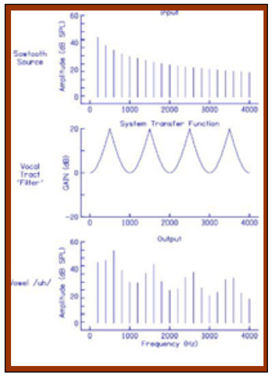
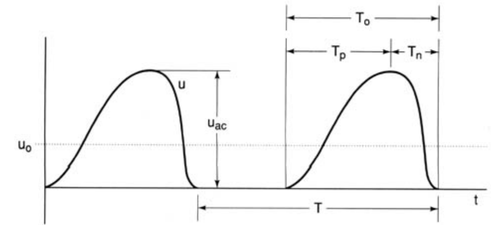
Introduction to Vowel acoustics
S674 D. Kewley-Port 2005

- Source-Filter Theory
- Vowel Formants
- Hillenbrand et al., 1995



SOURCE: Glottal waveform

- ***Airflow*** **(in cm^3/sec) over time**



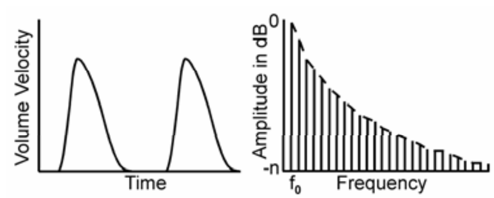
- Airflow pulses are source of energy voiced sounds
- <u>fundamental frequency</u>: F0 = 1/T

Acoustic Theory of Speech Production

- a.k.a. Source-Filter Theory
- Gunnar Fant, 1960
- Relates speech acoustics to vocal tract shapes
- Applies to disordered and normal speech
- INDEPENDENCE of source and filter

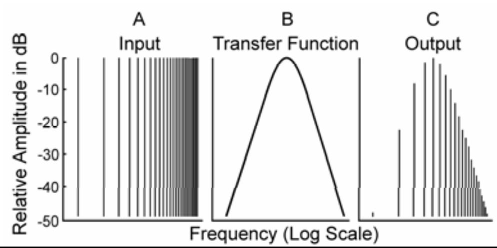
Spectrum of Glottal wave

- Waveform similar to sawtooth wave
- Spectrum has all harmonics of F0
- Envelope of spectrum decreases rapidly, depending on characteristics of talker



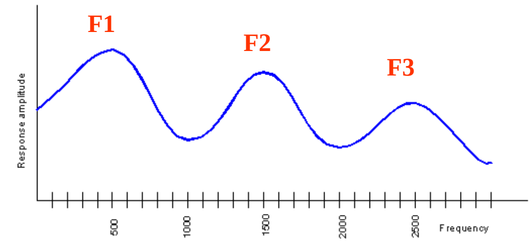
Filtering

- Input spectrum has many harmonics
- Transfer function specifies how input is altered by filter
- Output spectrum shows each harmonic with amplitudes attenuated by filter

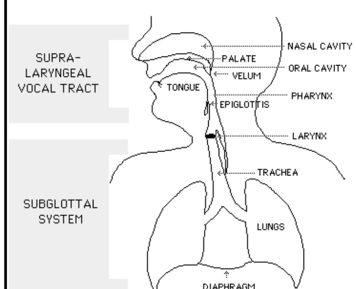


Open/closed tube model of /Ɨ/

- Plot of resonances or formants
- Broad bandwidths because soft walls vocal tract damp vibrations



Vocal Tract Model for /Ɨ/



- Vocal tract is a resonant system that acts as a filter of the glottal source
- /Ɨ/ is vowel with articulators in 'neutral' position

How to Calculate the Effects of Filtering

- Consider a relatively flat source spectrum (dB)
- Just one resonator in transfer function
- Combined spectrum: add dB effect of filter at each harmonic, A_H

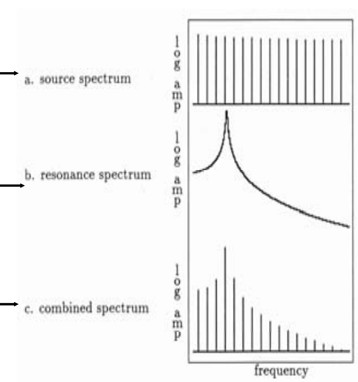
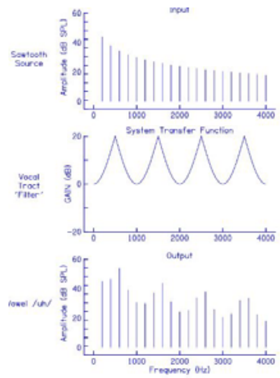


