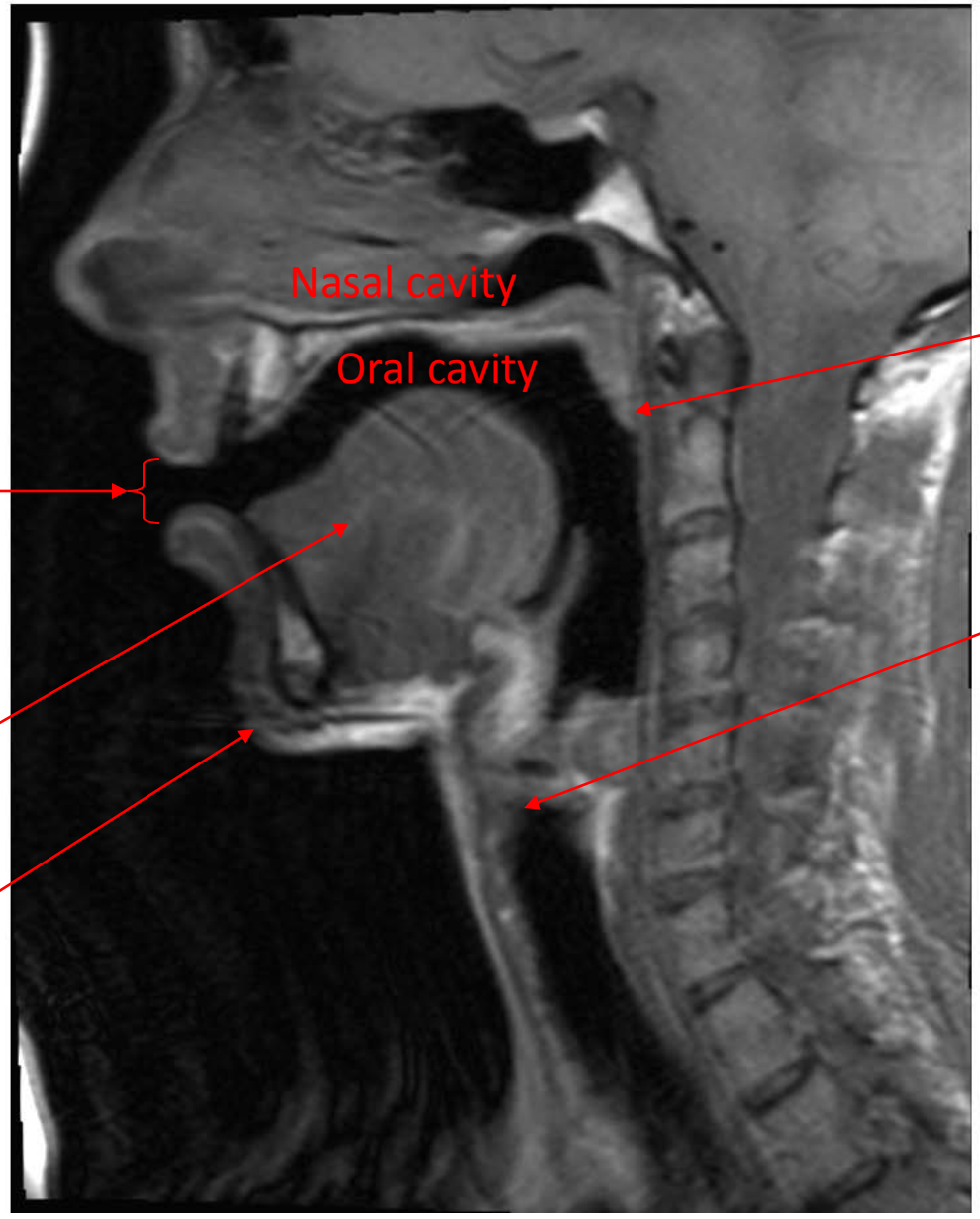
Not all Vocal Tracts are equal; but some are more equal than others

