More on Speech Perception

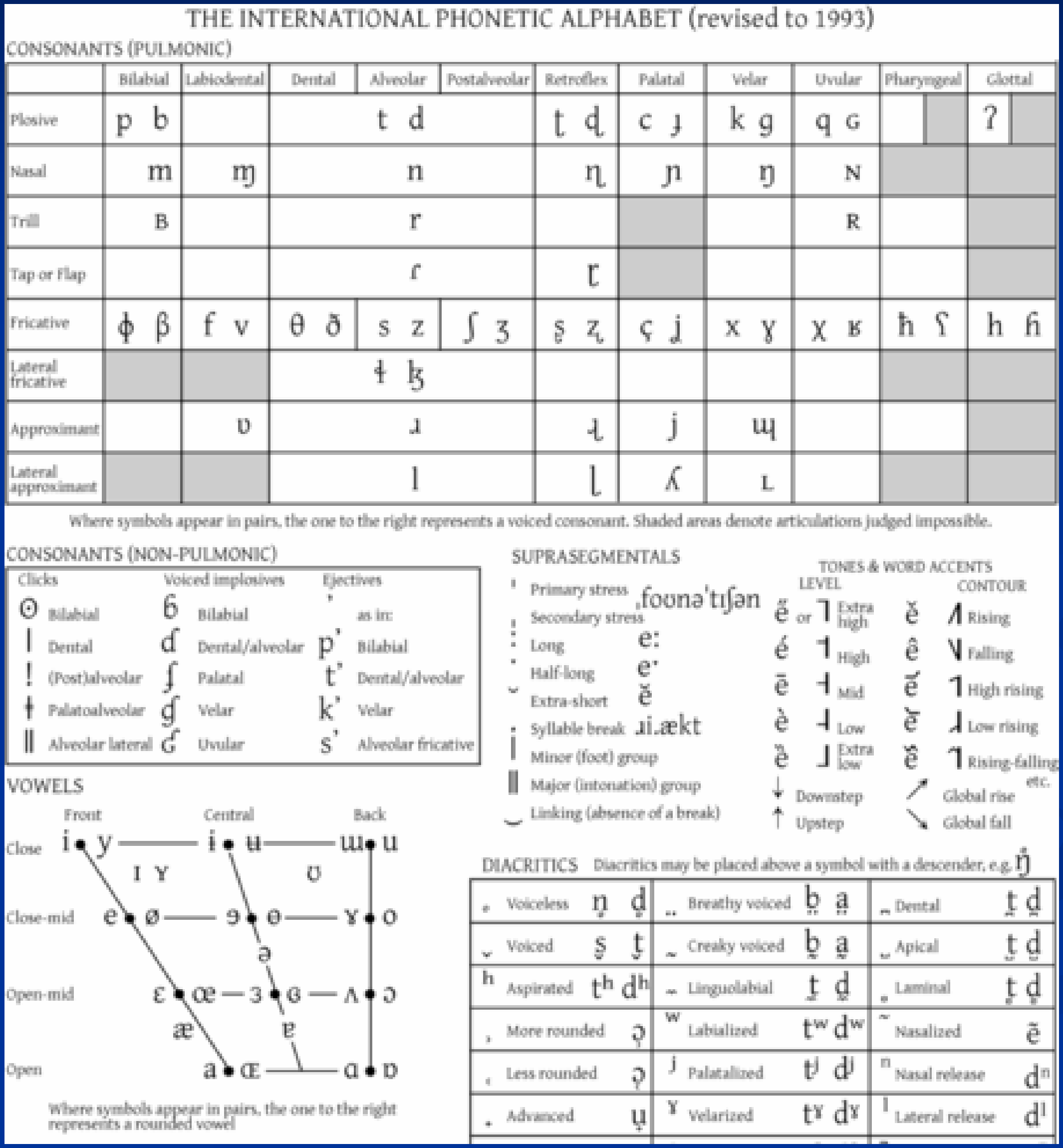
Phoneme Discrimination

LDER Chapter 2

What is a phoneme?