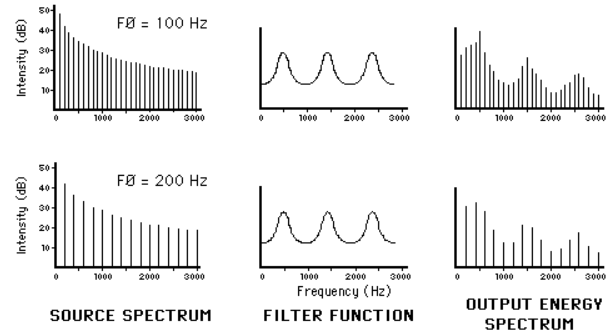
Figure 3.13: Sawtooth wave and a single resonance

Example: Source-Filter for /Ɨ/

- **Glottal spectrum, F0 = 200 Hz**
- **Resonances for /Ɨ/ at 500, 1500, 2500, 3500 Hz**
- **Add A_H for each harmonic to transfer function (here a 'gain') to obtain output spectrum /Ɨ/**



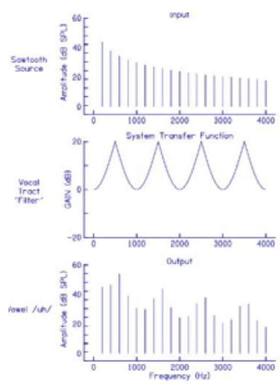
Same filter, F0 dense or sparse harmonics



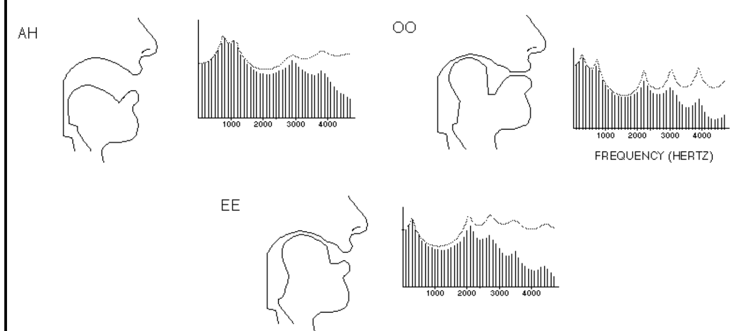
DKP vowels, 2005

Formant as Resonance

- **Note, output spectrum series harmonics different amplitudes.**
- **"Formants" in output visually imposed on harmonics.**
- **F0 & formants interact**



Same source, different resonances

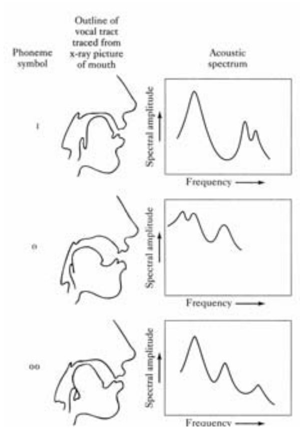


(Dashed lines shows transfer functions)

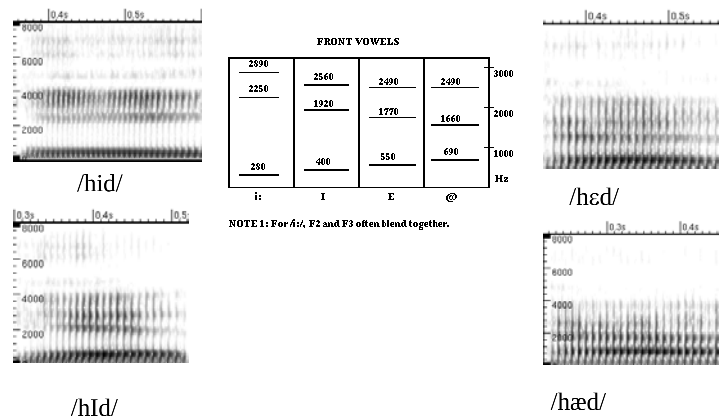
DKP vowels, 2005

How do we measure formant frequency?

