Tone-Vowel Interaction in Standard Chinese

Phil Hoole[1] & Fang Hu[2]

[1]*Phonetics Institute, Munich University*

[2]*Phonetics Lab, Institute of Linguistics, Chinese Academy of Social Sciences*

hoole@phonetik.uni-muenchen.de | hufang@cass.org.cn

Overall Question

To what extent is tone production reflected in supraglottal articulation?

Here we explore two main approaches to this question.

Topic 1: Does position of tongue (and jaw and lips) vary systematically with tone?

Differences have been found

(See Zee, 1980; Torng, 2000; Erickson et al., this meeting; Hu, this meeting)

BUT

Often rather sporadic
i.e not consistent over speakers and over the complete vowel system.

Topic 2: To what extent is tone production associated with *visible* movements?

Burnham et al. (2001) found evidence that limited tonal perception is possible based on visual information alone (see also Mixdorff & Charnvivit, this meeting).

But what visual information is actually involved here?

One candidate is **head movement**.

Yehia et al., 2002, found systematic relations between head-position and F0.

A third topic linking articulation and fundamental frequency: *Intrinsic Pitch*

What are the articulatory mechanisms underlying intrinsic pitch?

A digression to consider German.

Does Mandarin Chinese provide new insights?

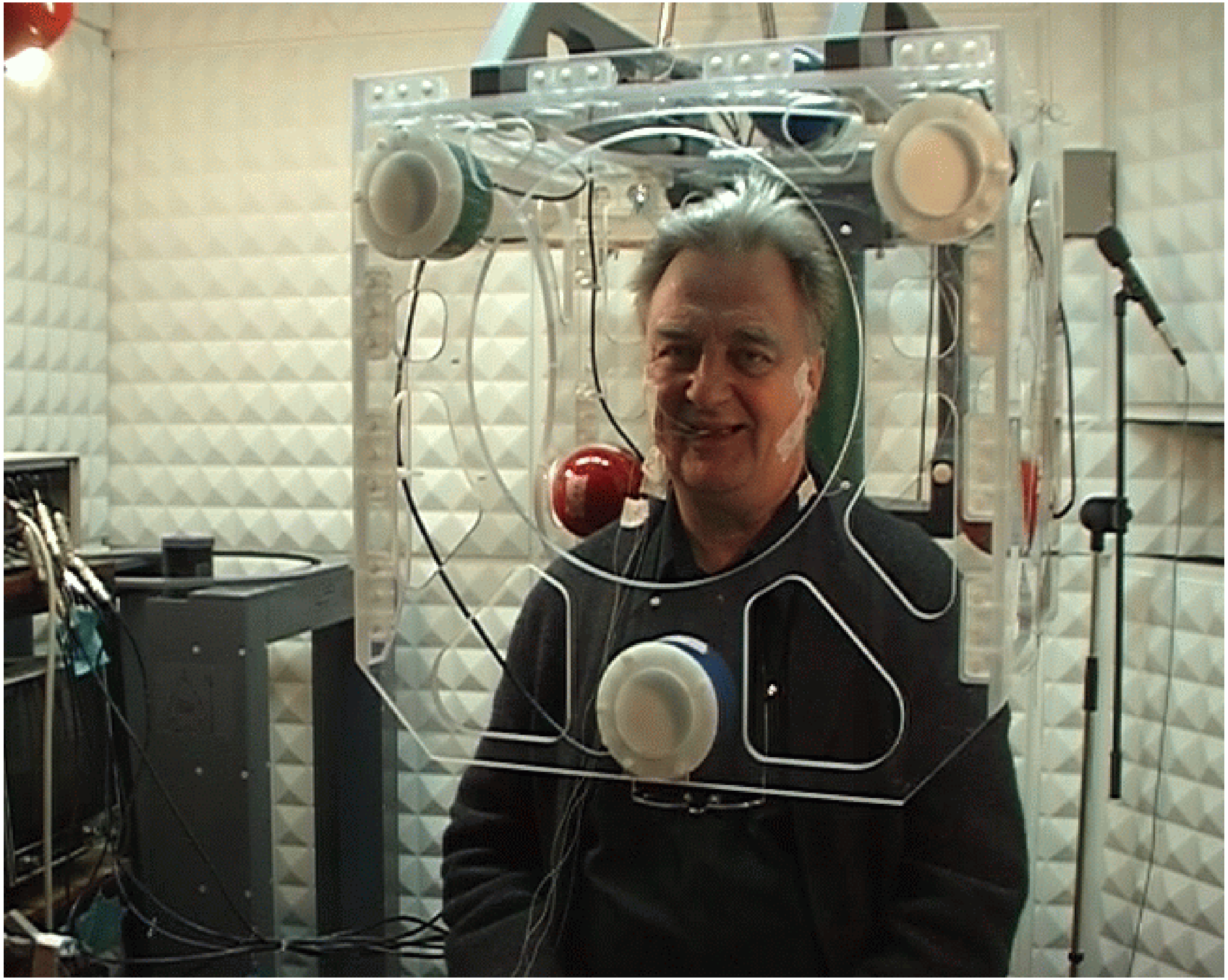
Method

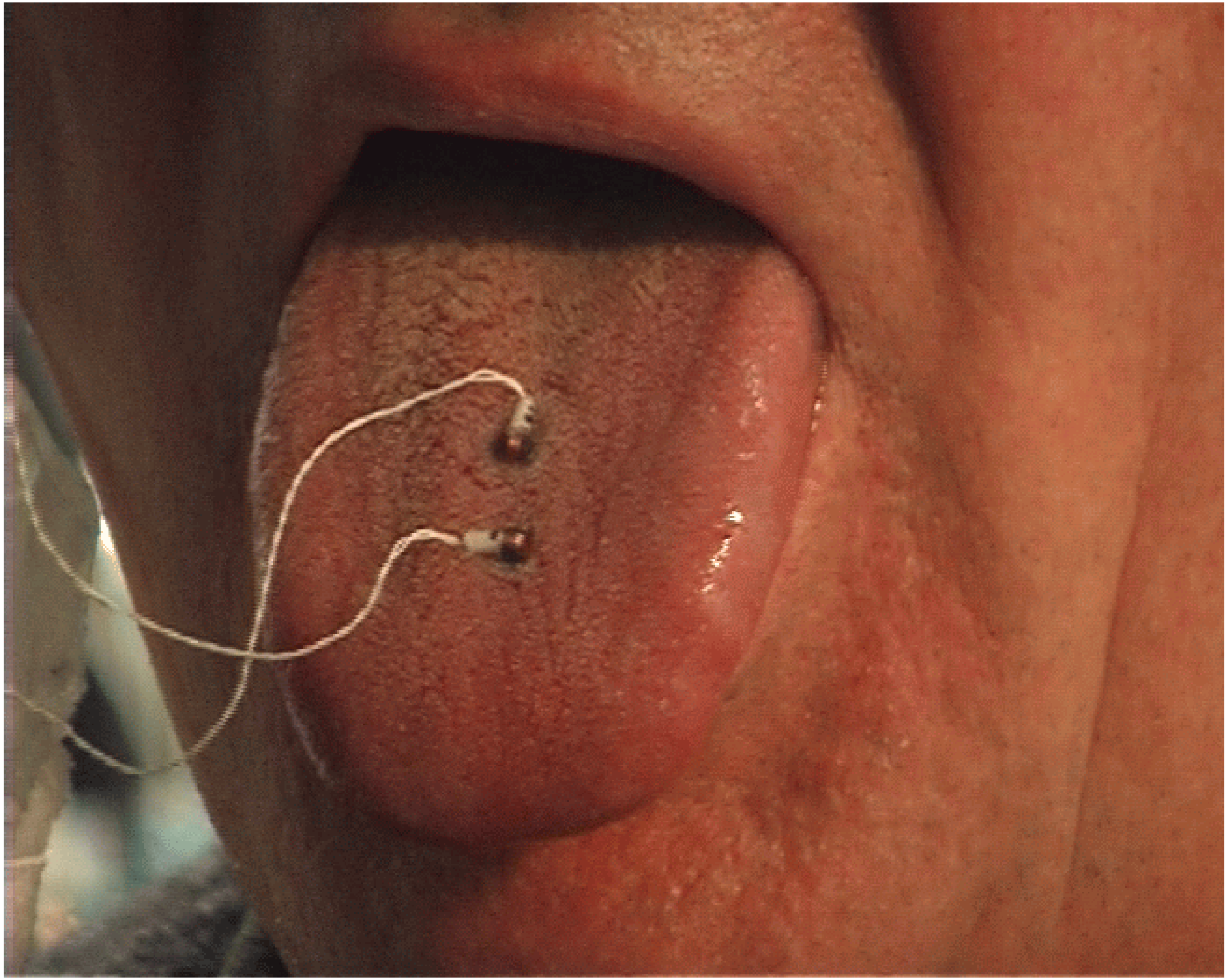
Monitoring of tongue, lip, jaw and head movement using three-dimensional electromagnetic articulography (EMA).

Better regarded as a **five**-dimensional system. Each sensor provides:

Three coordinates for position (x, y, z)

Two coordinates defining sensor orientation (azimuth, elevation).





Movie demos:

1. Articulatory movement including sensor orientation
   For this experiment three sensors on tongue; one each on lower-lip and jaw

2. Head movement monitored with two sensors and decomposed into the three translational and three rotational degrees of freedom.

Speech Material

Main material: The vowels /i/, /y/, /u/, /ɤ / and /a/.

Spoken in isolation on each of Tones 1 to 4 of Mandarin

In addition: Vowel /o/ spoken on Tone 1 only

Total of 12 repetitions of each combination of vowel and tone.