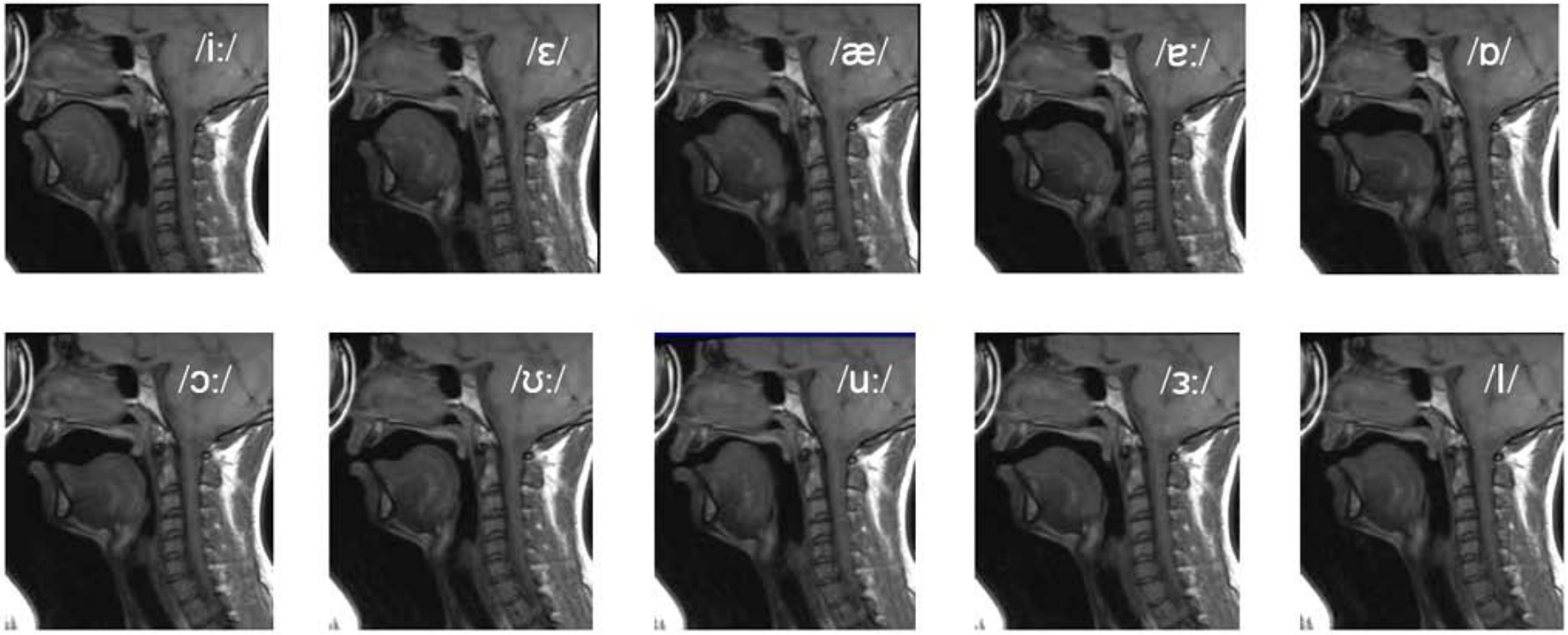
Catherine Watson

With thanks to Kalyan Chilukuri, and Helen Searle

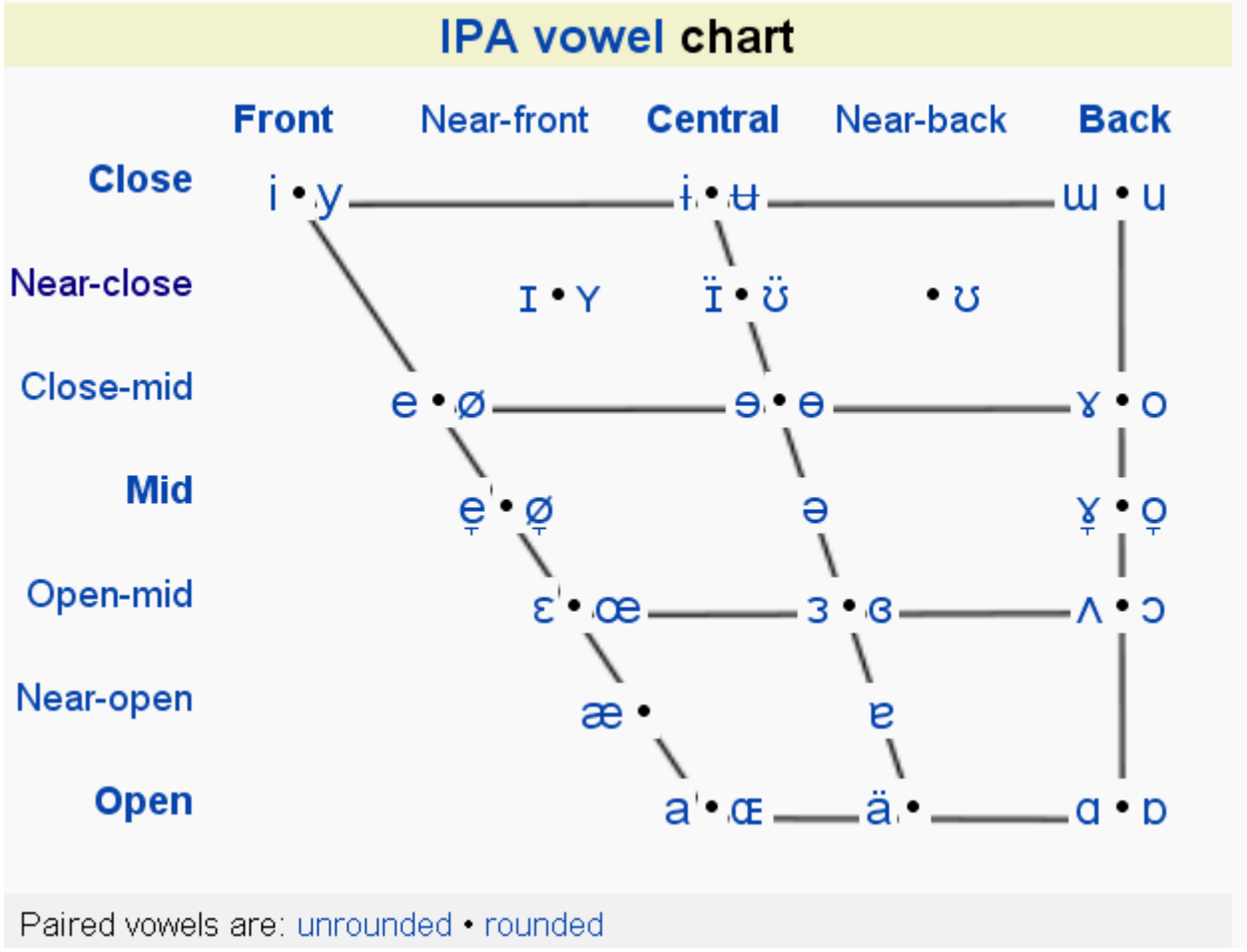
The Vocal Tract



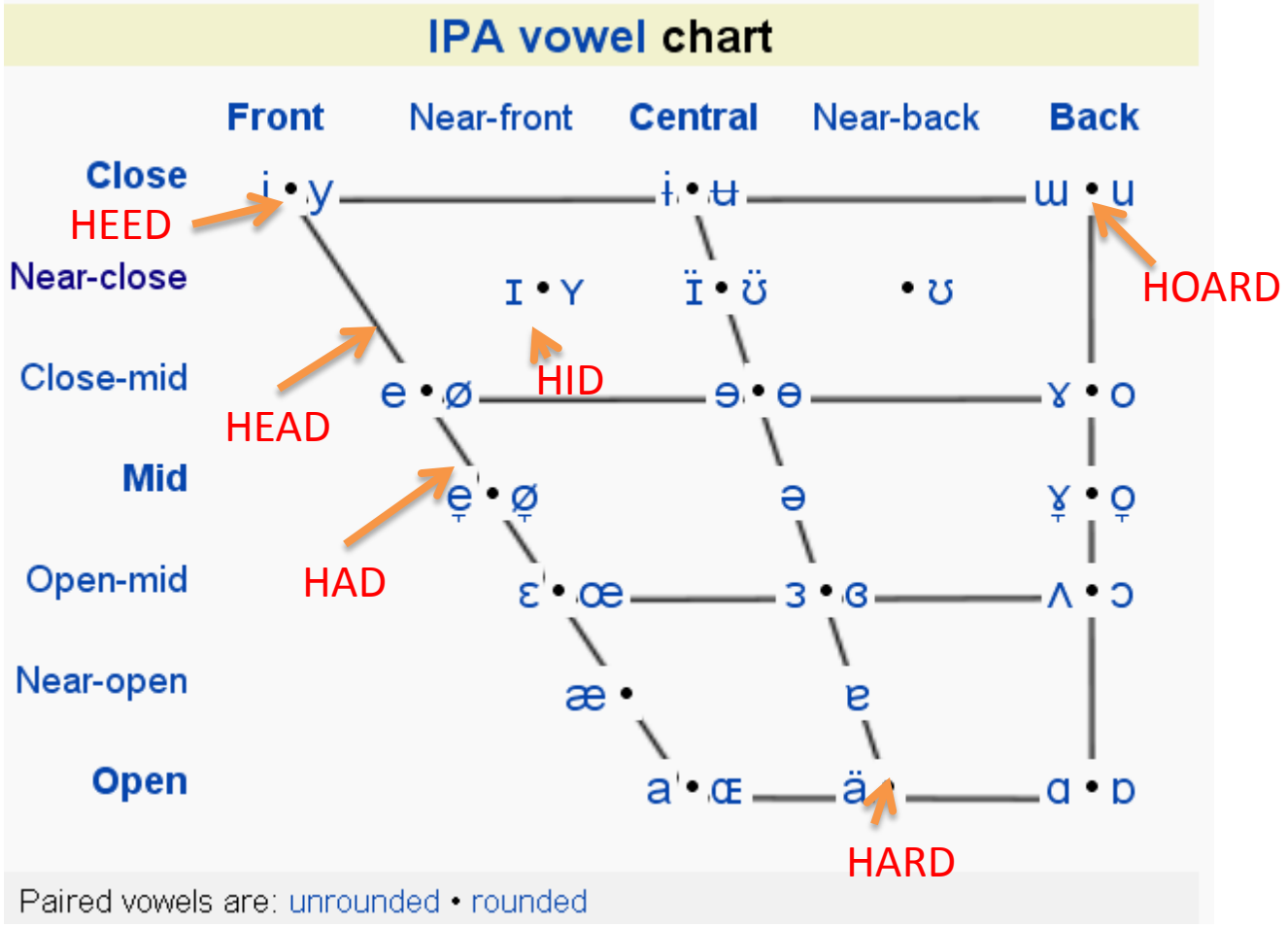
Mid-saggital cross-sections



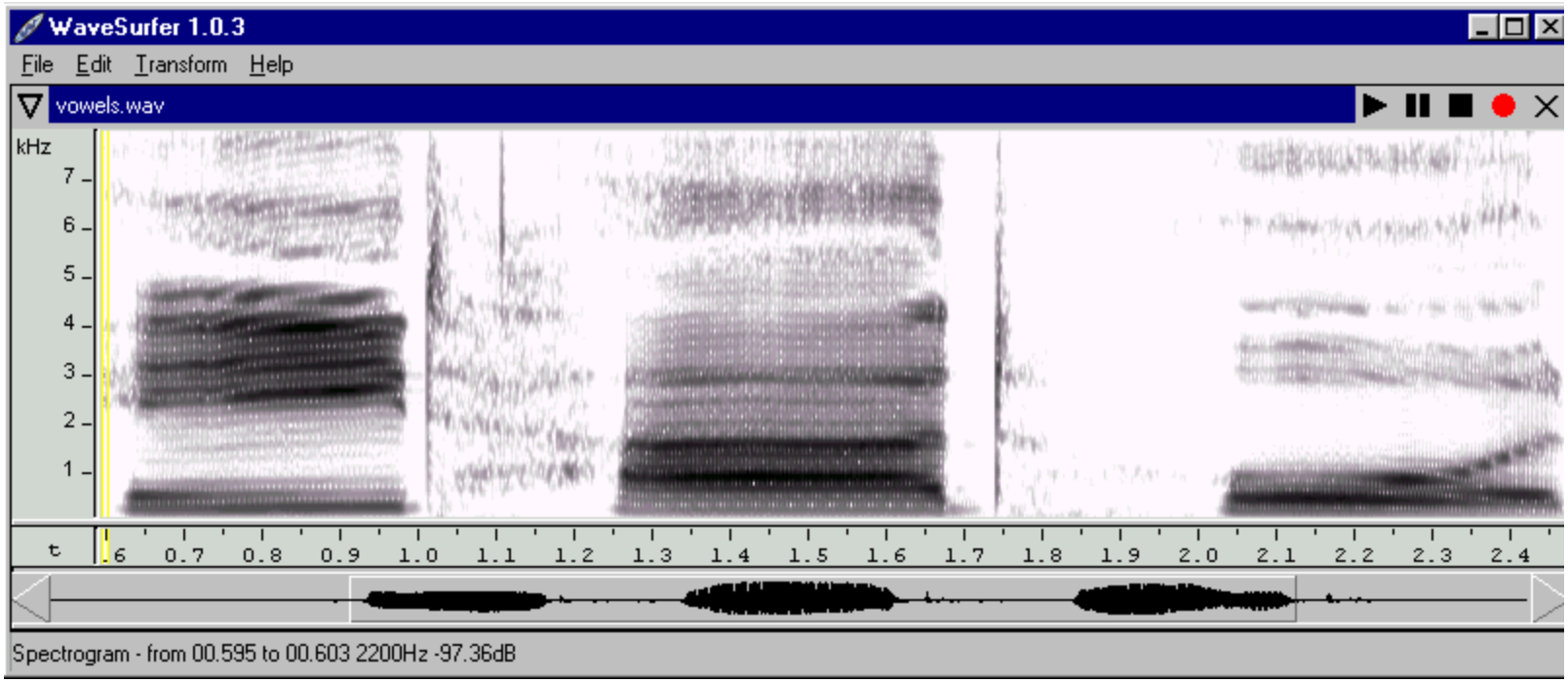
Phonetic representation of vowels



Phonetic representation of vowels



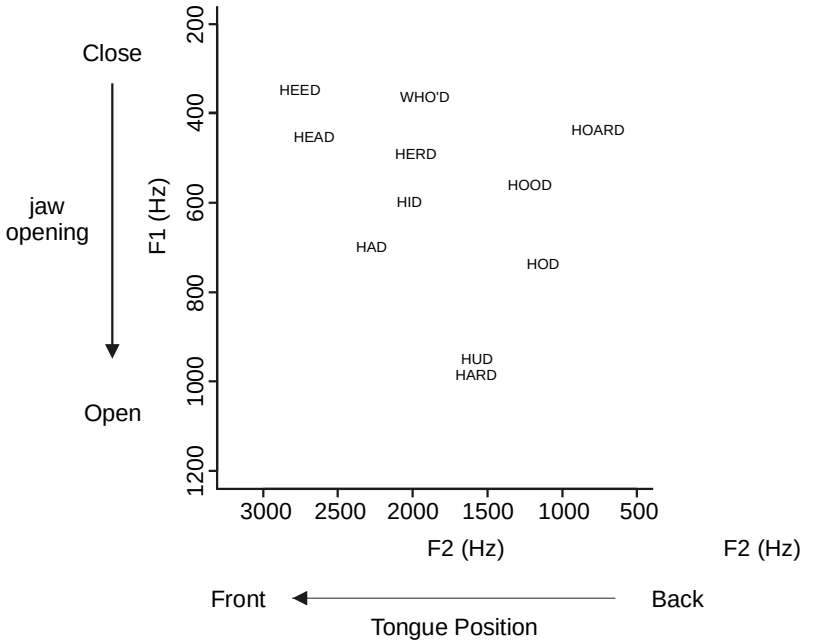
Spectrogram of some vowels



Heed Hard Hoard

