

# **A practical introduction to acoustic phonetics day 3: Vowels**

Wouter Jansen

Department of Linguistics

University College London

wouter.jansen@kuvik.net

<http://wouter.jansen.kuvik.net/teaching>

January 19, 2005

# The acoustic theory of speech production (once again)

- The speech signal is generated by vocal tract filtering of a **source** sound
- Hence **source-filter** theory of sound production
- Sound sources in speech production:
  - ◆ Vocal fold vibration (voicing)
  - ◆ Turbulence noise produced at the glottis/in the larynx/pharynx (whisper)
  - ◆ Turbulence noise generated elsewhere (plosive releases, fricatives)

# Vocal tract filtering

- Some components of the voice source (or noise source) spectrum are enhanced because they correspond to **resonance frequencies** of (parts of) the vocal tract
- These resonance frequencies change when the shape of the vocal tract changes, leading to different sound qualities
- In vowels (and sonorant consonants) these resonances, which manifest themselves as broad peaks in the power spectrum, are commonly known as **formants**

# Resonances in the vocal tract

- With a neutral configuration (no constrictions at any point) the vocal tract can be approximated as a single tube with a single set of resonances
- The constricted vocal tract can be approximated using a model consisting of two tubes, each with its own set of resonances
- To understand how vocal tract resonances arise it is important to understand a few things about **standing waves**

# Resonant frequencies ( $t = 1 - t = 4$ )

+

0                      +

-                      0                      +

0                      -                      0                      +  
-

# Resonant frequencies ( $t = 5 - t = 8$ )

+	0	-	0
		-	0

0	+	0	-
	-	0	+

-	0	+	0
-	0	+	0

0	-	0	+
0	+	0	-

# Resonant frequencies ( $t = 9 - t = 12$ )

+	0	-	0
+	0	-	0

0	+	0	-
0	-	0	+

-	0	+	0
-	0	+	0

0	-	0	+
0	+	0	-

# Resonant frequencies ( $t = 13$ )

+	0	-	0
+	0	-	0



# Non-resonant frequencies ( $t = 5 - 8$ )

+	-	+	-
		-	+

-	+	-	+
	-	+	-

+	-	+	-
-	+	-	+

-	+	-	+
+	-	+	-

# Non-resonant frequencies ( $t = 9 - 12$ )

+	-	+	-
-	+	-	+

-	+	-	+
+	-	+	-

+	-	+	-
-	+	-	+

-	+	-	+
+	-	+	-

# Non-resonant frequencies ( $t = 13$ )

+	-	+	-
-	+	-	+

# Resonances of tubes closed at one end and open at the other

- The lowest resonance of a tube that's closed at one end and open at the other has a wavelength of 4 times the length of the tube  $L$
- Higher resonances  $\lambda_n$  occur at  $\frac{4}{(2n-1)}L$ , where  $n$  is the resonance number  $(\frac{4}{3}, \frac{4}{5}, \frac{4}{7}, \frac{4}{9})$
- As the frequency of a wave travelling in a medium depends on the propagation speed  $c$ , the frequency of the first resonance is  $\frac{c}{\lambda_1} = \frac{c}{4L}$

# Resonances of tubes closed at one end and open at the other

- More generally:  $\frac{(2n-1)c}{4L}$ , which yields higher resonances at odd integer multiples of the first resonance:  $F_2 = 3F_1$ ,  $F_3 = 5F_1$ , etc.

# Schwa – ə

- Given the speed of sound in air (atmospheric pressure; 20° C) of approximately 35,000 cm/s
- and a vocal tract of 17.5 cm
- we get  $F_1 = \frac{35,000}{70} = 500\text{Hz}$  for the vowel [ə]
- ... an  $F_2$  of 1,500 Hz, and an  $F_3$  of 2,500 Hz

# Two-tube models of vowel production

- Vowels that are characterised by a (modest) constriction somewhere in the vocal tract can be approximated using **two-tube models**
- The back tube or cavity represents the portion of the vocal tract behind the constriction
- The front cavity represents the portion of the vocal tract in front of the the constriction
- Each of these cavities produce their own set of resonances, in addition to the **Helmholtz resonance** that arises from the **coupling** of the two cavities

# Two-tube models of vowel production

- The Helmholtz resonance is proportional to the volumes of the back cavity, the front cavity and the volume of the constriction itself: it represents the value of the **first formant** ( $F_1$ ) of vowels
- In relatively front vowels, the frequency of the first resonance of the **back cavity** is relatively low: it represents the **second formant** of these vowels
- In vowels with a back articulation, the frequency of the first resonance of the back cavity is higher, and represents the **third formant** of such vowels

# Two-tube models of vowel production

- In vowels with a front articulation, the first resonance of the **front cavity** is relatively high, and represents the  $F_3$  of such vowels
- In vowels with a back articulation the first resonance of the front cavity is relatively low, and represents  $F_2$



# The relation between articulation and acoustics in vowels

- $F_1$  corresponds roughly to (the inverse of) vowel height: high vowels have a low  $F_1$
- $F_2$  corresponds roughly to (the inverse of) vowel backness: back vowels have a low  $F_2$
- Vowel rounding depresses both  $F_3$  and  $F_2$

# References

Clark, J. & C. Yallop (1995) *An Introduction to Phonetics and Phonology*. 2nd ed. London: Blackwell.

de Boer, B. (2001) *The Origins of Vowel Systems*. Oxford: Oxford University Press.

Johnson, K. (1997) *Acoustic and auditory phonetics*. London: Blackwell.

Ladefoged, P. (2003) *Phonetic Data Analysis. An Introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell.

Ladefoged, P. & I. Maddieson (1996) *Sounds of the World's Languages*. Oxford: Backwell.

Liljencrants, J., & B. Lindblom (1972) Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* 48: 839-862.

Lindblom, B. (1986) Phonetic universals in vowel systems. In J. Ohala & J. Jaeger (eds.) *Experimental Phonology*. Orlando: Academic Press, 13-44.

Schwartz, J.-L., L.-J. Boë, N. Vallée and C. Abry (1997a) Major trends in vowel system inventories. *Journal of Phonetics* 25: 233-253.

Schwartz, J.-L., L.-J. Boë, N. Vallée and C. Abry (1997b)

The Dispersion-Focalization Theory of vowel systems.  
*Journal of Phonetics* 25: 255-286.

Stevens, K. (1998) *Acoustic Phonetics* Cambridge, MA: MIT Press.