- Smallest unit of language that signals a change in meaning
  - “pat” vs “bat”
- An abstract representation of actual sounds (phones)
  - Different instantiations of a phoneme are *allophones*
    - e.g., “water” -- [t], [?], or [ɾ]

Infants discriminate all sounds



THE INTERNATIONAL PHONETIC ALPHABET (revised to 1993)

CONSONANTS (PULMONIC)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap				ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

CONSONANTS (NON-PULMONIC)

Clicks	Voiced implosives	Ejectives
ʘ Bilabial	ɓ Bilabial	ʼ as in:
ǀ Dental	ɗ Dental/alveolar	pʼ Bilabial
! (Post)alveolar	ʄ Palatal	tʼ Dental/alveolar
ǂ Palatoalveolar	ɠ Velar	kʼ Velar
ǁ Alveolar lateral	ʛ Uvular	sʼ Alveolar fricative

SUPRASEGMENTALS

ˈ Primary stress ˌfoʊnəˈtɪʃən
ˌ Secondary stress
ː Long eː
ˑ Half-long eˑ
˘ Extra-short ĕ
. Syllable break ɹi.ækt
| Minor (foot) group
‖ Major (intonation) group
‿ Linking (absence of a break)

TONES & WORD ACCENTS

LEVEL

e̋ or ˥ Extra high
é ˦ High
ē ˧ Mid
è ˨ Low
ȅ ˩ Extra low
↓ Downstep
↑ Upstep

CONTOUR

ě ˩˥ Rising
ê ˥˩ Falling
e᷄ ˧˥ High rising
e᷅ ˩˧ Low rising
e᷈ ˧˦˧ Rising-falling etc.
↗ Global rise
↘ Global fall

VOWELS

Front, Central, Back

Close: i • y — ɨ • ʉ — ɯ • u

ɪ ʏ ʊ

Close-mid: e • ø — ɘ • ɵ — ɤ • o

ə

Open-mid: ɛ • œ — ɜ • ɞ — ʌ • ɔ

æ ɐ

Open: a • ɶ — ɑ • ɒ

Where symbols appear in pairs, the one to the right represents a rounded vowel

DIACRITICS Diacritics may be placed above a symbol with a descender, e.g. ŋ̊

̥ Voiceless n̥ d̥	̤ Breathy voiced b̤ a̤	̪ Dental t̪ d̪
̬ Voiced s̬ t̬	̰ Creaky voiced b̰ a̰	̺ Apical t̺ d̺
ʰ Aspirated tʰ dʰ	̼ Linguolabial t̼ d̼	̻ Laminal t̻ d̻
̹ More rounded ɔ̹	ʷ Labialized tʷ dʷ	̃ Nasalized ẽ
̜ Less rounded ɔ̜	ʲ Palatalized tʲ dʲ	ⁿ Nasal release dⁿ
̟ Advanced u̟	ˠ Velarized tˠ dˠ	ˡ Lateral release dˡ

Adults discriminate sounds in their language

THE INTERNATIONAL PHONETIC ALPHABET (revised to 1993)

CONSONANTS (PULMONIC)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d				k g			
Nasal	m			n				ŋ			
Trill											
Tap or Flap				ɾ							
Fricative		f v	θ ð	s z	ʃ ʒ						h
Lateral fricative											
Approximant				ɹ			j				
Lateral approximant				l							

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

What happened?

CONSONANTS (PULMONIC)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap				ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Where symbols appear in pairs, the one to the right represents a voiced consonant. S

THE INTERNATIONAL PHONETIC ALPHABET (revised to 1993)

CONSONANTS (PULMONIC)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d				k g			
Nasal	m			n				ŋ			
Trill											
Tap or Flap				ɾ							
Fricative		f v	θ ð	s z	ʃ ʒ						h
Lateral fricative											
Approximant				ɹ			j				
Lateral approximant				l							

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

The question:

- Infants distinguish minimally different phonemes
  - Even those not found in their native language
- Adults do not appear able to distinguish minimally different phonemes that are not in their native language
  - But adults are usually better than infants at tasks!
- What is responsible for this change from infant perception to adult perception?

How can we examine this?

- Need a procedure that is adaptable for both infants and adults!
  - Adults are not so good at sucking
  - Infants can’t push buttons
- Monitor continuous stream of syllables
    - /ba/ /ba/ /ba/ /ba/ /da/ /ba/ /ba/ /da/ /ba/ /ba/ /da/
  - For adults and older children
    - Press a button whenever you hear target: /da/
  - For infants
    - Conditioned head turning paradigm

Conditioned head turning

- Assistant shows infant toys to keep them occupied
- Speech stimuli presented over speaker
- Infant trained ("conditioned") to turn head towards speaker when stimulus changes
  - when infant turns head for a change, re-inforcer activates, displaying animated animals

