

Acoustics:	<u>Fundamental</u> (harmonics)	<u>Formants</u>
Perception:	Pitch	Timbre

Sound patterns as seen on spectrographic displays

An often used graphical representation of speech

- X Axis: time
- Y Axis: frequency
- Z Axis (grey scale): intensity

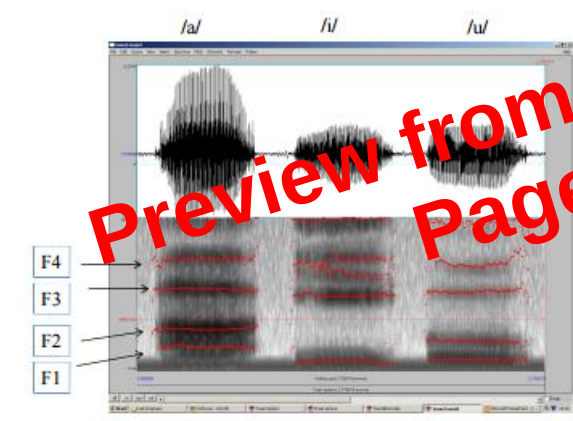
i.e. On a spectrogram, a darker trace corresponds to more intensity in the signal at that particular frequency, so would correspond to a higher peak in the spectrum.

Speech patterns in spectrograms:

- Spectrograms are produced using a set of band-pass filters that approximate how the ear analyses sounds
- Many of the speech patterns seen on spectrograms are relevant for how we perceive speech sounds in terms of their voicing, place and manner of articulation

Characteristic speech patterns for different classes of sounds (brief review):

Vowels



- Formants as seen in spectrograms are bands of high energy. They can be tracked automatically in software such as Praat or SFS, but not always accurately

Monophthong vs Diphthongs

- In monophthongs, the vocal tract configuration stays the same throughout the sound, so formants are steady. Therefore, there is no change in timbre
- In diphthongs, articulators move during the production of the vowel, so formants change in frequency, which results in changes in timbre

Plosives

- Articulation
 - o Involves a complete obstruction (manner of articulation) at a given place along the vowel tract (place of articulation)

Lecture 3 – Effect of Age on Speech Production

This lecture:

- What do we mean by 'development'?
- Stages of development in infant production in the first year of life
- Impact of social interaction on development
- Can infant vocalisations be used to diagnose language or other disorders?
- Later stages of development of speech production in older children
- Factors affecting speech production in adulthood
- Impact of ageing on speech production

Traditional views of speech development

- Based around concept of 'adult norm'
- Strong focus on early developmental stages
 - o Infant research
 - o Pre-school age
- What is the 'adult norm'?
 - o Typically: young adult university students

Why does this traditional view need to be reviewed?

- Ongoing developments occur into late adolescence
- Strong influence of social interaction on speech development
- Evidence of changes in perception/production throughout lifespan

Physiological constraints on early infant vocalisations

- Overlapped larynx and nasopharynx at the back of the vocal tract cause nasalised vocalisations in very young infants
- Small vocal tract cavity restricts tongue movement
- Larynx raised

Physiological constraints affect vocalisations especially until 7 months. By 7 months, rhythmic jaw oscillations emerge, so reduplicated canonical syllables become possible.

Methodological considerations – evaluating vocalisations in infants

Typical technique:

- Recording of infant's vocalisations at regular intervals
- Painstaking manual classification of:
 - o Types of vocalisations
 - o Sounds produced during vocalisations
- Checks of inter-transcriber consistency

What are the characteristics of early phonation in young infants? (*Burder et al., 2008*)

- Predominance of 'modal' phonation
- Many 'non-modal' types of infant phonation

- Reduction in temporal overlap when talking in presence of modulated maskers
- Shifts in spectral energy according to frequency of background noise
- Linguistic barriers
 - Infant-, child-, foreigner- and pet-directed speech show different combinations of acoustic-phonetic and linguistic characteristics

Real communications in difficult conditions: *Granlund (2015)*

Do normal hearing (NH) children make adaptations when speaking to hearing impaired (HI) children?
Are HI children able to make adaptations too when speaking to HI peers despite their delayed speech development?