(Considerable amount of additional material collected for diphthongs, apical vowels and rhotacized vowels collected, but not analyzed here)

1 female Mandarin speaker, native of Beijing

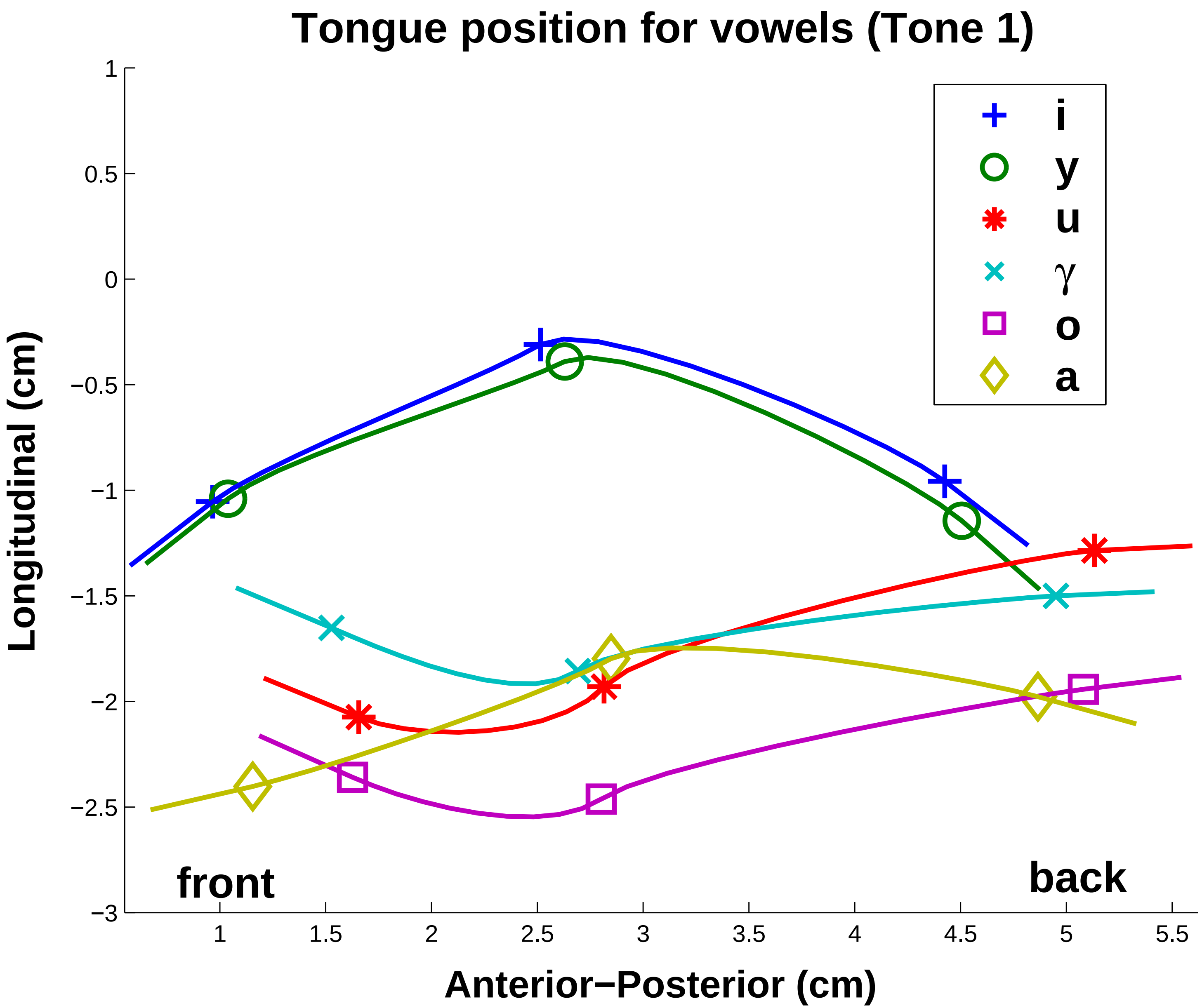
Results Topic 1
Articulator Movement

Setting the scene

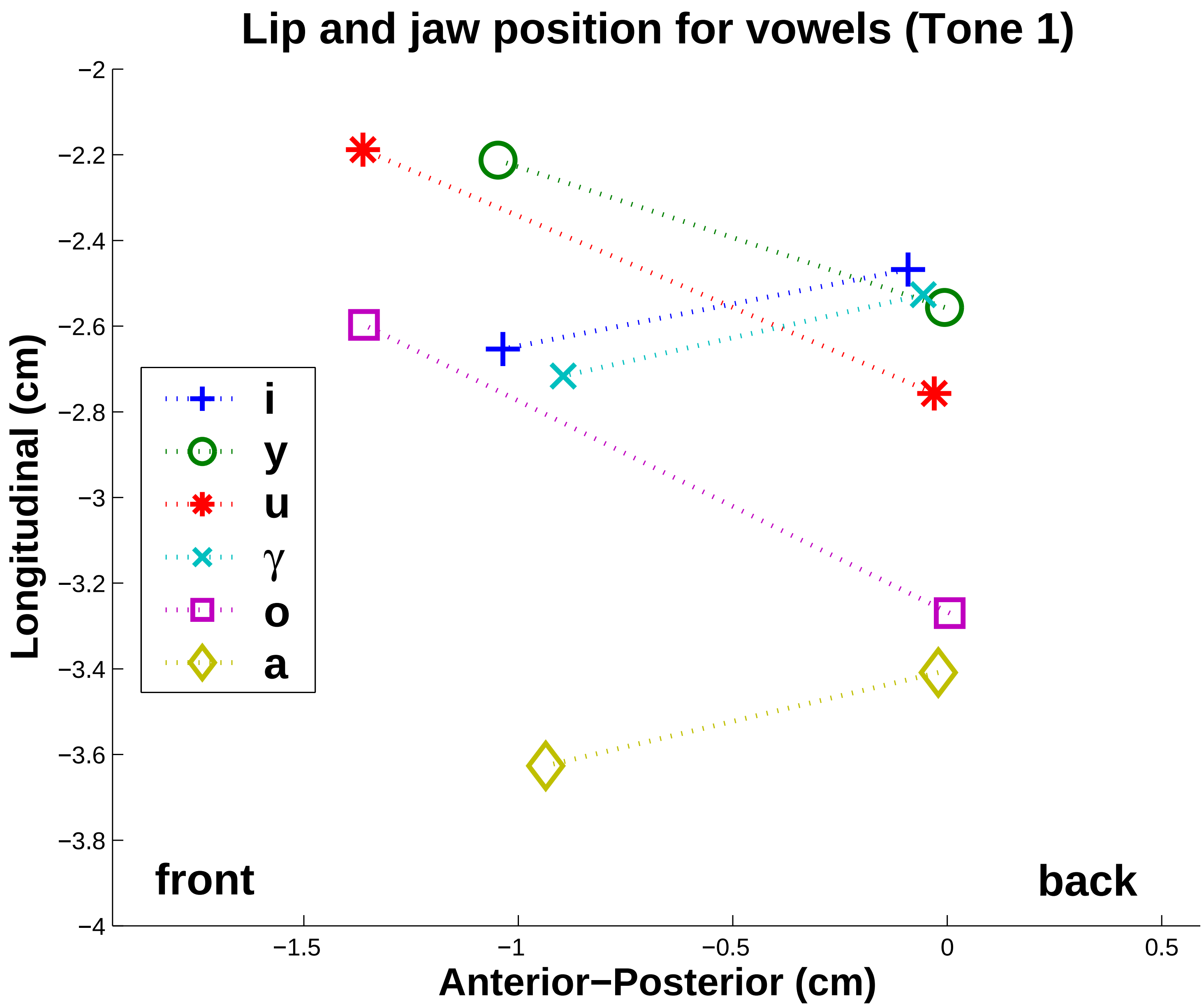
Sensor positions extracted at midpoint of each vowel

Overview of articulator configurations for the target vowels

Data for Tone 1 only, by way of example.



Tongue position for vowels (Tone 1)
i
y
u
ɣ
o
a
front
back
Longitudinal (cm)
Anterior−Posterior (cm)
1
0.5
0
−0.5
−1
−1.5
−2
−2.5
−3
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5


Lip and jaw position for vowels (Tone 1)
Longitudinal (cm)
Anterior−Posterior (cm)
−2
−2.2
−2.4
−2.6
−2.8
−3
−3.2
−3.4
−3.6
−3.8
−4
−1.5
−1
−0.5
0
0.5
i
y
u
γ
o
a
front
back

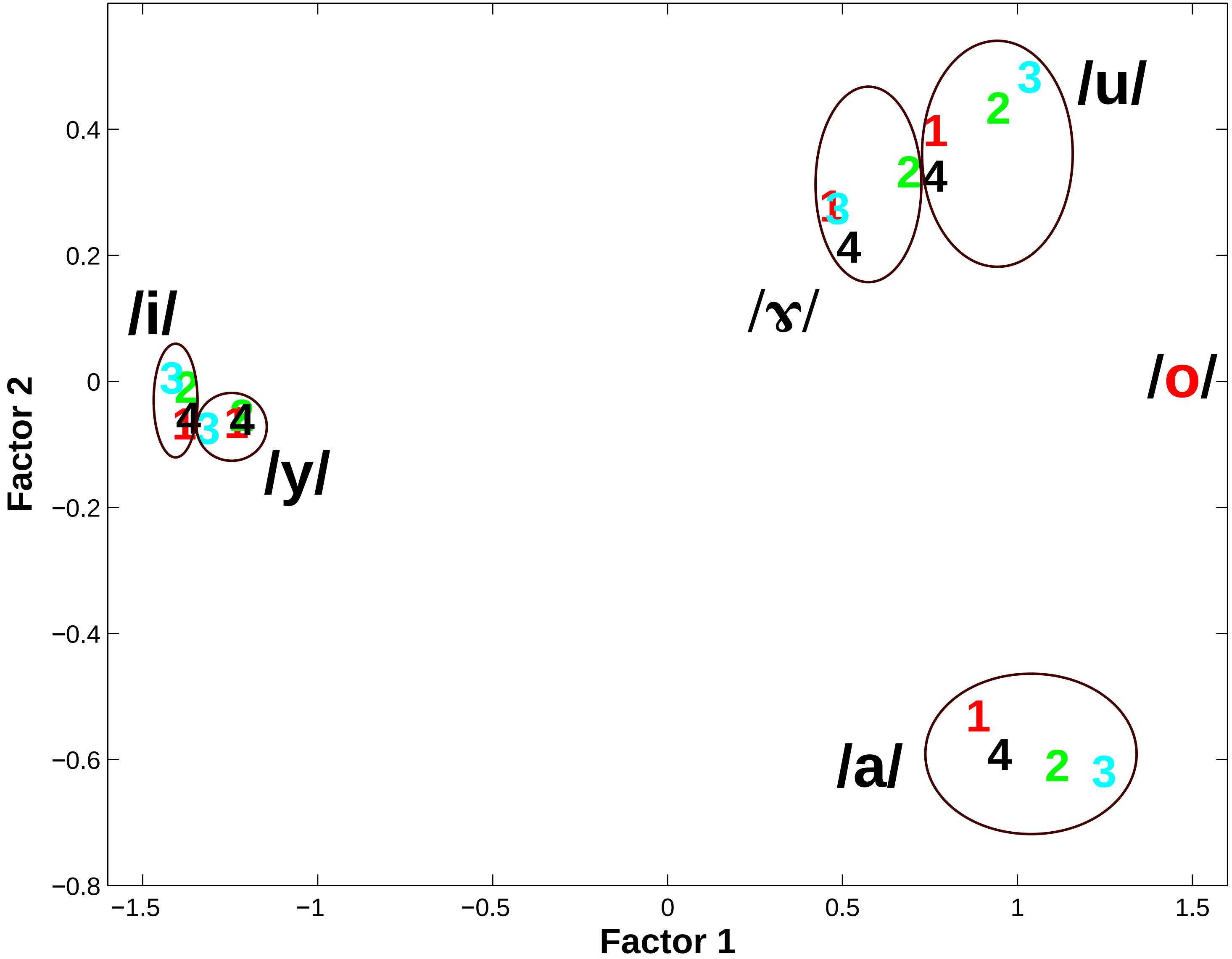
An intermediate step:

Principal component analysis of the anterior-posterior and vertical coordinates of the three tongue sensors.

