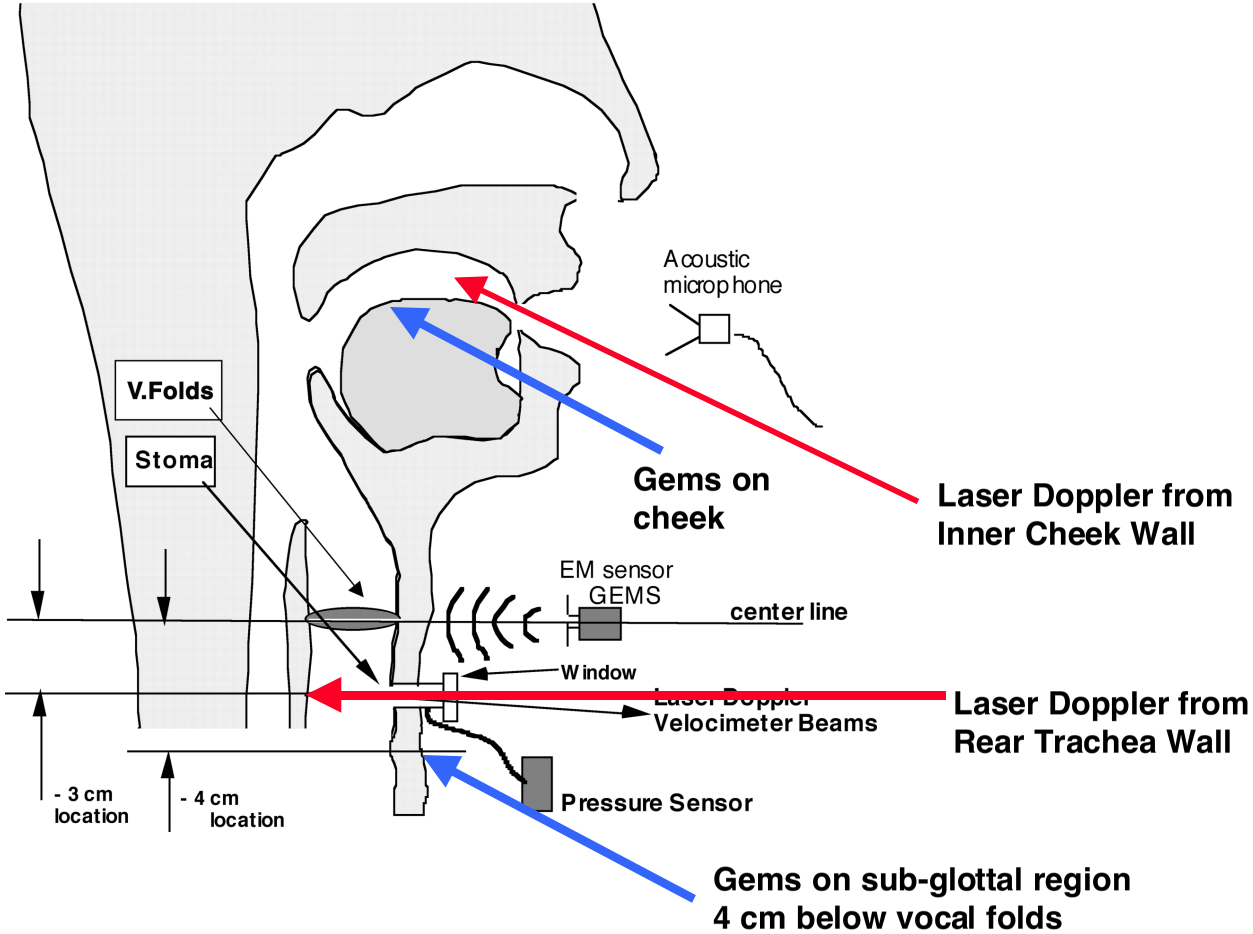
- (Macro-movements of jaw, tongue, lips, soft palate, etc. are being quantified)

What are the Sources of Radar-like Sensor Signals?

- **Experiment configurations-glottis, trachea walls, cheek wall**
- **Laryngeal Prominence Location Data**
- **Vocal Tract Wall Data**
- **Conclusions**

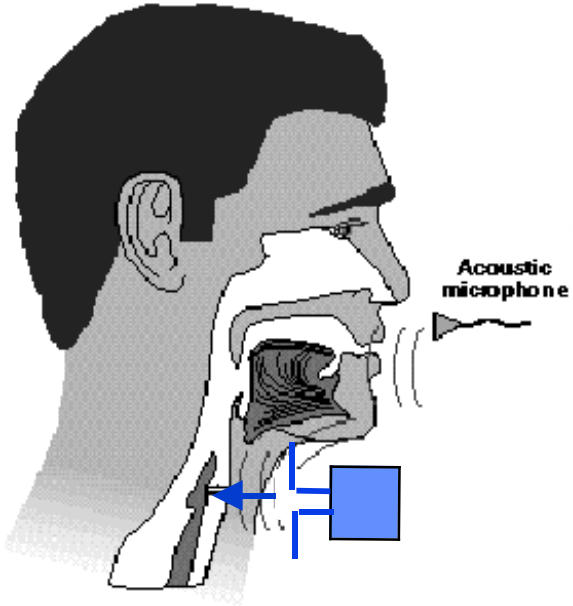


Laser doppler and EM sensor signals from a subject having a "stoma" in the neck are compared

