05/08/24

Articulatory phonetics

Phonology

two things when we talk about speech.

perception

Hearing

Acoustic phonetics → very close to speech perception.

instruments are used

when we have to talk about phonology.

correlation with mental organisation

methodologies are diff.

we ask for judgement / impression etc.

do you hear this or that

voicing.

vibration of the vocal cords of the larynx.

not a linguistic term

fundamental frequency associated (f_0)

modulation → frequency changes

pitch → is not an absolute no. (what we hear)

→ psychoacoustic impression

low pitch high pitch

fundamental frequency.

→ is an absolute no.

Some people have a very good perception of pitch.

freq.

→ not static

when we start ↑ when we end the ↓ sentence

vocal tract → resonating chamber.

(vocal cord)

analogy strings.

strings of an instrument only are not useful without a resonating chamber which amplifies the sounds when the strings vibrate.

vowels have diff. bands of frequency.

when we change the shape of the mouth.

frequency gets highlighted/changed.

closest approximation of human speech to the sinusoidal waveform is the child's speech. (rarely)

Normally human speech is not sinusoidal.

combination of diff. sine waves → complex sine waves.

larynx { vibrates at a produces certain periodicity. freq.

we look at a spectrogram on acoustic phonetics

stacking diff spectrum.

looking at the change over time

Anything which is voiced has periodicity.

Formants } (for vowels)

230 Hz.

270

3000 Hz

result of the resonating chamber.

quasi periodic

more than fundamental freq.

lowest harmonic → fundamental freq.

all harmonics upto ∞ integer multiple of fundamental freq.

Vocal tract decides which freq. will stand out. the harmonics which will be highlighted.

similar for voiced consonants.

eg aperiodic → noisy.
(no pattern)
fricatives

Spectrogram → also called as visible speech.

26/09/24

quasi periodic (* some harmonic basis.
sound waves
↳ not a regular repeating pattern

• Aperiodicity
↳ no harmonic basis
↳ not multiples of fundamental frequency.
integer

diff cavity of the vocal tract → diff natural frequencies
(resonance) amplification happens.
when matches with the external freq.

filtering the frequencies ⇒ idea of band pass filter.
happens.
(not the freq. of larynx)

More repeating pattern → voiced as approximant.

No repeating pattern. → consonant
or stoppage of repeating pattern.

praat
free available software

blue line → represents pitch
yellow line → intensity.

all speech events happens in those boundaries

law boundary when a person starts speaking

law boundary when a person ends a sentence.

dark bard. → formants, vowels.

some breaks in blue lines → (noise) no pitch.

our hear mechanism will compensate and we will hear the pitch throughout

voiceless sounds → no periodicity.

fricative or aspiration.

nasals → anti resonance
↳ because of the nasal cavity.

there are resonants but are attenuated

does not vibrate, like the vocal tract (cavity) vibrates.

break for periodicity, sudden release.

voiceless stop

there's lot of carry over of sounds
nasal n → a.

assimilation
one sound influences other sounds
ex voiceless becomes voiced

raised but have periodicity (you can have both also)

raised & does not have periodicity.

Languages do not prefer to have voiced fricatives.

noised & having periodicity.

27/09/24

speech sound.

sound wave → has component frequencies

(if the speech sample has a regular repeating pattern or not. we look at

however, spectrogram has more information. → harmonics, frequencies, pitch, intensity.

Entire speech mechanism starts at the larynx.

Everything related to larynx → is called laryngial

vocal tract → filter

generation of the sound (source) → glottis.

output from the lips has a diff freq than the freq which was produced at the glottal.

some freq. get attenuated
-"- get highlighted
(get more energy)

sound sources
- periodic → glottis.
- aperiodic
  ↳ fricatives, aspiration, noisy.

vowels → quasi periodic

→ sometimes we talk about f_3.

f_3
f_2
(f_1)

lower formants give most of the information.