Two factors explain about 96% of the variance
Factor 1: 87%, Factor 2: 9%

Statistical testing for effect of tone on articulator position based on these factor scores

Vowel−Tone combinations in the factor space



Influence of tone on tongue position:

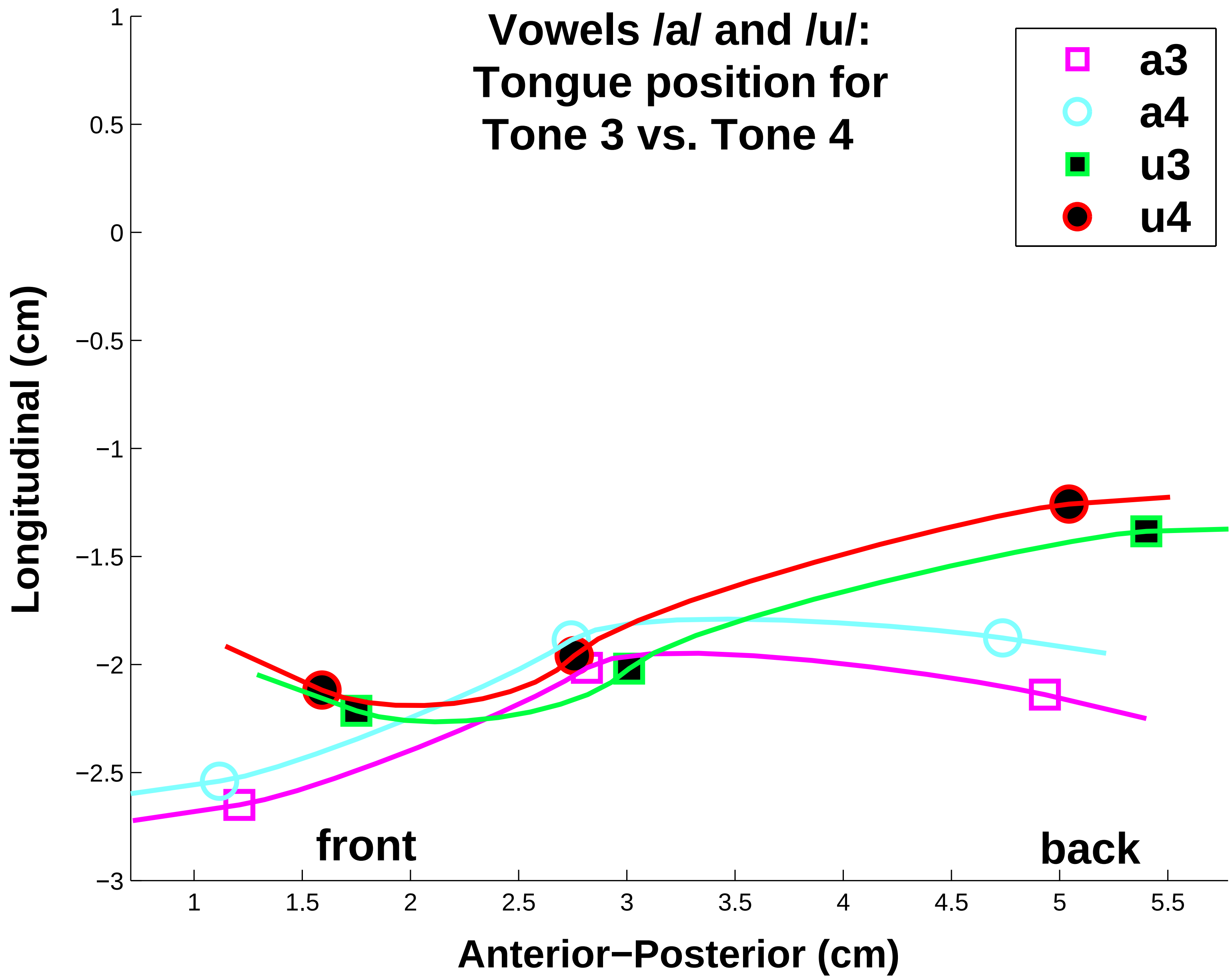
Clearly no common pattern for all vowels

But several individual contrasts were statistically significant ($p<0.01$)

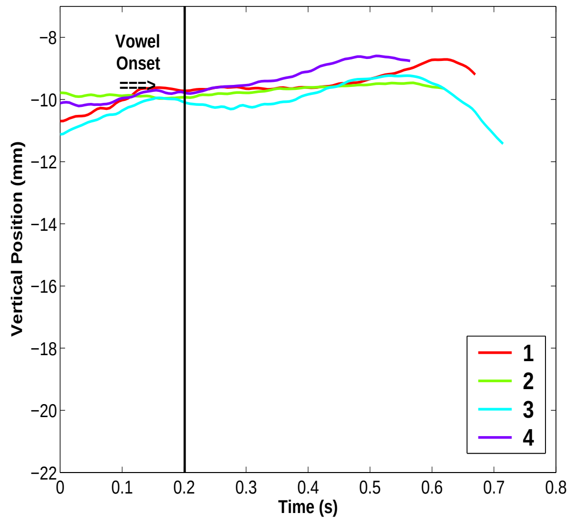
/a/ Factor 1: Tone 3 > Tone 1 and 4
Tone 2 > Tone 1

/u/ Factor 1: Tone 3 > Tone 4
Factor 2: Tone 3 > Tone 4

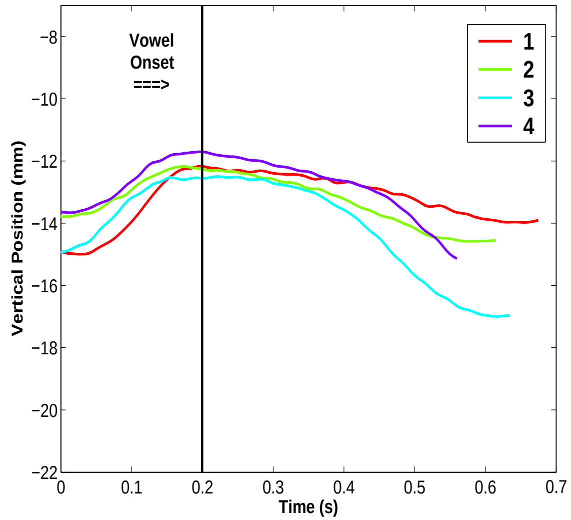
/y/ Factor 1: Tone 3 < Tone 1, 2 and 4


Vowels /a/ and /u/:
Tongue position for
Tone 3 vs. Tone 4
a3
a4
u3
u4
Longitudinal (cm)
1
0.5
0
−0.5
−1
−1.5
−2
−2.5
−3
front
back
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
Anterior−Posterior (cm)