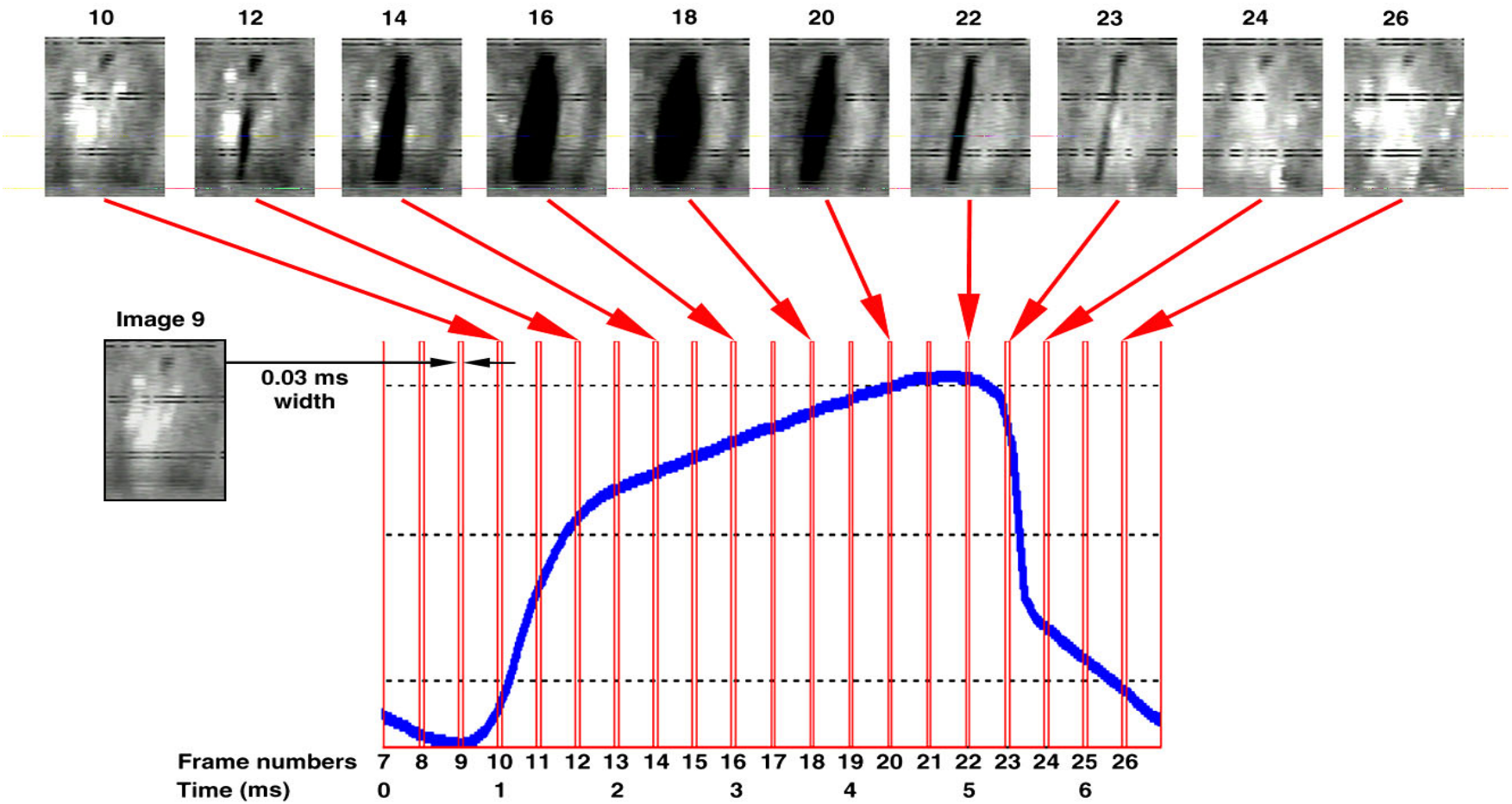


What are the Sources of Radar-like Sensor Signals?

- Experiment configurations
- Laryngeal Prominence Location Data
- Vocal Tract Wall Data
- Conclusions



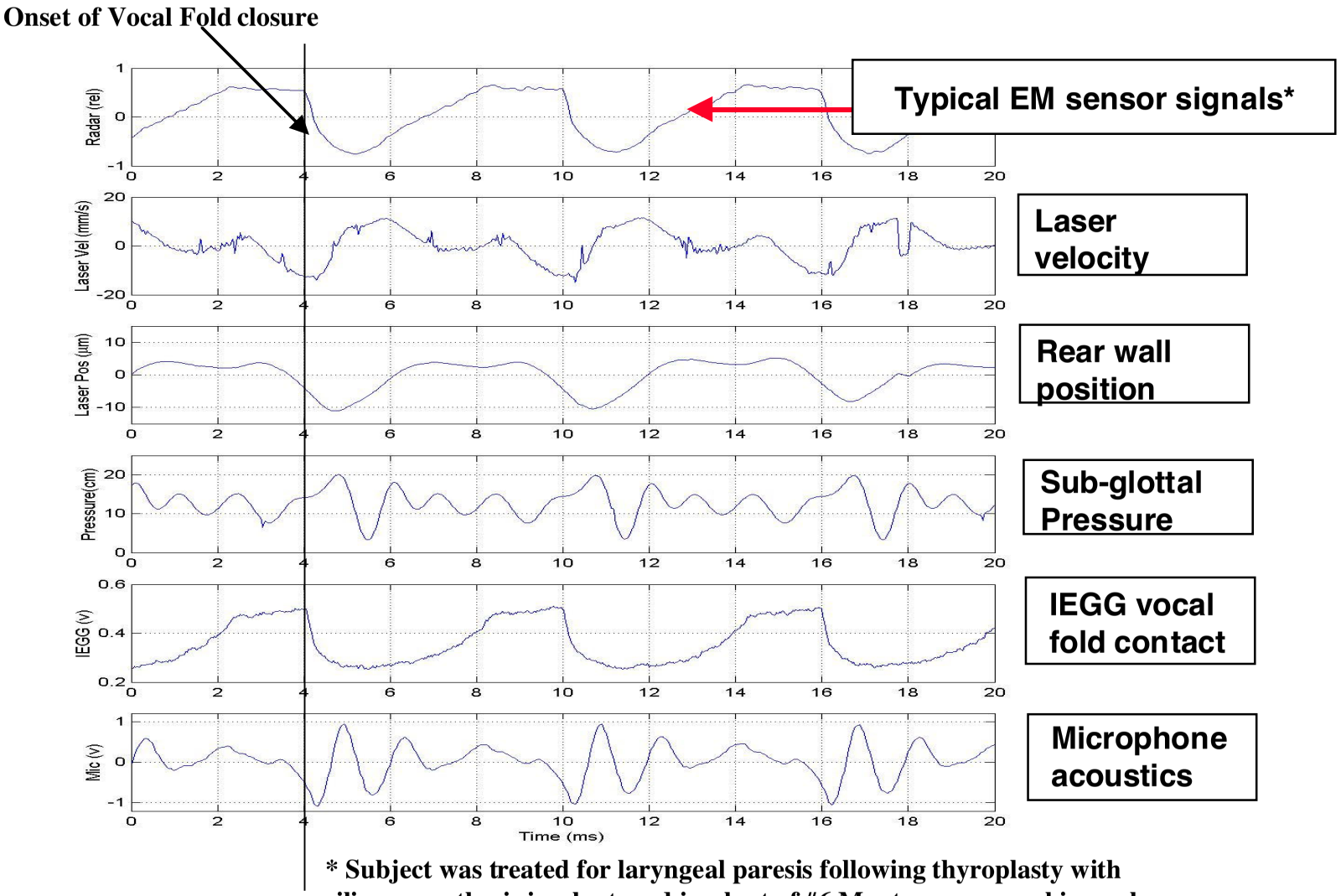
An EM sensor signal reflected from the glottal area, is correlated to an EM sensor signal using high speed video (3 per ms, 0.03 ms exposure)*



at-40443-u-jh-003

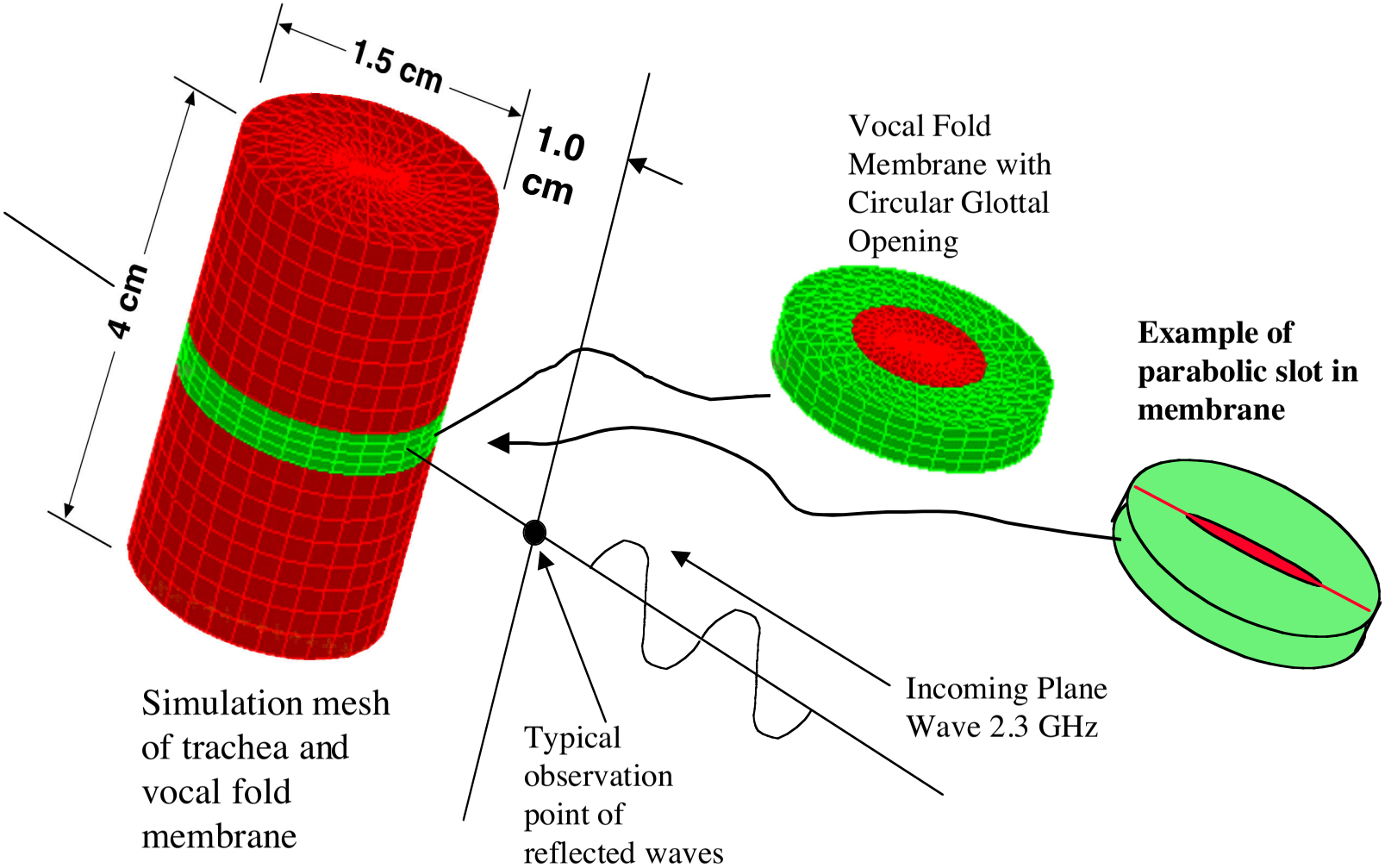
* G.C.Burnett "Thesis UC Davis"