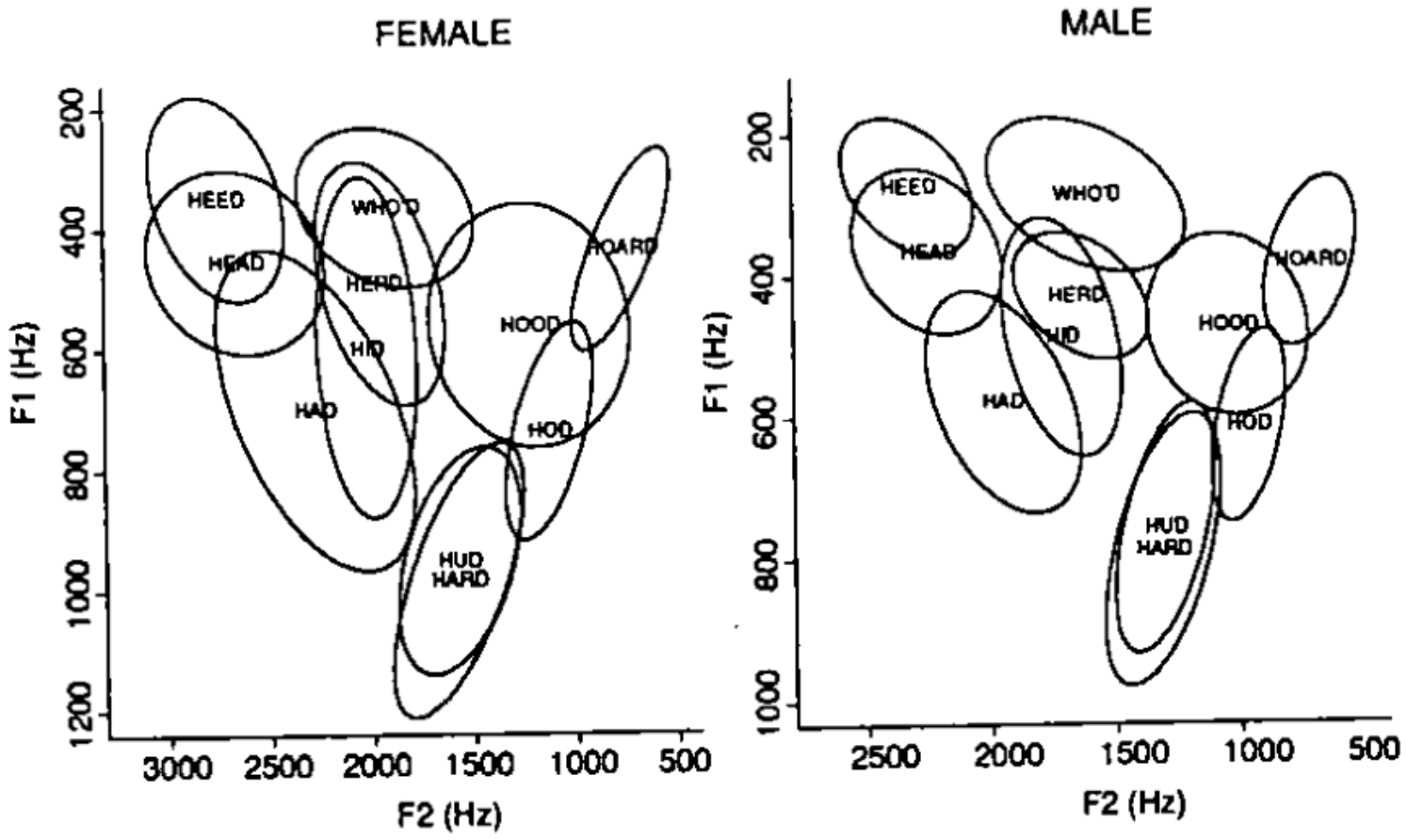
Formant Analysis

- It is possible to distinguish between any two vowels by the first two formants (F1 and F2) in their spectra.
- The F1 is inversely correlated with jaw height, and F2 is inversely correlated with tongue frontness.
- Thus plots F1 vs F2 show how vowels can differ both acoustically and in terms of articulation.

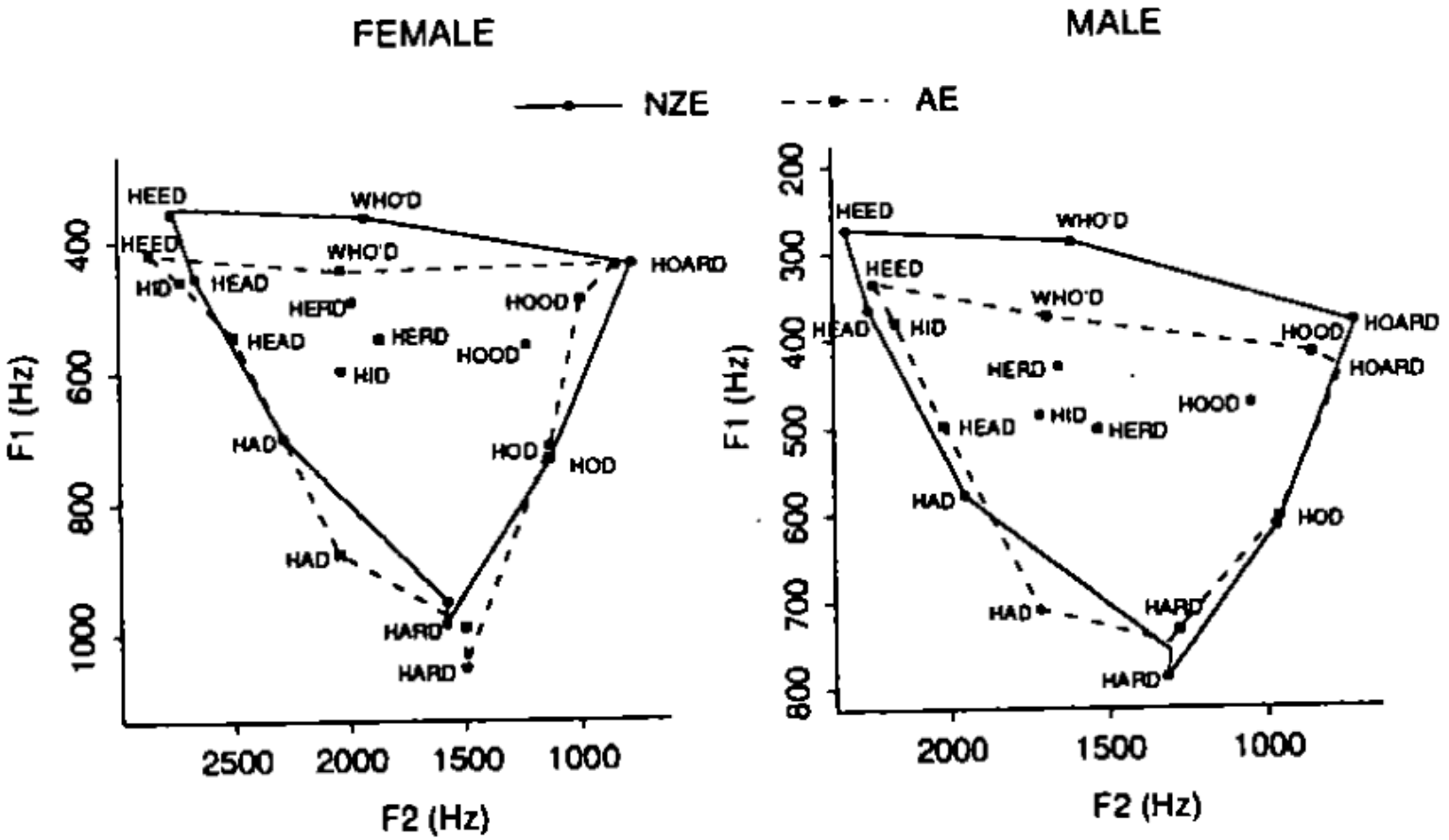


Adapted from Watson, Harrington and Evans (1998)

