

Lecture 11

Phonetics

- Describe the different vowel sounds in English and their formant frequencies
- Define what is meant by diphthongs and semivowels
- Identify the factors that distinguish consonants:
 - Place of articulation
 - Manner of articulation
 - Voicing
- Identify the difference between a phone and a phoneme
- Describe how the sound of a phoneme changes according to context

Vowel Sounds

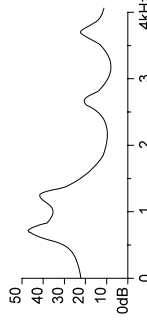
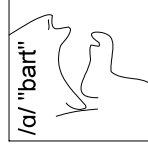
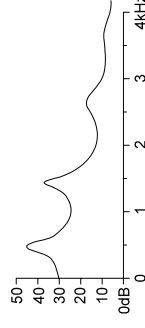
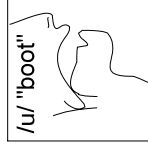
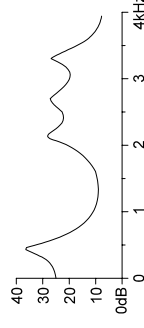
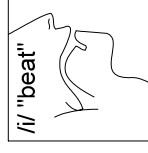
The tongue hump divides the vocal tract into two cavities.

The rear cavity determines f_1 , the first formant frequency:

- f_1 decreases as the tongue hump moves forwards or is raised higher

The front cavity determines f_2 :

- f_2 increases as the tongue hump moves forwards.



Vowel Formant Frequencies

Place: F=Front, C=Centre, B=Back of mouth
 Height: 1=low, 3=medium, 5=high

		f1	f2	f3
F5	i	280	2250	2890
F4	ɪ	400	1920	2560
F2	ɛ (E)	550	1770	2490
F1	æ (@)	690	1660	2490
C4	ɜ (R)	480	1250	1700
C3	ə (X)	500	1150	1650
C2	ʌ (A)	600	1200	2800
B5	u	310	870	2250
B4	ʊ (U)	450	1030	2250
B2	ɔ (C)	590	880	2540
B1	ɒ (O)	500	900	2300
B1	ɑ (A)	710	1100	2540
***	h			(Same as vowel)

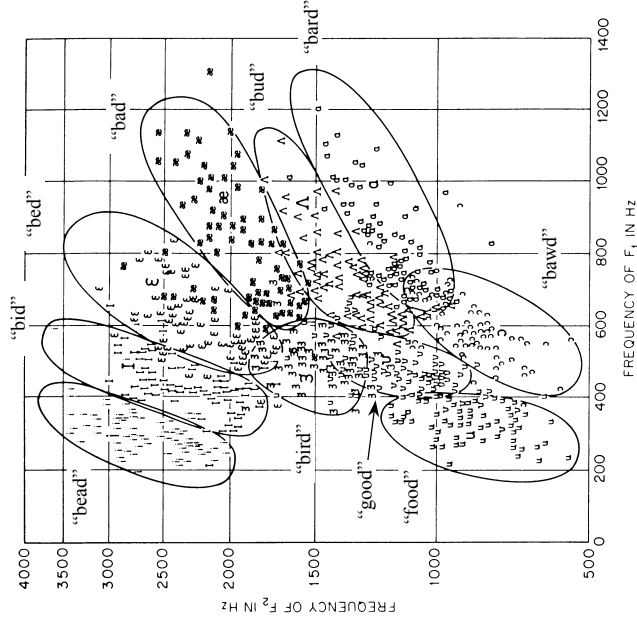
Vowel symbols are from the International Phonetic Alphabet (IPA): <http://www2.arts.gla.ac.uk/ipa/ipa.html>
 ARPABET symbols are in parentheses.

The neutral C3 vowel ə (X) is usually known as "schwa": it usually occurs in unstressed syllables.

f1, f2, f3 are the typical values for the first three *formants* (resonant frequencies) of the vocal tract.

Formant Frequency Variability

American Male speakers:



Diphthongs and Semivowels

Diphthongs are formed by a transition between two vowels within a single syllable

Place: F=Front, C=Centre, B=Back of mouth
 Height: 1=low, 3=medium, 5=high

Place	IPA	example
F3/F4	eɪ	"bay"
C3/B3	aʊ	"no"
C1/F3	aɪ (Y)	"buy"
C1/B3	aʊ (W)	"now"
B2/F3	ɔɪ (O)	"boy"
F3/C3	ɪə	"pier"
F2/C3	eə	"pair"
B4/C3	ʊə	"poor"

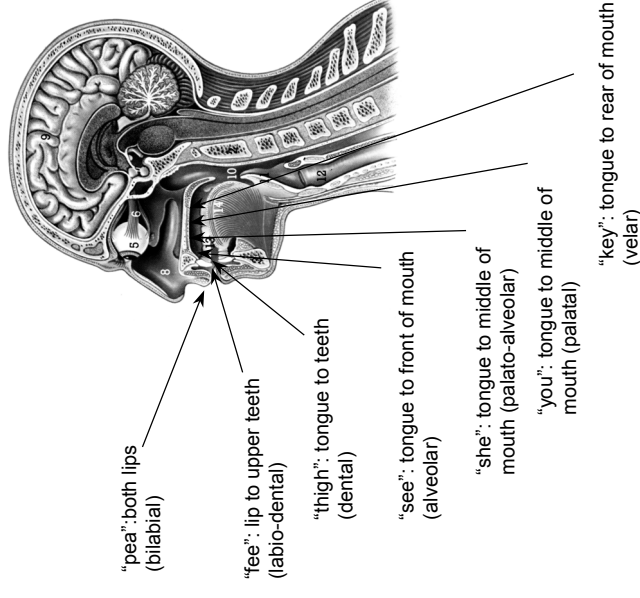
Semivowels or Approximants are like vowels but with the tongue or lips almost closing off the vocal tract completely

B4	w	"went"
F4	j (y)	"you"
F4	r	"run"
F4	l	"love"

Consonants

Consonant sounds use the tongue or lower lip to block the vocal tract almost completely ⇒ turbulent airflow.

