Acoustic Phonetics

David House

Speech physiology and speech acoustics

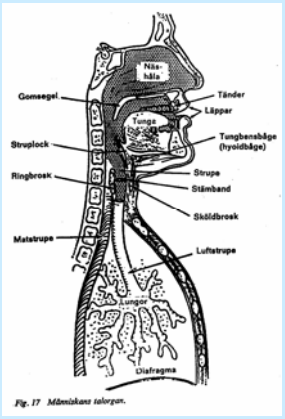
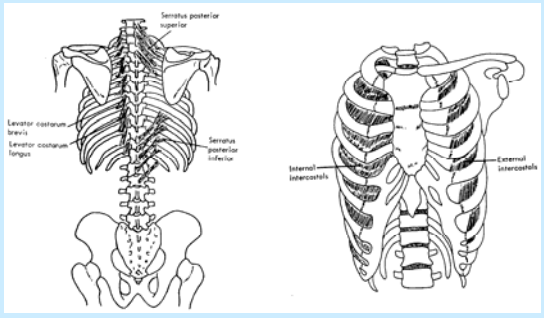
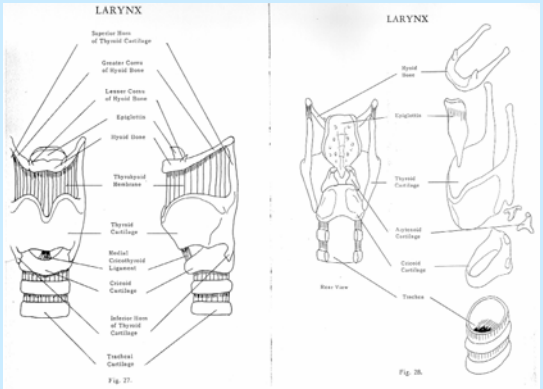


Fig. 17 Människans talorgan.

The lungs and the larynx

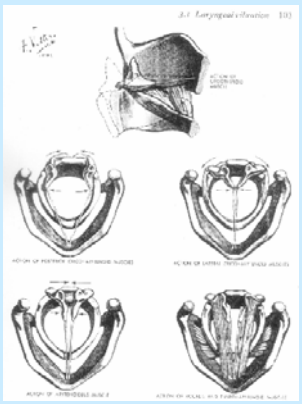
- Expiratory respiration – generate sound
- trachea *luftstrupen*
- larynx *struphuvudet*
  - cartilage, muscles and ligaments
  - glottis *röstspringan*
  - vocal folds *stämläpparna*
    - vocalis muscle, vocal ligament
- epiglottis *struplocket*





Voice

- Biological function of the larynx
  - Protect the lungs and airway for breathing
  - Stabilize the thorax for exertion
  - Expel foreign objects by coughing
- Phonation and voice source
  - Creation of periodic voiced sounds
  - Vocal folds are brought together, air is blown out through the folds, vibration is created



Muscular control of phonation

- Lateral control of the glottis
  - adduction (for protection and voiced sounds)
  - abduction (for breathing and voiceless sounds)
- Longitudinal control of the glottis
  - tension settings of the vocalis muscle
  - control of fundamental frequency (F0)

Voice quality

- Phonation type (lateral tension)
  - Tense (pressed) voice *pressad*
  - Normal (modal) voice *modal*
  - Flow phonation *flödig*
  - Breathy voice *läckande*
- Vocal intensity
  - Interaction between subglottal lung pressure and lateral (adductive) tension

Voice pitch

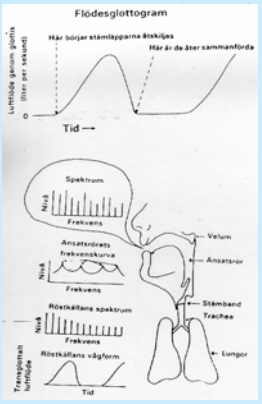
- Pitch level
  - high-pitched or low-pitched voice (average F0)
- Pitch range
  - large or small
- Register
  - modal
  - falsetto
  - creak *knarr*