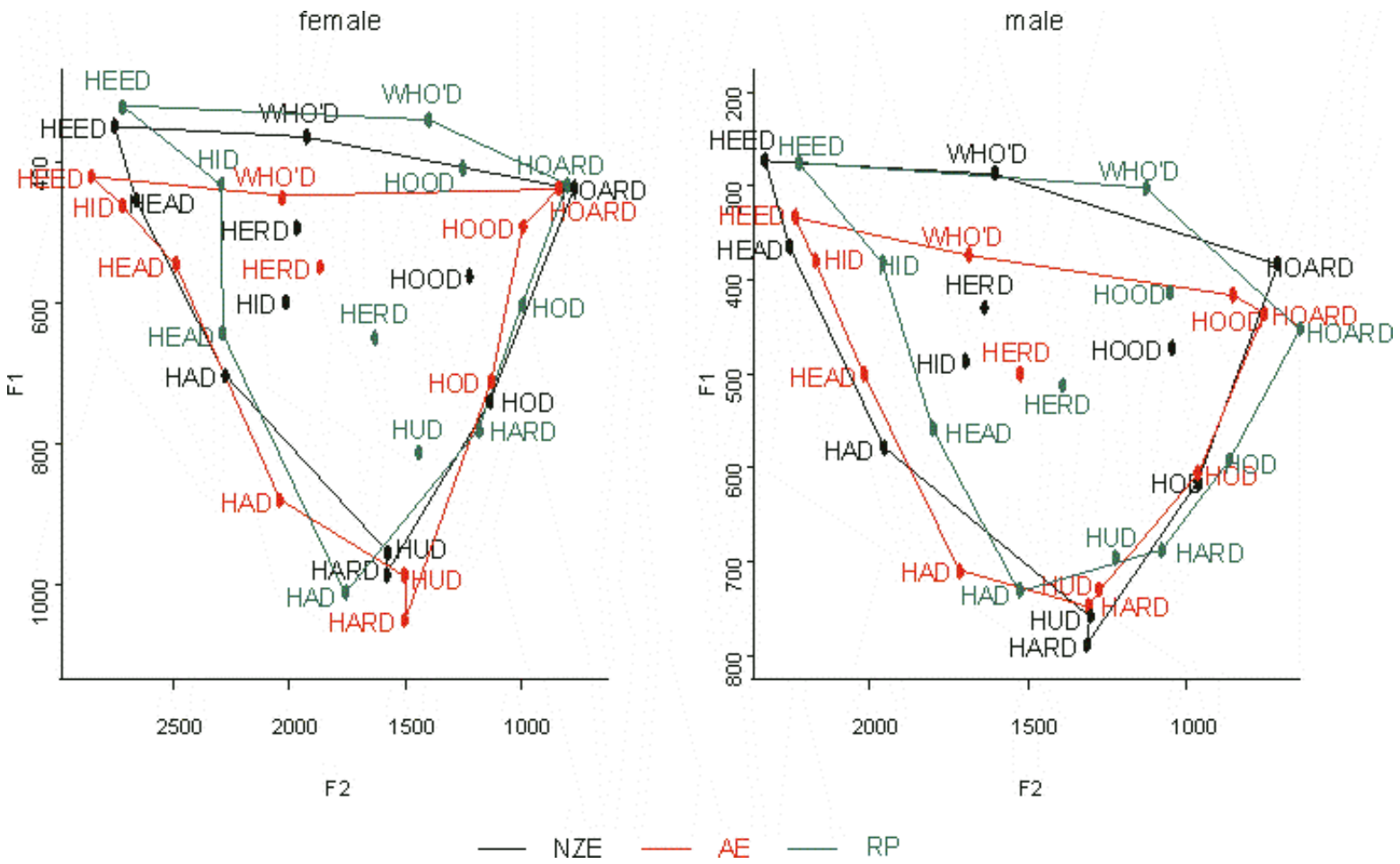
New Zealand English



Australian and New Zealand English

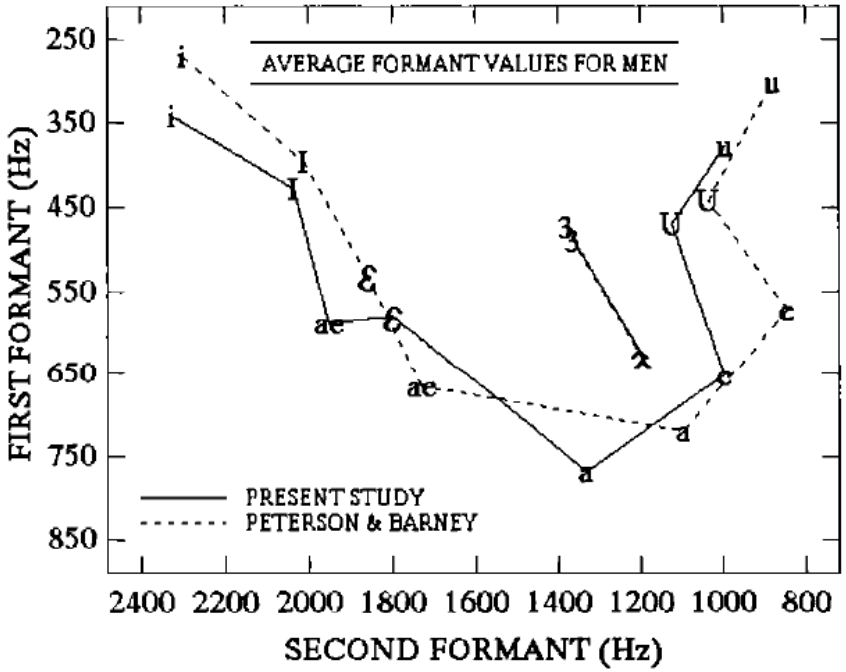


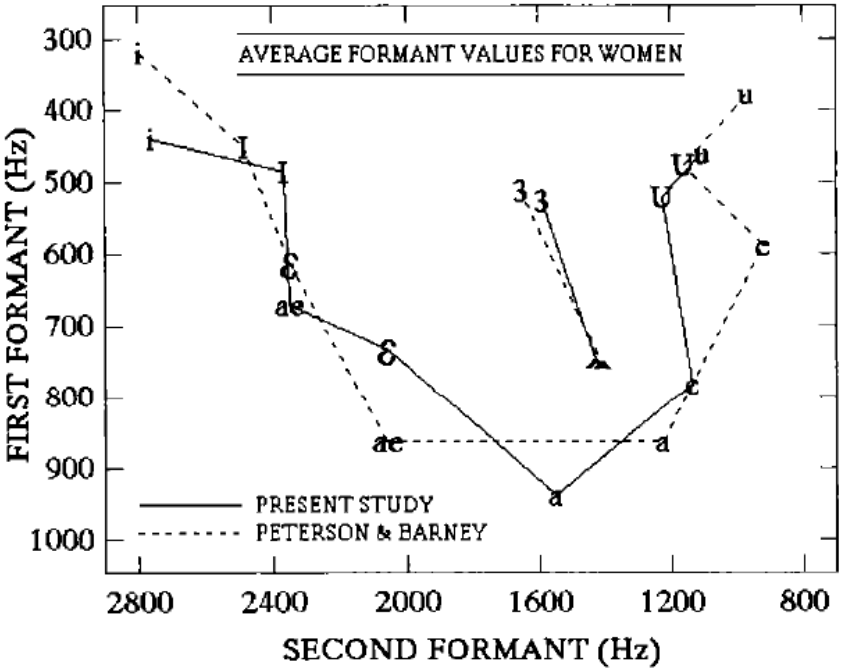
Vowel spaces for New Zealand English, Australian English and Received Pronunciation



Watson, Harrington, and Maclagan,1998

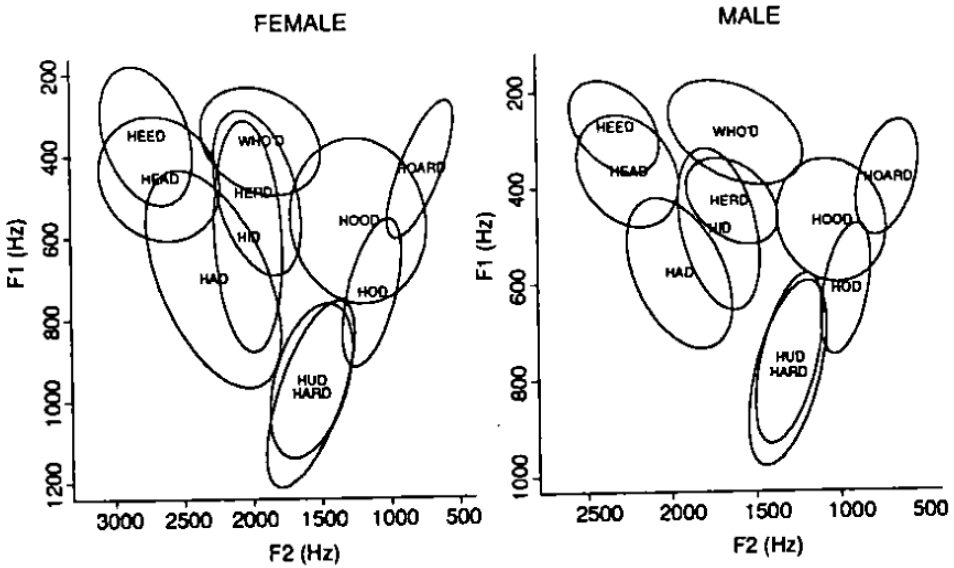
American English (Hillenbrand et al 1995)



