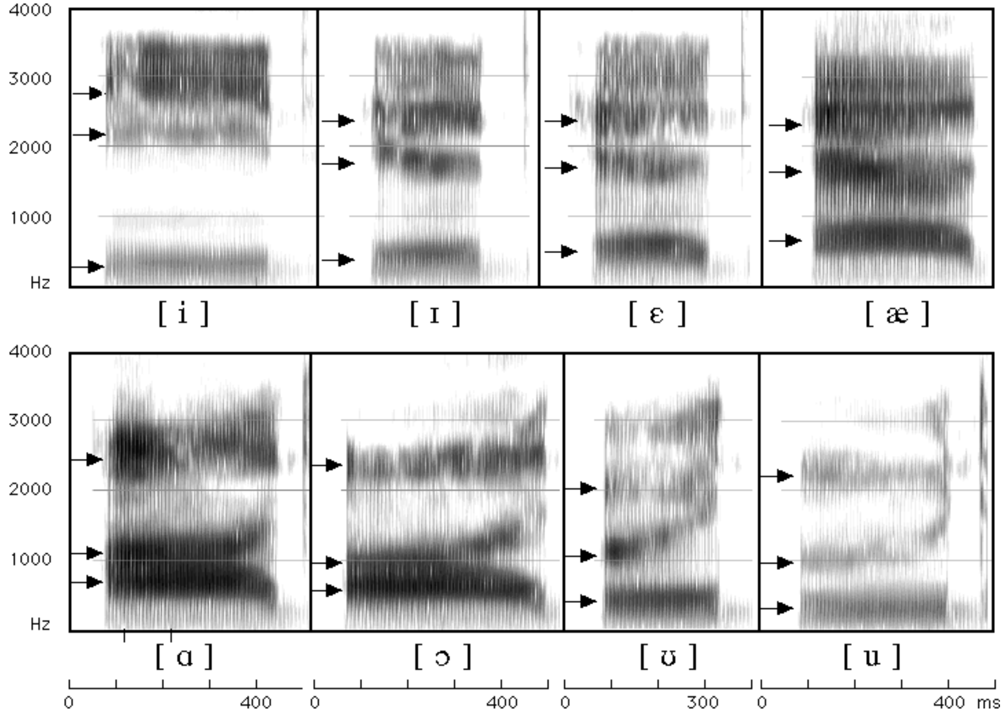
24.963

Linguistic Phonetics

The acoustics of vowels



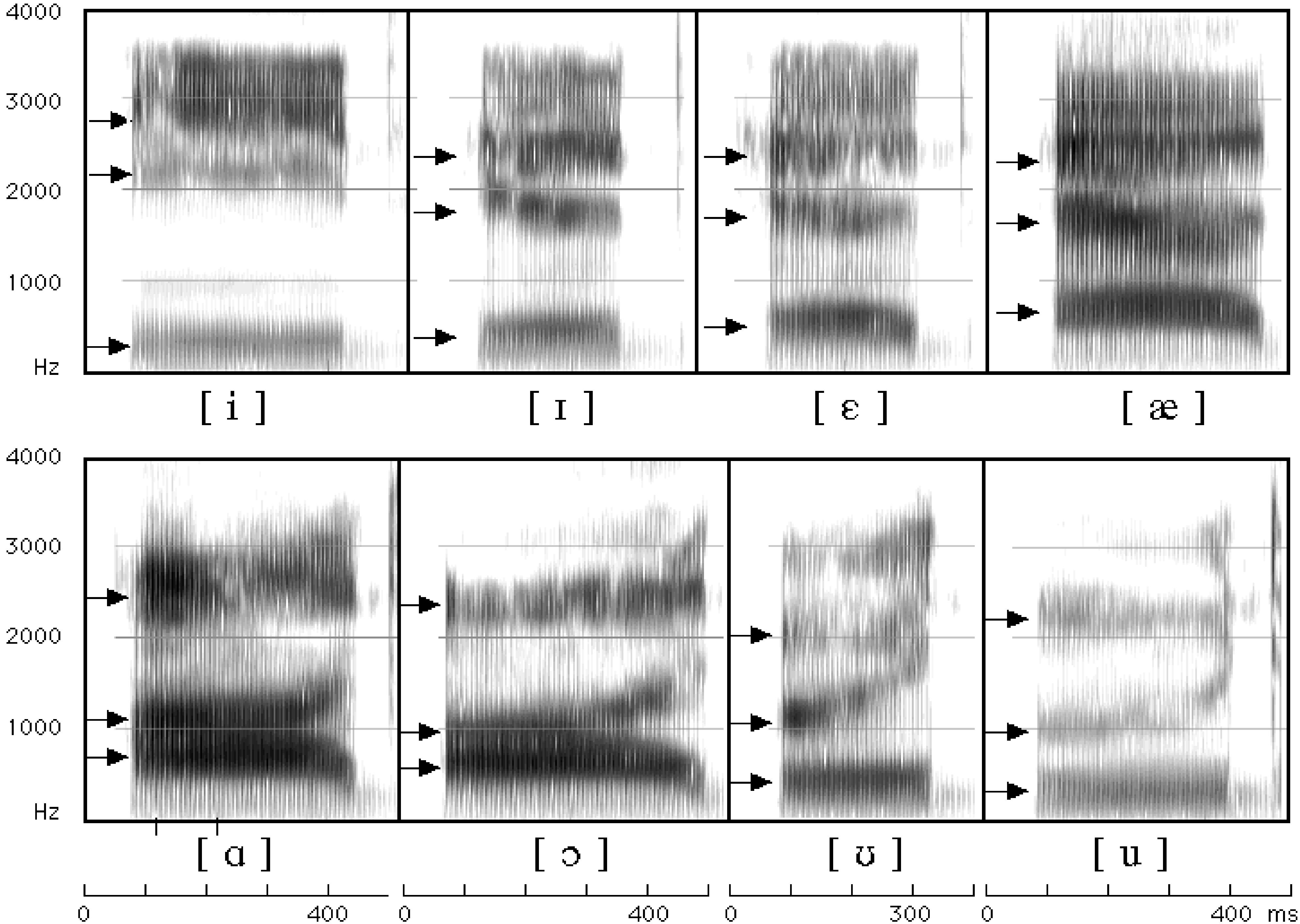
- No class on Tuesday 10/13 (Tuesday is a Monday)

Readings:

- Johnson chapter 6 (for this week)
- Liljencrants & Lindblom (1972) (for next week)

Assignment:

- Modeling lip-rounding, due 10/15



4000
3000
2000
1000
Hz
[i]
[ɪ]
[ɛ]
[æ]
4000
3000
2000
1000
Hz
[ɑ]
[ɔ]
[ʊ]
[u]
0
400
0
400
0
300
0
400 ms

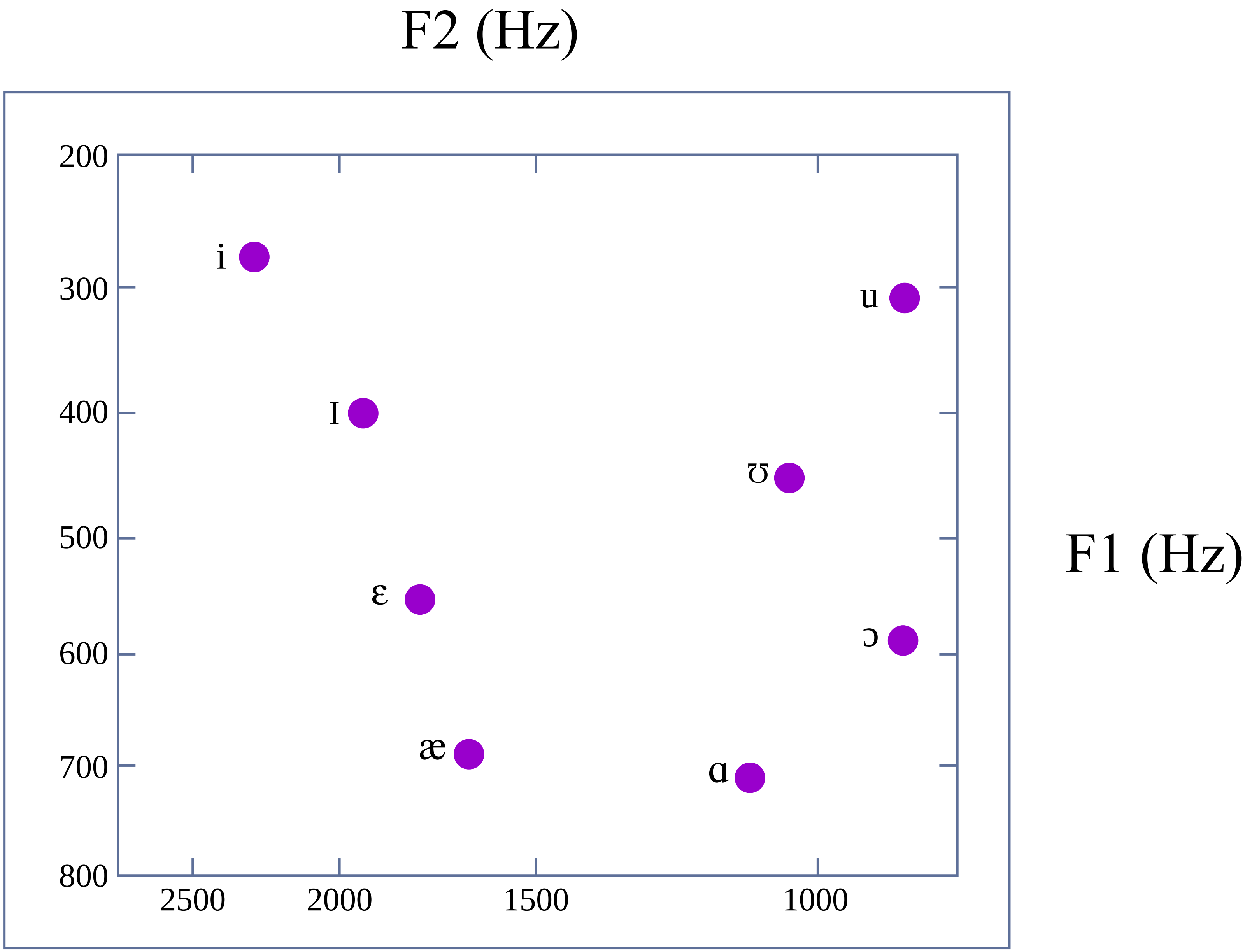


Image by MIT OCW.
Adapted from Peter Ladefoged. *A Course in Phonetics*. 5th ed. Berlin, Germany: Heinle, 2005.
ISBN: 9781413006889. Available at: https://www.phonetics.ucla.edu/course/contents.html.