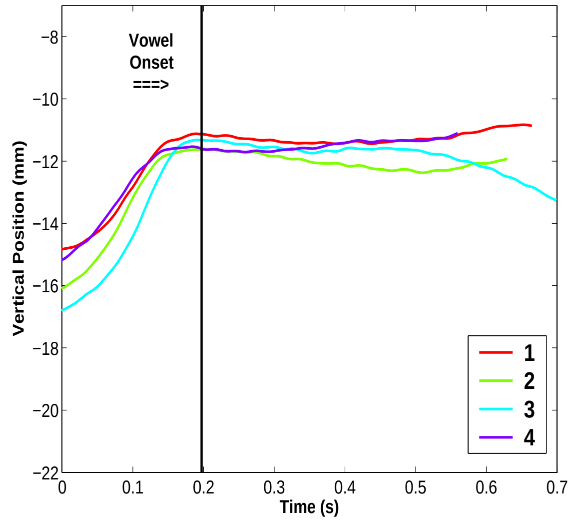
Vowel i



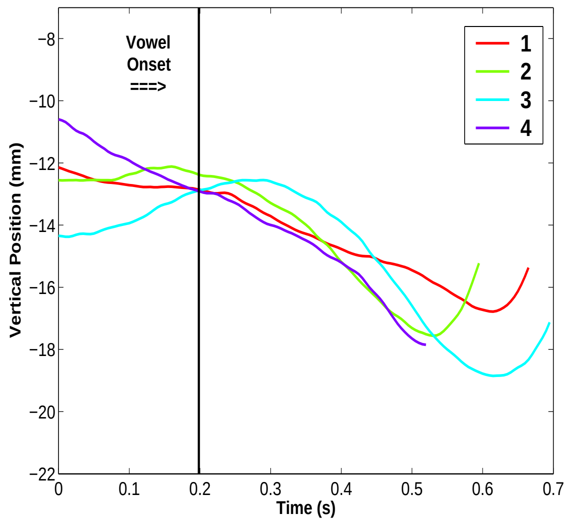
Vowel u



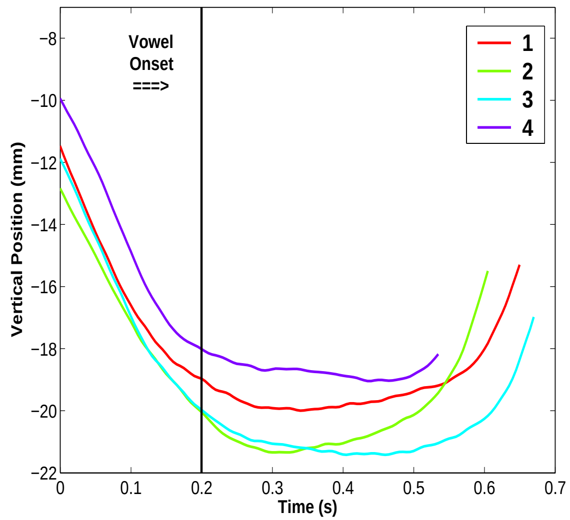
Vowel y



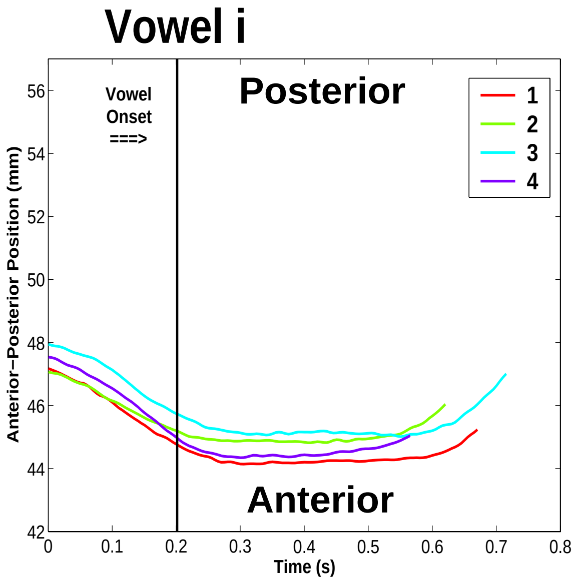
Vowel w

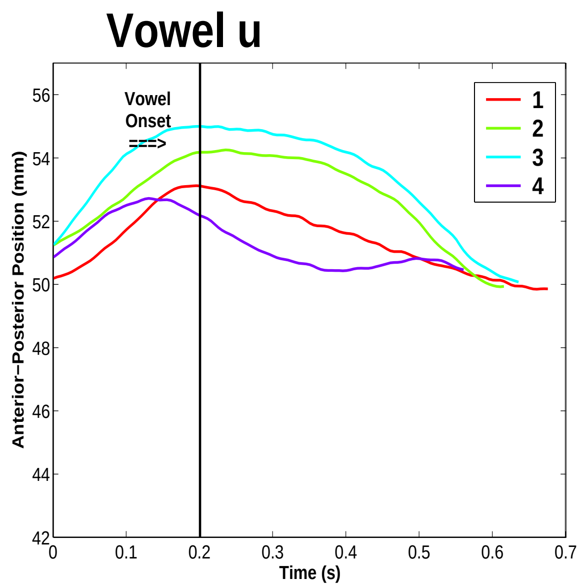


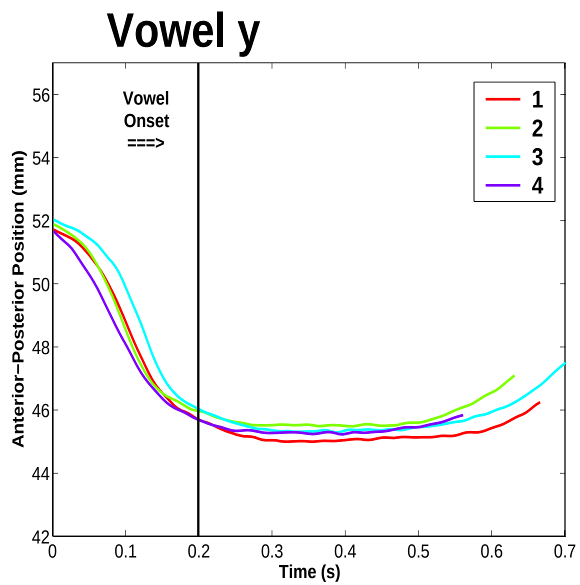
Vowel a

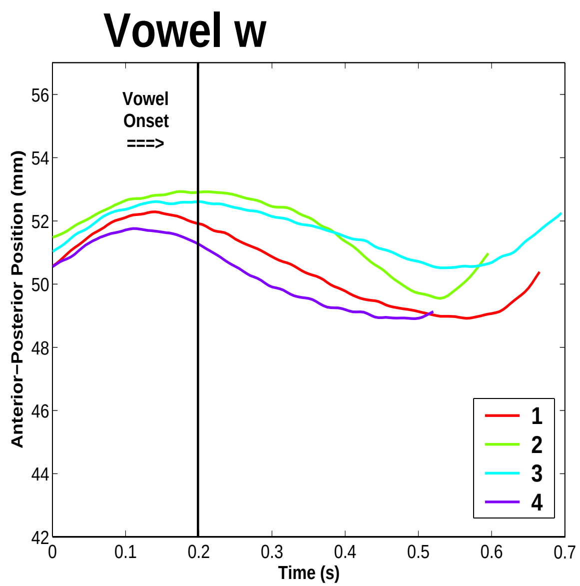


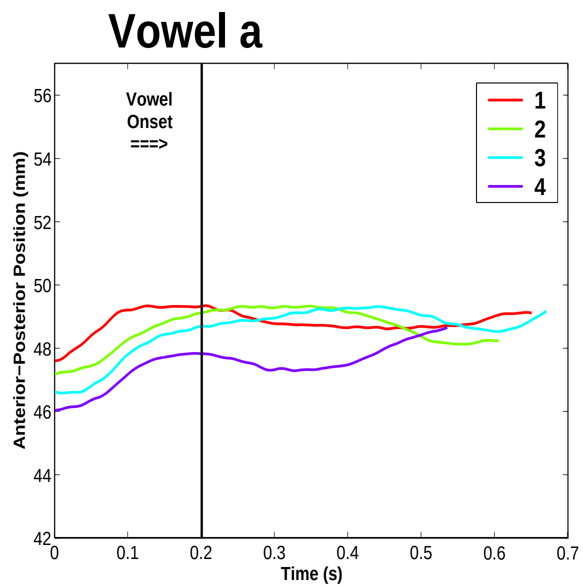
Vertical movement of rearmost tongue sensor for each tone











Anterior-posterior movement of rearmost tongue sensor for each tone

Jaw and lip position:

Are tongue-position differences attributable to jaw-position differences?

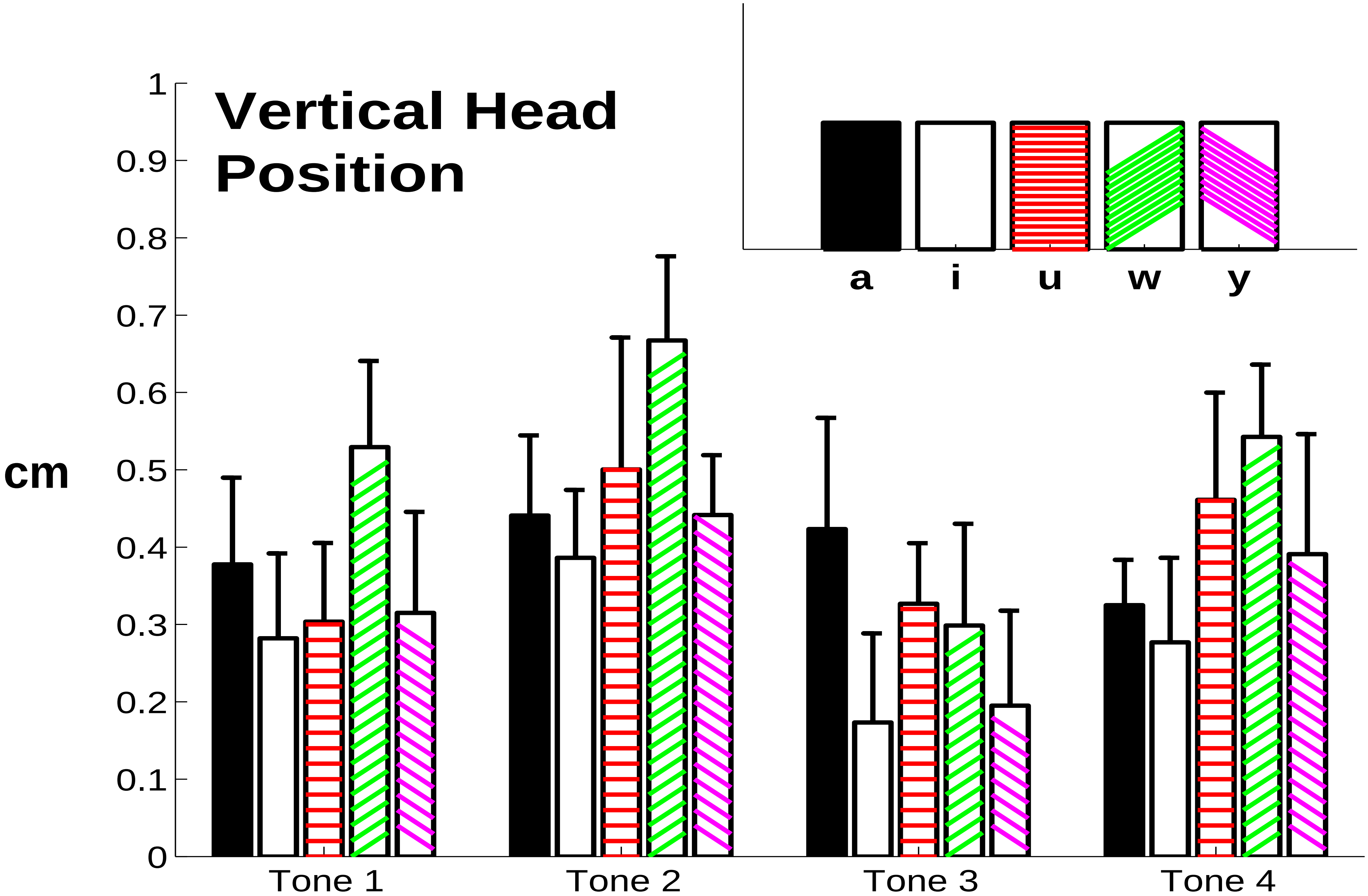
/a/ and /u/: Jaw more posterior for Tone 3 than Tone 4
/i/ and /y/: Jaw higher for Tone 3 than Tone 4

