Identifying Explainable Acoustic Features

Anna Barney and Atiyeh Alinaghi

University of Southampton

UNIVERSITY OF CAMBRIDGE

RELOAD

Aim

- Long term goal:
    To identify transparent features of speech sounds which can be used to predict Respiratory Tract Infection (RTI) progression

- Transparent = explainable in terms of the interaction between

    disease process, and

    acoustics of speech production.

- Critical for uptake of health applications that we are confident about the relationship between the classification or prediction and the disease both for safety and for equity.

Talk structure

- Production of voiced speech

- RTI symptoms

- RTI effect on speech sounds

- Candidate acoustic features

- Preliminary outcomes – sustained vowel classification

- Next steps

Voiced Speech Production

Voiced Speech Production

Nasal cavity
Hard palate
Alveolar ridge
Lips
Teeth
Esophagus
Trachea
Soft palate (velum)
Uvula
Tongue
Epiglottis
Vocal Cords
Glottis
Larynx

Epiglottis
Vocal cords
Vindpipe
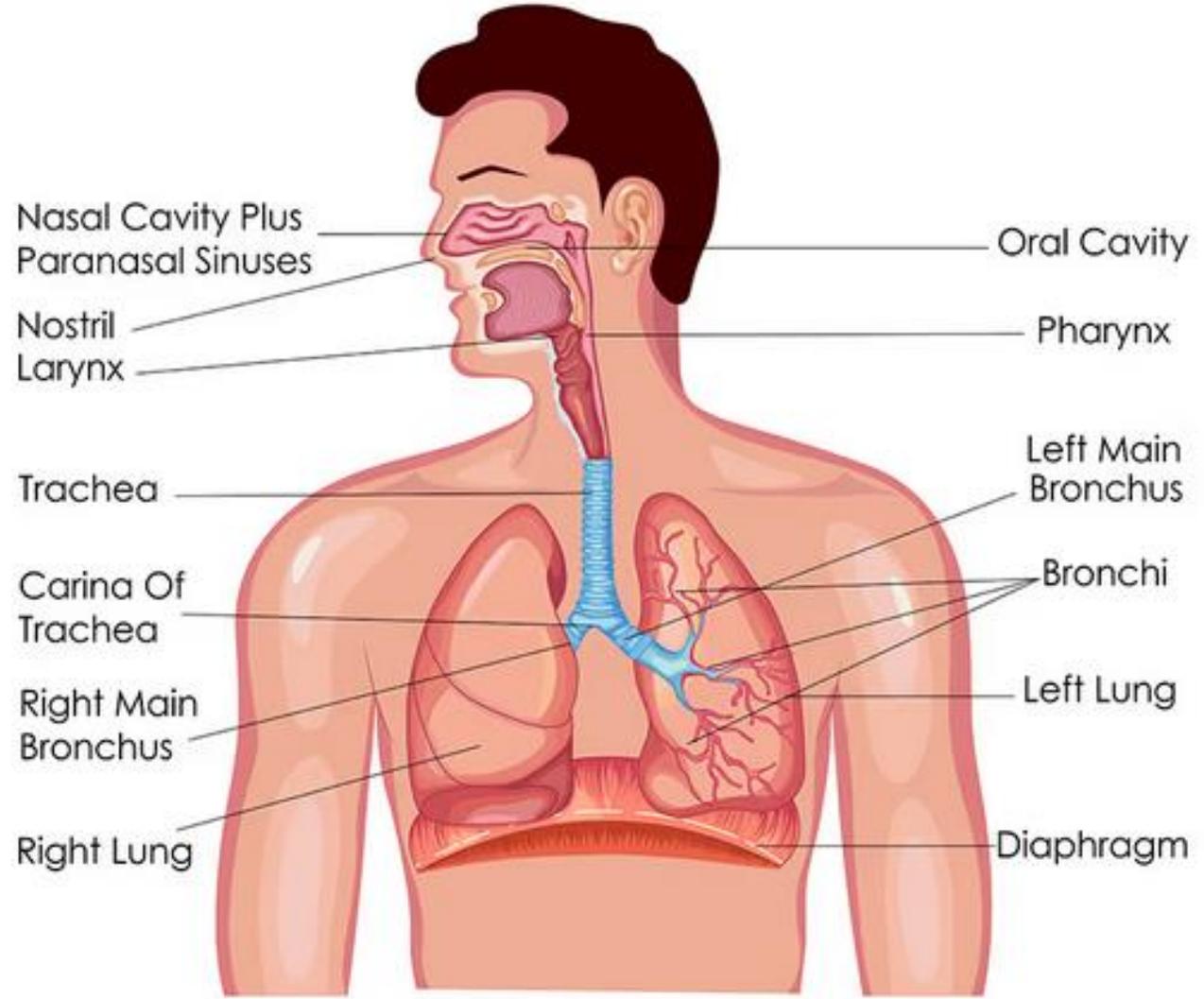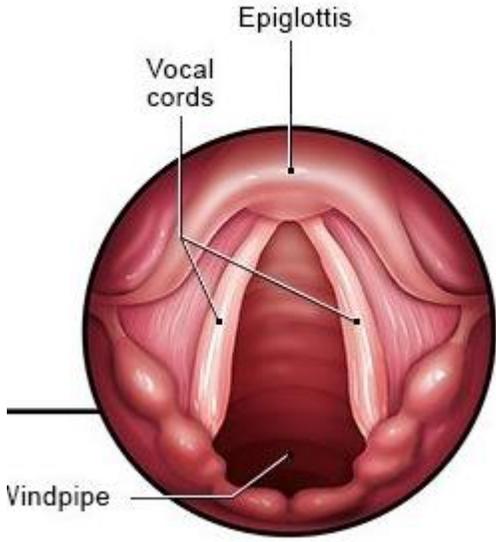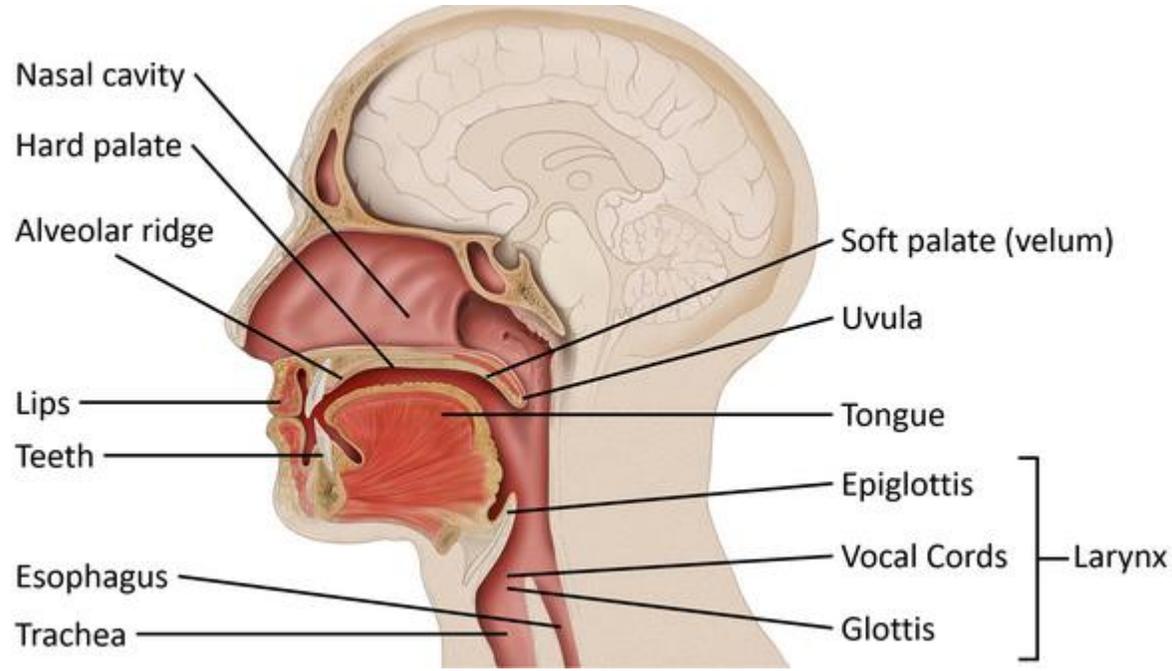
Speech Production