Use of voice in normal speech

- Boundary signalling
  - vocal intensity greatest at phrase beginnings
  - pitch generally higher at phrase beginning
  - creak as a signal of phrase endings
- Social marker
  - voice quality as a signal of group identity (dialect)
- Expression of attitude and emotion
  - happy or angry
  - serious or sensual

Source-filter theory

- Voice-source waveform (during phonation)
  - Transglottal airflow measurements
- Spectrum of the voice source
  - Decreases in amplitude with increasing frequency
- Vocal tract resonances
  - Dependent on position of the tongue and lips
- Spectrum of radiated sound
  - Sum of voice source and vocal tract resonances



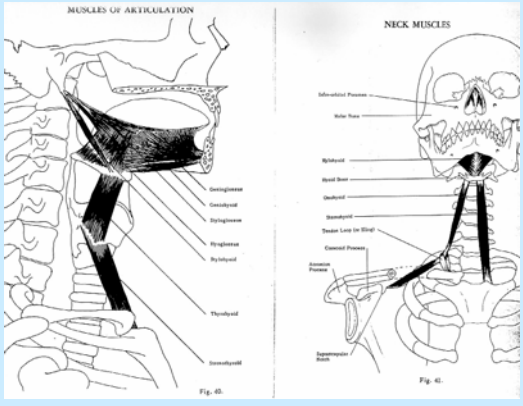
From Sundberg: Röstlära

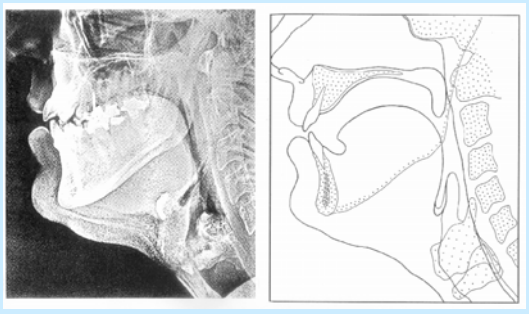
Vowels and consonants

- Speech production (phonetics)
  - Free air passage through the pharynx, mouth and the lips = vowel
  - Constricted or closed air passage = consonant
- Function (phonology)
  - Nuclear in the syllable = vowel
  - Marginal in the syllable = consonant
- Exceptions
  - Some voiced consonants (e.g. syllablic nasal)
  - Approximants or semi-vowels (e.g. [j] [w])

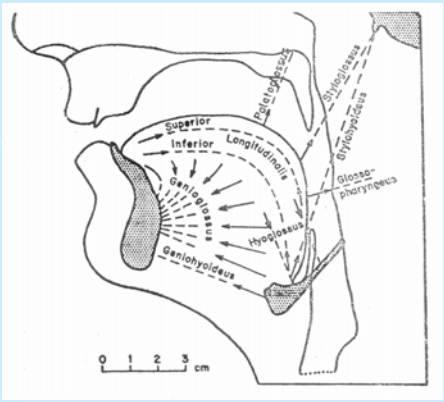
The vocal tract

- Throat, (svalget): *pharynx, faryngal*
- Oral cavity, (munhålan): *os, oral*
- Nasal cavity, (näshålan): *nasus, nasal*





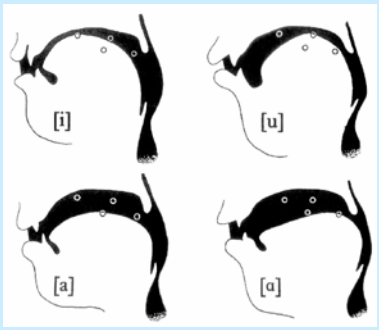
From Laver: Principles of Phonetics



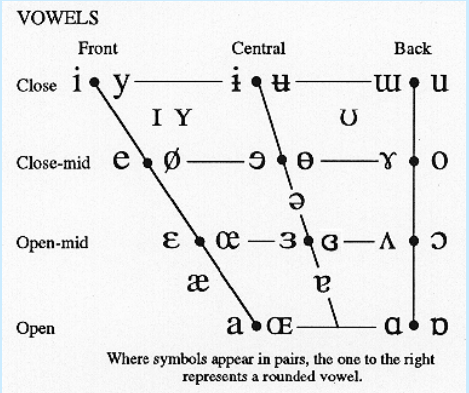
Vowel articulation

- Cardinal vowels
  - Reference vowels
  - Four corner vowels form the corners of the vowel chart
- Descriptive terminology
  - Close-open (high-low) *sluten-öppen*
  - Front-back *främre-bakre*
  - Unrounded-rounded *orundad-rundad*
  - Oral-nasal (e.g. French) *oral-nasal*