Lower-lip:
/a/ and /u/: More posterior for Tone 3 than Tone 4

Results Topic 2
Head Movement

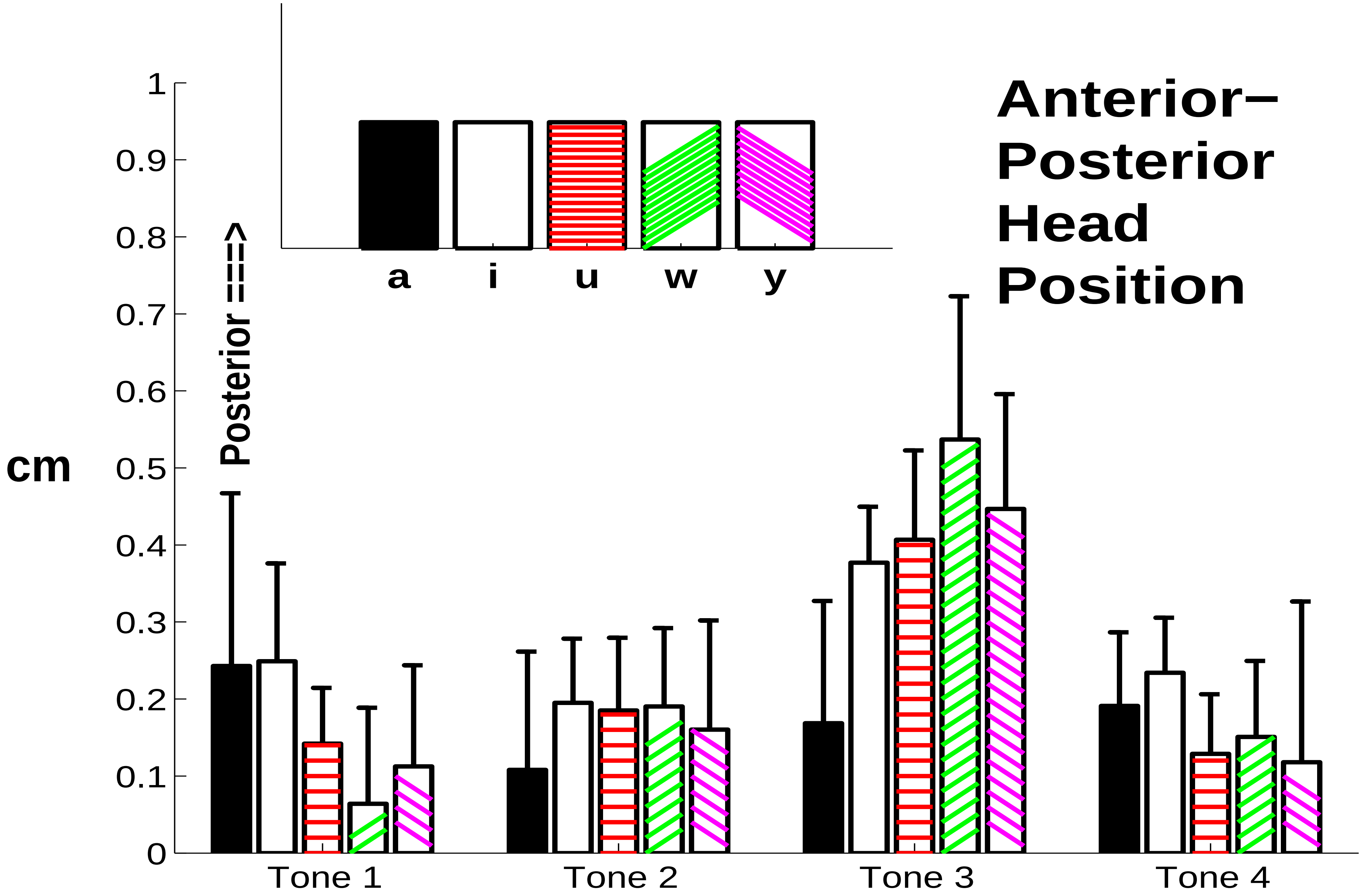
Preprocessing:

Raw position data converted to deviation from average position in each block of repetitions



Vertical Head Position
cm
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Tone 1
Tone 2
Tone 3
Tone 4
a
i
u
w
y

Anterior–Posterior Head Position

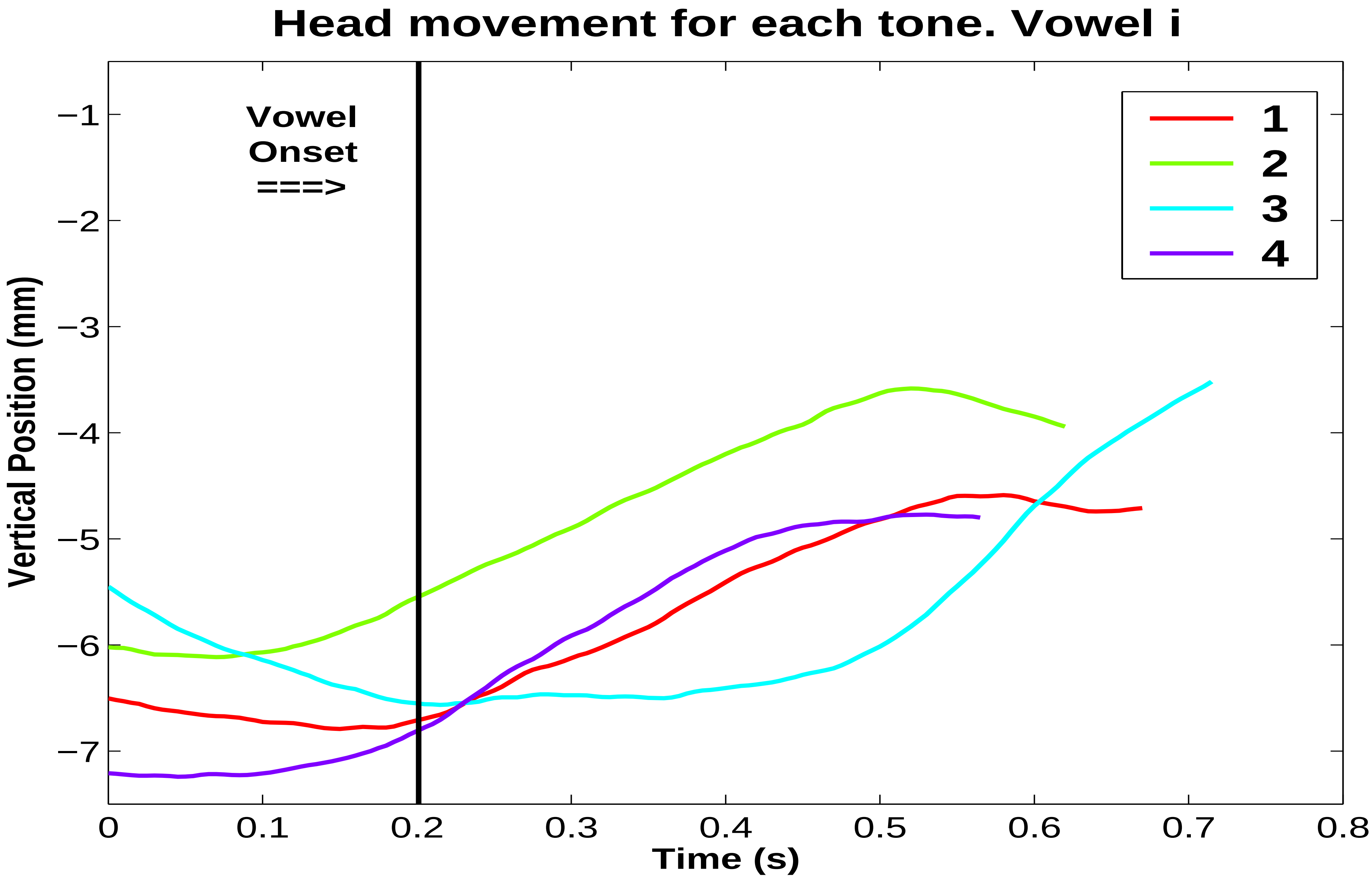


The clearest differences are between Tone 3 and Tone 2:

Head position lower and more retracted for Tone 3 than Tone 2 for all vowels except /a/.

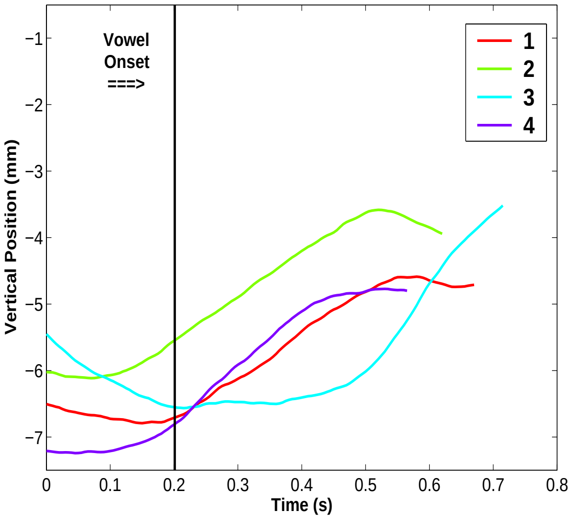
Magnitude of the differences is small, but clearly significant ($p<0.01$)

Does analysis at vowel mid-point miss anything?