The EM sensor laryngeal-prominence signal is from vocal fold opening and closing*, typ. 1 volt signals

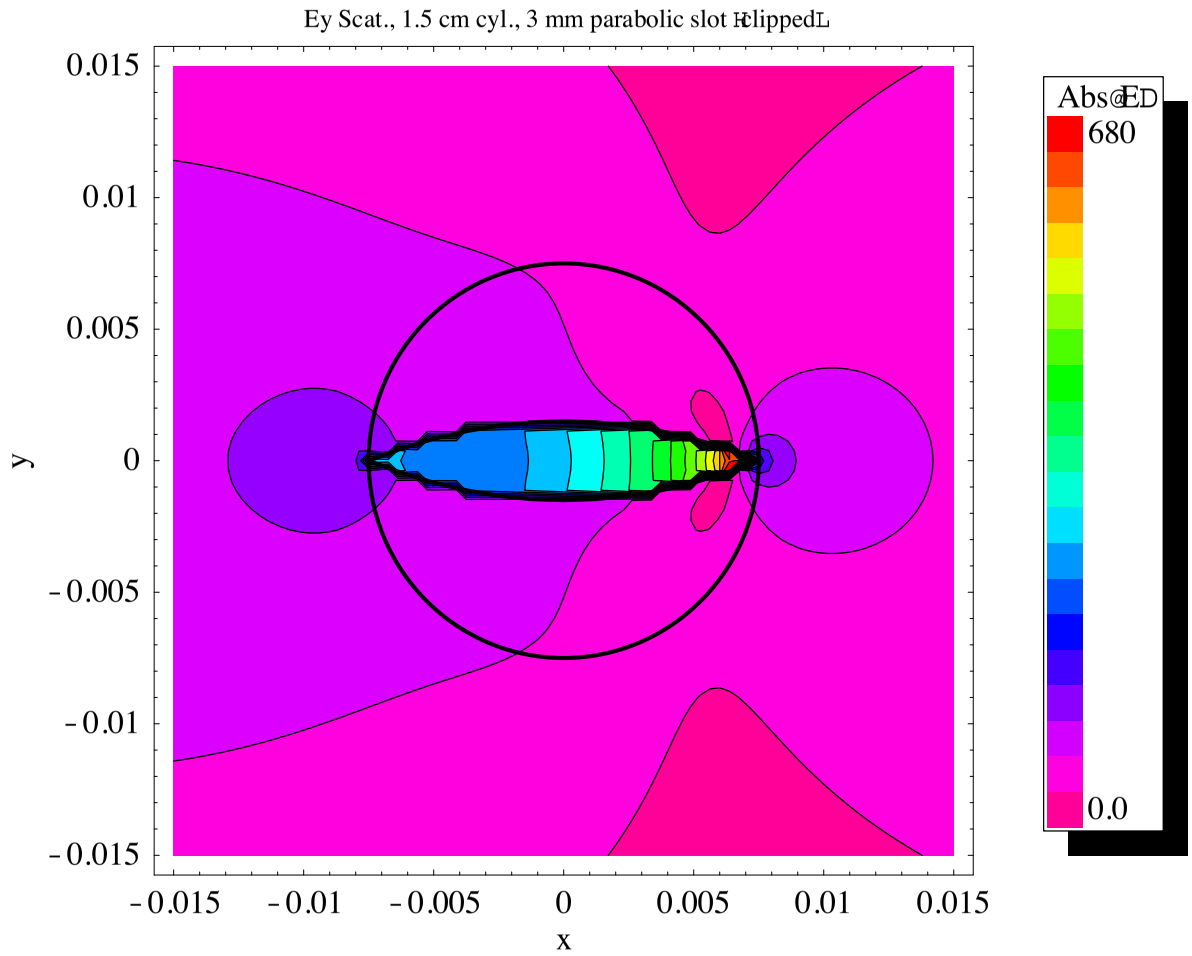


* Subject was treated for laryngeal paresis following thyroplasty with silicon prosthesis implant, and implant of #6 Montgomery speaking valve.

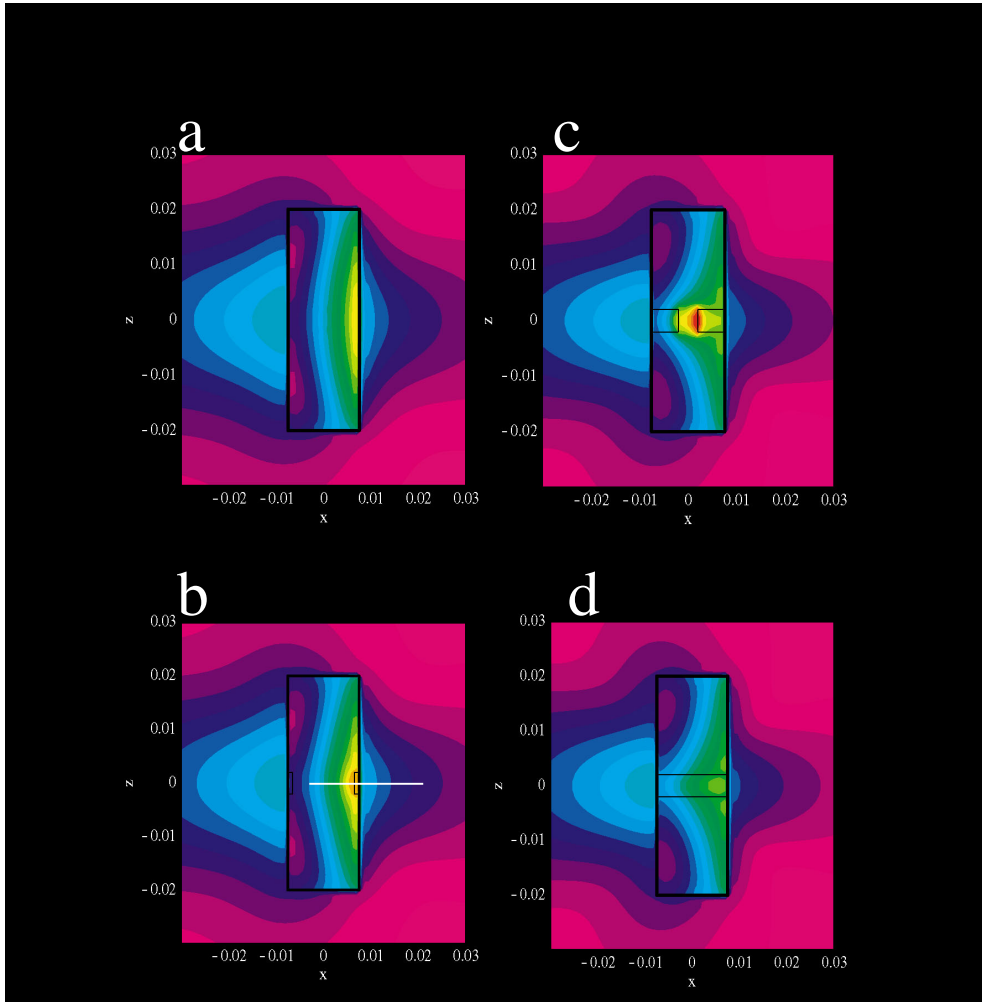
The neck model is an infinite dielectric (ε = 25), with an internal "soup can" air space divided by a 4 mm membrane



Extensive EM simulations were employed to obtain amplitude and phase of Reflecting EM waves



Simulated EM wave wave reflections from the glottal opening show origin of strong “glottal” signals



Sagittal slice through 1.5 cm dia air tube, with 4mm vocal fold membrane & circular hole.