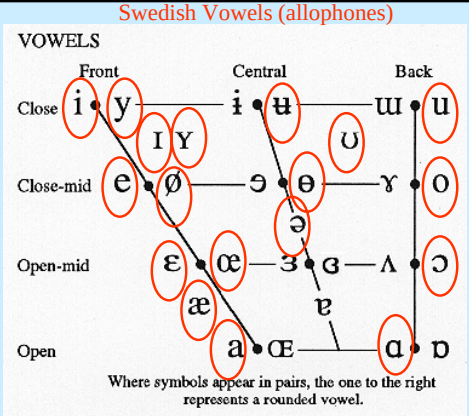
Tongue shapes of four of the cardinal vowels



From Elert: Allmän och svensk fonetik



Swedish Vowels (allophones)



Tabell 5.1 De långa och korta vokalerna i svenskt rikspråksuttal.

Långa vokaler		*Korta vokaler*	
fonetiskt tecken	*nyckelord*	*fonetiskt tecken*	*nyckelord*
[ɑː]	*mat*	[a]	*matt*
[eː]	*vet*	[e]	*vett, året*
[iː]	*vit*	[ɪ]	*vitt*
[uː]	*bo*	[ʊ]	*bott*
[ʉː] el. [ʉ̟ː]	*hus*	[ɵ]	*hund*
[yː]	*byt*	[ʏ]	*bytt*
[oː]	*gå*	[ɔ]	*gått*
[ɛː]	*säl*	[ɛ]	*vätt*
[æː]	*här*	[æ]	*kärr*
[øː]	*hö*	[øˠ]	*höst*
[œː]	*hör*	[œ]	*förr*

From Elert: Allmän och svensk fonetik

Phonological features

- Binary features for vowels
  - e.g.
    - ± high
    - ± low
    - ± back
    - ± round
- Feature matrix
  - Feature specification for each phoneme

Vokalfonem

A. Svenska

		förträngningens läge			
		främre			bakre
tungkroppens läge	högt	i	y	ʉ	u
	mellan	e	ø		o
	lågt	ɛ			ɑ
		orun-dade	utrun-dade	inrundade	
		läppartikulation			

B. Finska

		förträngningens läge		
		främre		bakre
tungkroppens läge	högt	i	y	u
	mellan	e	ø	o
	lågt	ɛ		ɑ
		orun-dade	rundade	
		läppartikulation		

From Gårding: Kontrastiv fonetik och syntax med svenska i centrum

Vowel acoustics

- Spectrogram
  - Narrow band spectrogram
  - Wide band spectrogram
- Formants (F1, F2, F3, F4)
- Acoustic vowel diagram (F1, F2)
- Formant transitions

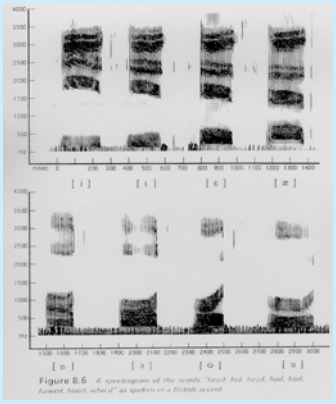
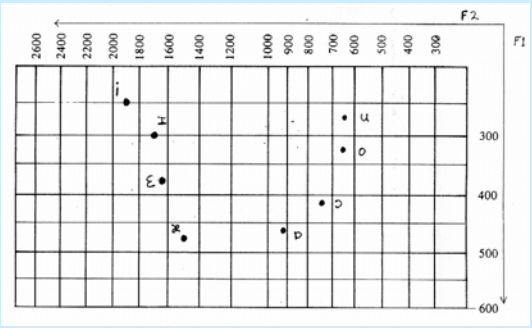
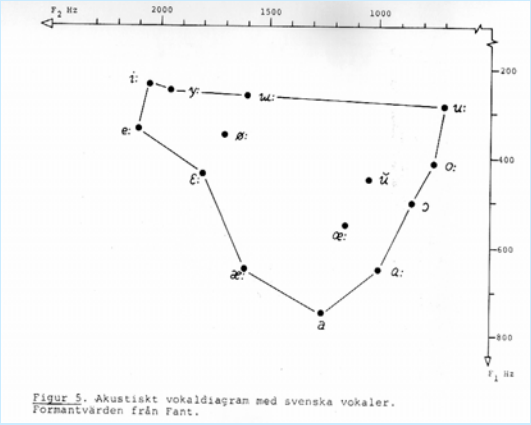