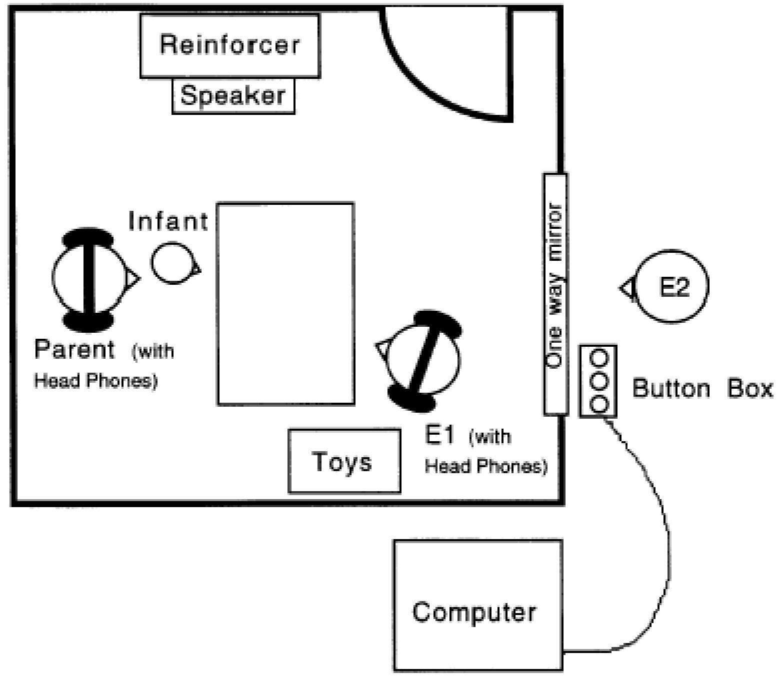
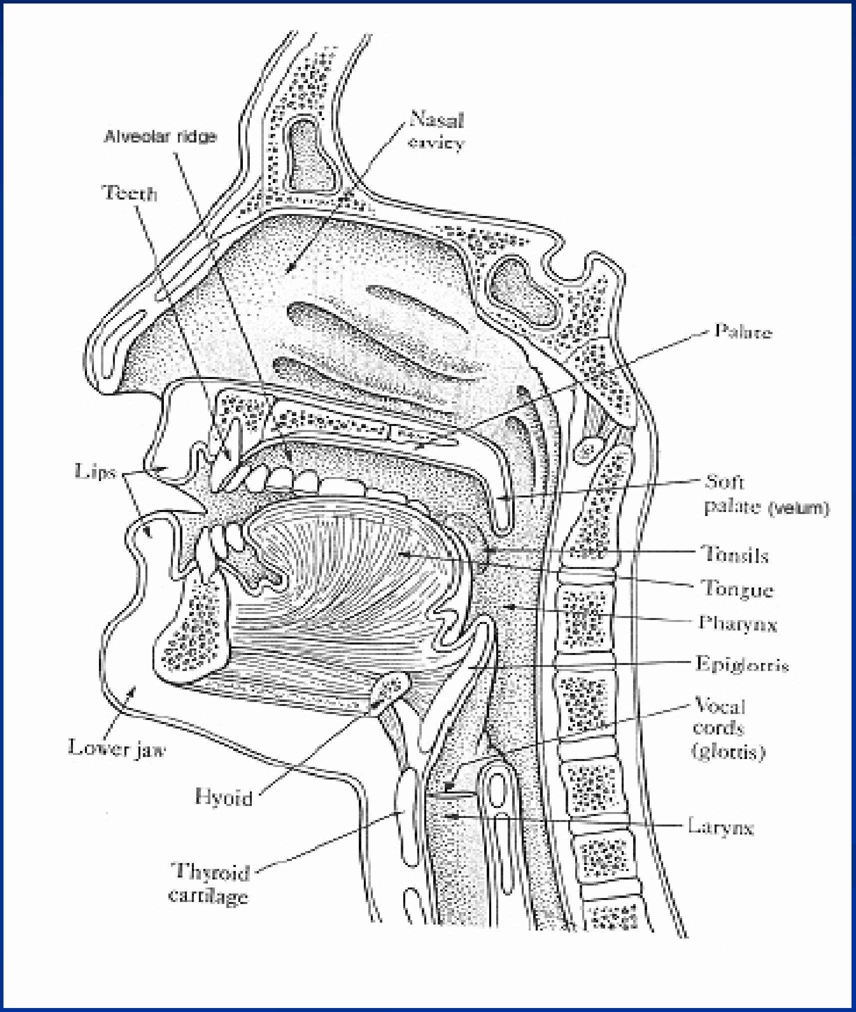


Figure 1. A schematic diagram of the Conditioned Head Turn procedure.