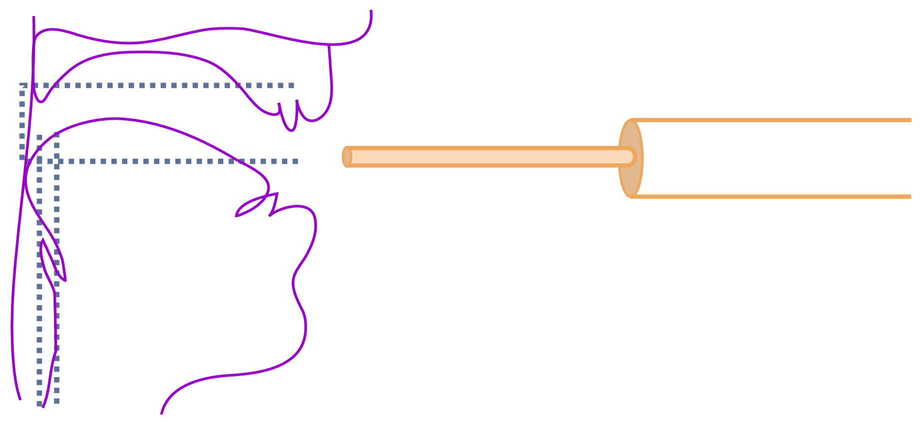
The Acoustics of Vowels

Source-Filter models:

- Source: voicing (usually)
- Filter characteristics can be given a basic but useful analysis using simple tube models.
- Tube models can be supplemented by perturbation theory for approximate analysis of the effects of wide constrictions.

Low vowels [ɑ, a, æ]

- Pharyngeal constriction



The shape of the vocal tract in the vowel [ɑ] as in father schematized as two tubes.

Image by MIT OCW.

- Since the back tube is much narrower than the front tube, each can reasonably be approximated by a tube closed at one end and open at the other.
- The resonances of the combined tubes deviate from the values we would calculate for these configurations in isolation because the resonators are acoustically coupled.
- The degree of coupling depends on the difference in cross-sectional areas.

Low vowels [ɑ, a, æ]

$$F_n = \frac{(2n-1)c}{4L}$$

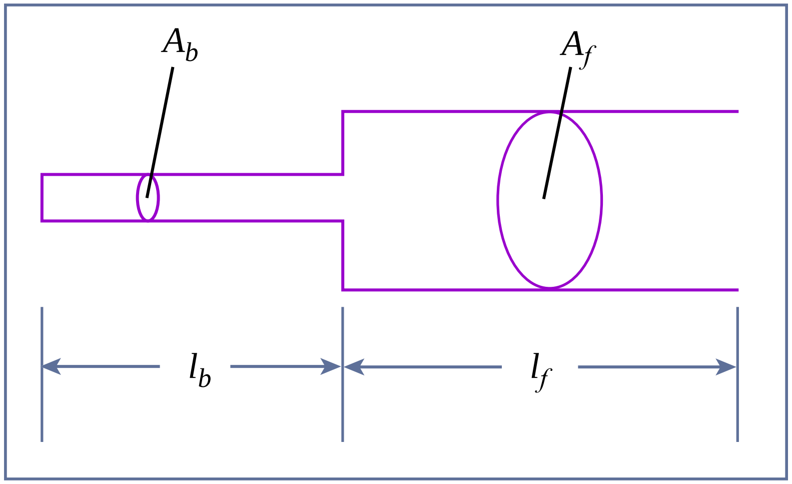


Image by MIT OCW.
Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics.* Malden, MA: Blackwell Publishers, 1997. ISBN: 9780631188483.

nomogram

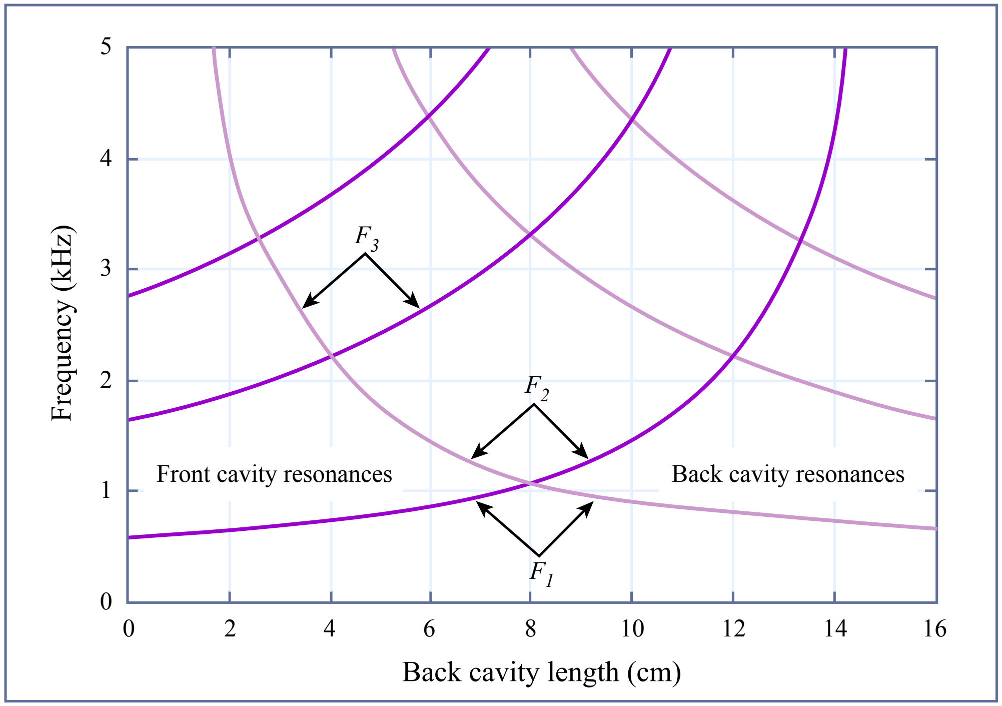


Image by MIT OCW.
Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997. ISBN: 9780631188483.

Non-low vowels (e.g. [i, e])

- Short constriction in the mouth

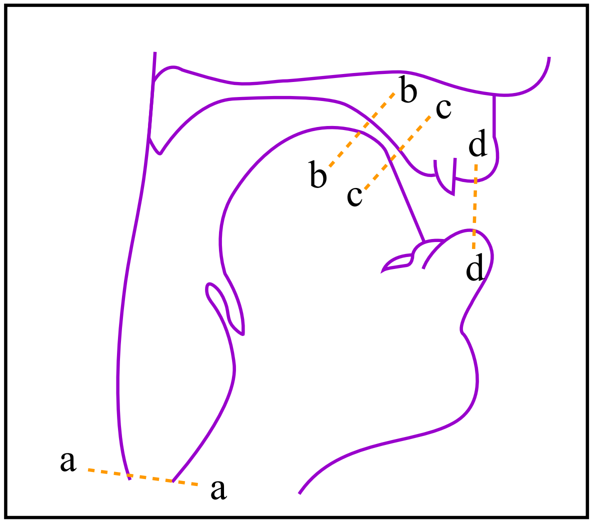


Image by MIT OCW.
Adapted from Ladefoged, Peter. *Elements of Acoustic Phonetics*. 2nd ed. Chicago, IL: University of Chicago Press, 1996.

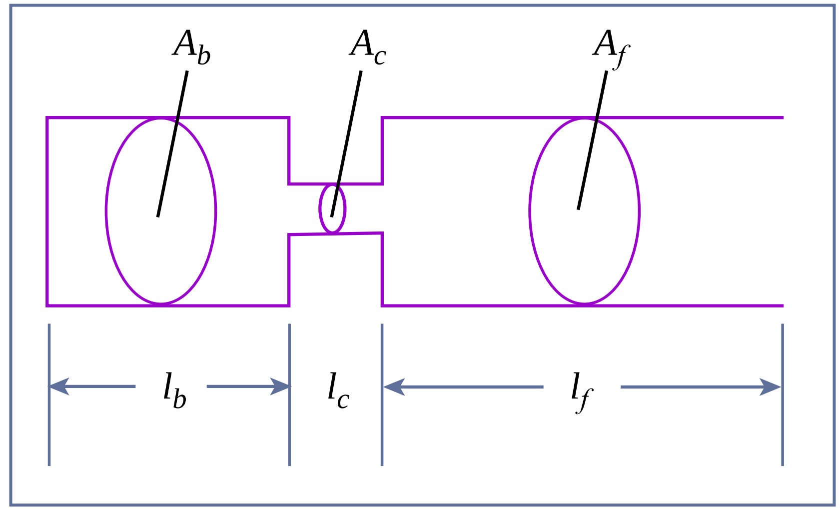


Image by MIT OCW.
Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997. ISBN: 9780631188483.