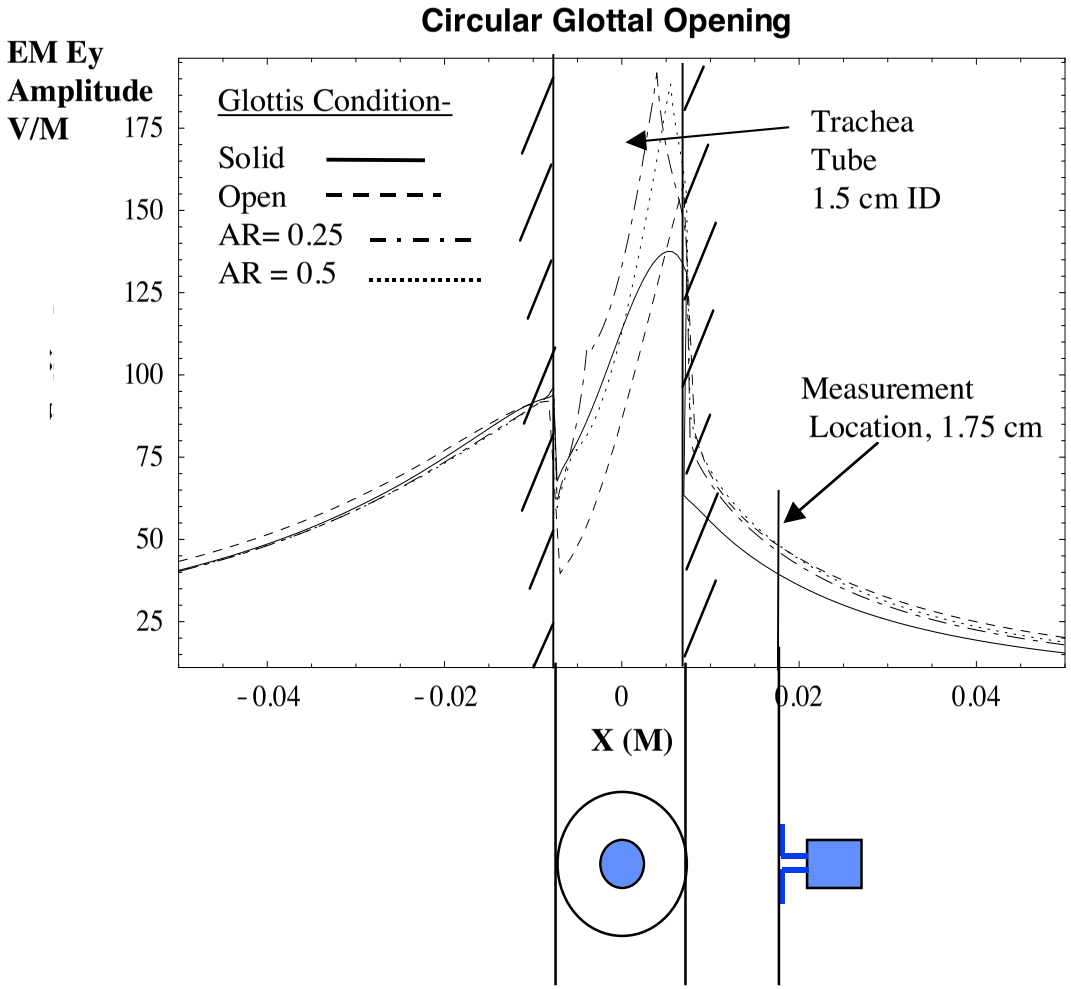
A) open tube, wave reflects from first surface

B) fold adducting, carried wave into tube

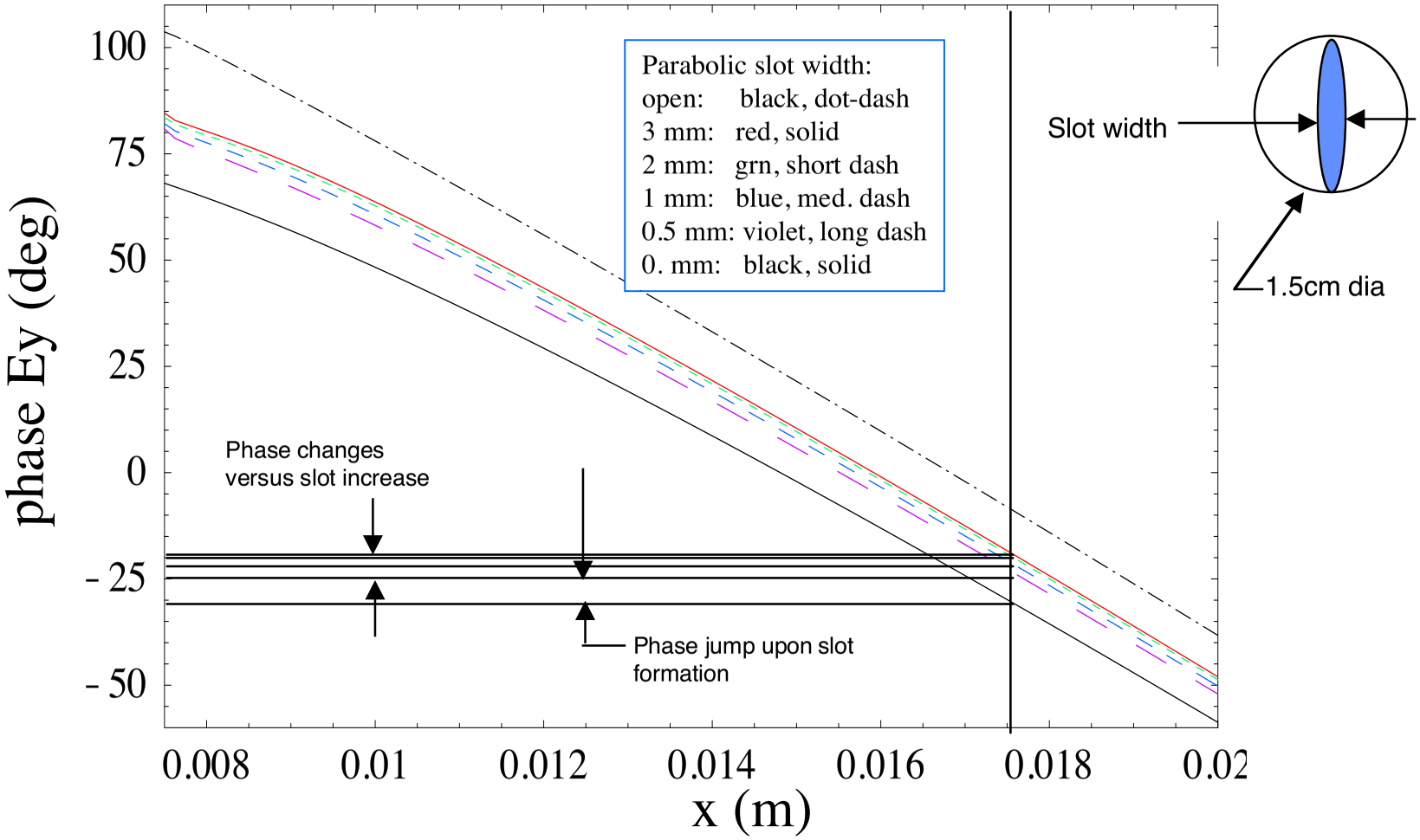
C) nominal opening in folds, showing reflection locations