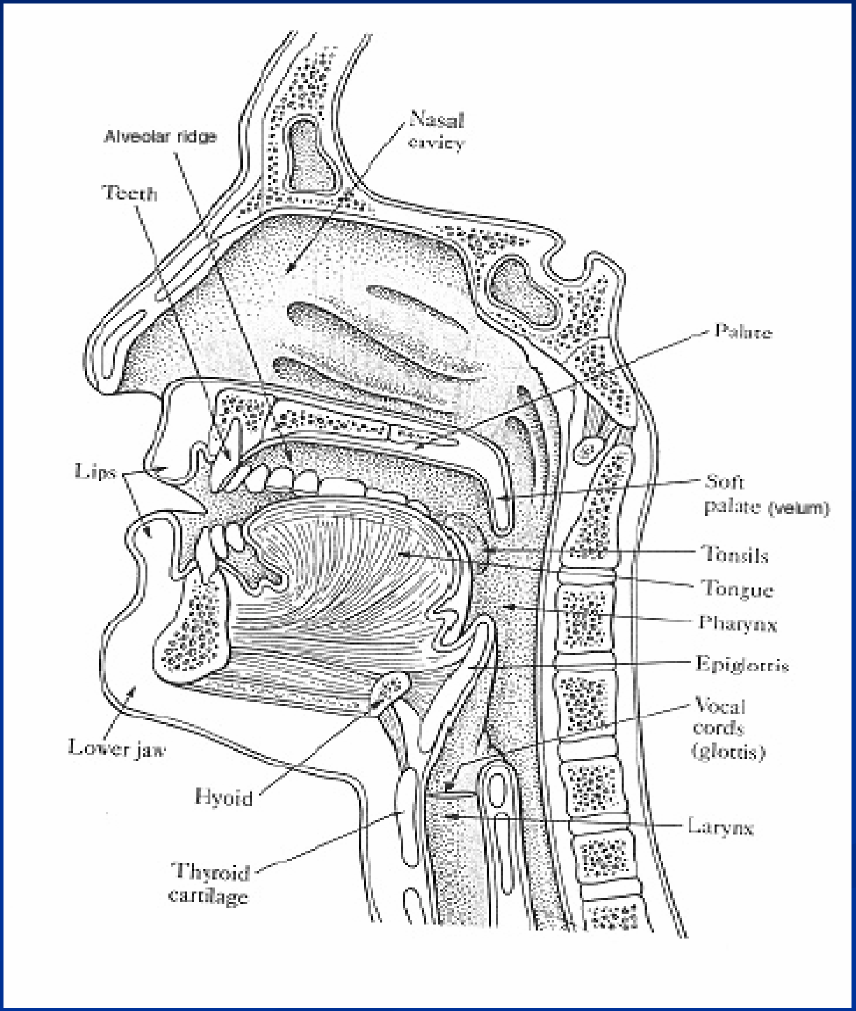
What did they test?

- First Contrast
  - Hindi /Ta/ vs. /ta/
  - /Ta/ has retroflex stop
  - /ta/ has a dental one
  - Both sound like /ta/ to English speakers
    - English /t/ is alveolar



What did they test?

- Second Contrast
  - Hindi /t^h/ vs. /d^h/
  - /t^h/ is voiceless aspirated
  - /d^h/ is voiced aspirated
  - Both sound like /t/ to English speakers
    - English /t/ is aspirated



Dissecting a stop consonant

- Silence
- Burst (release/aspiration)
- Vowel

- When does the vowel start?
- Voice onset time (VOT)
  - Voiced stops: 10-30 ms
  - Voiceless: 40-100 ms

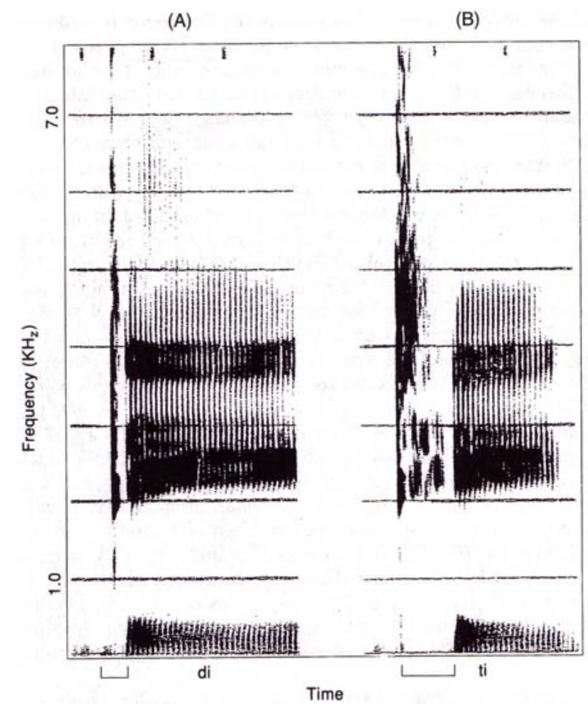
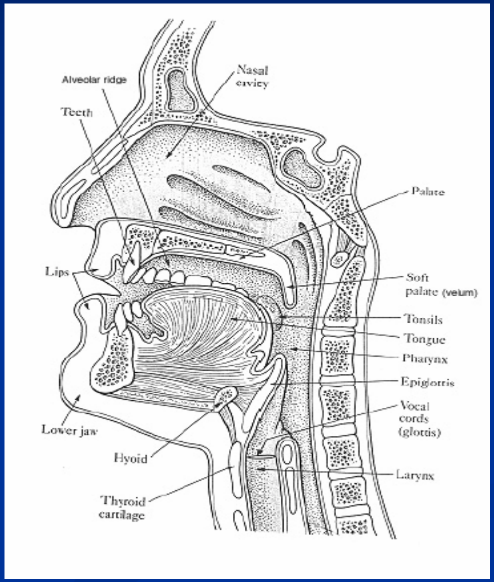