- Speech produced with communicative intent (dialogue)
- 'Ecologically-valid' situation: problem-solving task imposing a certain cognitive load
- Involves speakers who are used to regularly communicating with HI peers
- Direct comparison of NH and HI children on same task

Grid task (*Granlund, 2015*)

- To elicit several repetitions of keywords in a communicative situation
- Production and perception of target sounds
 - Problems will be more explicit
 - Problems will need to be resolved
- To elicit misunderstandings

2. Grid task



Global acoustic-phonetic strategies:

- Hi speakers:
 - Speak more slowly
 - Speak more loudly
 - Have a wider F0 range than NH speakers
- In HI directed speech, both NH and HI speakers:
 - Slightly decrease speech rate
 - Increase the intensity of their speech (more so for NH speakers)
 - Increase F0 range

Lecture 8 – ‘Perceptual Warping’ and Linguistic Effects

Summary of how speech perception is challenging...

- In short, phonetic cues are spread out over time, are highly variable and can be obscured by noise

How we meet the challenges:

Phonetic Information > Perception > Lexicon

- This session:
 - o Use of linguistic structure
 - o ‘Perceptual warping’ of the acoustic cues
- Last session:
 - o Use of redundant multimodal phonetic cues

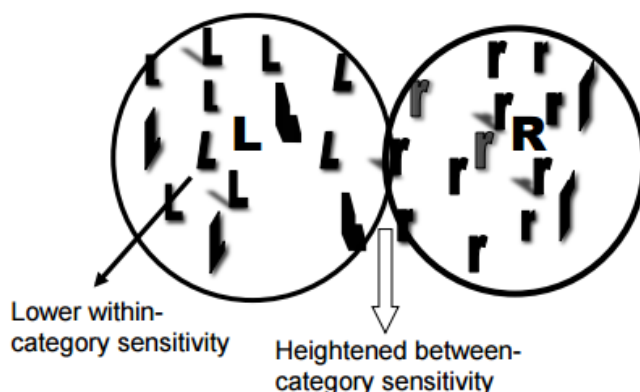
How we meet the challenge of speech perception?

Part 3: Perceptual Warping

Speech has high phonetic variability

- Some of this is “good” variability
 - o It can cue the differences between phonemes
 - o e.g. The different burst frequencies of /b/ and /d/
- Some of this is “bad” variability
 - o A lot of variability is irrelevant to phonemic categorization
 - o e.g. Differences in formants between adults and children
 - o Some of this “bad” variability is actually useful for other things (e.g. indexical cues)
- Fortunately, our auditory system is tuned to be highly sensitive to the “good” variability and less sensitive to the “bad” variability

Notion of perceptual warping



- Equal physical steps perceived as unequal
- Certain differences have much greater effect on perception than others
- Tested using *discrimination experiments*

Sentence rating for L2 speakers with different ages of arrival - Flege (1998)

- Lose the ability to learn a second language at puberty
- Loss of plasticity is gradual
- Adults can continue to learn
 - o Long-term experience can help
 - o Computer based training works
- Possible social factors
- Some (?) changes in normal plasticity
- And L2 phonological systems

Do all second language phonemes get more difficult with age? Best et al. (1988)

Study on Zulu Click Discrimination by English-Speaking Adults and Infants

Table 4
Mean Percentage of Correct Performance on the Minimal-Pair Click Contrasts

Feature contrast	Contrast type		
	Place of articulation		
	Apical vs. palatal	Apical vs. lateral	Palatal vs. lateral
Voicing category			
Voiceless unaspirated	97.7	80.6	95.8
Voiceless aspirated	97.2	82.9	94.4
Voiced	92.6	86.6	96.8
	Voicing		
	Voiceless aspirated vs. unaspirated	Voiced vs. voiceless aspirated	Voiceless aspirated vs. voiceless unaspirated
Place category			
Apical	82.4	90.1	88.0
Palatal	88.8	88.8	89.4
Lateral	89.4	92.7	89.4

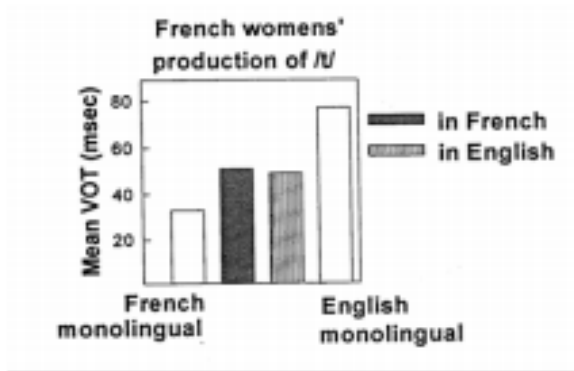
More broadly...