Figure 8.6 *A spectrogram of the words "heed, hid, head, had, hod, hawed, hoed, who'd" as spoken in a British accent.*

From Ladefoged: A course in phonetics

Acoustic vowel diagram (F1, F2)





Figur 5. Akustiskt vokaldiagram med svenska vokaler. Formantvärden från Fant.

Consonant articulation

- Voiceless or voiced
  - fortis or lenis
  - aspirated or unaspirated
- Manner of articulation
  - How is the sound produced?
- Place of articulation
  - Where is the constriction or closure located?

Manner of articulation

- Fricatives *frikativor (spiranter)*
- Stops, plosives *klusiler, explosivor*
  - aspiration
  - unreleased
  - affricates (stop + fricative) *affrikator*
- Liquids *likvidor*
  - laterals *lateraler*
  - trills *tremulanter (vibranter)*
- Nasals *nasaler*

The tongue: *lingua*

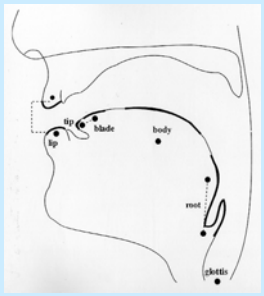
- Tongue tip: *apex, apikal*
- Tongue blade: *predorsum, predorsal* (also *corona, coronal*)
- Tongue back: *dorsum, dorsal*
- Tongue root: *radix*

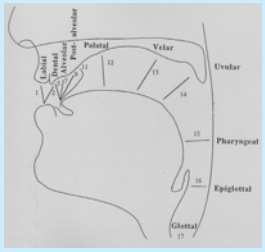
The palate

- Alveolar ridge (tandvallen) *: alveoli, alveolar*
- Hard palate (hårda gommen): *palatum, palatal*
- Soft palate (mjuka gommen): *velum, velar*
- Uvula (tungspenen)*: uvula, uvular*

The teeth and lips

- teeth: *dentes, dental*
- lips: *labia, labial*
  - rounded - *labialised*
  - unrounded - *delabialised*

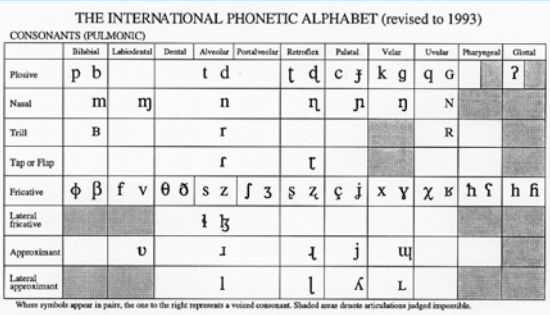




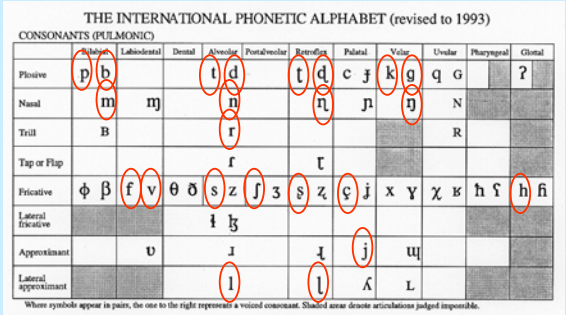
From Ladefoged: A course in phonetics

Place of articulation (IPA)