Formants are the harmonics which are highlighted (get more energy)/standing out from the vocal tract (resonating chamber)

f_1 in "i" is low. (constrict that cavity being narrower)

f_1 in "a" is higher.

higher the vowel, lower the f_1

f_1 of "u" is also low.

f_2 is much higher for u than i.

the other cavity is not much constrained (oral)

there are typical avg f_0 values in

in some cultures - male voice is considered to be much lower in pitch

voice can be modulated

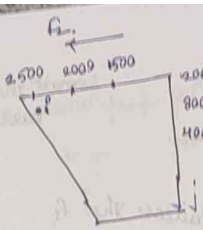
→ formant formula.

consider open tube

schwa → supposed to be most stable vowel.

most uniform/normal vocal tract ← a

e → vocal tract is more extended.



axis are reversed because
F1 is higher for lower vowels
& vice versa.

(low for high vowels)

slight separation in F1 & F2 values
on the graph

↓
æ

back vowels → F1 & F2 values come together a lot
on the graph.

(back-low vowel) ɒ

high vowel → less separation btw F1 & F2

↓
u

F1
↓
pharyngeal
or back
cavity

F2
↓
front
cavity.

(constrained
or not
constrained)
will determine.

03/10/24

connected speech
↳ entire sentence

shape of the vocal tract changes as we speak
we have diff filtering
↳ some highlighted & some attenuated

source - glottis
filter - vocal tract

Periodicity - Larynx
Aperiodicity - other resonating chambers

when we plot, we plot inversely.
so that higher the vowel, lower f_1

for singing, higher formants also become important
Ex. f_3

speaking and singing are diff respectively

some people (singers) undergo operations for the larynx because their career is dependent on it.

[a is a more neutral vowel]
changes from one vowel to another.
lengthens or contracts the vocal tract.

schwa has the most uniform cross section of the vocal tract. ∴ formula is simple to get the formants

suppose i has helmholtz resonance
diff vowels will have diff formula.

front vowels f_1 & f_2 are much separated
back vowels f_1 & f_2 are close

i highest front vowel
ɑ lowest back vowel
} acoustically very different.

04/10/24

Source filter theory.

periodicity → glottis source (synonymously used: larynx, centre part)
aperiodicity → oral source
non glottis source

lot of sounds are combination of (.)

(ex. voiced fricatives + z, v) — contrast with the voiceless

voicing → periodicity of glottis source

fricatives → noise
→ (slow release)
∴ aperiodicity.

Human sounds (complex) → {Regular + non regular} patterns

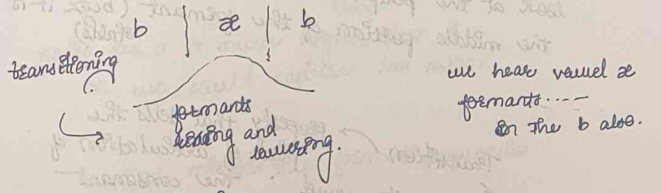
---

due to conflicting periodicity
↳ contrastive pairs may not be there in some languages because it may not be perceptible to some people.

θ ð
s z.

f_3 → also goes with F_2.
↳ tells about the back vowels

I (ɣ)
↳ rounded vowels will have lower formants than their unrounded counterparts.

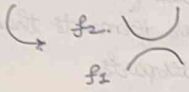


{Mirroring}

aspirated sounds → lot of noise. (similar to fricatives)

Rhotics → no clear distinction between vowel and consonants

Velar pinch



when we calculate the formants we don't look at the peripheries, we look at the middle portion of the formant (because it is stable)

Transition → property that tells the place of articulation of the consonant

consonant → vowel
release

that makes consonants perceptible

mostly F2.

coronals → having the property of F3 going up.

closure — release
gap

if we have a gap. ⇒ plosives

without the release (transition) of consonants onto a vowel we could not know the consonant (perceptible)

aspirated consonants are much longer because of the burst of release of air.

obey
noise for s is much higher than ʃ
so we could capture less noise

h → mirrors the formant frequencies
↓
fricative → not attenuated (unlike other consonants)

e ɛ are allophones in most languages but acoustically they are different.

semivowels → gliding.
- ~~[illegible]~~
↳ f_2 much higher.

f [illegible]

nasals. (attenuated)
soft mucus is not of much resonating.
(nasal tract)
→ (dampened) {unlike the vocal tract much resonating}

s → has higher frequencies

front a } diff
back a }

f → dense energy.