- The back cavity can be approximated by a tube closed at both ends.
- The front cavity is approximated by a tube closed at one end.
- Neglects coupling. The degree of coupling depends on the cross-sectional area of the constriction.
- How do we account for the F1 of high vowels?

$$F_n = \frac{nc}{2L}$$

$$F_n = \frac{(2n-1)c}{4L}$$

Helmholtz resonators

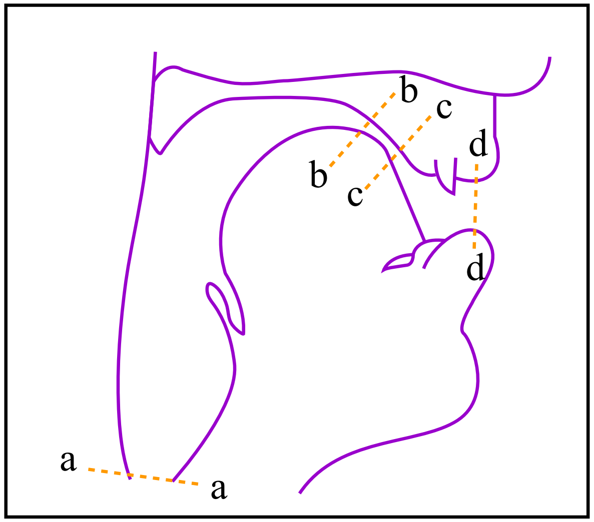


Image by MIT OCW.
Adapted from Ladefoged, Peter. *Elements of Acoustic Phonetics*. 2nd ed. Chicago, IL: University of Chicago Press, 1996.

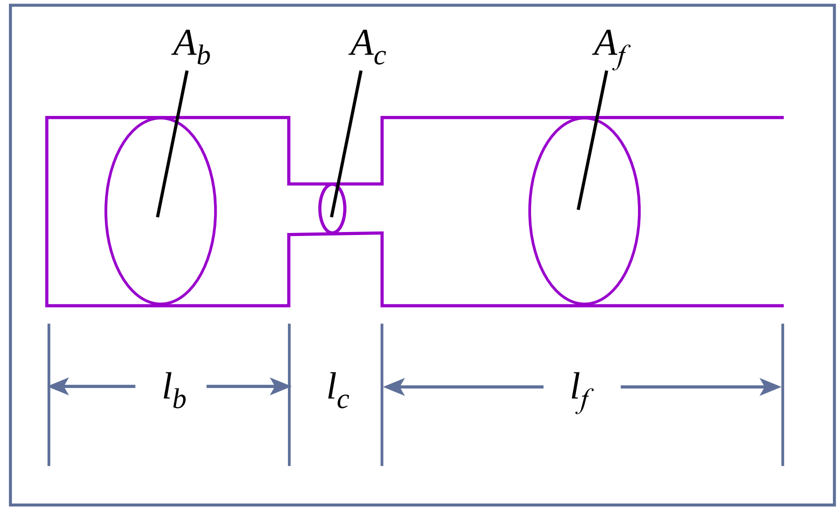


Image by MIT OCW.
Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997. ISBN: 9780631188483.

- The back cavity and the constriction together form a resonant system called a Helmholtz resonator.
- If the length of the constriction is short, the air in it vibrates as a mass on the 'spring' formed by the air in the back cavity.
- Resonant frequency, $f = \dfrac{c}{2\pi}\sqrt{\dfrac{A_c}{Vl_c}} = \dfrac{c}{2\pi}\sqrt{\dfrac{A_c}{A_b l_b l_c}}$

Non-low vowels - nomogram

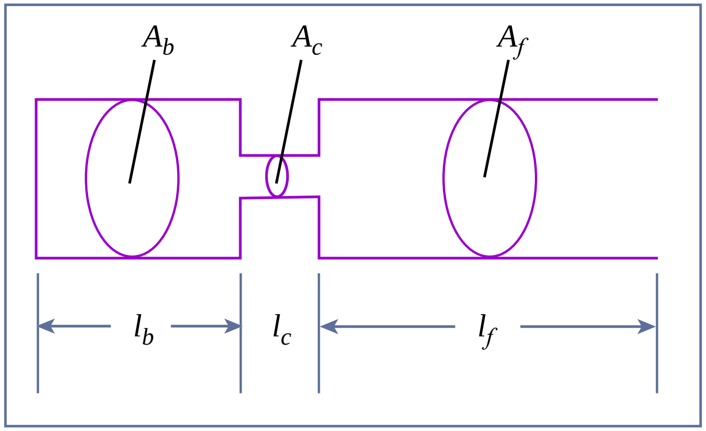


Image by MIT OCW.
Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997. ISBN: 9780631188483.

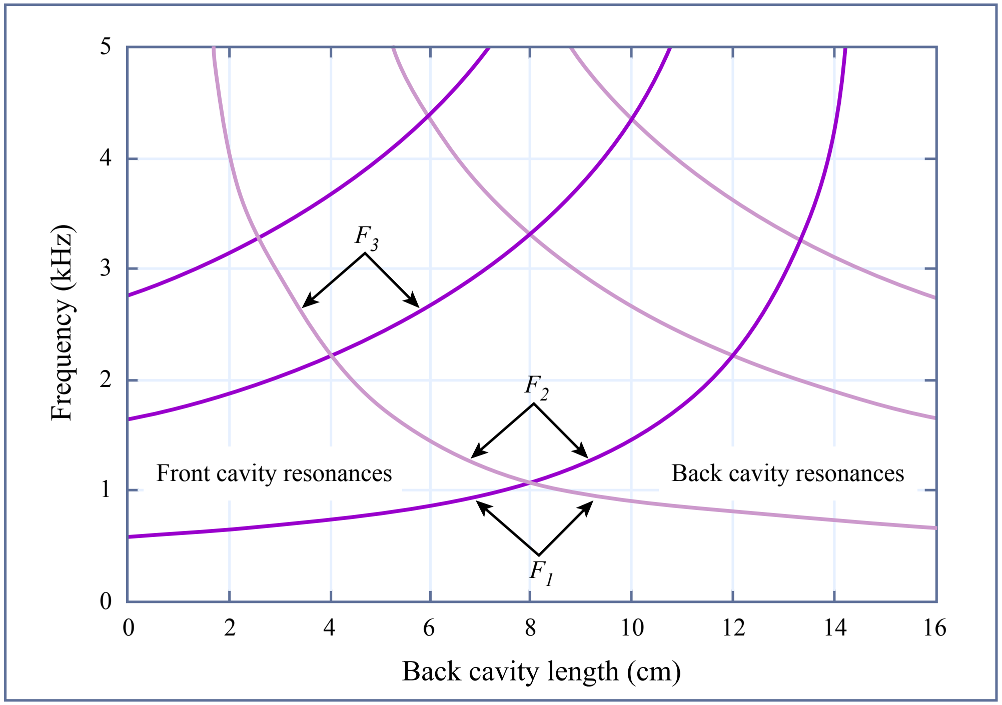


Image by MIT OCW.
Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997. ISBN: 9780631188483.

front cavity $$F_n = \frac{(2n-1)c}{4L}$$

back cavity $$F_n = \frac{nc}{2L}$$

back cavity + constriction

$$f = \frac{c}{2\pi}\sqrt{\frac{A_c}{A_b l_b l_c}}$$

- How would you model a mid vowel?

Perturbation Theory (Chiba and Kajiyama 1941)

- Constriction near a point of maximum velocity (V_n) lowers the associated formant frequency.
- Constriction near a point of maximum pressure raises the associated formant frequency.