FIGURE 3.8

Spectrograms of the syllables [di] and [ti] to illustrate voice-onset-time (VOT). In each syllable the burst is seen as the dark vertical line to the left. The onset of voicing is indicated by the onset of the first formant (start of the segments that have regular striations). VOT is the distance between the burst and the onset of voicing (horizontal bracket starting at the burst of each syllable).

What did they test?

- Third Contrast
  - /ba/ vs. /da/
  - Difference in phonemic in both Hindi and English
  - Cued by place of articulation difference



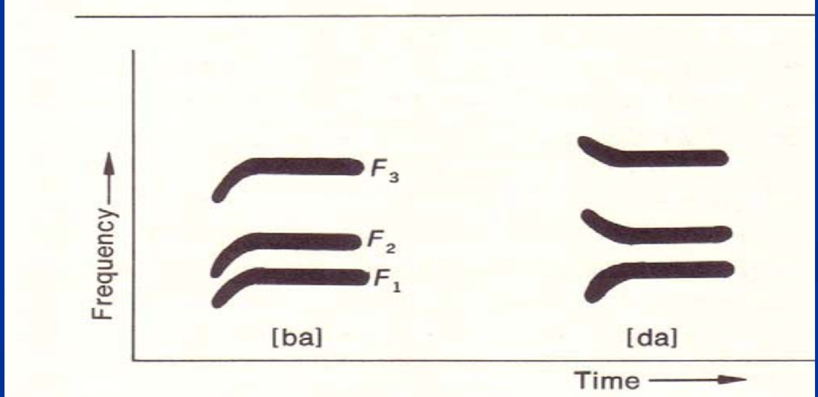


FIGURE 7-1.
Formant frequency patterns for the sounds [ba], [da],

Who did they test?

- English learning infants (aged 6-8 months)
- English speaking adults
- Hindi speaking adults

- Is this sufficient?
- Why not test Hindi-learning infants?