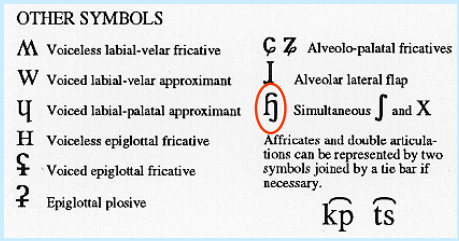
- Bilabial
- Labiodental
- Dental
- Alveolar
- Postalveolar
- Retroflex
- Palatal
- Velar
- Uvular
- Pharyngeal
- Glottal (laryngeal)



Swedish consonants



SJ-LJUDET I SVENSKA



Phonological features

- ±consonant
- ±sonorant
- ±obstruent
- ±anterior
- ±coronal
- ±continuant
- ±voice

Konsonantfonem

A. Svenska

Artikulationssätt				Artikulationsställen: lab	lab-dent	dent alv	pal vel	glott
egentliga konsonanter	klusiler		tonlösa	p		t	k	
			tonande	b		d	g	
	frikativor		tonlösa	ɧ	f	s	ɕ	h
			tonande		v		j	
vokallikn. konsonanter	likvidor	later				l		
		vibr				r		
	nasaler			m		n	ŋ	

B. Finska

egentliga konsonanter	klusiler			p		t	k	
	frikativor		tonlösa			s		h
			tonande		v		j	
vokallikn. konsonanter	likvidor	later				l		
		vibr				r		
	nasaler			m		n	ŋ	

From Gårding: Kontrastiv fonetik och syntax med svenska i centrum

Consonant acoustics (1)

- Fricatives
  - Noise frequency
  - Formant transitions in adjoining vowels
- Stops
  - Occlusion phase (silence)
  - Plosive release
  - Aspiration
  - Formant transitions in adjoining vowels

Consonant acoustics (2)

- Liquids
  - Laterals
    - Formants similar to vowels, lower intensity
    - Formant transitions
  - Trills
    - Quickly repeated stops
    - Short vowel-like pulses
    - Formant transitions

Consonant acoustics (3)

- Nasals
  - Vowel-like with lower intensity
  - Nasal resonances (nasal formants)
  - Formant transitions in adjoining vowels

Prosody

- Suprasegmental speech characteristics
  - Temporal relationships
  - Stress patterns
  - Speech rhythm
  - Intonation
- Functions of prosody
  - Lend prominence (emphasize, de-emphasize)
  - Grouping function (combine, separate)

Prosodic categories

- Stress (syllable)
  - Speech rhythm, alternating stressed-unstressed
- Word accent (word)
  - accent I (acute), accent II (grave)
- Focus (phrase accent)
  - Emphasis, contrastive emphasis
- Juncture (phrase, utterance)
  - Boundary signals and connective signals

Acoustic features of prosody

- Time (quantity)
- Fundamental frequency (F0) (pitch, intonation)
- Intensity (loudness)

References

- Elert, Claes-Christian (1995) Allmän och svensk fonetik. Norstedts Förlag, Stockholm
- Ladefoged, Peter (1982) A course in phonetics. Harcourt Brace Jovanovich, New York
- Laver, John (1994) Principles of phonetics. Cambridge University Press, Cambridge
- Sundberg, Johan (1986) Röstlära. Proprius, Stockholm

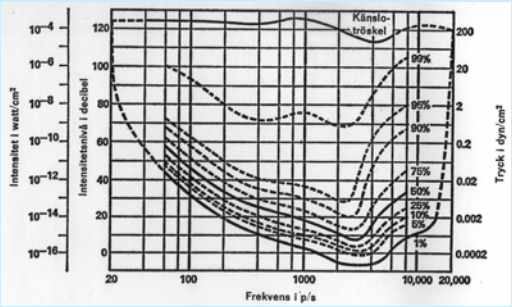
Psychoacoustics and speech perception

David House

Hearing acuity

- Sensitive for sounds from 20 to 20 000 Hz
- Greatest sensitivity between 1000-6000 Hz
- Non-linear perception of frequency intervals
  - E.g. octaves
    - 100Hz - 200Hz - 400Hz - 800Hz - 1600Hz
  - 100Hz - 800Hz perceived as a large difference
  - 3100Hz - 3800 Hz perceived as a small difference