- Use computer techniques such as LPC (linear predictive coding) to calculate transfer functions as estimates of formant frequency

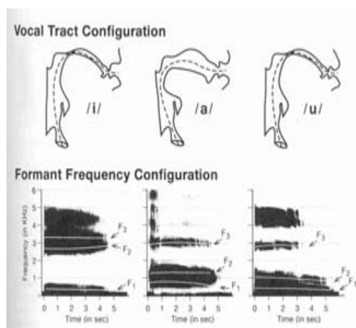


Front Vowel Formants



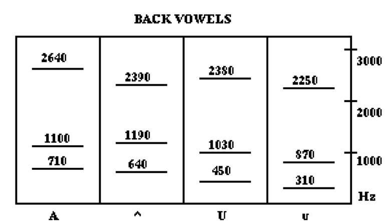
Or Measure Formants from Spectrograms

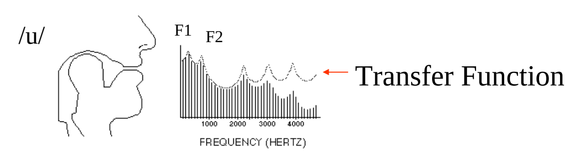
- Measure 'center frequency' of formants
- Often only 'steady-state' frequencies
- First two formants most important



Back Vowel Formants

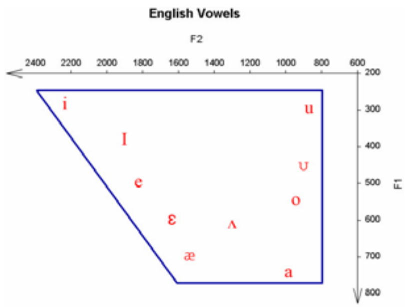
F1 & F2 Low in frequency for back vowels





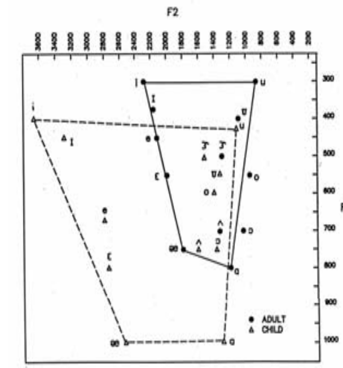
Organized in F1 X F2 Vowel Plot

- 0,0 in upper right corner, F2 horizontal and F1 vertical
- Similar to IPA phonetic chart
- Acoustics, not articulation best definition of relationship of vowels



Vocal Tract Size & Acoustics

- Peterson & Barney, 1952 measured vowels from 76 men, women & children
- For women & children, frequencies increased
- Vowel pattern remained the same

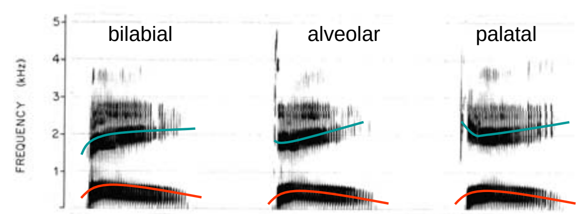


"Talker" Variability

- But vowel resonances influenced by length L of vocal tract.
- And vocal tracts come in vastly different sizes, and yet we all understand one another
- For vowels, how?
- Pattern in F1 X F2 plots

Formant Dynamics

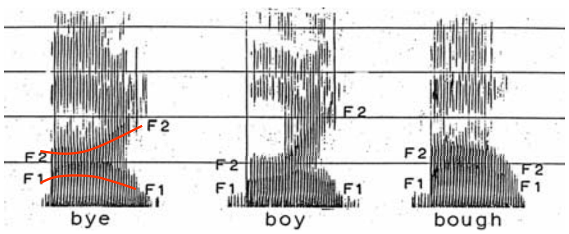
- Formants are not steady state
- Both consonantal transitions and inherent vowel spectral change



bay day gay

Diphthongs

- Diphthongs: combination of two vowels:
  - / ai / /o i / / a u /
- Acoustics highly dynamics
- Describe as formant change between high, mid & low vowel concentrations

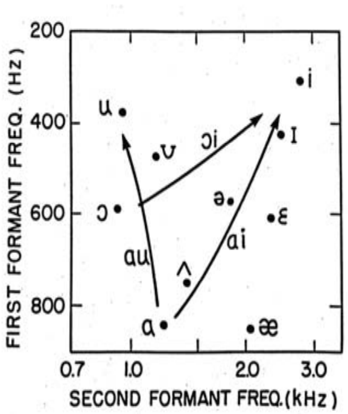


Sources of Vowel Variability

- Vocal tract length
- Formant dynamics
- Coarticulation: “tone” is [t õ n]
- Number of syllables: please, pleasing, pleasingly
- Rate of speech
- Dialect & gender