9/10/24

sibilant fricatives
↳ air pushed against the front teeth
↳ s S difference is the frequency one is at the high freq.
you won't be able to produce them without front teeth.

channel fricatives as non sibilant fricatives
→ air bouncing off & creating turbulence.

medal phonation
proper periodicity.
repeating patterns
↓
if not there.
↓
then. creaky phonation

some languages have different sounds for creaky phonation & non creaky phonation.

modal
creaky
breathy
stop
- phonation range

some languages have tone & property phonation
low, high
3rd distinction may be creakyness
Ex. chinese.

fo important for modal phonation.

affecting effects are diff. for nasals,

liquids → sole

glides (ya) sounds like
↳ like a vowel
but a consonant

(a university
not
an university.)

n
↳ stop & (consonants)

we can humm with all the vowels
and consonants
humming (n, mmm - - -)
↳ sounds like.

ph & h.
aspirated ↗ ↘ fricative.
source spectra. → similar

ya & v
glides
movements of the formants.

Real glides are diphthongs.

17/10/24

quasi periodic

vowel → most periodic sounds

F_2 transition → tells place of the articulation of the consonant

articulation ---→ perception

(structure of ear plays an imp role in perceptual abilities)

impressionistic judgement → qualitative study.

(what one hears/ did not hear)

units of speech perception { syllable, phoneme, phonological features (voiced/voiceless)

linguists. → study → impressionistic judgement.

others → signals

(speech technology) → TTS text to speech

sounds to produce more human like.

signals are continuous, not categorical. (no breaks)

(but are perceptual ability is such that we listen & distinguish btw words---)

for listeners, it's not a problem.

segmenting

(one phoneme sound comes into another phoneme)

(as such they are no breaks but we perceive them)

coarticulation -

production of one segment affects production of another segment

voicing starts?

before the stop is release, point of release, after the release

makes the diff. b/w p, b
t, d

p → immediately after the release, a vowel follows.

variants → f2 transition are different

DEEP DECK DOG DATE

(transition from consonant to vowel.)

change of formant at the release is diff from at the mid point (vowel is carrying some properties of consonants)

there can be variations of b

stocking tokens of b.

our brain of speech perceptual ability issue we will hear all of them as b

~~sound~~: p b ; t d

we ~~can~~ can hear p as b or vice versa depending on the ~~linguistic~~ categories in your languages ~~can~~ (but nothing in between)

(if there's no vowel following a consonant)

other ques
diff. in ampli
release duration
onset duration.

captain

primary/main que
voice onset time.

19/10/24

~~cater~~

~~leads to dl~~

what our perceptual system requires
minimum information

← cues

phonetic boundary

25-50ms

till here
we will keep on hearing the same sound

- - - -

~~listener has to recover~~.

t th

allophones in their language ←

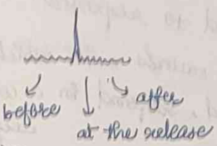
babies can distinguish b/w Hindi sounds.

but English speaking adults find it difficult
bcoz

(grown in English speaking environment)

voice onset time

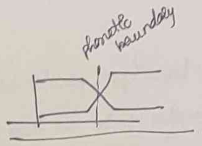
start of voicing



voicing can start

identification test

b or p.



discrimination test.
function/ experi
(more convoluted)

Same or different ← they say

(tockens) ← to hear.
listeners are given participants

{ listeners may think that the b they hear is not the b which they know }

chinchillas ⇒ shockers.

animals associate sound to response.
If they hear the sounds with which they are habituated, respond to a certain way.

23/10/24

voice onset time → unambiguous.
stable acoustic cue.

f2 transition → varies across languages (as a boundary)

French has a lot of voicing. (Also Indian languages)
before release there is voicing

- English has a limited voicing.
  English. (voiceless stops are always aspirated
  voicing may be/ not be that intense)

(generally no voicing before release) (English speakers use that aspiration as a cue.)

f2 transitions → vacuards.

Filtering system ⟷ cochlear system. (cochlea) captures all the filtering.
(lot of similarity)
there has to be a match.

In auditory phonetics we look/study at the ear.