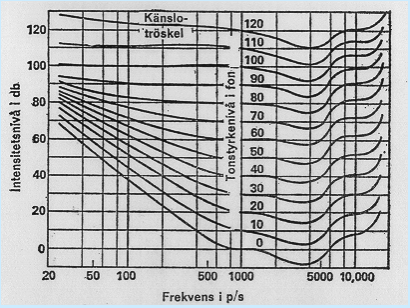
Absolute auditory threshold



Demo: SPL (Sound pressure level) dB

- Decreasing noise levels
  - 6 dB steps, 10 steps, 2*
  - 3 dB steps, 15 steps, 2*
  - 1 dB steps, 20 steps, 2*

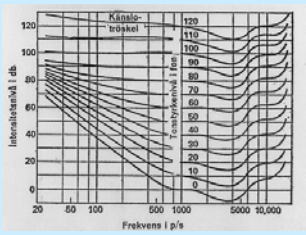
Constant loudness levels in phons

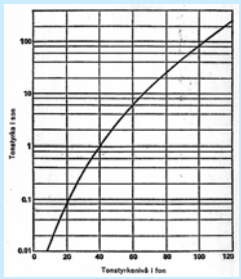


Demo: SPL and loudness (phons)

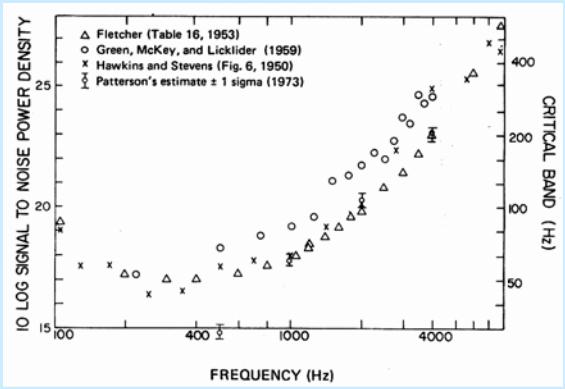
- 1) 50-100-200-400-800-1600-3200-6400 Hz
  - 1a: constant SPL 40 dB, 2*
  - 1b: constant 40 phons, 2*
- 2) 125-250-500-1000-2000-4000-8000 Hz
  - Decreases by 5dB in 10 steps at each freq.
  - Count how many steps you hear at each frequency

Relationship phon - sone





Demo: 0,01-0,1-1-10 son (2)*

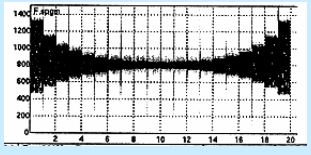


Critical bands

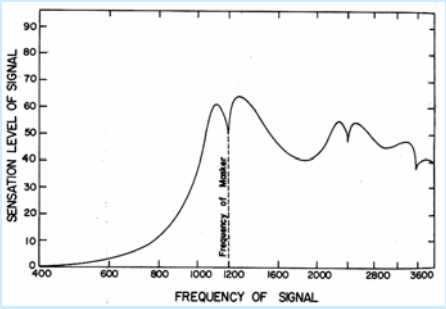
- Bandwidth increases with frequency
  - 200 Hz (critical bandwidth 50 Hz)
  - 800 Hz (critical bandwidth 80 Hz)
  - 3200 Hz (critical bandwidth 200 Hz)

Critical bands demo

- Fm=200 Hz (critical bandwidth 50 Hz)
  - B= 300,204,141,99,70,49,35,25,17,12 Hz
- Fm=800 Hz (critical bandwidth 80 Hz)
  - B=816,566,396,279,197,139,98,69,49,35 Hz
- Fm=3200 Hz (critical bandwidth 200 Hz)
  - B=2263,1585,1115,786,555,392,277,196,139,98 Hz