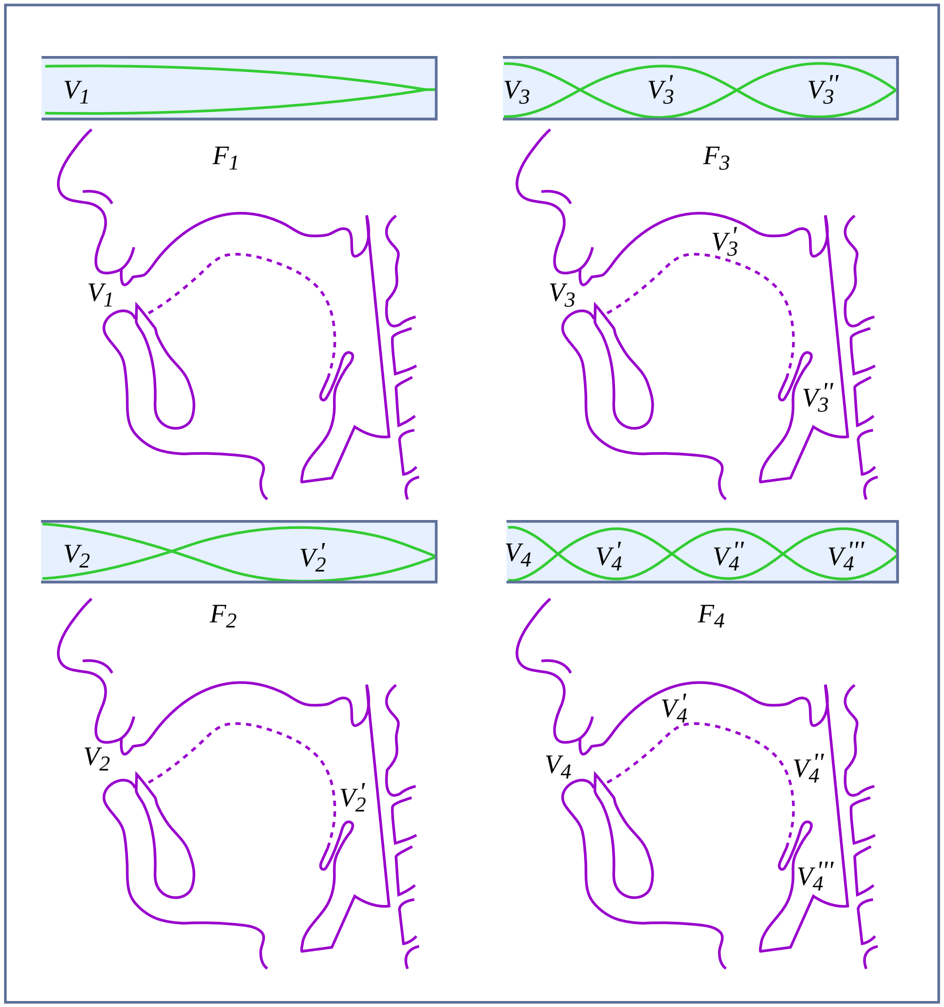


Image by MIT OCW.
Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997. Based on Chiba and Kajiyama 1941.

Perturbation Theory (Chiba and Kajiyama 1941)

- What is the effect of a pharyngeal constriction?
- Does this correspond to the tube model above?
- How do you raise F2 maximally?

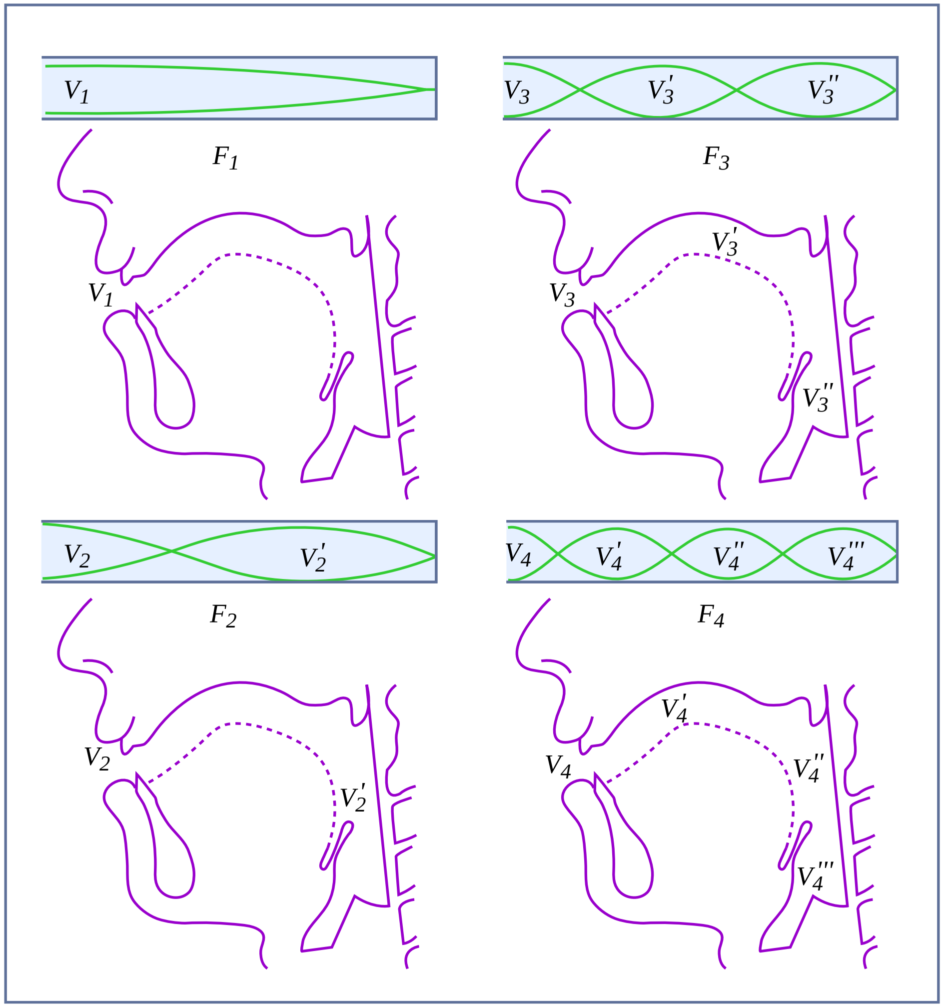


Image by MIT OCW.
Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997. Based on Chiba and Kajiyama 1941.

Perturbation Theory vs. two-tube models

- Our simple tube models ignore acoustic coupling and are therefore most valid where constrictions are narrow.
- Perturbation theory accounts for the effects of small perturbations of a uniform tube, and thus is most accurate for open constrictions.
- Mrayati et al (1988): perturbation theory is generally valid for constrictions greater than 0.8 cm^2, and two-tube models are valid for a constriction of 0.05 cm^2 or less, with a transitional region in between.

- Mrayati, Carré & Guérin (1988). Distinctive regions and modes. *Speech Communication* 7, 257-286.

American English [ɹ]

- American English [ɹ] is characterized by an exceptionally low F3 (<2000 Hz).

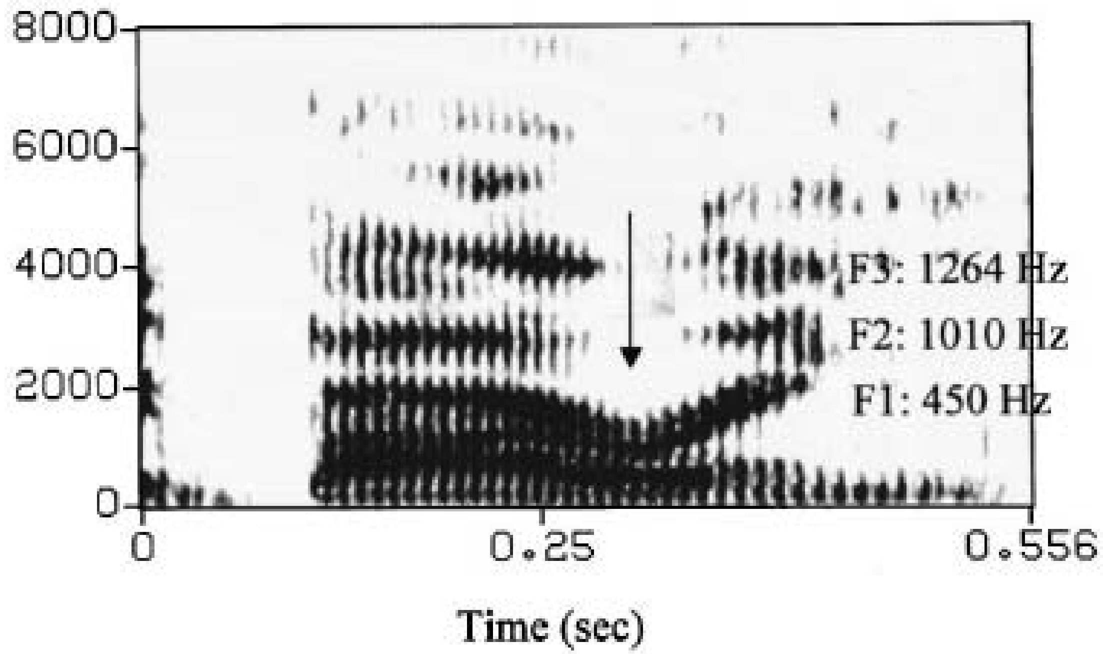
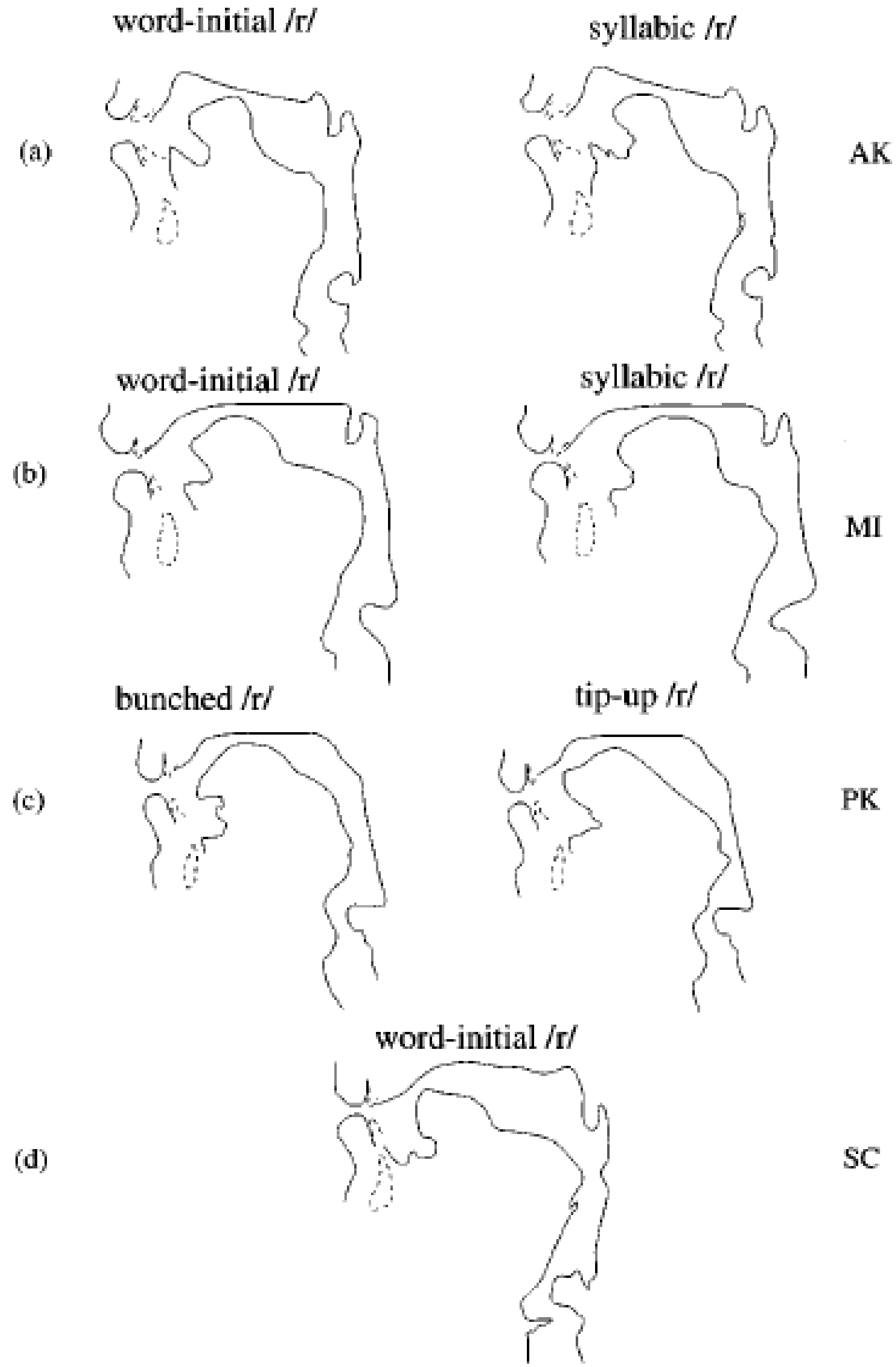