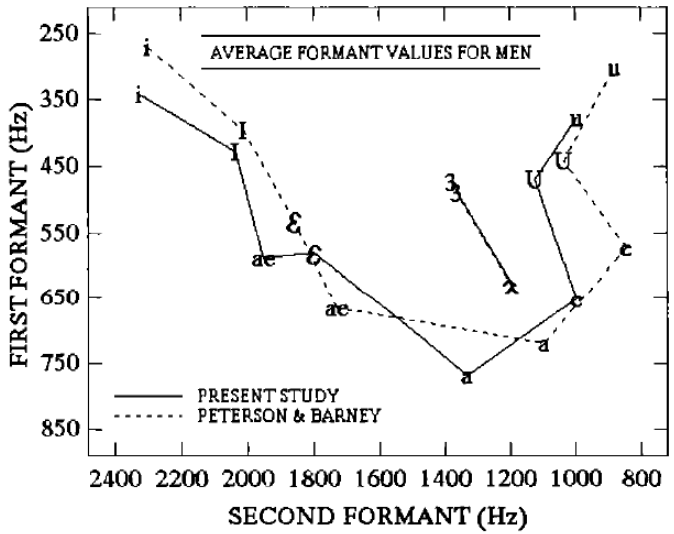


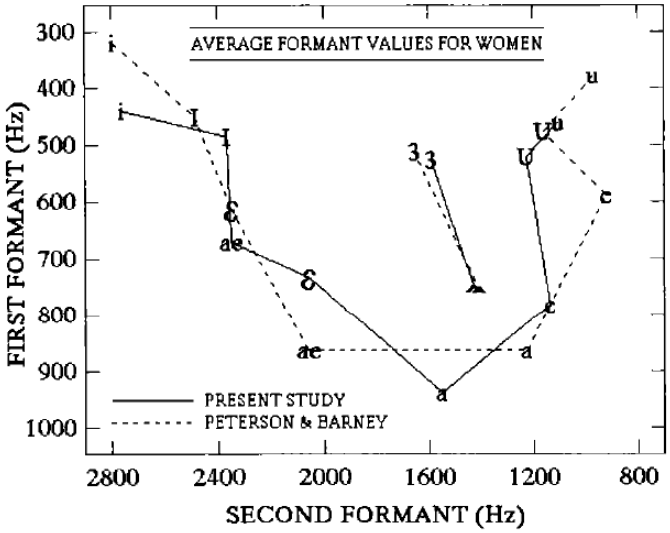
i-HEED, I-HID, ɛ-HEAD, æ-HAD, ʌ-,HUD, ɑ-HOD, ɔ-HOARD, ʊ-HOOD, u-WHO'D, ɜ-HERD

NZE vs. AmE

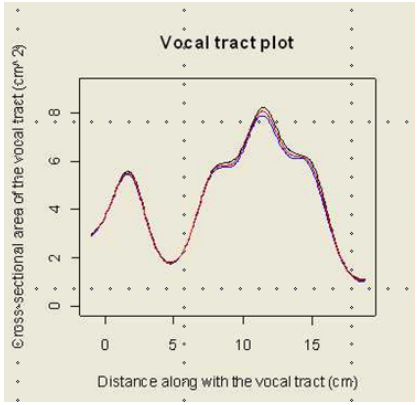


i-HEED, I-HID, ɛ-HEAD, æ-HAD, ʌ-,HUD, a-HOD, ɔ-HOARD, ʊ-HOOD, u-WHO'D, ɜ-HERD

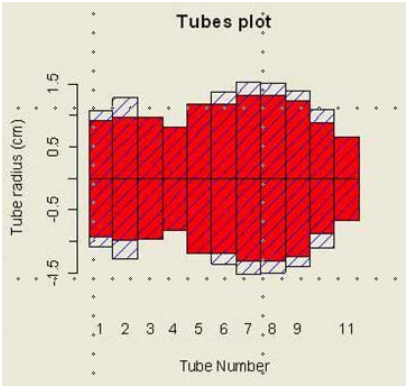




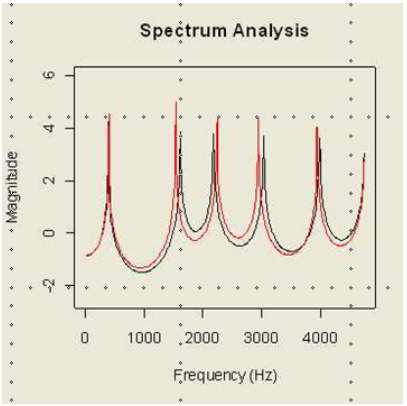
Acoustic Tube Model of the vocal tract



Obtain the cross-sectional area of the vocal tract



Represent it as a series of loss-less acoustic tubes of varying diameter, but equal width

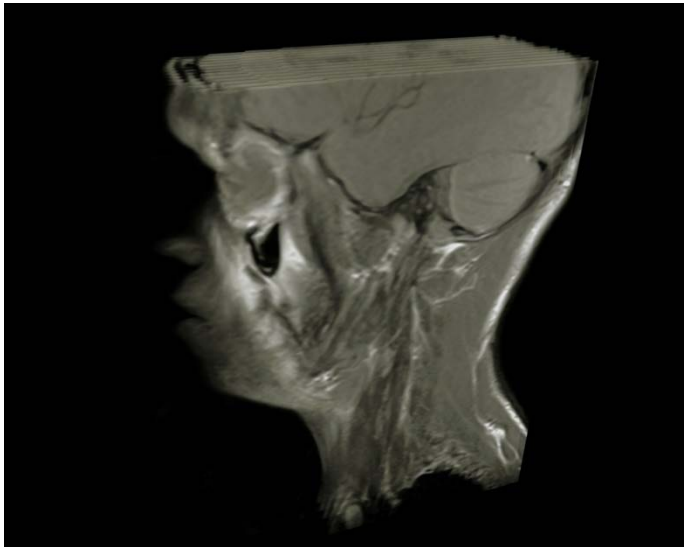


Find the resonances of the acoustic tube. First two resonances are strongly related to the vowel formants

COMPARISON OF NZE AND AMERICAN ENGLISH VOCAL TRACTS

NZE Vocal Tract Analysis using MRI

- 4 MR images were acquired with a 1.5T Siemens Magnetom Avanto MRI scanner in 2010 and 3 with the 3T version in 2011
- 7 NZE speakers (2 20-25 yrs, three 40-50 yrs, two 65+ yrs)
- The scans were done in parallel sagittal planes, each slice was 6mm, with no gaps and the images were T1-weighted.
- Speakers pronounced sustained vowels in “hVd” context: ie HEED, HID, HEAD, HAD, HARD, HUD, HOD, HOARD, HOOD, WHO’D,HERD.
- Each vocal tract shape took ~15 second to acquire, and each vowel had 13 slices.
- All participants were supine.

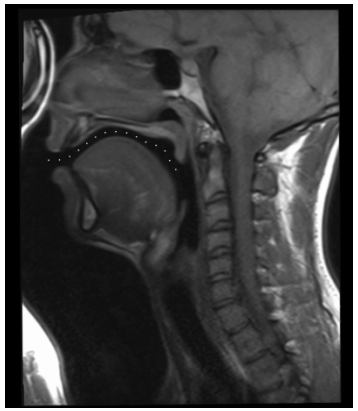


Thanks to Robert Bowmaker for the image.

American English MRI Data

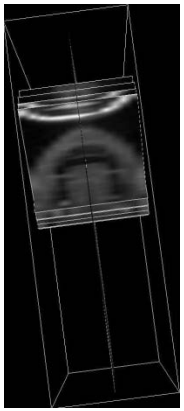
- Obtained from MRI studies by Story et al, published in 1996,1998 and 2008 respectively.
  - 1 man (done twice, 1994, 2002) 1 woman (1998)
  - MRI Lateral slices 5 mm apart
  - Images took 4 mins to acquire
  - 10 vowels said in isolation, (vowels in “HEED”,”HID”,”HEAD”,”HAD”, ‘HUD”,”HOD”,”HOARD” “HOED”,”HOOD”,’WHO’D”, sound in “HERD” not considered a traditional vowel”
- Data presented in tables of cross-sectional area vs. distance (the message used to obtain the cross-sectional areas different than approach we took).