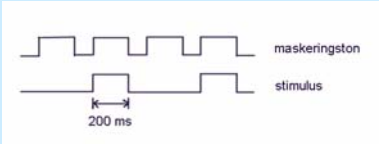


Effects of masking



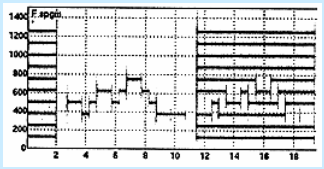
Effects of masking

- Low frequencies more effectly mask high frequencies
- Demo: how many steps can you hear?
  - a) masking tone 1200 Hz, stimulus 2000 Hz
  - b) masking tone 2000 Hz, stimulus 1200 Hz



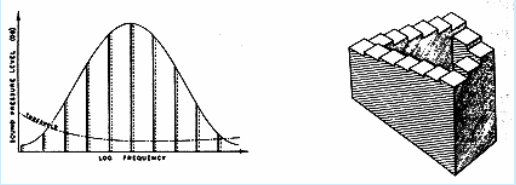
Holistic vs. analytic listening

- Demo 1: audible harmonics (1-5)
- Demo 2: melody with harmonics
- Demo 3: vowels and audible formants



Circularity in pitch

- R N Shepard
- J-C Risset
- J Liljencrants



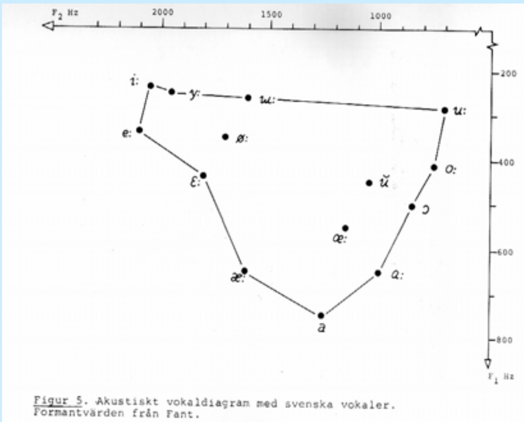
http://asa.aip.org/sound.html

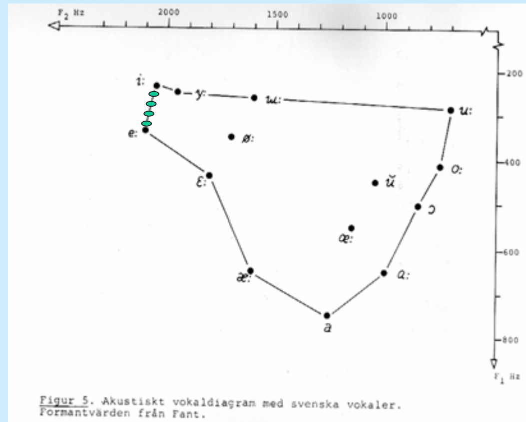
Perception of vowels

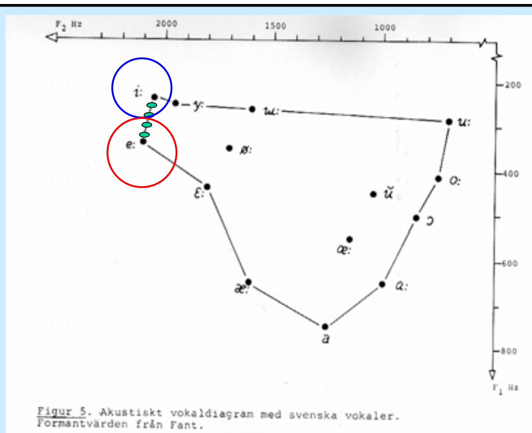
- Formants (general relationship acoustics-articulation)
  - F1: information on jaw opening
    - higher F1= more open
  - F2: information on front-back
    - higher F2=more front
  - F3: information on lip rounding
    - lower F3=more rounded

Perception of vowels

- Identification
  - Perceive which vowel is pronounced
- Discrimination
  - Hear that two vowel sounds are different
- Categorical perception
  - Difficult to discriminate within a category
  - Easy to discriminate between categories