- Modal voice

Opening phase
Subglottal pressure forces folds apart

Closing phase
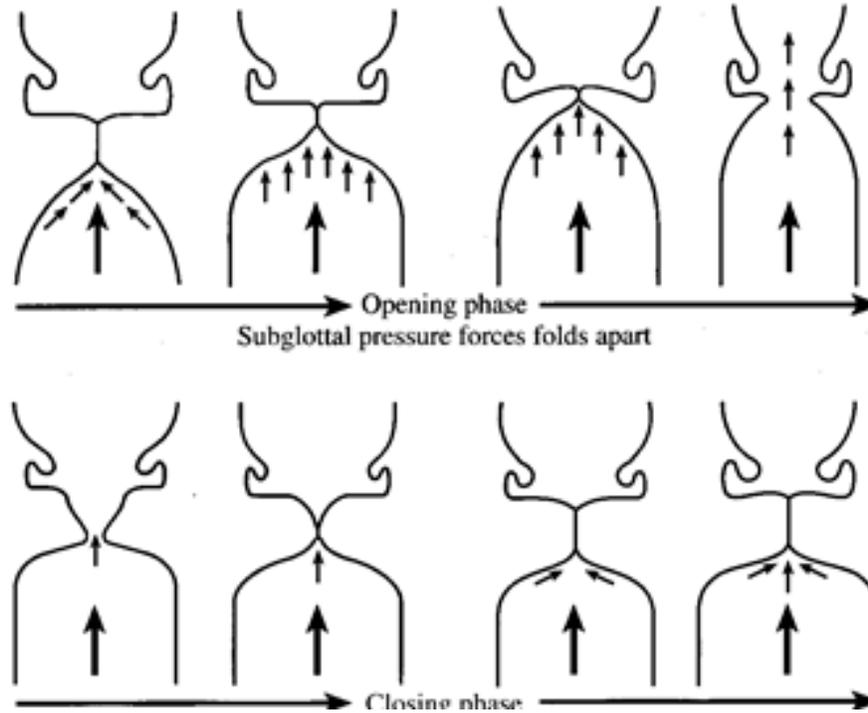
RTI symptoms

- Upper RTIs include: common colds, sinus infections, tonsillitis, laryngitis, flu

- Lower RTIs include: bronchitis, pneumonia, bronchiolitis, flu

Symptom	URTI	LRTI
cough	x	x
Mucous (phlegm)	x	x
Blocked or runny noes	x	x
Sore throat	x	x
Breathlessness or tight chest		x
High temperature	x	x
Hoarse voice	x	x
High respiratory rate		x
Feeling generally unwell	x	x

Symptoms and speech

Symptom	Affects	Perceived as	Measure
Breathlessness/tight chest	Vocal fold vibration	Difficulty starting or sustaining voicing	Number of voice breaks
Hoarse voice	Vocal fold vibration	'creaky' or 'breathy' voice, reduced f0 variance	HNR, H1:H2, jitter, shimmer, stdv(f0)
Sore throat	Upper vocal tract	Voice quality change	Increased spectral tilt, formant frequencies, formant bandwifths
Mucous viscosity	Upper vocal tract, vocal fold vibration	Voice quality change, difficulty starting or sustaining voice	Number of voice breaks, changes in spectral tilt, formant frequencies
Blocked nose	Nasalisation	Reduced b/m and d/n contrast	Spectral energy, formant frequencies,

Progress

- Identified 45 acoustic features which are good candidates for linking symptoms to speech changes.

- Extracted features from 408 examples of sustained vowel /a/ for RTI/~RTI classification using SVM.

- Accuracy 72%; True positive rate 79%, AUC 80%.

Confusion Matrix		Prediction	
		noRTI	hasRTI
Actual	noRTI	103	62
	hasRTI	50	193

Receiver Operating Characteristic Example

ROC curve (area = 0.80)
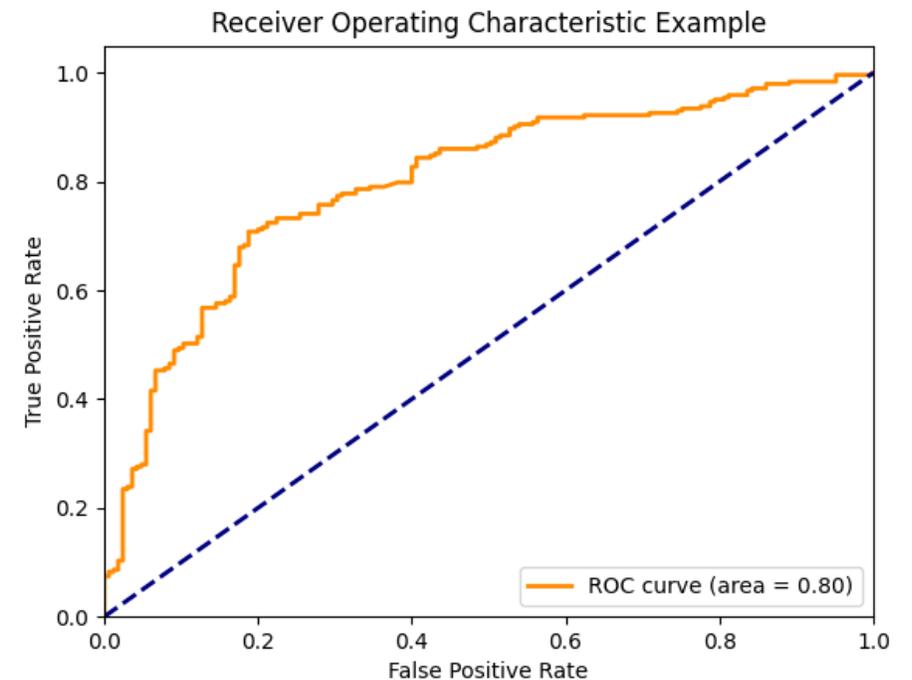
Next steps:

- Short term
  - Explore predictive properties of features
  - Application to running speech using a phonetically designed sentence.

- Longer term
  - Application to ambient speech
  - Effect of underlying chronic conditions (eg COPD, asthma)

Thank you for listening.
Questions and comments welcome.