Results

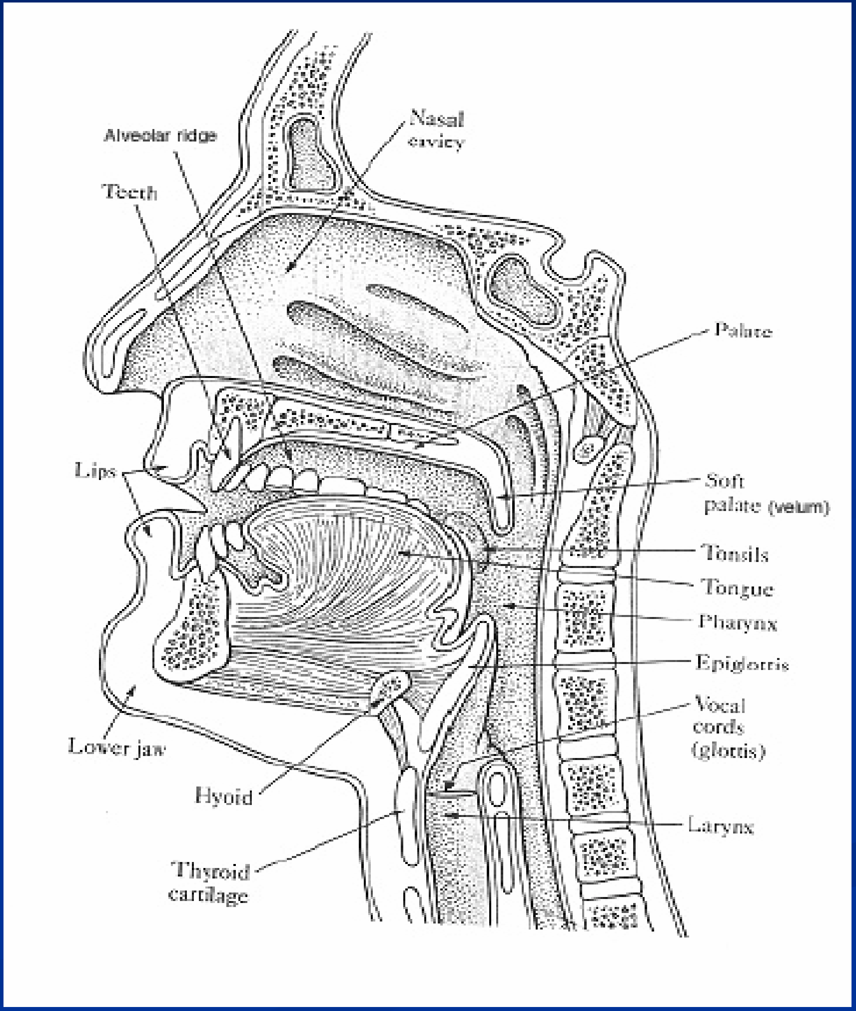
- All three groups could discriminate /ba/ from /da/
- English learning infants and Hindi speaking adults could discriminate the two contrasts found in Hindi but not English
- English speaking adults had trouble with the two Hindi contrasts
  - After training, improved on voicing contrast but not retroflex/dental contrast

When do infants grow up?

- At what age does this change?
  - "Critical Period Hypothesis"?
  - Much earlier!
  - English speaking children aged 12, 8 and even 4 showed same pattern as English speaking adults
  - Hindi children aged 4 could discriminate Hindi contrasts when tested with the same paradigm

Before 4?

- Infants between 6 and 12 months tested:
  - On retroflex/dental Hindi contrast
  - On a new contrast from Nthlakampx (aka Thompson)
    - Glottalized velar /k'/
    - Glottalized uvular /q'/
    - Both sound like /k/ to English speaking adults