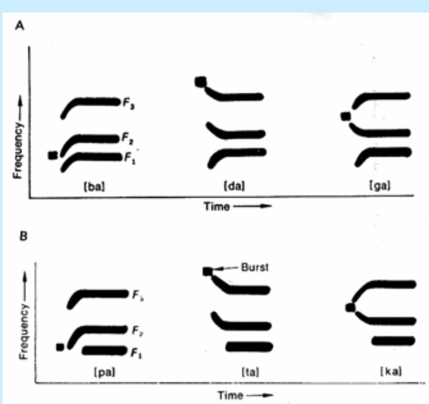






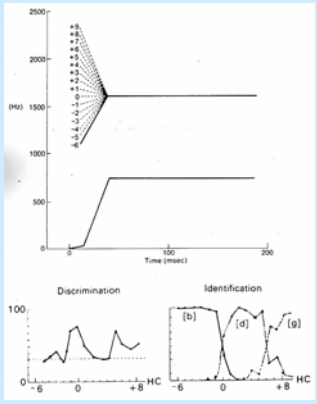
Perception of stops

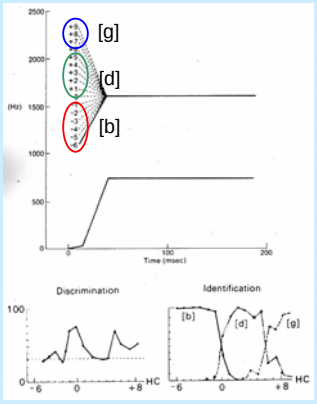
- Frequency of the burst release
  - Provides information on place of articulation
- Formant transitions in adjoining vowels
  - Also information on place of articulation
- Voiced occlusion or aspiration
  - Provides information on manner of articulation



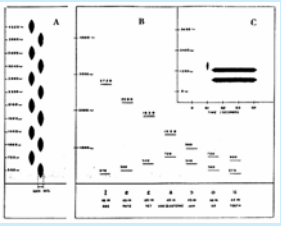
Perception of stops

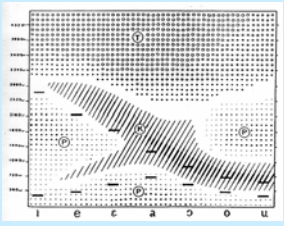
- Early experiments with speech synthesis
  - Formant transitions alone were sufficient to identify place of articulation (ba-da-ga)
  - Identification and discrimination of stops
- Categorical perception of stops
  - Difficult to discriminate within a category
  - Easy to discriminate between categories





Early perception experiment: noise burst preceding different vowels

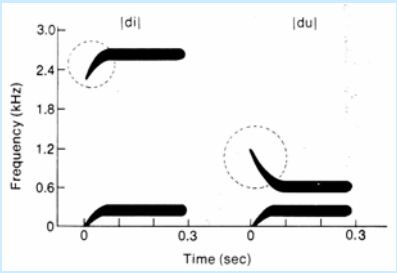




stimuli results

Liberman et al. JASA 24, 1952, 599

Invariance and segmentation problem



Invariance and segmentation problem

- The same phoneme has different cues in different contexts, e.g. F2-transitions for [di] [du].
- Where are the segment boundaries?
- Problem is a result of coarticulation
- Problem has inspired the classic perception theories

Classic theories of speech perception

- Invariance theory
  - The acoustic signal is the most important (invariant)
- Motor theory
  - Speaker's nerve impulses for speech motor control are calculated by the brain by analysing the acoustic signal.
  - Articulation is the most important
- Direct perception
  - The speaker's articulatory movements are directly perceived by the listener

Cognitive theories

- Top-down speech processing
  - Expectation and linguistic knowledge set the frame
  - Incoming words are compared to hypotheses
- Bottom-up processing
  - Acoustic signal is transferred to words
  - Message formed from words

Psycholinguistics

- The mental lexicon
- "Top-down" perception and context
  - experiments with filtered speech
  - experiments with phoneme detection (e.g. [s])
    - "They had been up all night and needed to sleep"
    - "They didn't know if they would be able to sleep"

Speech acquisition theories

- Innate
  - Possible psychophysical limits
  - e.g. the number of vowels that can be discriminated
- Acquired
  - Language-specific categories
    - Several high, front vowels in Swedish: language categories develop making use of psychophysical limits
    - One high front vowel in Japanese: category differences are lost