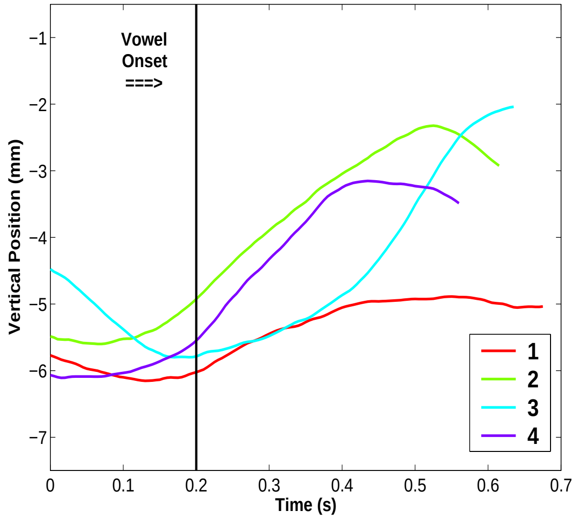


Head movement for each tone. Vowel i
Vowel
Onset
===>
1
2
3
4
−1
−2
−3
−4
−5
−6
−7
Vertical Position (mm)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Time (s)

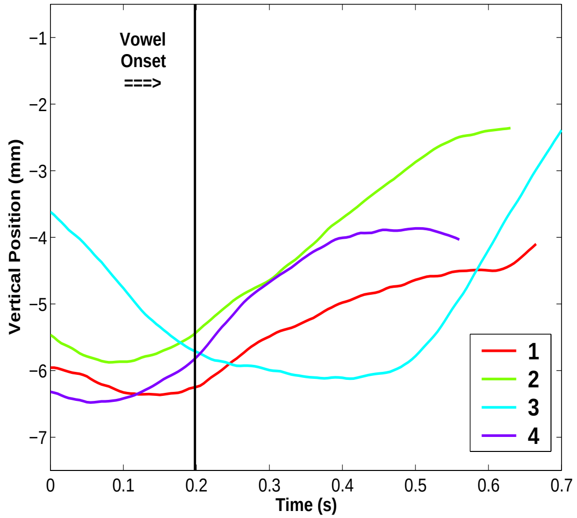
Vowel i



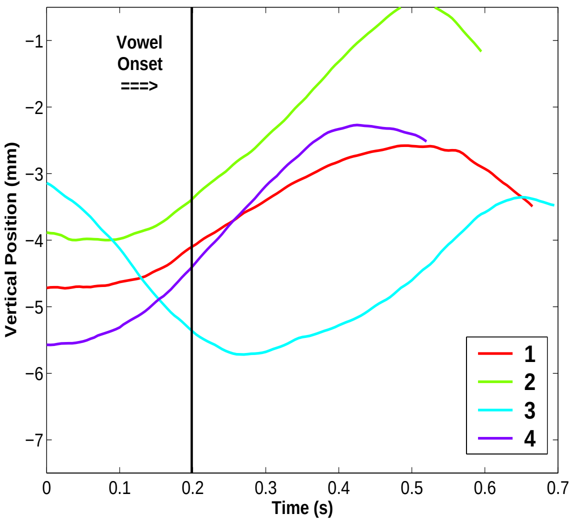
Vowel u



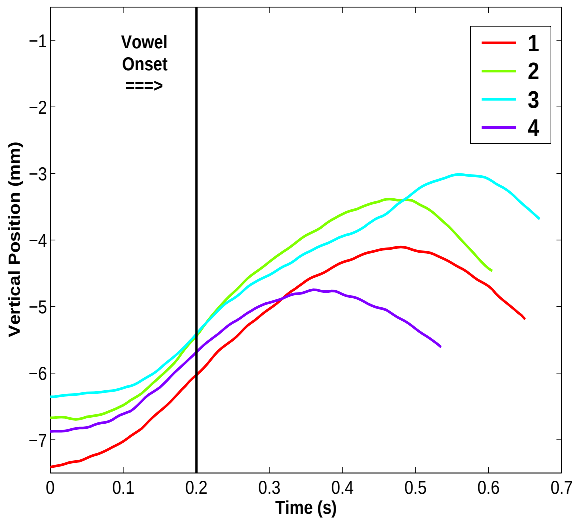
Vowel y



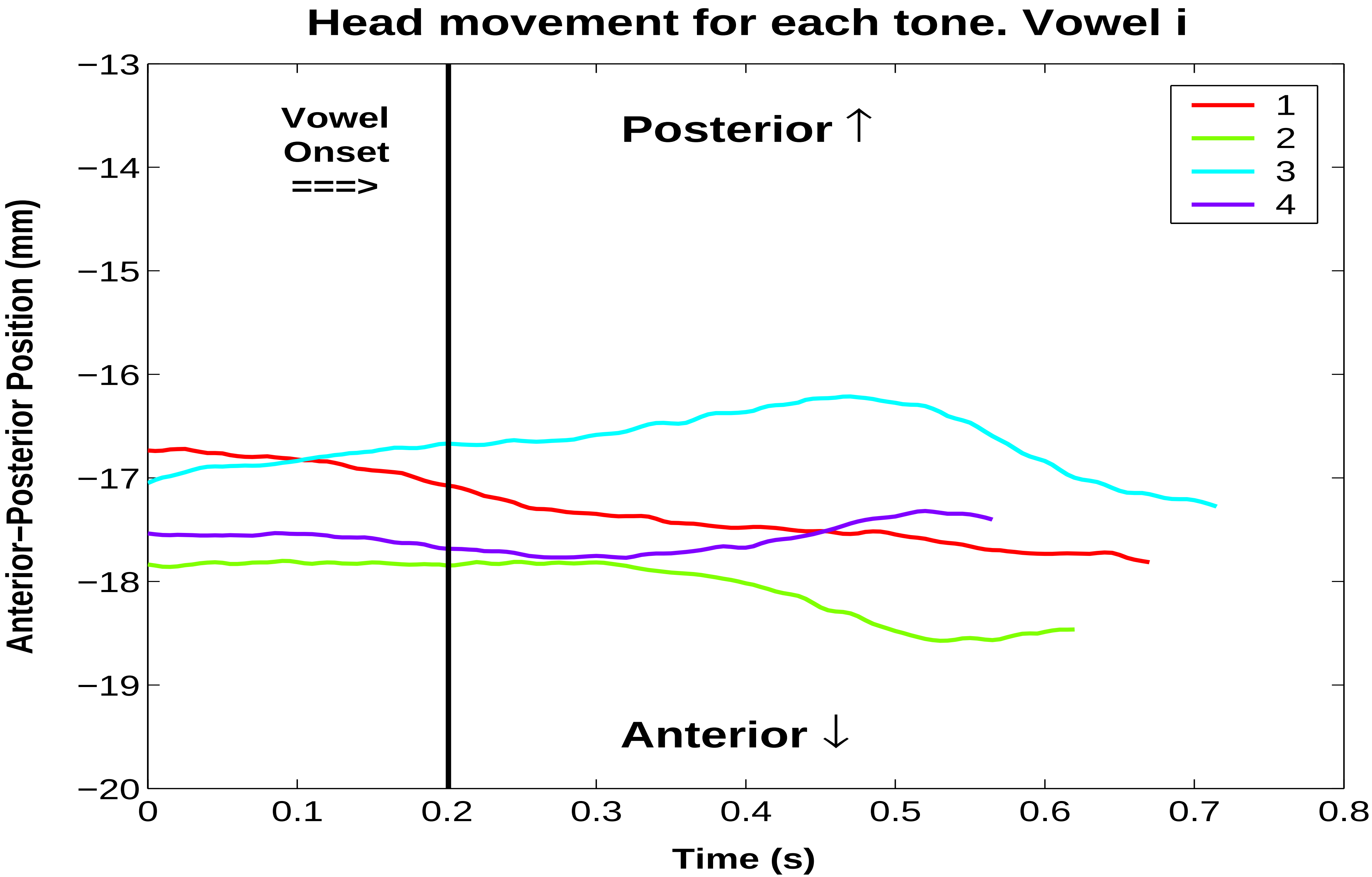
Vowel w

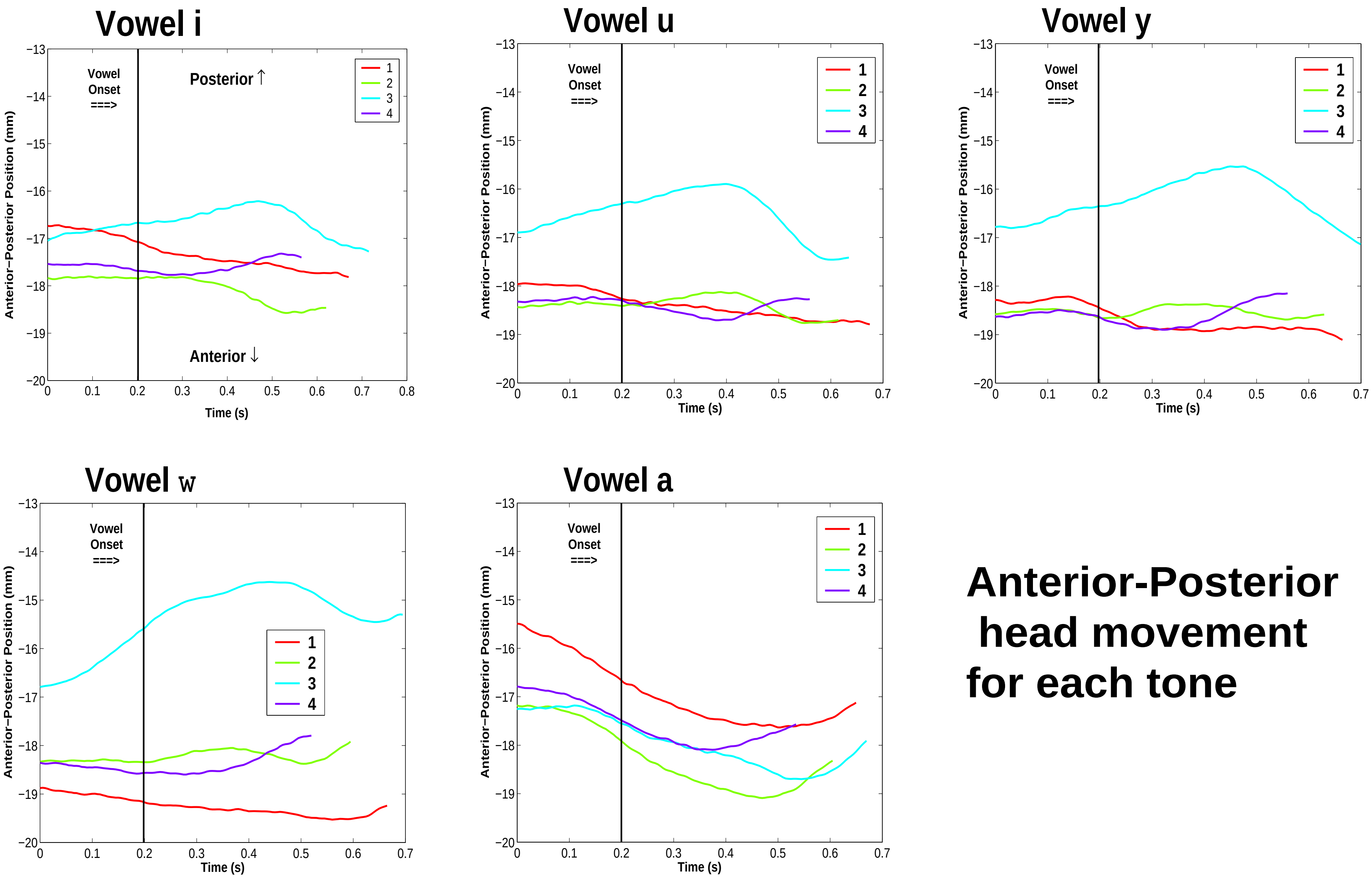


Vowel a



Vertical head movement for each tone


Head movement for each tone. Vowel i
Vowel
Onset
===>
Posterior ↑
Anterior ↓
1
2
3
4
Anterior–Posterior Position (mm)
Time (s)
−13
−14
−15
−16
−17
−18
−19
−20
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8



Anterior-Posterior head movement for each tone

Overall time-course of head movement is rather different for Tone 3

This is also reflected in velocity values:

Velocity of upward and forward movement higher for Tone 3

Velocity patterns are attractive because they may be communicatively more robust than subtle differences in position (cf. Keating et al., 2003).