Diphthongs in F1 X F2 plots

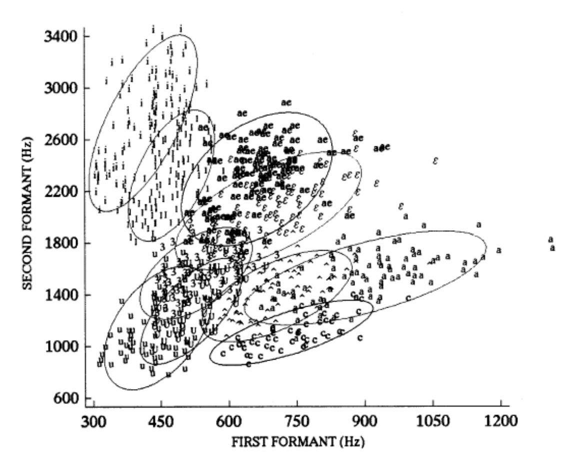
- Measure starting and ending value of F1 & F2
- Show as dynamic arrows in F1 X F2 plots



Hillenbrand, Getty, Clark & Wheeler, 1995

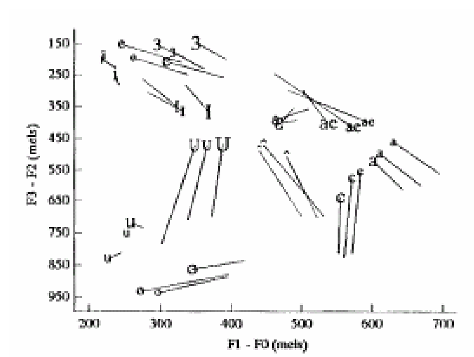
- Update and enhancement of Peterson & Barney, 1952 study
- 45 men, 48 women, 46 10-12 yrs
- 12 vowels, /hVd/
- Measured F0, F1, F2 ,F3, F4 & duration
  - at 20%, Steady-state, 80% of duration
  - Used LPC, FFT, grey-scale spectrograms
  - good reliability on formant values, 1-2%

- **Results: Overlap of vowels similar to other reports**



Spectral Change patterns

- **Symbol at 80% starting from 20%**



Comparison to PB 1952

- **F1 x F2 acoustic spaces similar, but higher than PB**

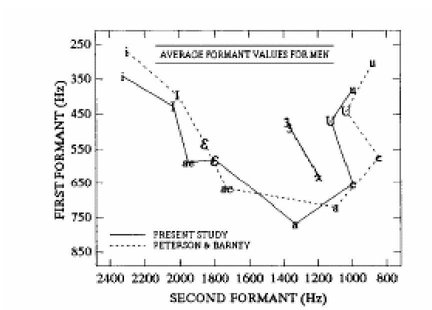


FIG. 5. Acoustic vowel diagrams showing average formant frequencies for men from the present study and from Peterson and Barney ("ae"=/æ/, "a"=/ɑ/, "c"=/ɔ/, "ʌ"=/ʌ/, "з"=/ɝ/).

Identification

- **Very similar, high identification vowels**
  - **PB 95.4%**
  - **HGCW 94.4**
- **Confusions mostly on diagonal of Table VII**
- **Used discriminate analysis with various sets pf measurements to determine which set predicts classification the best.**

Classification Results

- F1 & F2 at two times (20% & 80%) best.
- Duration adds small benefit

	All tokens					
	One sample		Two samples		Three samples	
Parameter set	NoDur	Dur	NoDur	Dur	NoDur	Dur
$F1,F2$	68.2	76.1	87.9	90.3	87.7	90.4
$F1,F2,F3$	81.0	84.6	91.6	92.7	91.8	93.1
$F0,F1,F2$	78.2	82.0	91.6	92.5	91.0	92.6
$F0,F1,F2,F3$	84.7	87.8	93.6	94.1	92.8	94.8
	Well identified tokens only					
	One sample		Two samples		Three samples	
Parameter set	NoDur	Dur	NoDur	Dur	NoDur	Dur
$F1,F2$	71.4	80.0	90.8	93.6	90.7	93.3
$F1,F2,F3$	85.3	89.1	95.4	96.2	95.3	95.8
$F0,F1,F2$	82.3	86.3	95.5	96.3	94.8	96.0
$F0,F1,F2,F3$	88.7	91.6	97.3	97.8	96.6	97.3

Study Guide 1

- 1. What were the problems with the Peterson and Barney (1952) study, and did Hillenbrand et al. solve them?
- 2. What does "steady-state" vowel mean and how was it determined here?
- 3. What measures performed best in the discriminate analysis? What implications does this have for models of vowel perception?
-
- General
- 1. What is the importance of vowels in speech intelligibility?
- a) For normal-hearing people
- b) For hearing-impaired people

Summary

- Vowel acoustics well understood
- Two studies have confirmed that F1 & F2 most important acoustic information to classify vowels, with small benefit from F0 and duration (secondary cues) when measured meticulously
- But acoustics for vowels has many sources of variability
- Information used by listeners to accurately identify vowels in ordinary speech not well understood.