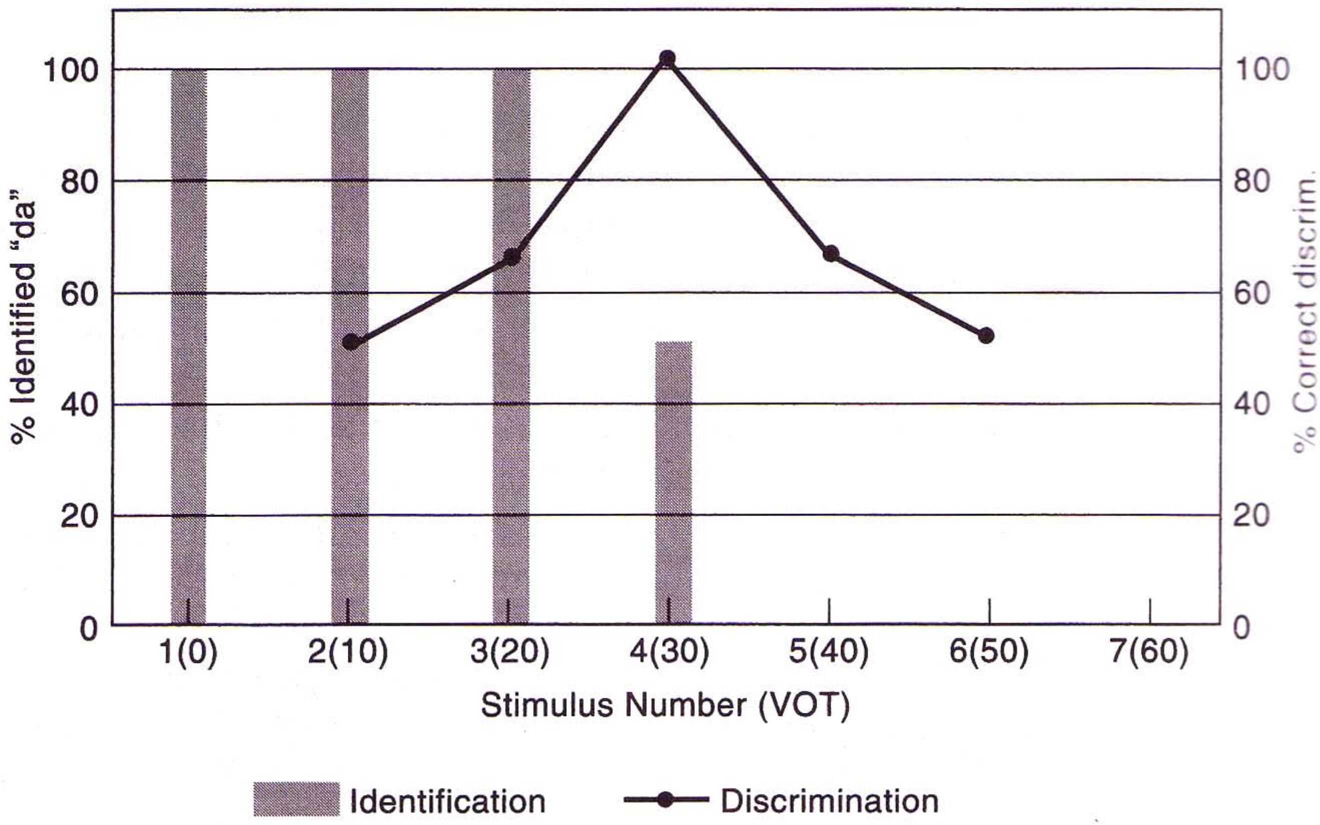
Between 6 and 12 months

- English learning infants could discriminate both contrasts
  - At 6-8 months old
  - But not at 10-12 months old
- What about Hindi and Nthlakampx infants?
  - 11-12 month old infants in both groups could distinguish contrasts in their native language
- Perceptual loss not just an aging effect
- Reflects language-specific experience!

Categorical Perception

FIGURE 3.9

Identification (bars) and Discrimination (line) functions for alveolar stop consonants varying on voice-onset-time (VOT). The illustration is for a 7-step continuum, varying from [da] to [ta]. The horizontal axis contains the number and VOT value (in brackets) of each stimulus in the continuum. The left vertical axis is for the identification results and is expressed in percent identified as "da." The right vertical axis is for the discrimination results and is expressed in percent correct.



Discrimination is % different over 20 ms intervals (1 vs 3; 2 vs 4; 3 vs 5; etc.)

Categorical Perception: /ba/ vs /da/

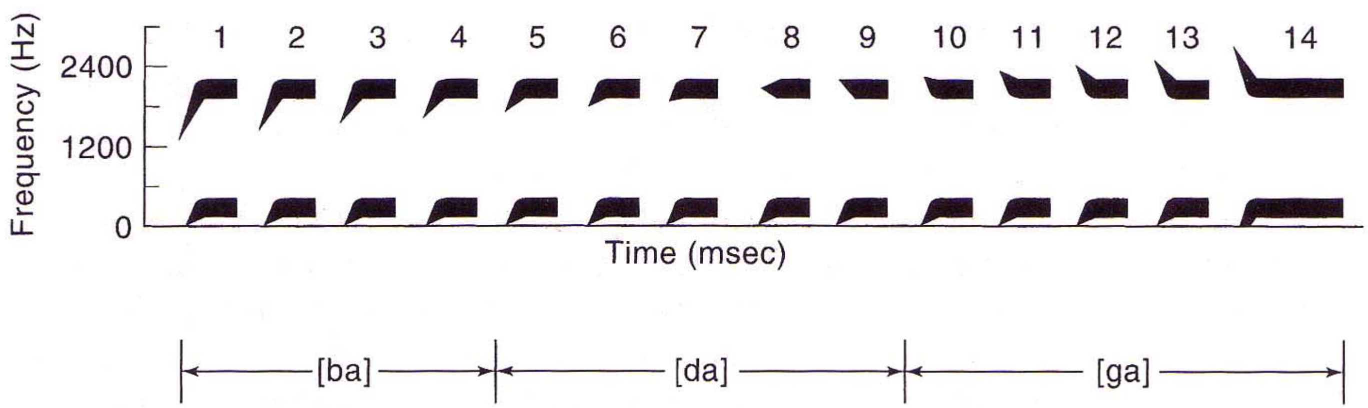


FIGURE 3.10

Two formant syllables produced on the Pattern Playback synthesizer. The illustration is for a 14-step continuum. Each stimulus is numbered. The stimuli vary in the onset and direction of the second formant transition. This acoustic cue signals place of articulation for the stop consonants [ba], [da], and [ga]. The divisions marked by arrows at the bottom of the figure mark the phoneme boundaries.