The spectrum of the sound produced is affected by the portion of the vocal tract between the constriction and lips:



Consonant Classification

Consonants are classified as:

- *nasal*: mouth cavity blocked and velum open
- *stop or plosive*: mouth cavity blocked then released
- *fricative*: mouth cavity not fully blocked
- *affricate*: a stop followed by a fricative

Many consonants have voiced ("zzz") and unvoiced ("sss") forms:

	Place	Voiced	Unvoiced
Nasal	B	m	"me"
	A	n	"knee"
	V	ŋ (G)	"rang"
Stop	B	b	"bee"
	A	d	"dee"
	V	g	"guy"
Fricative	L	v	"vee"
	D	ð (D)	"thy"
	A	z	"zoo"
	PA	ʒ (Z)	"vision"
	A	dʒ (J)	"jive"
Affricate	A	tʃ (C)	"chime"

Place: B=Bilabial, L=Labiodental, D=Dental, A=Alveolar, PA=Palato-alveolar, V=Velar
 ARPABet transcriptions are in parentheses.

Phones and Phonemes

Words consist of a sequence of sounds called phones

- "spit" = [s] + [p] + [ɪ] + [t]
- "pit" = [pʰ] + [ɪ] + [t]
- "bit" = [b] + [ɪ] + [t]

the /p/ in "pit" is aspirated (accompanied by a strong puff off air) whereas the /p/ in "spit" is not.

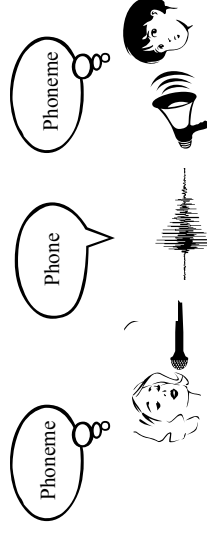
[p] and [pʰ] are examples of the /p/ phoneme.

A phoneme is a range of sounds that:

- are felt to be equivalent by speakers
- are interchangeable as far as word meaning goes: [b] and [pʰ] can't be the same phoneme because "bit" ≠ "pit"

Thus a phone is the name of a sound while a phoneme is the name of the underlying concept in the speaker's brain.

The set of phonemes for a particular language is the smallest number of different sounds needed to distinguish all its words. English needs about 42.



Coarticulation

Phonemes alter in sound because of neighbouring sounds

- Gross changes are called allophones
 - the vowels in unstressed syllables often move towards the neutral schwa vowel /ə/ (X):
 - "vanilla" → [və'nɪlə]
- Minor changes are coarticulation
 - Anticipatory Coarticulation: Consonants often change towards sounds occurring later in the sentence:
 - "I went by bus" → "I wemp by bus" in which /n/ → /m/ and /t/ → /p/ because of the bilabial /b/ of "by".

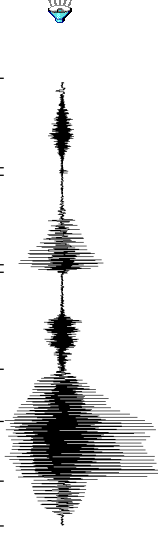
Conversational Speech

In normal fluent speech other variations may occur:

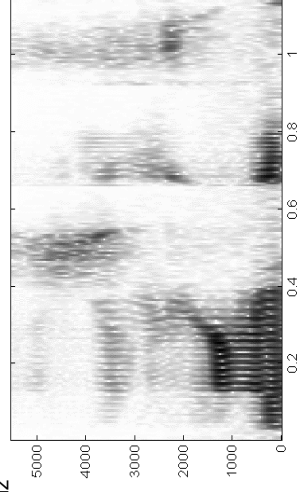
- Some words change if the following word starts with a vowel:
 - "the cat" versus "the ant"
- Some unstressed words are shortened
 - "you and me" → "you an me"
 - "I must see" → "I ms see"
 - "It would do" → "It 'd do"

"my speech"

m a I s p i t j



Hz



- Formant frequencies vary smoothly as the tongue and jaw change position.
- Amplitude varies continuously within phonemes
- Speech varies most at the boundary between phonemes

Comments on Spectrogram

- m Very little sound output because lips are closed
- aɪ As the tongue moves forwards in the middle of the diphthong f1 changes from 1000 Hz to 500 Hz while f2 changes from 1200 Hz to 2000 Hz
- s No vertical stripes since this is an unvoiced sound generated by turbulence. The sound source is close to the front of the mouth so the energy is concentrated at high frequencies around 4500 Hz
- pi Complete silence before the /p/ is exploded. Then a broad-band impulse followed by the vowel
- tʃ Again complete silence before the /t/ is exploded followed by a broad-band impulse. The /j/ has the tongue further back in the mouth than /s/ did and so the energy extends to lower frequencies.