D) closed glottis shows little reflection, wave passes through

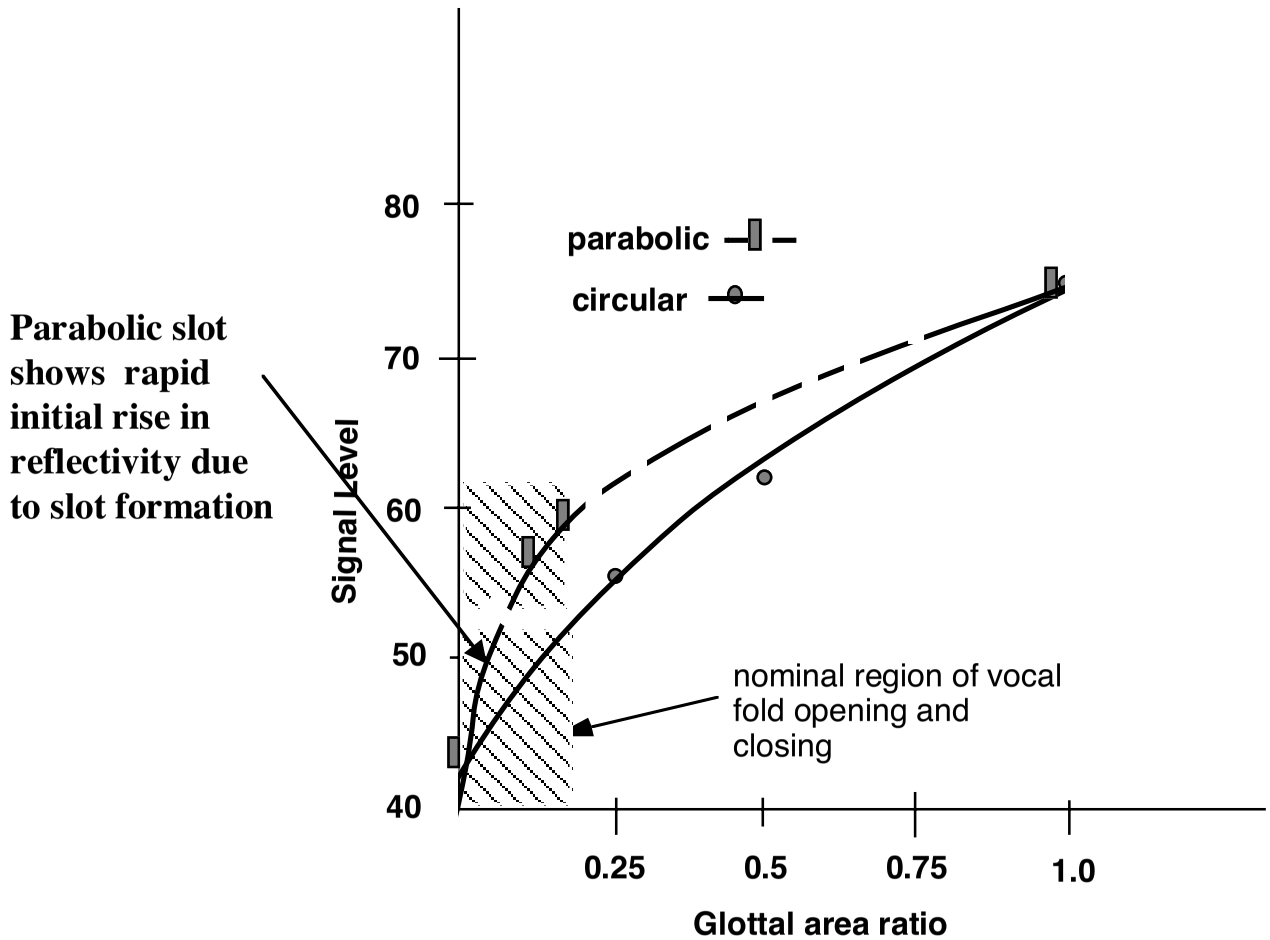
From the simulations we obtain reflected EM wave Amplitudes and Φs, for an open tube, a solid membrane, circular and slot openings in the membrane



Both phase and amplitude contribute to the EM sensor signal, A cos Φ, between configurations. A phase and amplitude jump occurs upon initial fold separation.

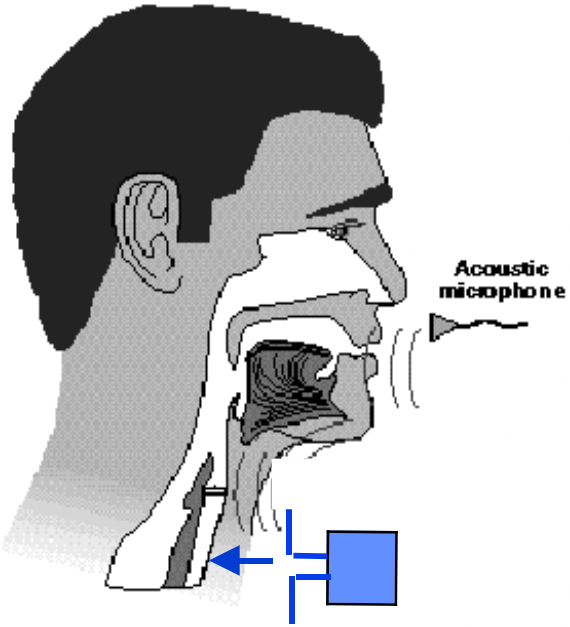


An air flow function can be generated by associating the EM sensor signal with glottal area (In Progress)

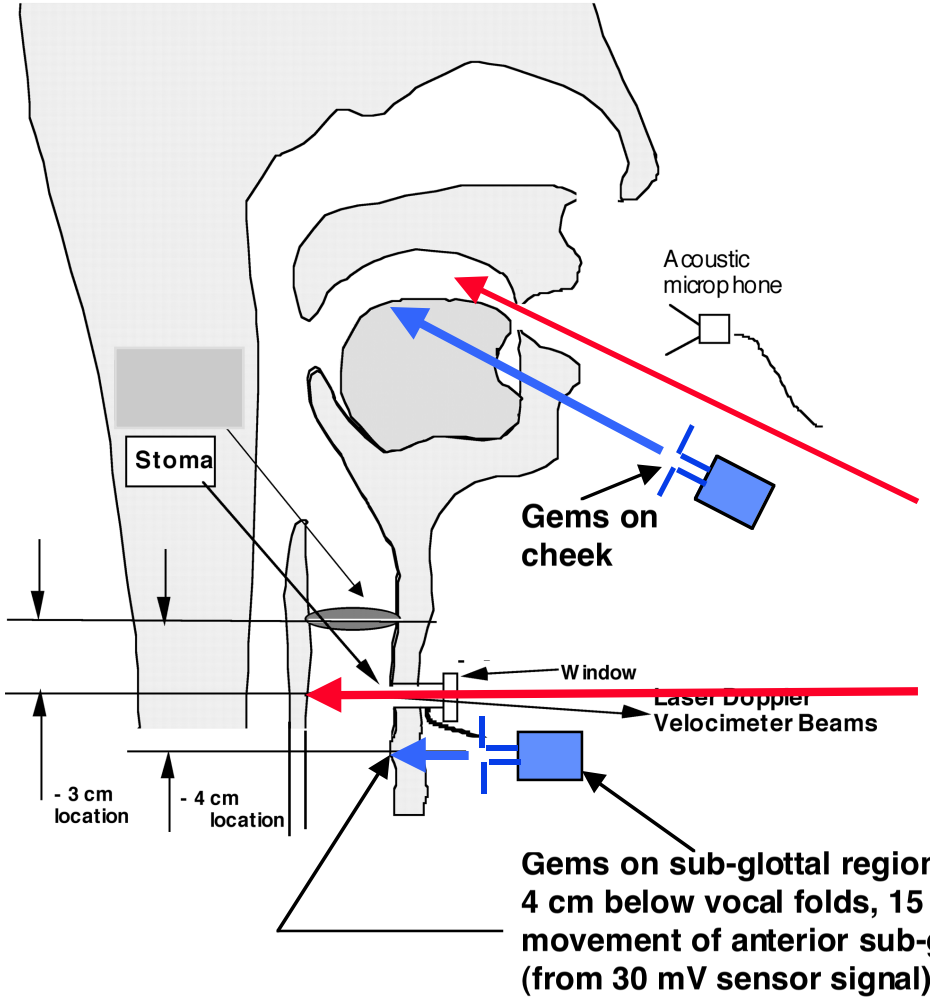


What are the Sources of Radar-like Sensor Signals?