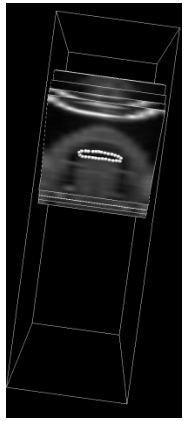
Obtaining the vocal tract data using CMGUI



1. Identify the vocal tract mid-line

2. For each node obtain plane Perpendicular to the centre line .



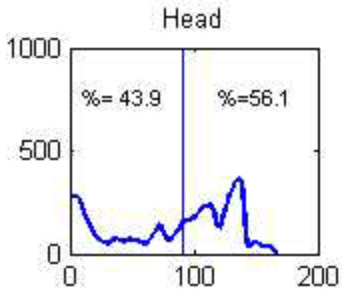


3. Mark out the vocal tract

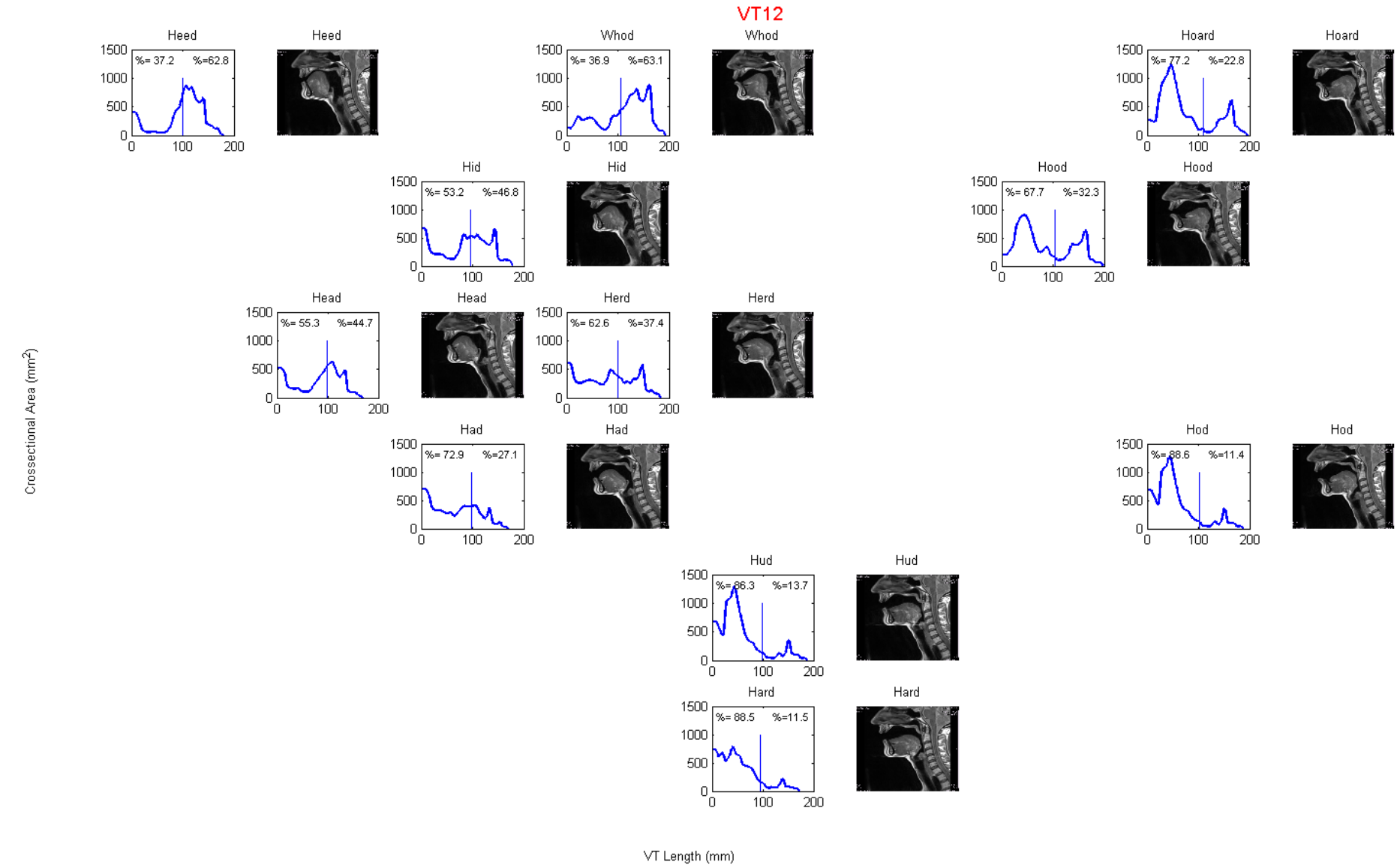
4. Repeat steps 1-3 for each of the 15 planes in the oral cavity

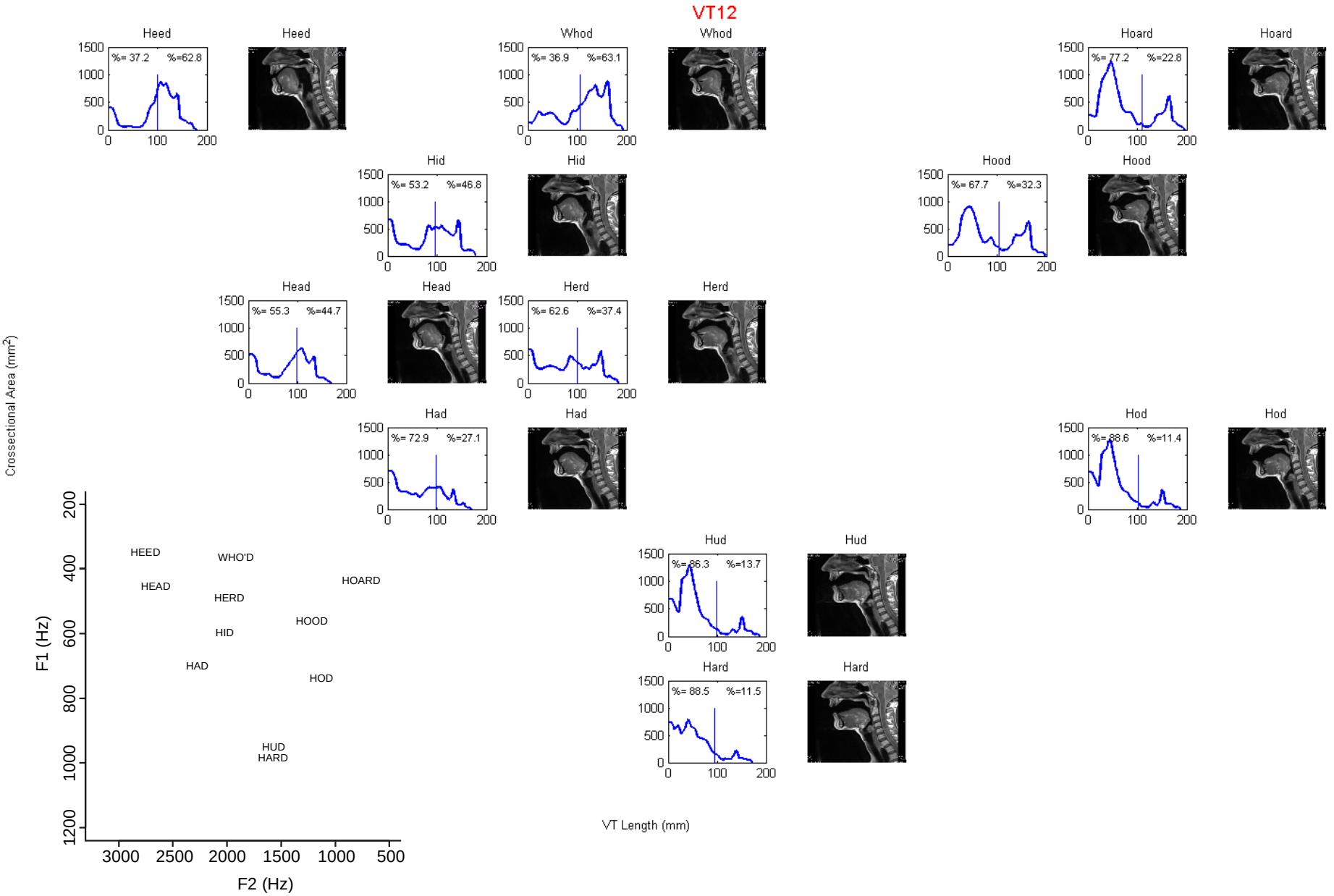
5 Calculate the area of each segment (perl script)