FIG. 1. Spectrogram of the word "barring" spoken by a male speaker.

Reproduced from Espy-Wilson, Carol Y., Suzanne E. Boyce, Michel Jackson, Shrikanth Narayanan, and Abeer Alwan. "Acoustic modeling of American English/r." The Journal of the Acoustical Society of America 108, no. 1 (2000): 343-356. doi: https://doi.org/10.1121/1.429469, with the permission of the Acoustical Society of America.

- American English [ɹ] is produced in a variety of ways across speakers and contexts (Alwan et al 1997 *JASA*, Westbury et al 1998, *Speech Comm*.).
- A basic distinction that is often made: ‘bunched’ vs. ‘retroflex’.
  - But there appears to be a continuum of variants.



Reproduced from Narayanan, Shrikanth S., Abeer A. Alwan, and Katherine Haker. "Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals." The Journal of the Acoustical Society of America 101, no. 2 (1997): 1064-1077. doi: https://doi.org/10.1121/1.418030, with the permission of the Acoustical Society of America.

Perturbation Theory (Chiba and Kajiyama 1941)

A nice story about Am. Eng. [ɹ]

- Three constriction: labial (lip protrusion/rounding), palatal (bunching or retroflexion), and pharyngeal.
- All 3 are near velocity maxima for F3, hence very low F3.
- But Espy-Wilson et al (2000) argue actual constrictions are in the wrong place

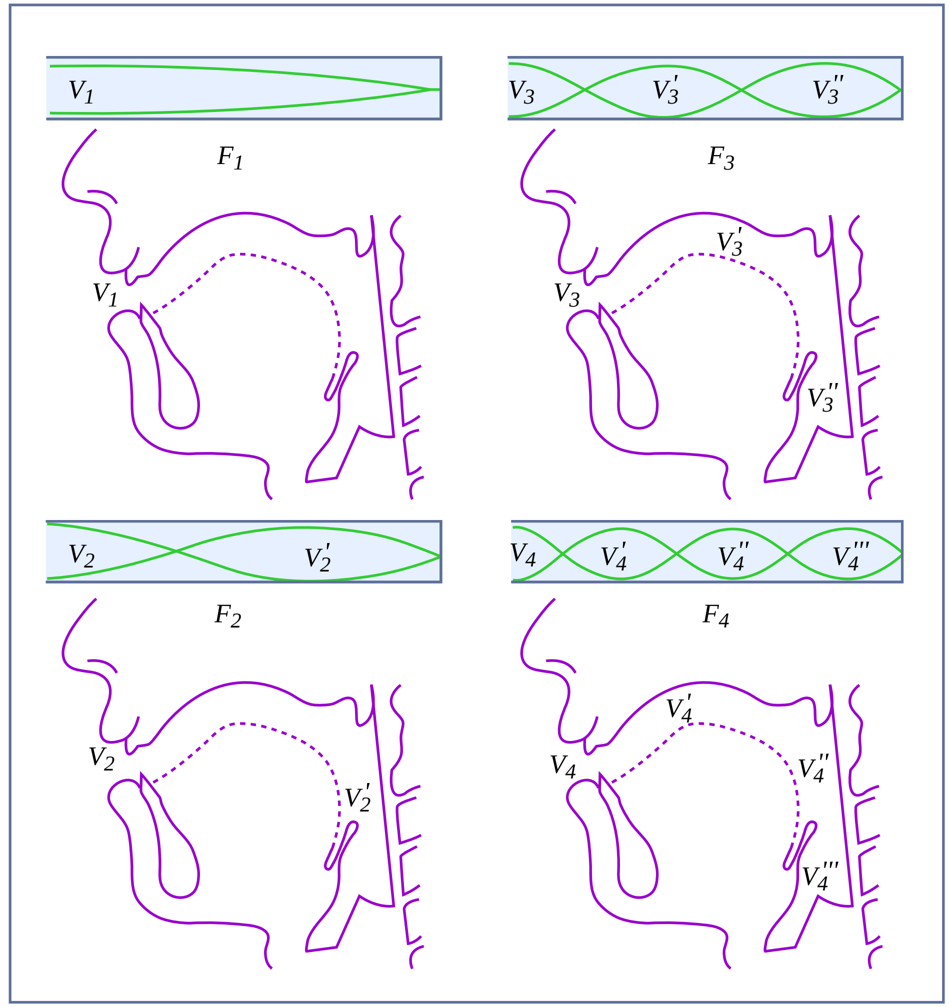
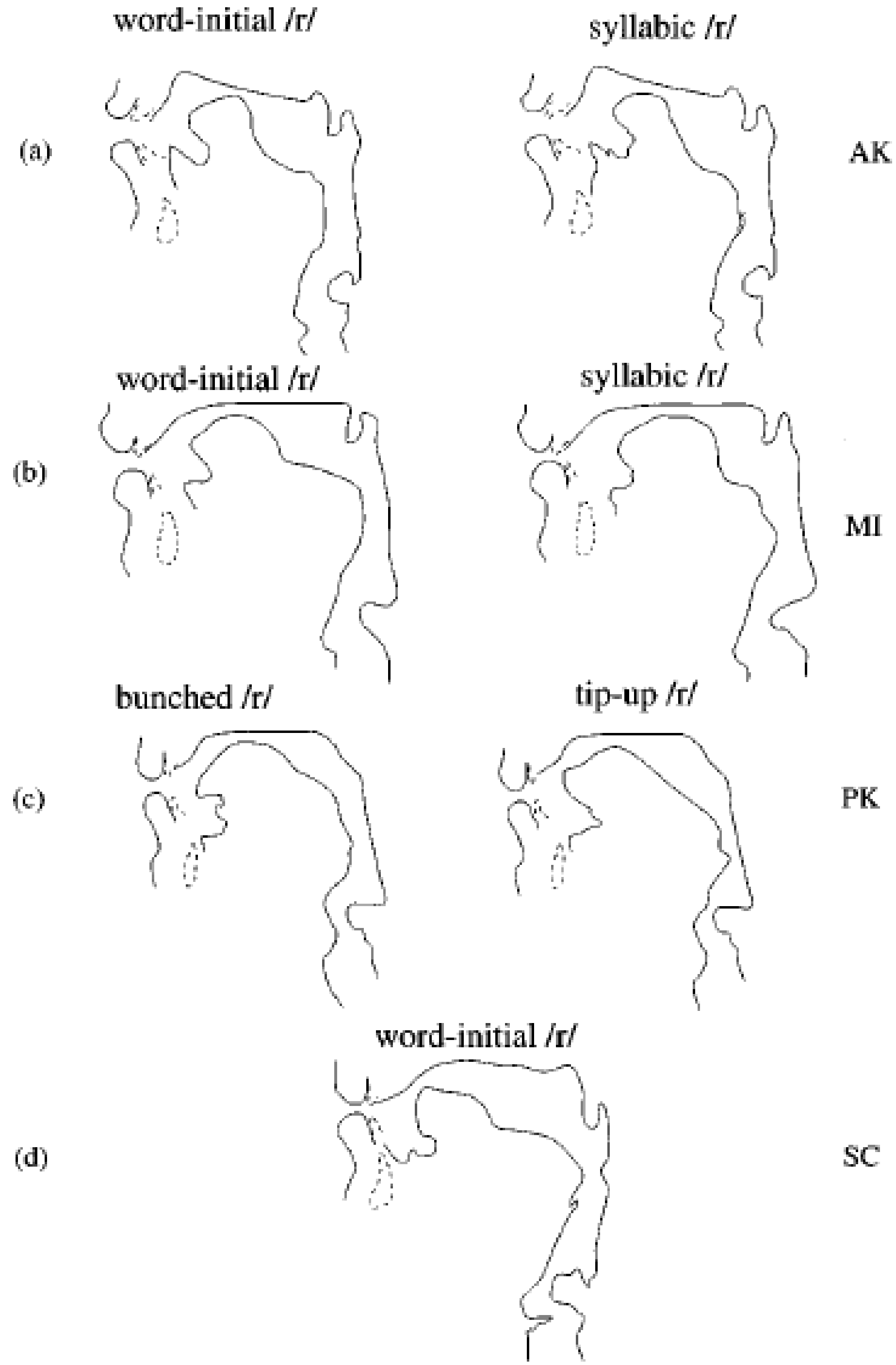


Image by MIT OCW.
Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997. Based on Chiba and Kajiyama 1941.