Head-movement certainly does not simply mimic the F0-contour.

Click on red box for animation of head movement for tones 2 and 3

Results Topic 3
Intrinsic Pitch

First a digression into German

Where does Mandarin come in?

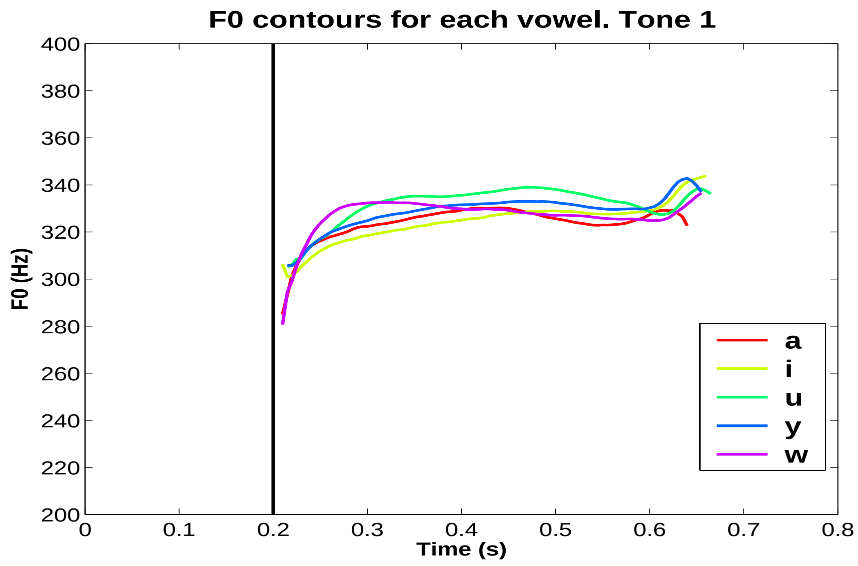
The German data indicated that intrinsic pitch could not be very directly related to jaw position.

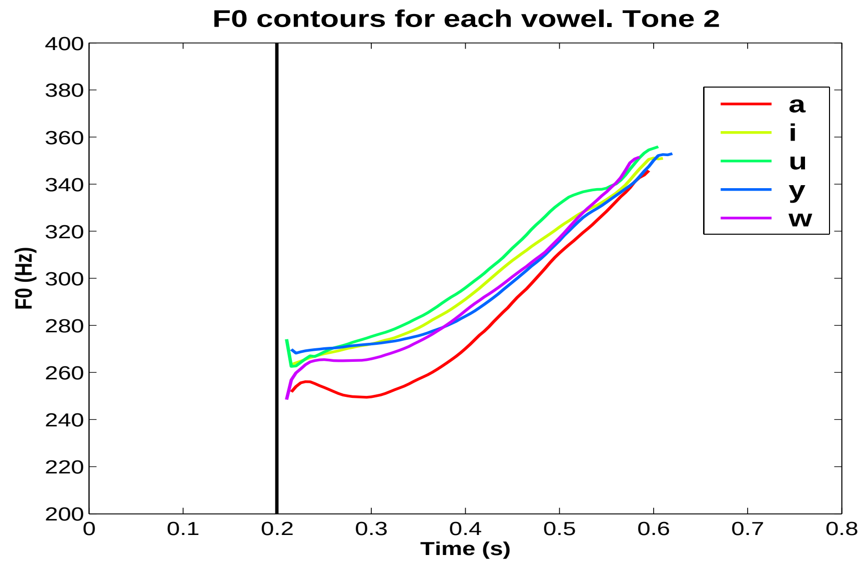
However, recent findings by Torng et al. (2001) for Mandarin suggested a reappraisal of the role of the jaw.

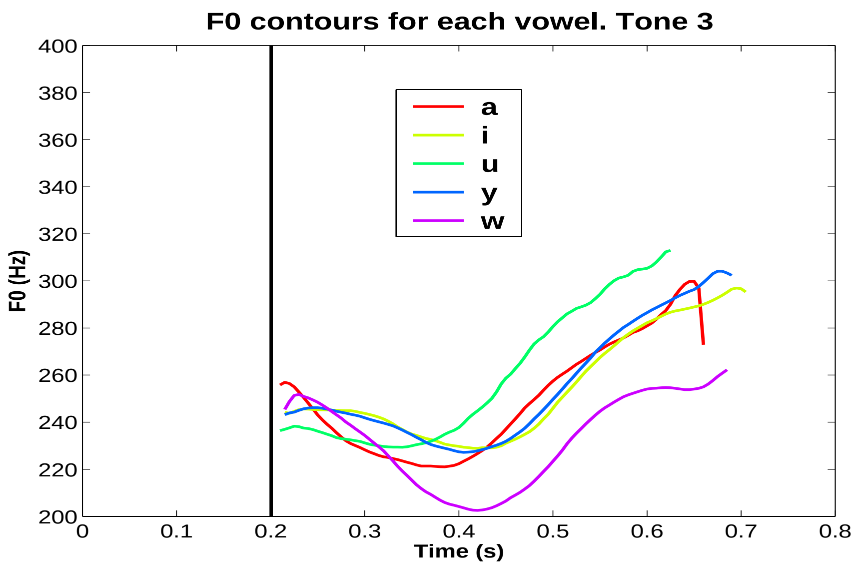
Their findings:

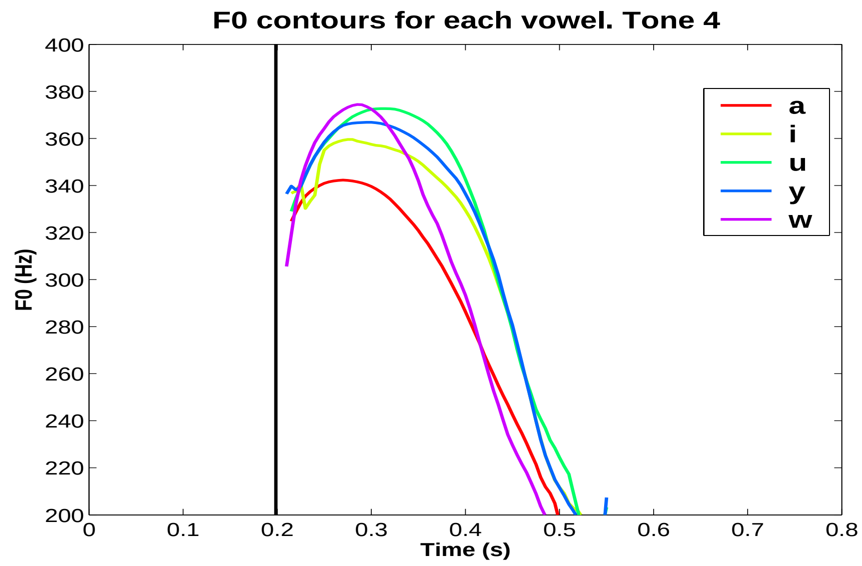
- Suprisingly low F0 for /i/, coupled with low jaw position
- Suprisingly high F0 for /o/, coupled with high jaw position

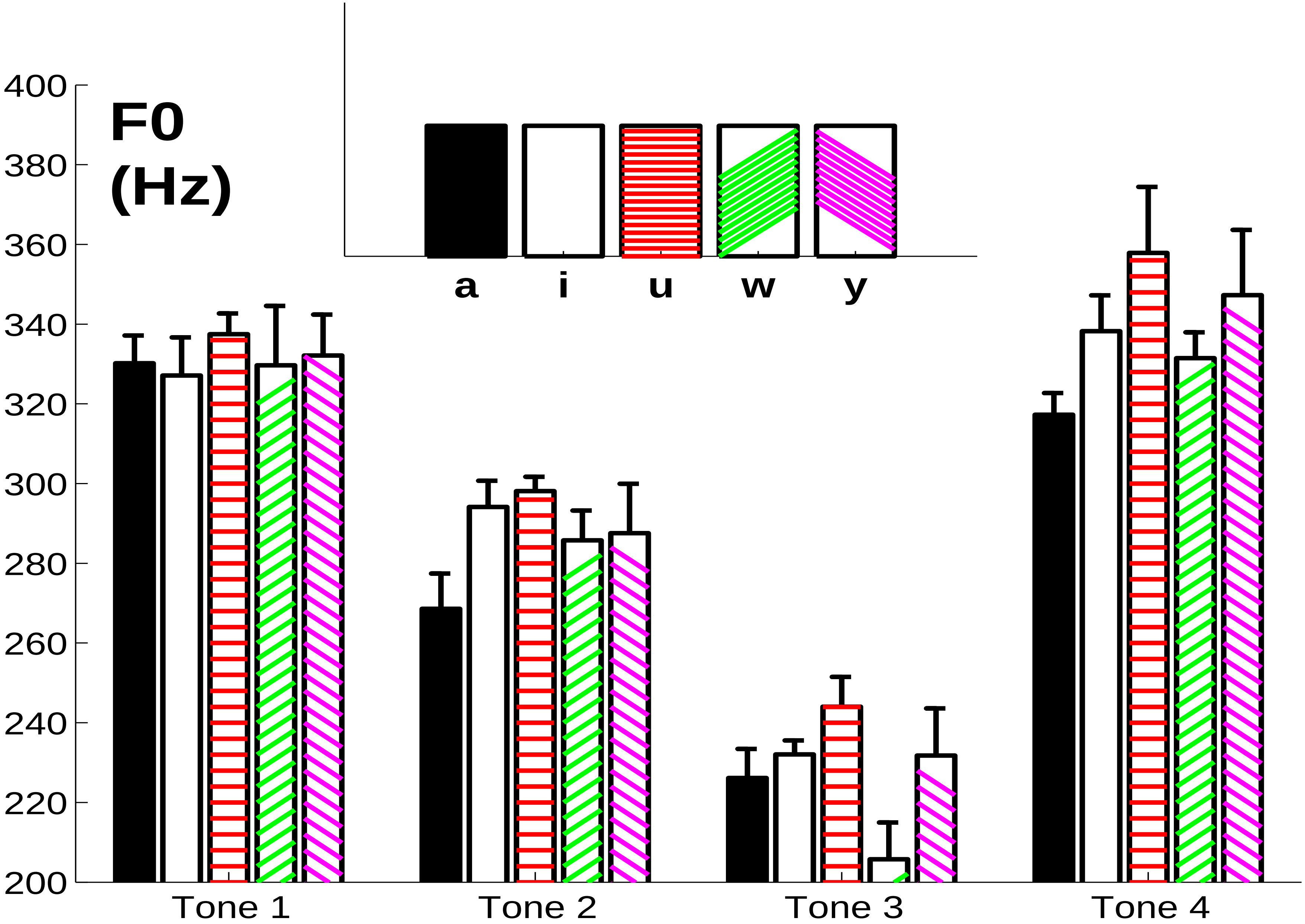
Can these findings be replicated?