6. Repeat steps 1-5 for the pharyngeal cavity

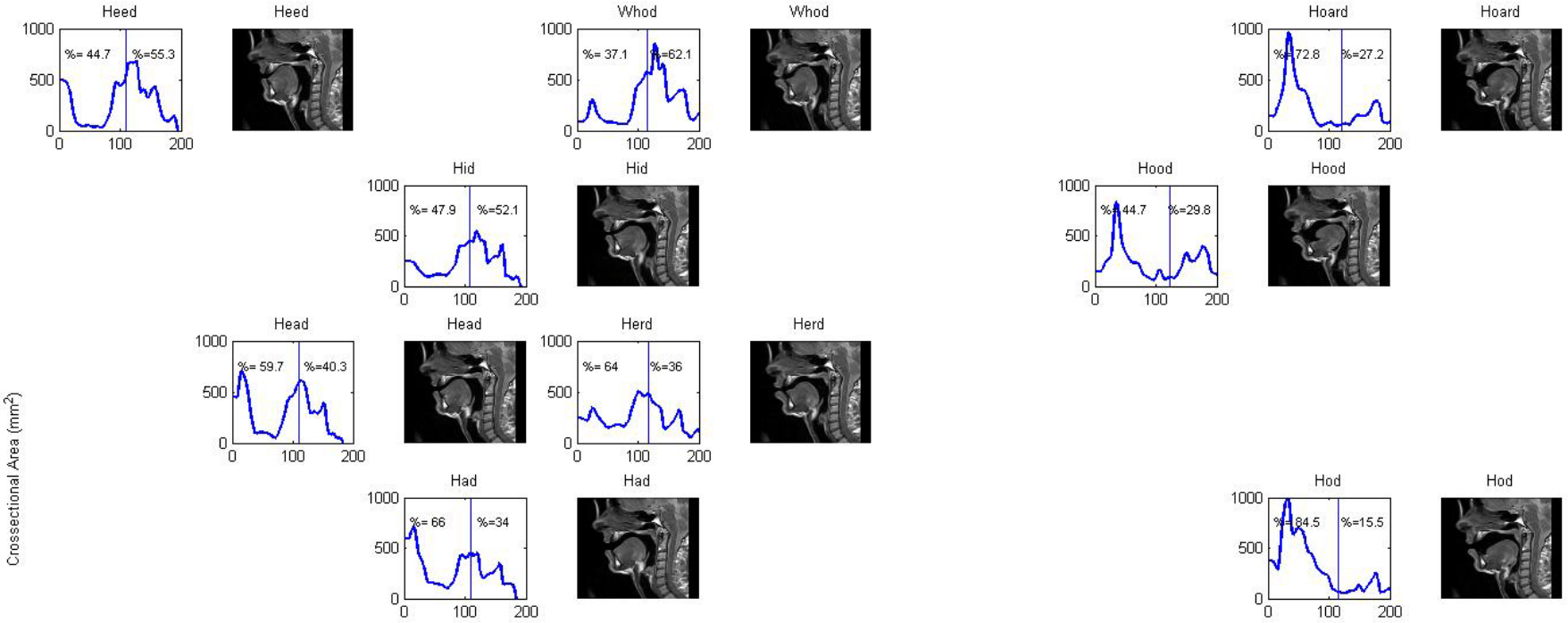


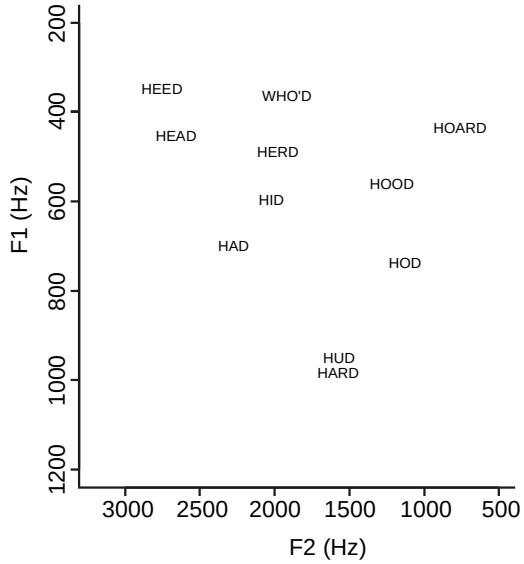
7. Combine oral and pharyngeal data, for plot of cross-sectional areas from lips to glottis

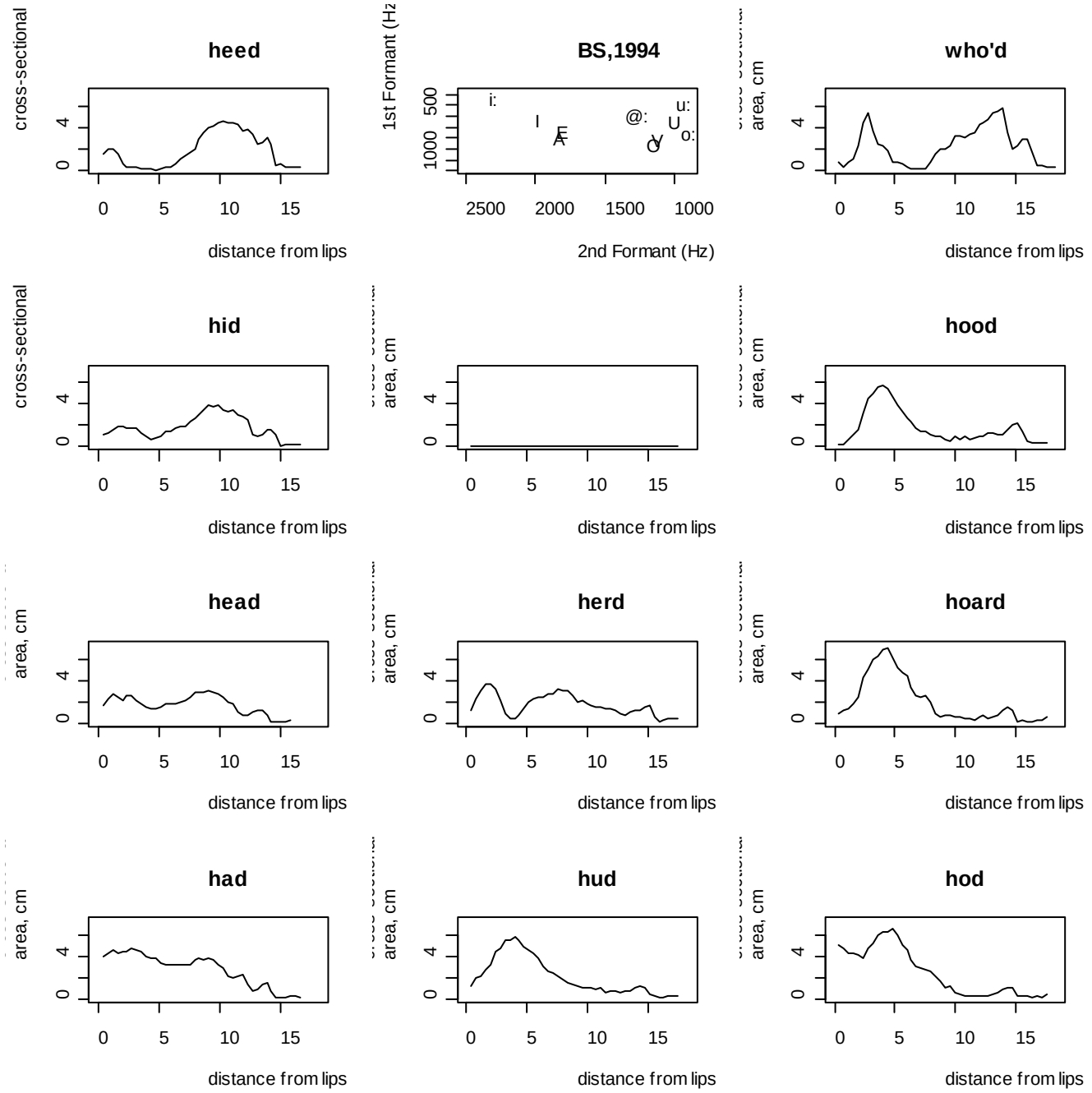

VT12
Heed
%= 37.2
%=62.8
Heed
Whod
%= 36.9
%=63.1
Whod
Hoard
%= 77.2
%=22.8
Hoard
Hid
%= 53.2
%=46.8
Hid
Hood
%= 67.7
%=32.3
Hood
Head
%= 55.3
%=44.7
Head
Herd
%= 62.6
%=37.4
Herd
Had
%= 72.9
%=27.1
Had
Hod
%= 88.6
%=11.4
Hod
Hud
%= 86.3
%=13.7
Hud
Hard
%= 88.5
%=11.5
Hard
1500
1000
500
0
0
100
200
Crossectional Area (mm²)
VT Length (mm)