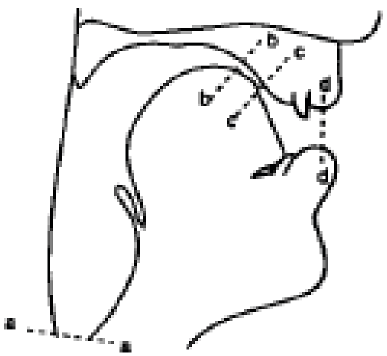
Espy-Wilson et al (2000) argue from MRI data that:

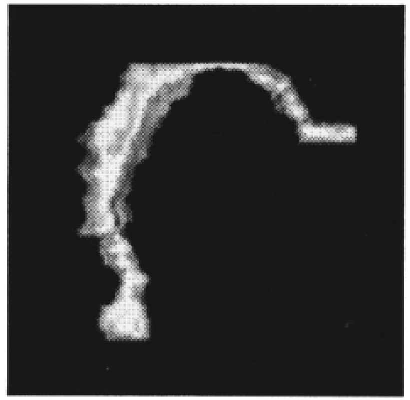
- Actual constrictions are in the wrong places, e.g. pharyngeal constriction is too high.
- Constrictions are too narrow to apply perturbation theory.
- Argue that F3 is a front cavity resonance.
- Low due to length (bunched) or sub-lingual cavity (retro) + lip constriction. (How long?)
- Or: lip constriction is narrow enough for the front cavity to form a Helmholtz resonator.



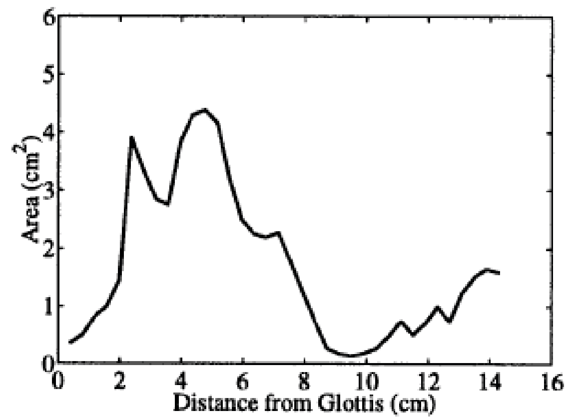
Reproduced from Narayanan, Shrikanth S., Abeer A. Alwan, and Katherine Haker. "Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals." The Journal of the Acoustical Society of America 101, no. 2 (1997): 1064-1077. doi: https://doi.org/10.1121/1.418030, with the permission of the Acoustical Society of America.

Constriction locations and area functions for [i] vowels



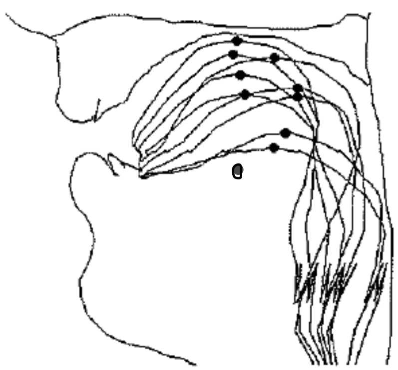


i_{mri}



Story et al (1998), MRI

Source: Story, Brad H., Ingo R. Titze, and Eric A. Hoffman. "Vocal tract area functions for an adult female speaker based on volumetric imaging." The Journal of the Acoustical Society of America 104, no. 1 (1998): 471-487.



Ladefoged & Maddieson (1996) – mean tongue positions

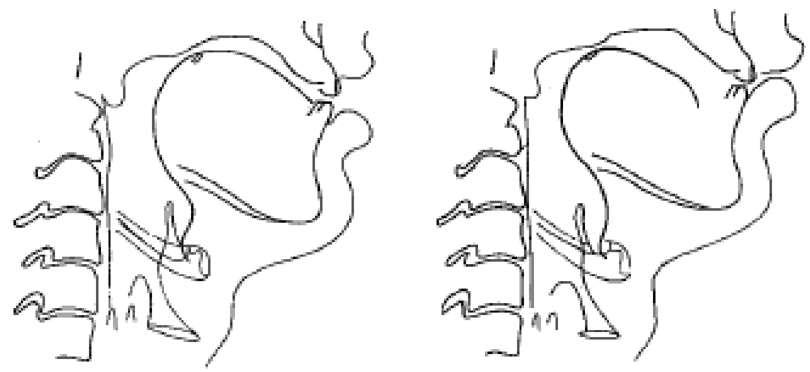
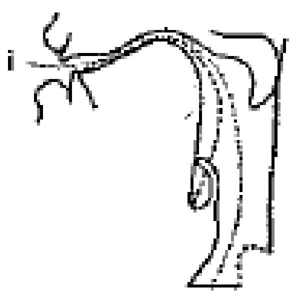
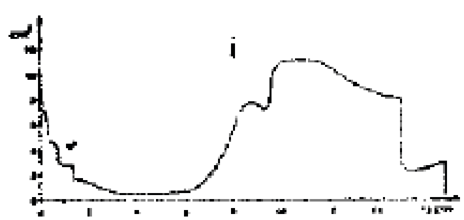


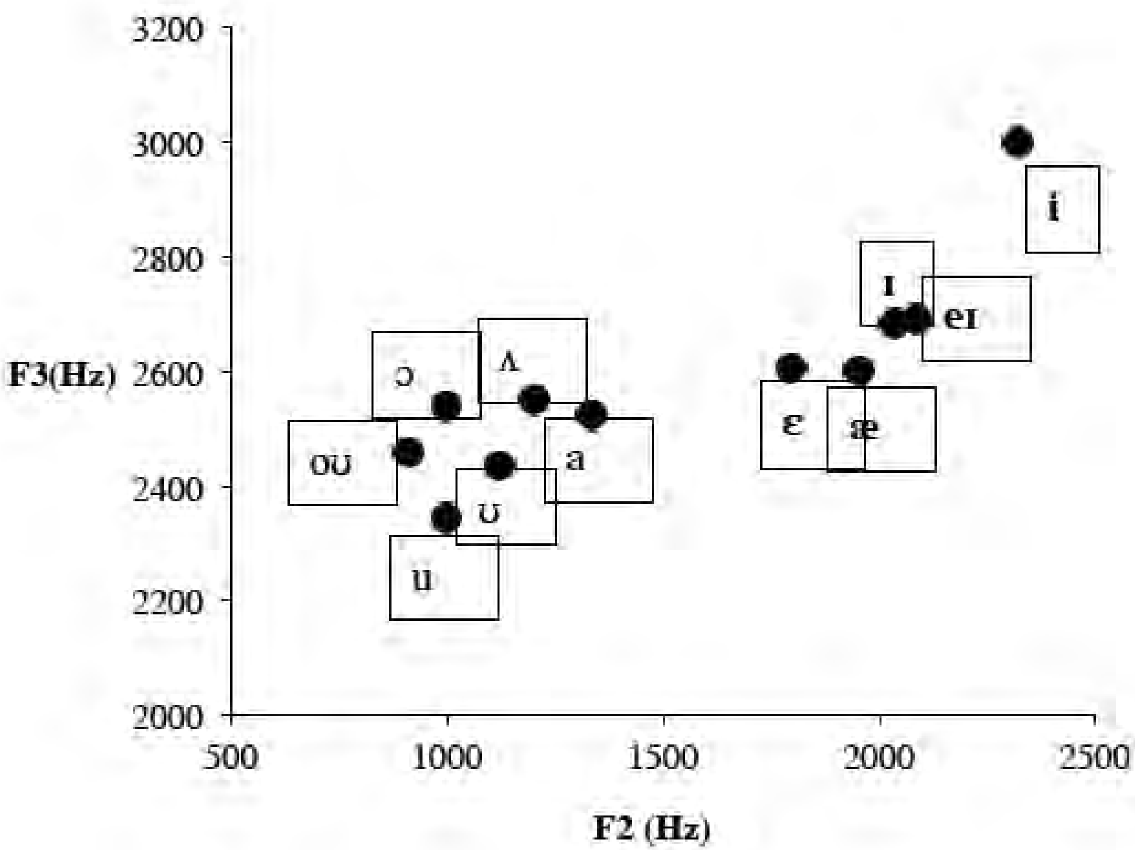
Figure 6.2 Midsagittal vocal tract configurations for the high vowels /i/ (left) and /u/ (right). Adult male speaker of English. (From Perkell, 1969.)