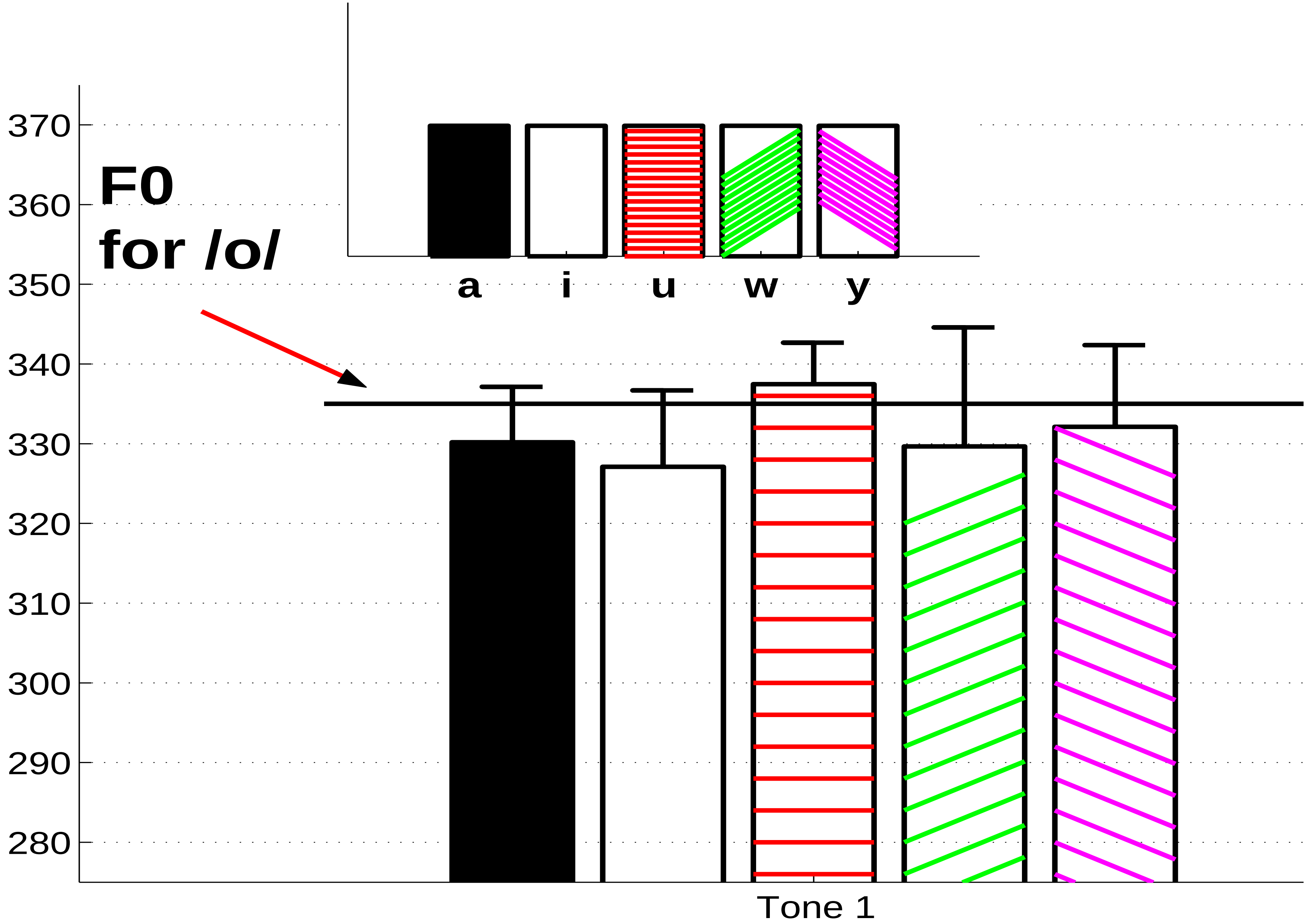

F0 contours for each vowel. Tone 1
F0 (Hz)
Time (s)
a
i
u
y
w


F0 contours for each vowel. Tone 2
F0 (Hz)
Time (s)
a
i
u
y
w


F0 contours for each vowel. Tone 3
F0 (Hz)
Time (s)
a
i
u
y
w


F0 contours for each vowel. Tone 4
F0 (Hz)
Time (s)
a
i
u
y
w

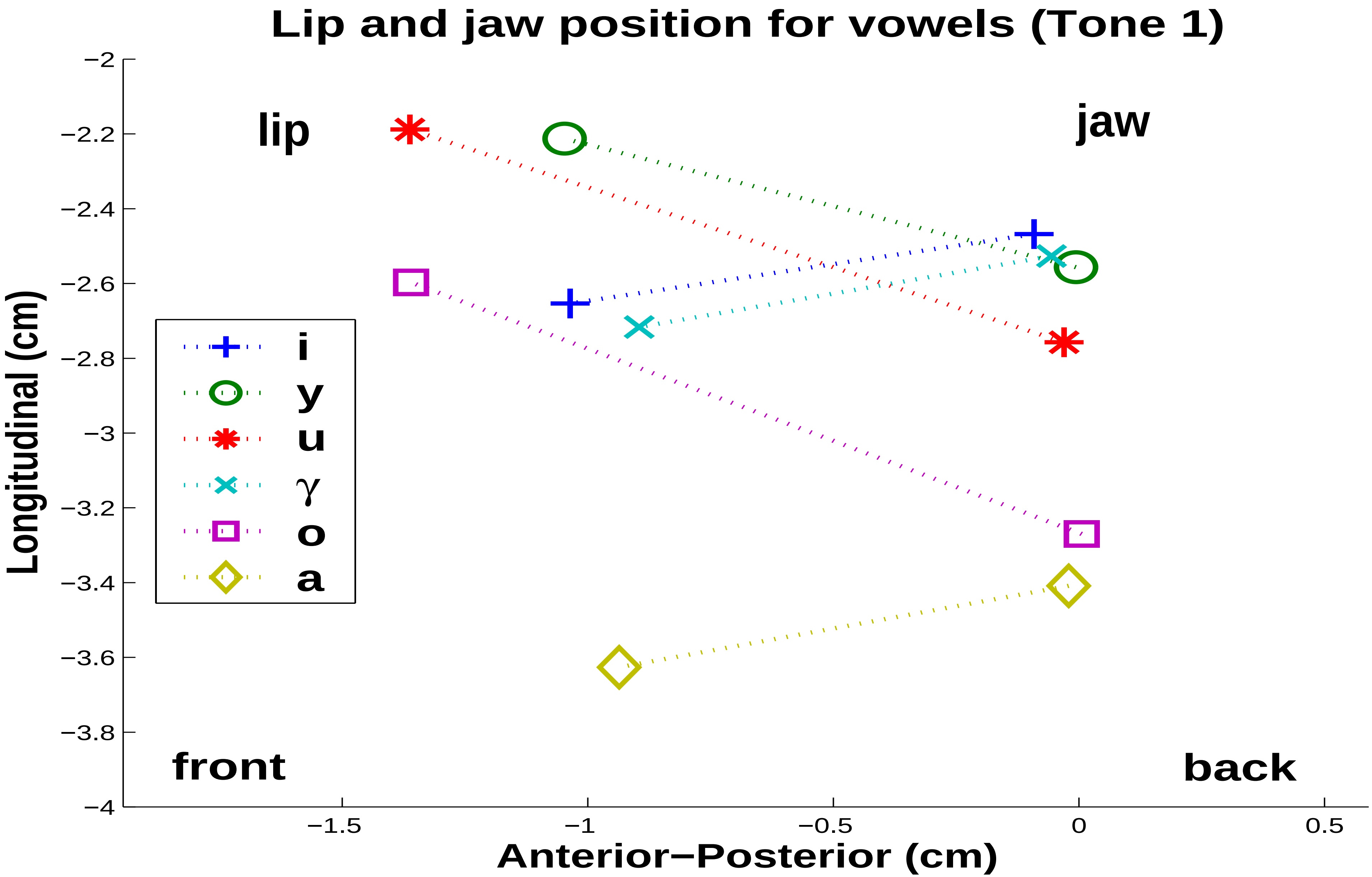

F0
(Hz)
400
380
360
340
320
300
280
260
240
220
200
a
i
u
w
y
Tone 1
Tone 2
Tone 3
Tone 4


370
360
350
340
330
320
310
300
290
280
F0
for /o/
a
i
u
w
y
Tone 1

No remarkable features of fundamental frequency for /i/ or /o/

(Material admittedly not ideal for studying intrinsic pitch)

What about the jaw?



Lip and jaw position for vowels (Tone 1)
lip
jaw
front
back
Longitudinal (cm)
Anterior−Posterior (cm)
−2
−2.2
−2.4
−2.6
−2.8
−3
−3.2
−3.4
−3.6
−3.8
−4
−1.5
−1
−0.5
0
0.5
i
y
u
γ
o
a

Again, no remarkable features:

jaw neither very low for /i/ nor high for /o/

Thus no confirmation for strong influence of jaw on intrinsic pitch

But closer study could be worthwhile:

Unlike European languages, Mandarin /i/ has no front vowel neighbours.

So greater scope for coarticulatory effects on jaw position in /i/?

General Conclusions

Articulatory differences between tones were found:
    Mainly involved retraction and lowering of tongue for back vowels for Tone 3

    Related to larynx lowering for low pitch?

    But effects do not generalize:
        Weak for front vowels (or in a different direction)

Head movement differences also found:
    Again, mainly involved Tone 3
    Audiovisual gestural enhancement of clear speech?