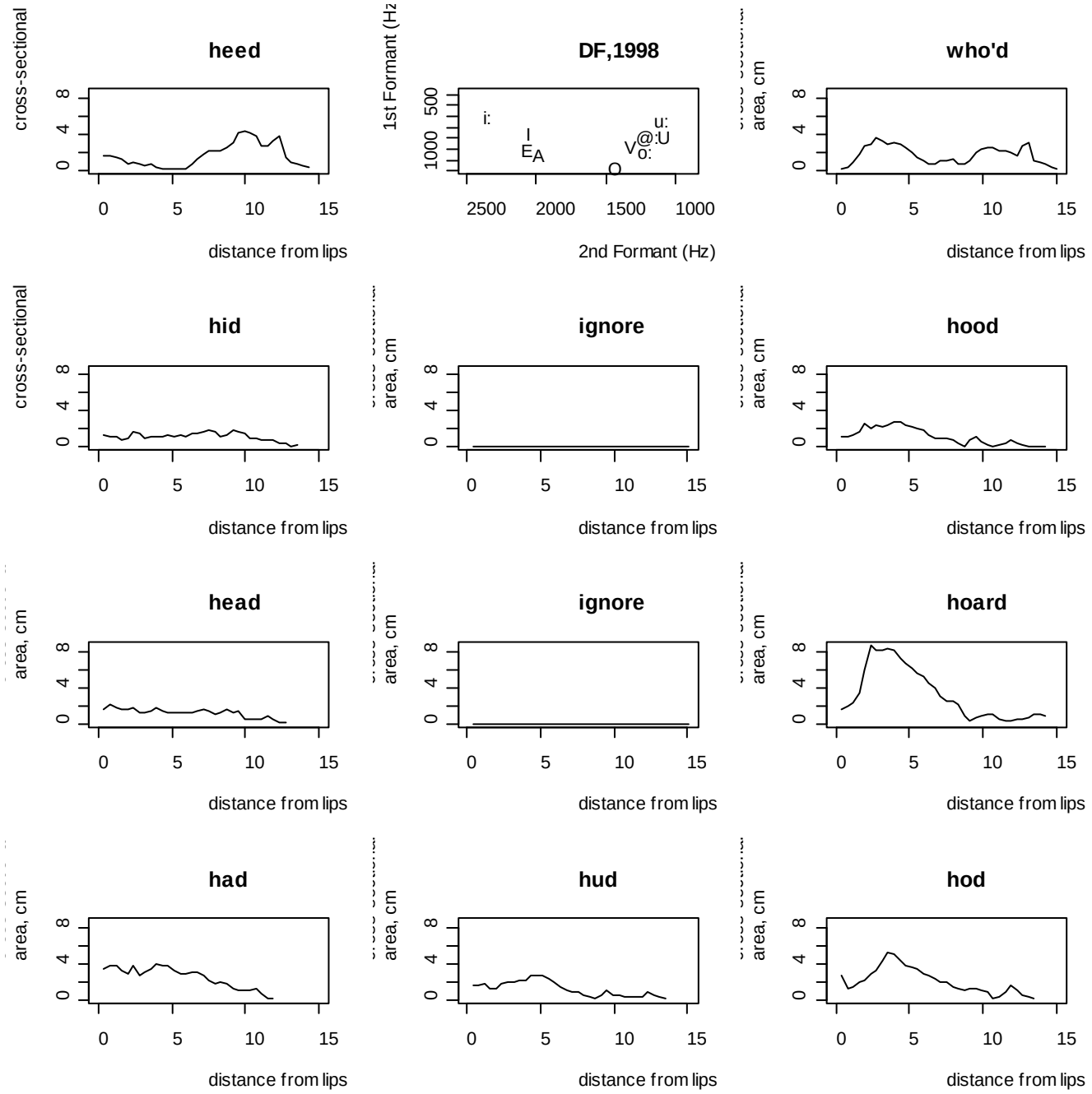

VT12
Heed
%= 37.2 %=62.8
Heed
Whod
%= 36.9 %=63.1
Whod
Hoard
%= 77.2 %=22.8
Hoard
Hid
%= 53.2 %=46.8
Hid
Hood
%= 67.7 %=32.3
Hood
Head
%= 55.3 %=44.7
Head
Herd
%= 62.6 %=37.4
Herd
Had
%= 72.9 %=27.1
Had
Hod
%= 88.6 %=11.4
Hod
Hud
%= 86.3 %=13.7
Hud
Hard
%= 88.5 %=11.5
Hard
HEED
WHO'D
HOARD
HEAD
HERD
HOOD
HID
HAD
HOD
HUD
HARD
F1 (Hz)
F2 (Hz)
Crossectional Area (mm²)
VT Length (mm)

VT09


Heed
%= 44.7
%=55.3
Whod
%= 37.1
%=62.1
Hoard
%= 72.8
%=27.2
Hid
%= 47.9
%=52.1
Hood
%= 44.7
%=29.8
Head
%= 59.7
%=40.3
Herd
%= 64
%=36
Had
%= 66
%=34
Hod
%= 84.5
%=15.5
Hud
%= 78.8
%=21.2
Hard
%= 86.5
%=13.5
Crossectional Area (mm²)
VT Length (mm)


HEED
WHO'D
HOARD
HEAD
HERD
HOOD
HID
HAD
HOD
HUD
HARD
F1 (Hz)
F2 (Hz)


heed
BS,1994
who'd
hid
hood
head
herd
hoard
had
hud
hod
cross-sectional
area, cm
1st Formant (Hz)
2nd Formant (Hz)
distance from lips
i:
I
E
A
@:
V
O
U
u:
o:
0
4
500
1000
0
5
10
15
2500
2000
1500
1000


heed
DF,1998
who'd
hid
ignore
hood
head
ignore
hoard
had
hud
hod
cross-sectional
area, cm
1st Formant (Hz)
2nd Formant (Hz)
distance from lips
i:
I
E
A
O
V
o:
@:
U
u:
0
4
8
500
1000
2500
2000
1500
5
10
15

NZE vocal tract length analysis

VT lengths, total length, in mm

	VT01	VT03	VT06	VT08	VT09	VT11	VT 12
HEED	187	165	153	141	194	184	180
HID	187	164	153	148	191	183	179
HEAD	188	169	153	150	191	179	171
HAD	178	167	153	140	184	177	170
HARD	183	161	158	138	197	181	173
HUD	180	168	154	141	202	178	177
HOD	181	166	157	145	212	181	188
HOARD	185	168	159	148	215	184	196
WHO'D	192	172	162	155	210	185	193
HERD	184	167	158	156	205	182	184
HOOD		168	159	155	217	186	198

Maximum length

Minimum length

- Vocal tract length=distance from lips to glottis
- Vocal tract lengths vary on a vowel by vowel basis
- Maximum vocal tract length – lip rounded vowels –typically WHO'D or HOOD.
- Minimum vocal tract length – open vowel – usually HAD or HARD
- Those with phonetics/singing training had the largest range of vocal tract lengths.

American English vocal tract length analysis

	BS, Gen Am	BS, Gen Am	DJ, Texas	VT02, AuE
	1996	2002	1998	2010
had	166.7	161.1	119	159
hard				158
head	158.8	173.0	127.0	154
heed	166.7	169.0	142.9	152
herd	174.6			160
hid	166.7	165.5	134.9	158
hoard	174.6	174.0	142.9	166
hod	174.6	170.9	134.9	167
hood	174.6	193.4	142.9	167
hud	174.6	171.4	134.9	163
who'd	182.5	195.9	150.8	172

NZE Oral Cavity Volume (as a % of total volume)

	VT01 %Oral	VT03 %Oral	VT06 %Oral	VT08 %Oral	VT09 %Oral	VT11 %Oral	VT12 %Oral
had	69.71064	52.15489	53.24878	66.91056	66.02024	54.88414	72.93169
hard	81.47909	62.56494		81.38633	86.53252	65.79099	88.50877
head	58.41326	43.88434	50.94779	55.84268	59.73428	51.65989	55.26764
heed	37.31436	27.49841	35.36114	47.04006	44.67762	34.92442	37.15744
herd	70.15052	57.34196	62.08471	50.65281	64.02861	51.01412	62.57381
hid	40.34088	42.96775	56.72259	50.05812	47.85999	47.17128	53.22545
hoard	84.19224	75.31994	72.31753	76.31848	72.80729	59.18025	77.15414
hod	83.27853	72.07845	78.57388	81.33628	84.48437	66.11407	88.59002
hood		45.99178	59.3262	66.11307	59.96857	63.5	67.67995
hud	78.69142	65.80489	71.72311	79.74097	78.80381	67.63535	86.2869
who'd	33.17369	27.41514	40.69692	36.16991	37.37681	31.99693	36.8691

Maximum

Minimum

- Note can't look at actual values, as can't guarantee actual vocal effort same for each vowel
- WHO'D has the lowest % volume in the oral cavity for all speakers with the except of VT06
- The next lowest % volume in the oral cavity is HEED, followed by HID, for all speakers except VT06, where its WHO'D followed by HEED.
- The top four largest % volume in the oral cavity are HARD, HUD, HOD, and HOARD, although the exact order changes for the different speakers.

The range of the smallest % oral cavity varied from 27% to 37%

The range of the largest % oral cavity varied from 68% to 87%.

American Eng. Oral Cavity Volume (as a % of total volume)

	BS, Gen Am	BS, Gen Am	DJ, Texas	VT02, AuE
	1996	2002	1998	2010
had	78.7	81.1	75.6	59.0
hard				72.9
head	69.0	68.2	69.2	54.2
heed	37.7	36.1	32.0	29.3
herd	73.3			48.3
hid	57.4	48.6	62.2	32.1
hoard	87.9	93.1	89.8	66.1
hod	91.5	92.6	81.2	73.3
hood	75.2	77.4	86.4	51.8
hud	85.5	88.8	83.0	72.6
who'd	42.3	49.7	56.9	42.5

American English Speakers:

BS1:HEED,WHO'D,HID,HEAD,HERD,HOOD,HAD,HUD,HOARD,HOD

BS2:HEED,HID,WHO'D,HEAD,HOOD,HAD,HUD,HOD,HOARD

DJ: HEED,WHO'D,HID,HEAD,HAD,HOD,HUD,HOOD, HOARD

Australian English Speaker

HEED, HID,WHO'D,HERD,HOOD,HEAD,HAD,HOARD,HUD,HARD,HOD

Comparison of mean NZE and Am E Vocal tract shape

% volume in oral cavity
44%

% volume in oral cavity
73%