- Experiment configurations
- Laryngeal Prominence Location Data
- **Vocal Tract Wall Data**
- Conclusions



Vocal tract wall experiments,e.g., trachea and cheek, show 5-10 micron motions versus 5-10 cm H_2O pressure cycles



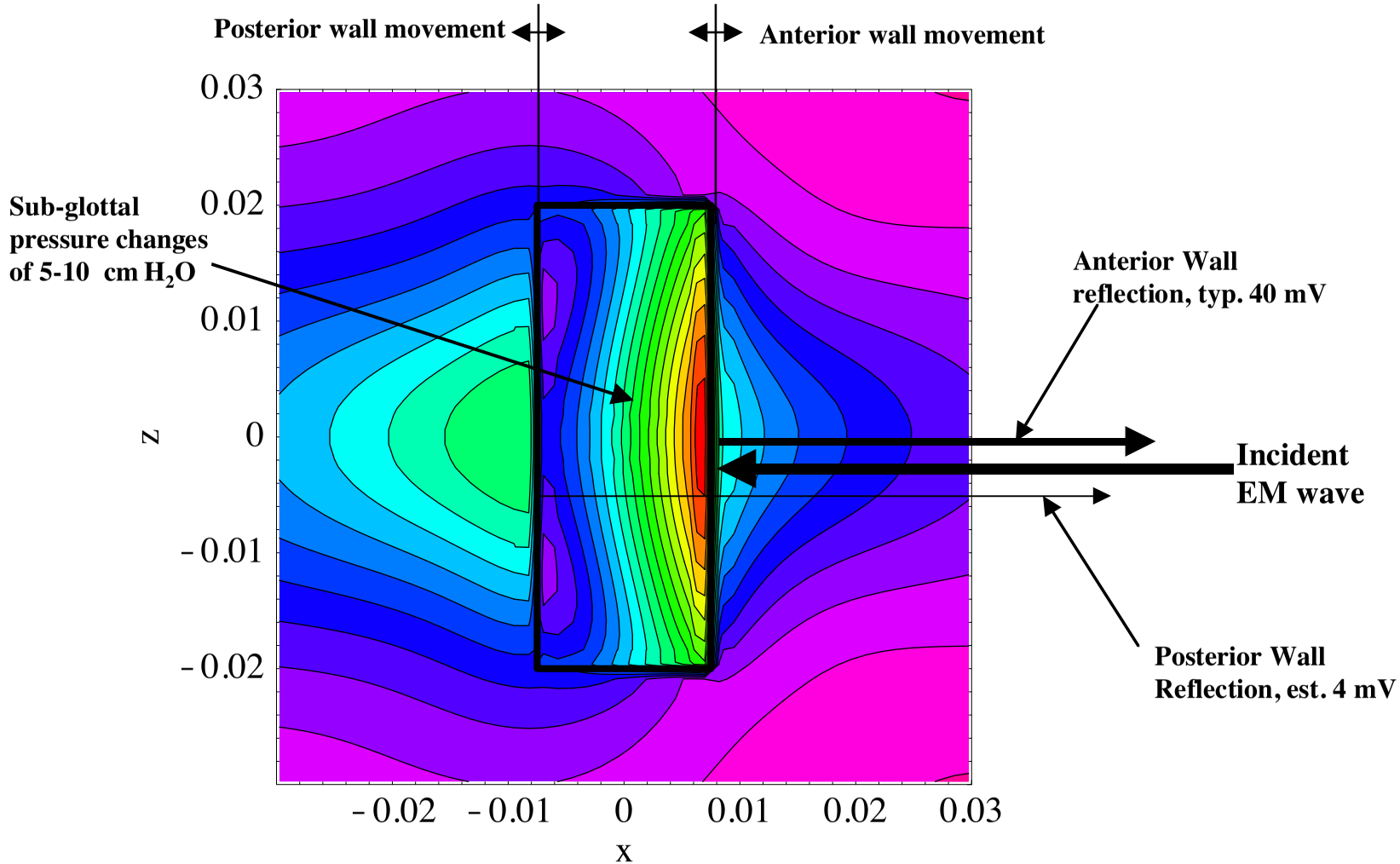
Employed an EM sensor, laser, pressure sensor, EGG, and microphone.

Laser Doppler from Inner Cheek Wall, 15 micron movement due to 5 cm H_2O Δ-pressure

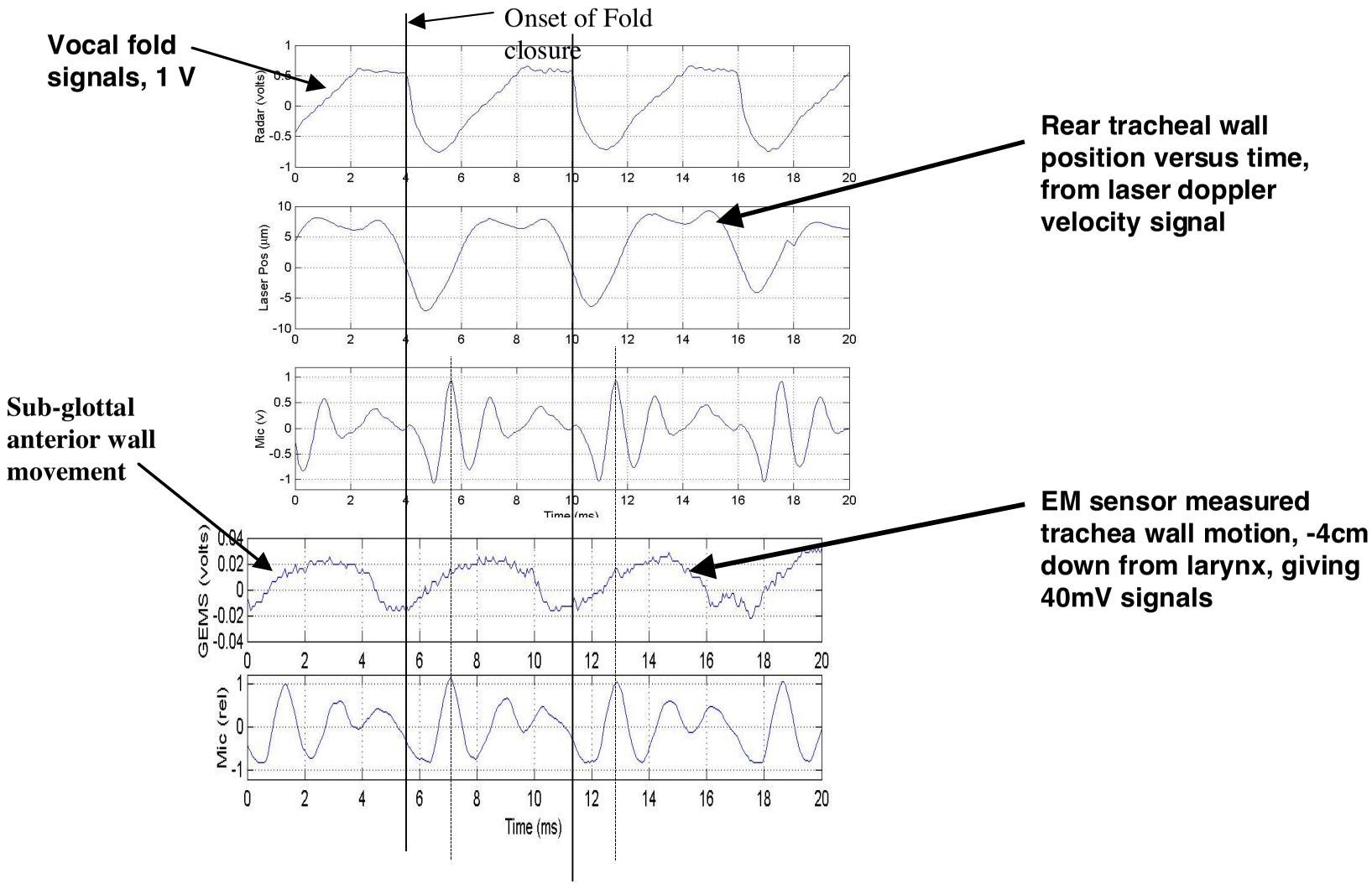
Laser Doppler from Rear Trachea Wall, 14 micron movement due to 7 cm H_2O Δ-pressure