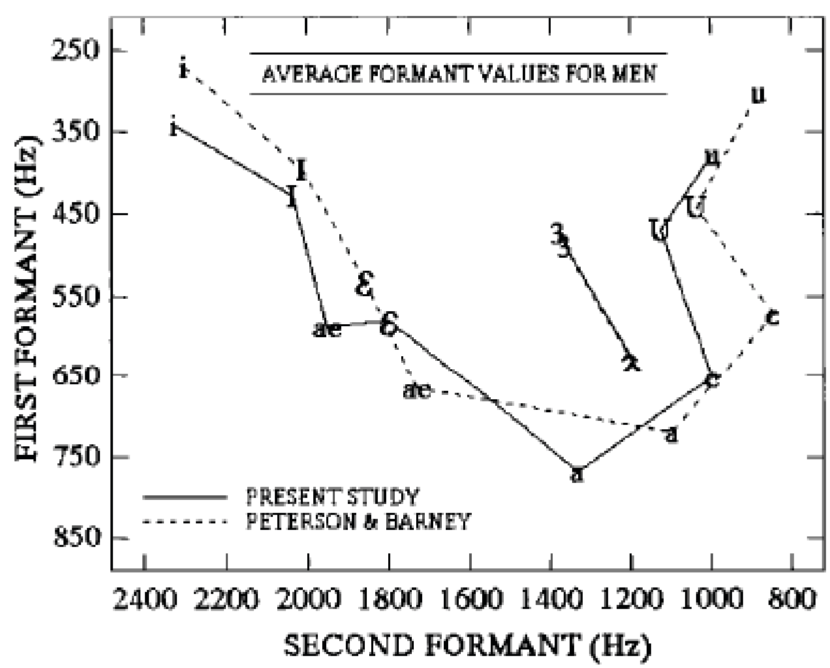




Fant (1960), Russian [i]
F2 2250 Hz, F3 3200 Hz

Hillenbrand et al (1995) – Michigan English vowel formants





Courtesy of The Acoustical Society of America. Used with permission.
Source: Hillenbrand, James, Laura A. Getty, Michael J. Clark, and Kimberlee Wheeler. "Acoustic characteristics of American English vowels." The Journal of the Acoustical society of America 97, no. 5(1995): 3099-3111.

Lip rounding

- Lip-rounding also involves lip protrusion so it both lengthens the vocal tract and introduces a constriction at the lips.
- Perturbation theory: All formants have a velocity maximum at the lips, so a constriction at the lips should lower all formants.
- Lengthening the vocal tract also lowers formants.
- Tube models: The effect of a constriction at the lips is equivalent to lengthening the front cavity. Protrusion actually lengthens the front cavity.
- This lowers the resonances of the front cavity - in front vowels the lowest front cavity resonance is usually F3, in back vowels it is F2.

Lip rounding

- Tube models 2: Fant (1960) suggests the front cavity plus lip constriction can form a helmholtz resonator.

Fant's (1960) nomograms

- A more complex tube model for vowels:

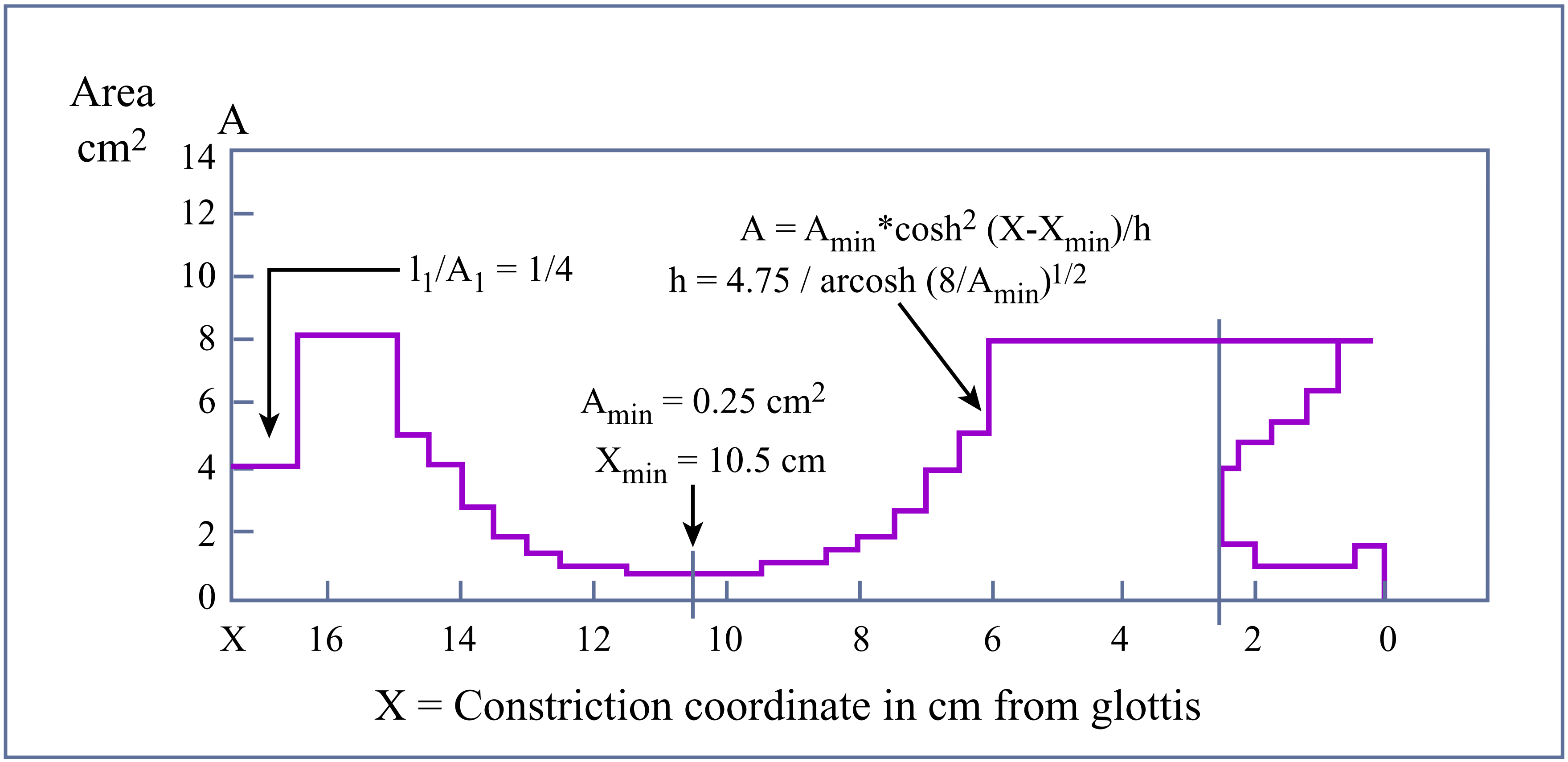
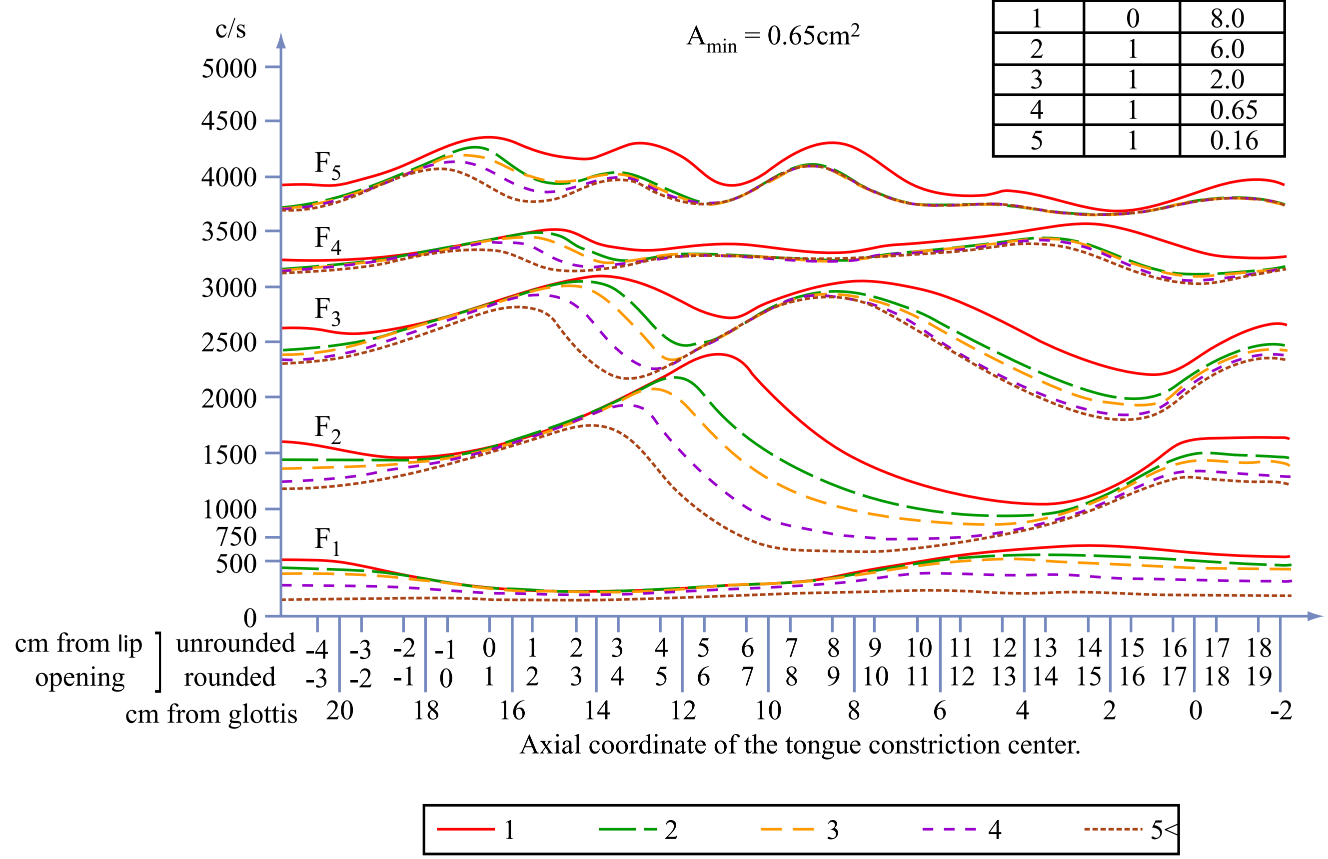