- Formant transitions vary continuously as a function of place of articulation

Infant Categorical Perception

- Create continuum of sounds (artificially) that varies between
  - Bilabial /ba/ - dental /da/ - retroflex /Da/
  - English learning infants aged 6-8 months
    - Distinguished proper boundaries between
      - /ba/ and /da/; /da/ and /Da/
  - English learning infants aged 10-12 months
    - Distinguished boundaries between
      - /ba/ and /da/ but NOT between /da/ and /Da/

What's the explanation?

- Maintenance/Loss
  - Only phonemic contrasts present in the native language will be maintained, others are lost permanently
    - Loss may reflect developmental changes in the brain
  - Maybe this is too strong…

Do adults remember anything?

- Perceptual Assimilation Model
  - Non-native contrasts that assimilate into a single native category are lost
    - Hindi /t/ and /T/ both map to English /t/
  - Non-native contrasts that don't assimilate well into a native category may be easier to discriminate
  - Non-native contrasts that are not remotely close to native categories should be well discriminated

Zulu

- Zulu clicks
  - [|] (tsk-tsk)
  - [||] (horse sound)

- Adults (and infants of all ages) easily discriminate Zulu clicks
  - but only younger infants discriminate Zulu contrasts that are closer to English sounds
- Loss of phonemic contrast discrimination may reflect experience after all, and not necessarily brain-specific changes

What about vowels?

- In German contrast between:
  - /but/ and /büt/
    - (high back rounded vs. high front rounded)
  - Adults make this discrimination easily
  - For infants, experience seems to play a role earlier for vowels than for consonants
    - 6-8 month old infants discriminate the vowels, but not as well as they discriminate non-native consonant contrasts
- Vowels may be somewhat different than consonants

Summary

- Within the first year of life infants –
  - are learning the phonemes of their language
  - grouping them into categories
    - grouping phone into phonemes
  - become less well able to discriminate non-native phonemic contrasts
    - For consonants when non-native sounds are similar to native ones
    - For vowels at a slightly earlier age

Language Discrimination

LDER Chapter 3

What’s special about babies?

Is there a difference?

- Human infants
  - Discriminate phonemes categorically
  - Are sensitive to the rhythm of speech
  - Process natural speech differently than backwards speech
- Other species
  - insects, birds, primates, mammals
  - Perceive their own species-typical sounds categorically
  - Some also perceive human speech categorically

How to test for a difference?

- Same problem as for testing infants vs adults, but worse
- Need a paradigm that is comparable for both babies and monkeys
  - Babies: high-amplitude sucking
  - Monkeys: head-orientation response (similar to conditioned head turn paradigm)

Language Discrimination Task

- Tested Japanese vs Dutch sentences
- Test sentences were read by 4 native (adult female) speakers of each language

- Contrasted
  - Language: Japanese vs Dutch
  - Speaker: within each language
  - Forwards vs. backwards speech

Human Infants

- High-amplitude sucking procedure
- Native French speaking infants
  - Language change
    - Habituate to 2 speakers of one language (Japanese or Dutch)
    - Switch to 2 speakers of the other language
  - Speaker change
    - Habituate to 2 speakers of one language
    - Switch to the other 2 speakers of the same language
- Greater increase in sucking for language change than speaker change indicates newborns distinguish the two languages

Human Results

- The infants did not discriminate the two languages…
  - But, shouldn’t they have?
  - Yes, but…
  - Speaker variability seems to impair language discrimination ability of infants (this susceptibility goes away after a few months)
- With only a single (synthesized) voice
  - Preserves prosody; removes some phonetic detail
  - Infants did discriminate the two languages!
  - But only forwards, not backwards
    - Backwards speech may eliminate cues necessary to distinguish the two languages

Cotton Top Tamarins

- Head orientation response
- Native Cotton-Top speaking Cotton-Tops
- Tested initially on their own species specific vocalizations, to ensure that test procedure worked
- Tested on same language stimuli as human infants
- Habituation – recovery of head orientation to loudspeaker indicative of detection of difference

Monkey Results

- With natural speech
  - Tamarins dishabituated in the
    - language change condition
    - more than in the speaker change condition
- With synthesized speech
  - Tamarins did not dishabituate more for language change than speaker change

  - But only for forwards speech!
  - Language change not detected with backwards speech

Summary

- Both human infants and cotton-top tamarins could distinguish Japanese and Dutch
- Speaker variability problematic for young infants but not cotton-top tamarins
  - Monkeys able to extract abstract linguistic information from a variable natural signal (babies catch up)
  - Monkeys handled synthetic speech less well than human infants (tamarins more sensitive to phonetic than prosodic contrasts?)
- Inability to distinguish languages when played backwards same for humans and monkeys
  - Suggests sensitivity to important aspects of speech
  - Low level details similar forwards and backwards

What about backwards?