- Spanish speakers have particular difficulty with English /i/-/I/ even though there are many new vowels in English
- Japanese speakers have a very marked difficulty with English /r/-/l/
- Anyone who has tried to learn an L2 finds that some sounds are harder than others...

Interference between L1 and L2 phonemes

Hazan et al.

In English, voicing contrast is longer than in French. In French, the /d/ is often pre-voiced, and the /t/ is often aspirated (i.e. shorter VOT).

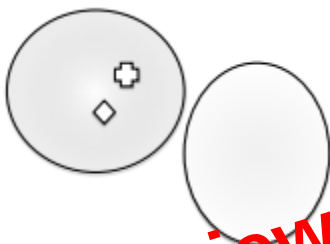


Categorical Perception

- Categorical perception more accurate at discriminating stimuli at boundaries than within categories
 - o Almost as if we perceive the stimuli in terms of their category labels
- Opposite of *perception*
 - o Same sensitivity for all acoustic differences

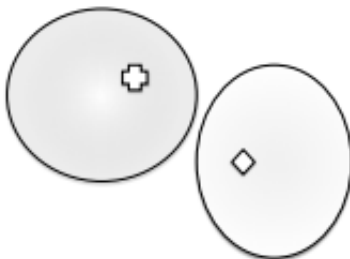
Assimilation types: Single Category

Discrimination: Poor



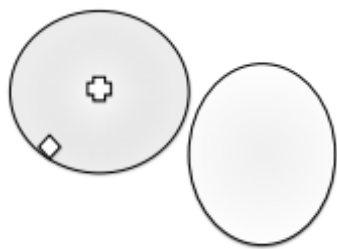
Two Category

Discrimination: Excellent



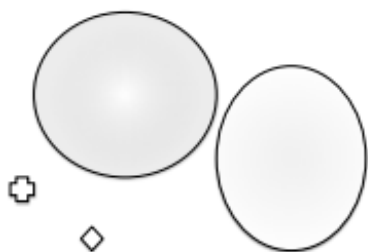
Category Goodness

Discrimination: Moderate to very good

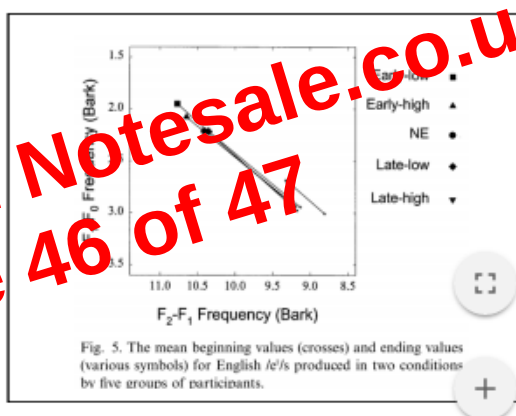
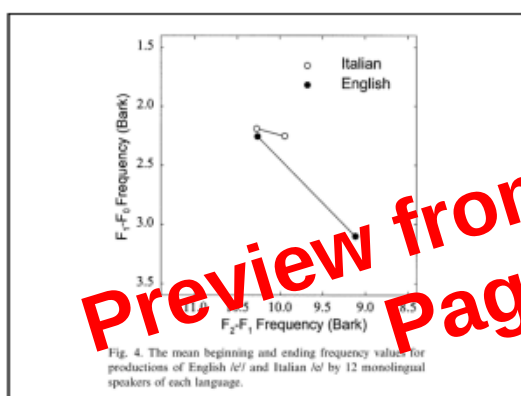


Non Categorical

Discrimination: Poor to very good



Interaction between the native and second language phonetic subsystems – *Flege et al.*



Effect of length of L2 exposure

Adults able to learn quite quickly, but slower than children, as the brain is less plastic.

See: *Lively et al. (1993)*

and *Pisoni et al. (1996)*

- Naturalistic variability and ID training improves...
 - o Identification of stimuli from new talkers and words in multiple phonetic positions
 - o Production
 - o And improvements are retained over time

CONCLUSIONS