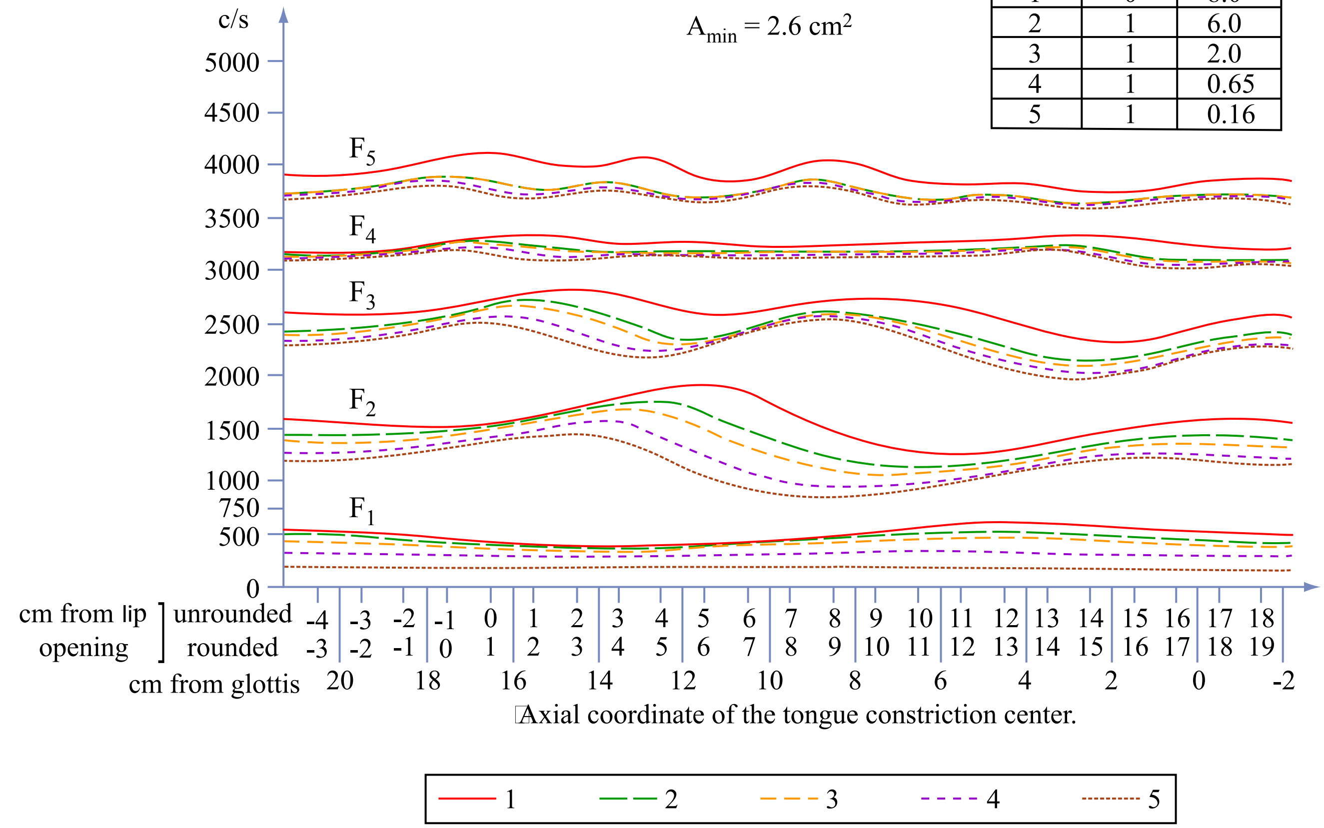
Image by MIT OCW.
Based on Fant, Gunnar. *Acoustic Theory of Speech Production*. The Netherlands: Mouton De Gruyter, 1960.

Nomogram showing variation in constriction location and lip-rounding - narrow constriction ($A_{min} = 0.65\ cm^2$)



Curve	L_1 cm	A_1 cm^2
1	0	8.0
2	1	6.0
3	1	2.0
4	1	0.65
5	1	0.16

Nomogram showing variation in constriction location and lip-rounding - wider constriction ($A_{min} = 2.5\ cm^2$)



Curve	L_1 cm	$A_1\ cm^2$
1	0	8.0
2	1	6.0
3	1	2.0
4	1	0.65
5	1	0.16

Image by MIT OCW.

Based on Fant, Gunnar. *Acoustic Theory of Speech Production*. The Netherlands: Mouton De Gruyter, 1960.

Nomogram showing variation in constriction location and degree.

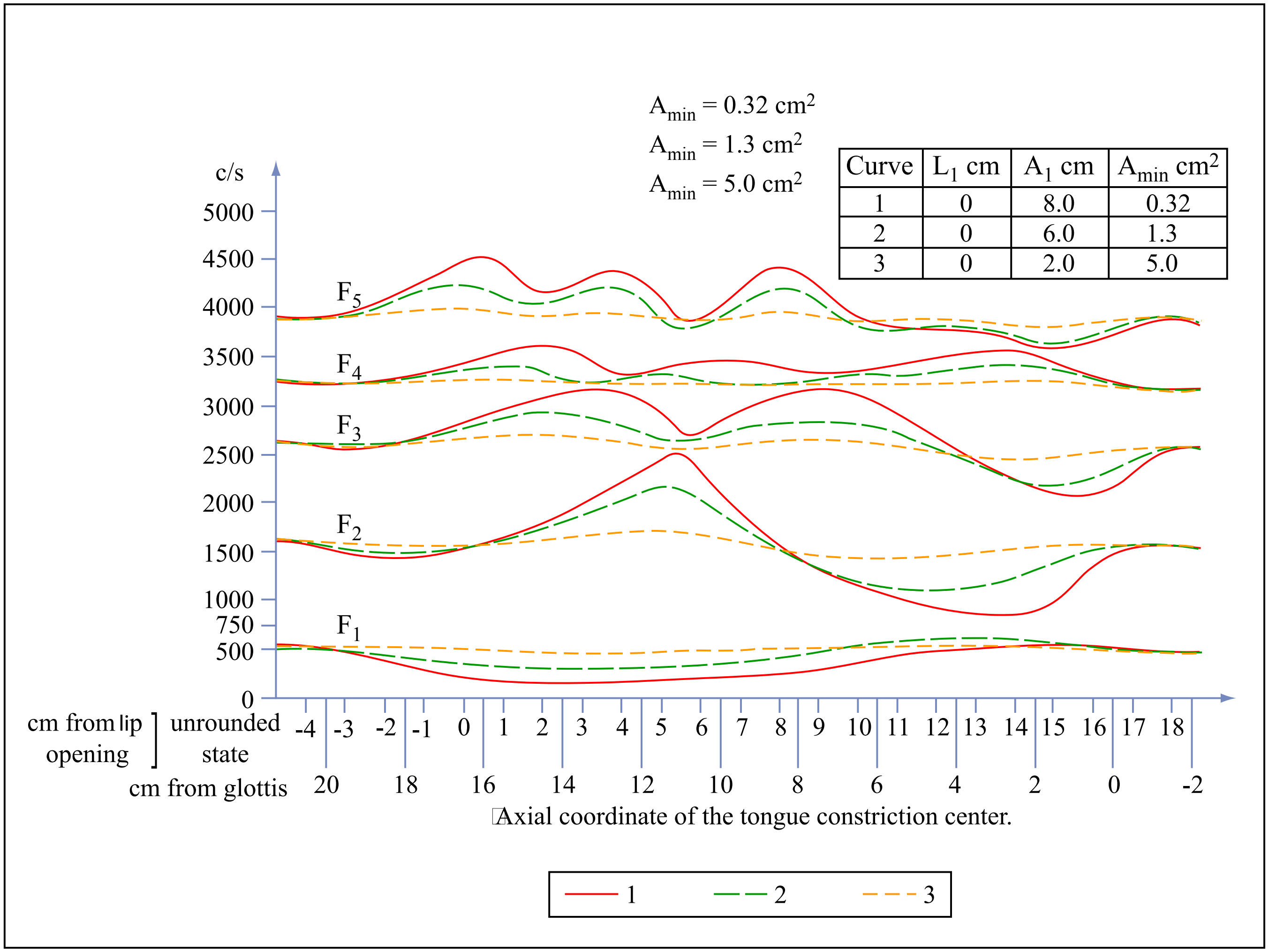


Image by MIT OCW.