Gems on sub-glottal region 4 cm below vocal folds, 15 microns movement of anterior sub-glottal wall (from 30 mV sensor signal)

EM sensor signal amplitude and shape can be estimated by simulating changes in position of posterior and anterior surfaces of the tracheal tube

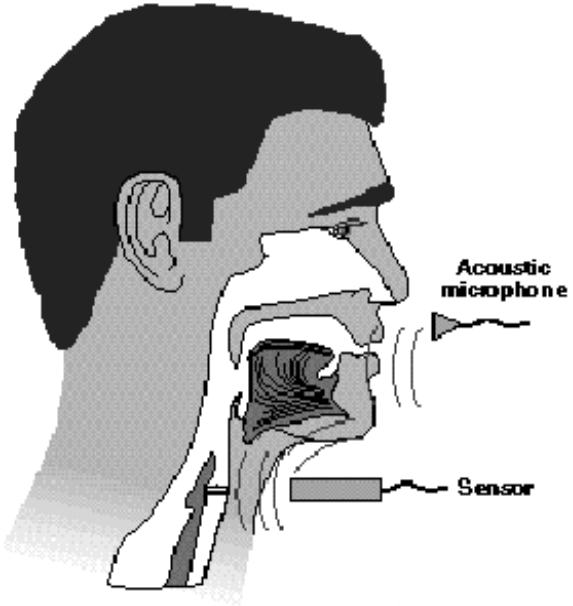


EM sensed subglottal trachea wall signals show anterior wall "ballooning" versus pressure



What are the Sources of Radar-like Sensor Signals?

- Experiment configurations
- Laryngeal Prominence Location Data
- Vocal Tract Wall Data
- Conclusions



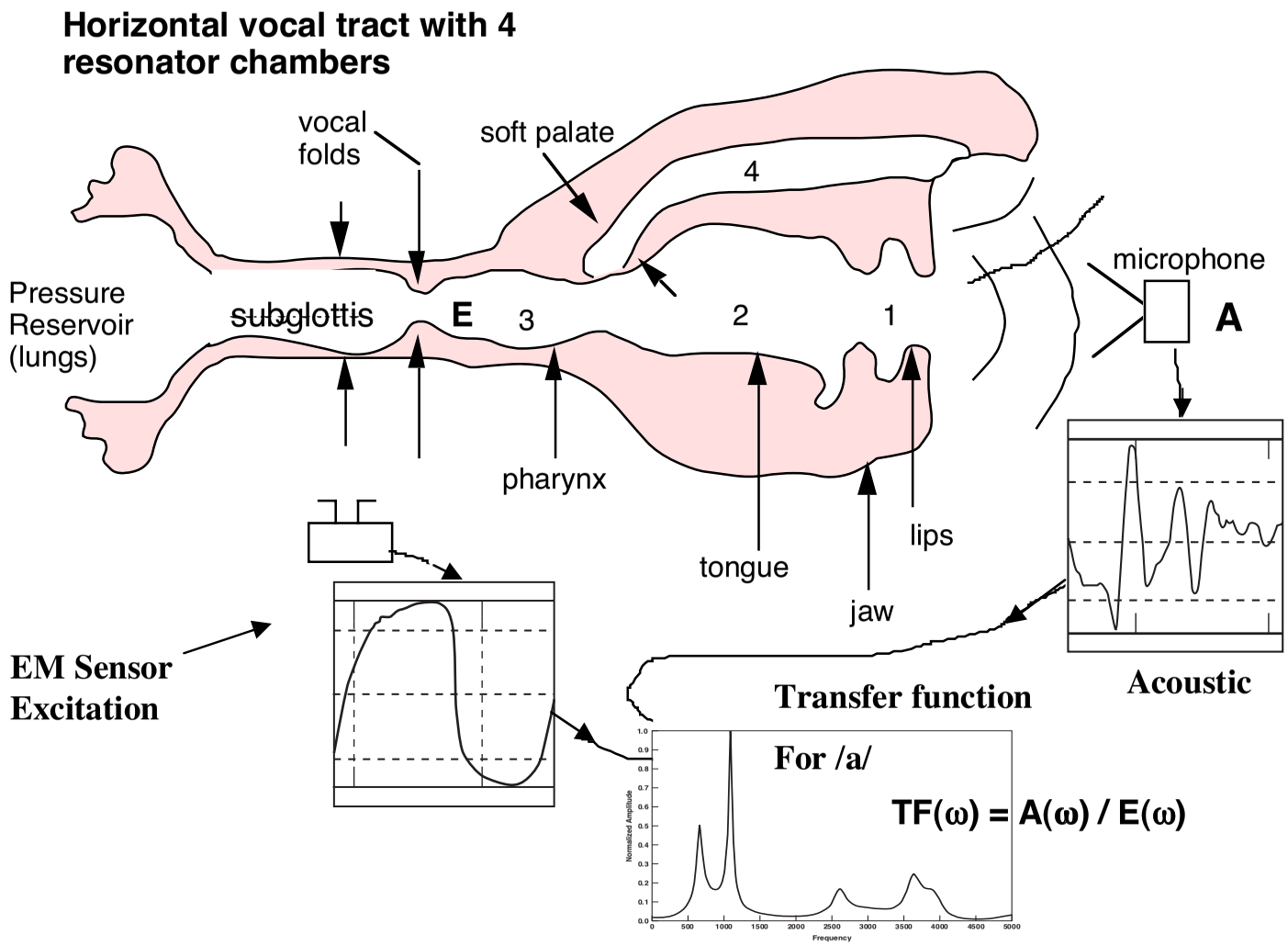
Conclusion: Low power, interferometric EM wave sensor signals are understood and very useful for real time speech processing

- We can measure Vocal-Tract tissue interface motions as small as $\Delta = \pm 1\ \mu m$ using < 0.5 mW interferometric sensors:

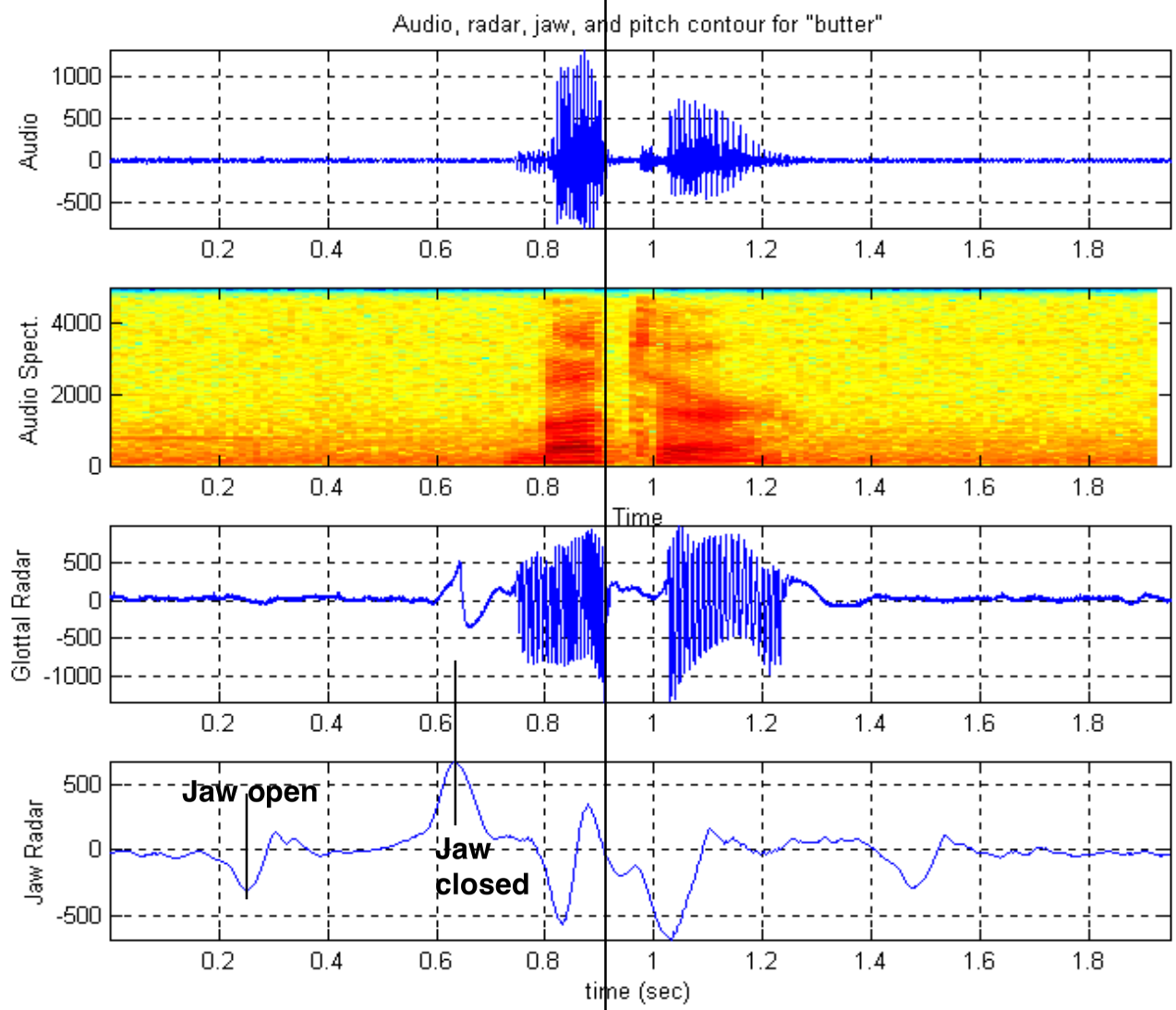
- Two types of Glottal related signals (100-200Hz) are obtained from the neck and head region:
  - Direct measurement of vocal fold cycle, $\Delta x = 0.1 - 1$ cm
  - Air pressure induced vocal tract wall movement, $\Delta x = 5\text{-}15\ \mu m$

- Low power EM sensor signal data are enabling tremendous improvements in human speech characterization for many applications
  - Low Bandwidth Vocoding, < 300 Hz
  - Denoising
  - Speaker Verification
  - Speech Recognition

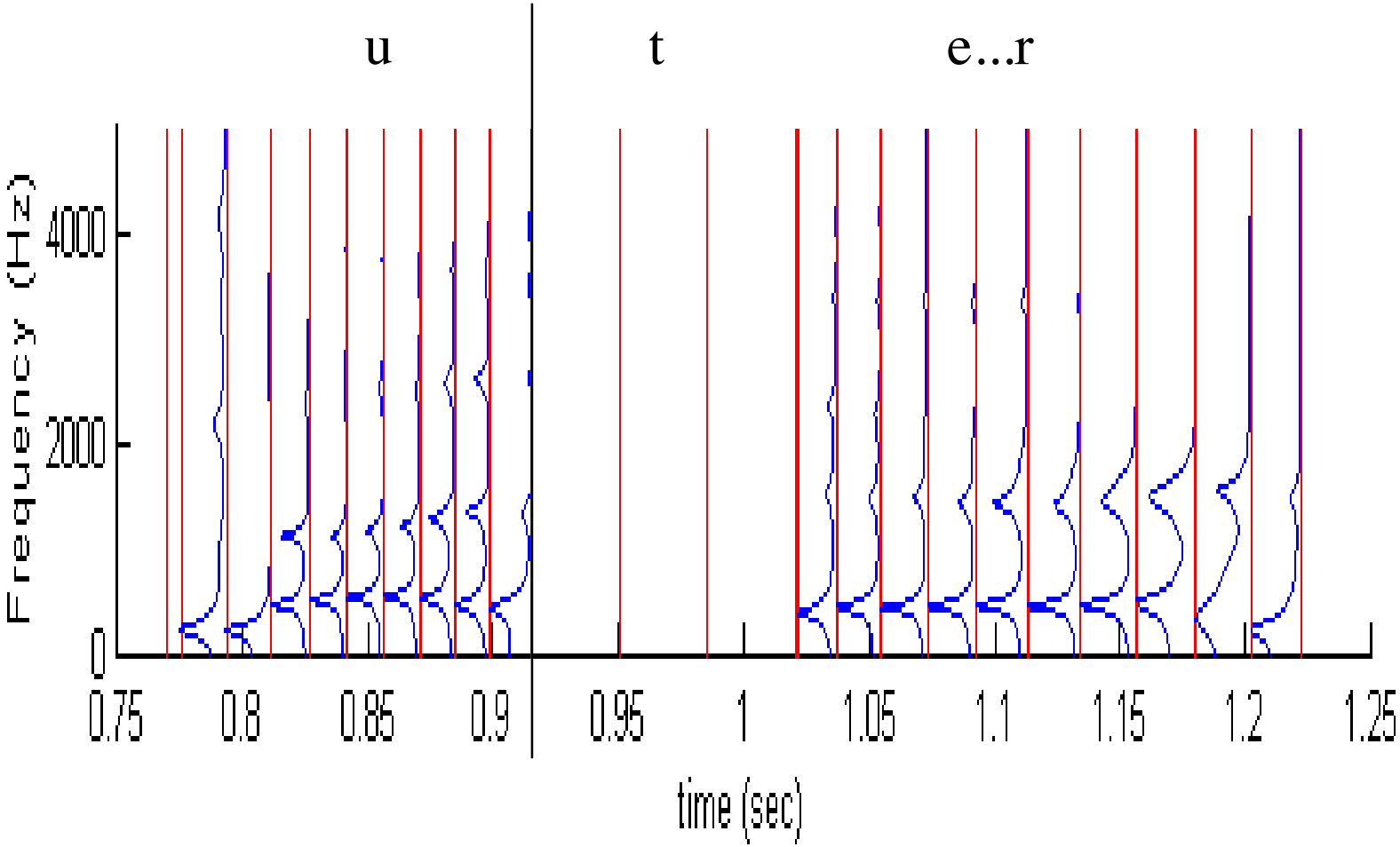
The human vocal tract can be well characterized if a sufficiently good excitation function can be obtained



2 EM sensor measurements for “butter”, glottis and jaw



Glottal-synchronous transfer functions for "butter" using EM sensor excitation & ARMA



Additional information is available at the following:

- See web sites for additional information:
  - http://speech.llnl.gov (many reports and papers)
  - http://www.aliph.com/main/sounddemos.htm
    (LLNL licensee: commercial applications and demonstrations)
- Recent working draft, available as preprint, describes details of EM wave interaction with glottal and tracheal structures (in review for publication) . Titled:

  "EM Wave Measurements of Glottal Structure Dynamics"
  Holzrichter, Ng, Burke,Champaign, Kallmann, Sharpe --LLNL
  Kobler, Rosowski, and Hillman -- Mass. Eye and Ear Infirmary
  available as Livermore Report number: UCRL -JC-14775 and
  above llnl.gov web site.

Major collaborators have been

- Prof. Robert Hillman - Harvard and Mass. Eye & Ear Infirmary
- Profs. Neville Luhman and Richard Freeman – UC Davis
- Drs. Greg Burnett and Todd Gable – Former students:
  Lawrence Livermore and UC Davis/ LLNL campus
- Prof. John Ohala – UCB
- Dr. Rebecca Leonard – UCD Hospital
- Prof. Ingo Titze – U of Iowa
- Dr. Wayne Lea – Speech Sciences Institute