/pa/
voicing will start for a.
vowel final consonants → may have very little release.

slavic languages
↳ very little vowels

Our perceptual system recovers the articulatory information cause
contrast is the central part/trait of a language.
otherwise how do we differentiate btw languages

b has voicing even before the release
it has periodicity (which we linguistically call it as voicing)

(range: ms)
prototypical for speaker.
stands for everything.

24/10/24

Procedures to collect responses from Infants.

response when these Infants have registered something ← Increase in the rate of moving the lips (sucking)

when they see some old they move their head, to see something. (head turn procedure)

Infants hear intermediates ex. thru p & b. changes.

they can distinguish thru language (contrasts) perceptual abilities are mindboggling for Infants.

English has robust voicing
∴ it has some additional properties such as aspirated consonants.

Some languages with less voiced consonants; aba speakers/people have this problem

but could not perceive bbb series of voiced consonants or clusters

Cues (minimal recoverable acoustic information)

1) VOT
2) duration of vowel
3) release burst of voiced & voiceless consonants.
4) amplitude

categorical perception
p ✕ b
either p or b.

Tone (aspect of pitch) high fall, low fall ...

minimal pairs: ma ma ma ma. Ex. Mandarin Chinese.

diff from.

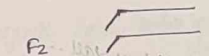
Information → is about the entire sentence
this is a duster.
this is a duster?
↳ rise in pitch ...

Infants have all these abilities of speech perception

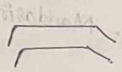
Mexican tone ability. ↑
English speaking infants loose after 9 months

← Discovery of language.
book by Peter.
experiments on infants.

Acoustic cue



F_2 transition, if we remove there, it is difficult to identify the place of articulation
change



} increasing effect

ex bab. ← when same consonants.

nasal attenuated formants { zero → anti formant property
pole → involves formant calculation.

t k stops

s S fricative sibilant fricatives

nasalization ←
our perception is higher for, we can hear the nasalized utterance also

VOT for fricatives is more complicated than the stops (to identify the boundary)

(vowel longer for voiced fricative than voiceless fricative)

28/10/24

• A unfamiliar foreign language seems continuous. /sounds.

because we don't know the words, we don't know where to put the breaks.

VOT → time delay b/w when a sound starts and when the voicing begins

categorical perception → listener can only discriminate b/w sounds they would identify as different categories is not intermediate categories

+VOT
-VOT → voicing starts before the release

delay for the voiceless one is much more
for ta is more
ta has more aspiration.

discrimination → based on. (recalling the token the listeners heard)

same or different to the sound they heard before

irrelevant variation
→ intermediate categories.

categories { voicing
place of articulation
manner of articulation

exemplar → something of the same category

30/10/24

McGurk

clashing ← { visual stimuli / speech stimuli }

our perceptual system interprets something for blw.

↓

Speech perception is multimodal

language has gestures also which are important. (for the articulation)

the participant hears a with eyes close ~~interpret~~ relies only on speech perception

also happens when the dubbing is not in sync

Top down approach

knowledge of the language

helps for Segmentation

points

---

bottom up (acoustic signal)

↳ {cues}

(weird)

bat — immediately pressing the button

bat — some time is required

non word

(we are like) what did I hear?

- we like to listen familiar voices

Amitabh Bachan voice is everywhere (phonepay) :)

shadowing → repeat aloud.

for linguistics, we follow people's judgement rather than grammar books

anomalous sentences. → for. for Black Americans double negative need not be a positive.

(affects the top down) Speak fluently but gibberish ← (temporal lobe) is affected of →

(affects the speech) problem in speaking but ← (frontal lobe) is affected speech perception is intact of →

plasticity loses with age.

1/11/24

phonetic properties

- phonetic diversity across languages

- there is a limit our vocal tract can produce no. of sounds.

(can also be a language) dialect (variety of language)
↑↓
language
(can also be a dialect)

due to some political reasons some dialects are standardized.

mutually intelligibility → understanding of different dialects (even if with differences)

If we lose languages → we lose knowledge

tone languages requires some specific cognitive abilities, we might lose them

(place of articulation)

9th position → retroflexes
in Indian languages most

the tongue moves rapidly from one position to other position
∴ we refer to the term gestures.

coronal are the grp of sounds produced by the tongue tip and tongue blade.

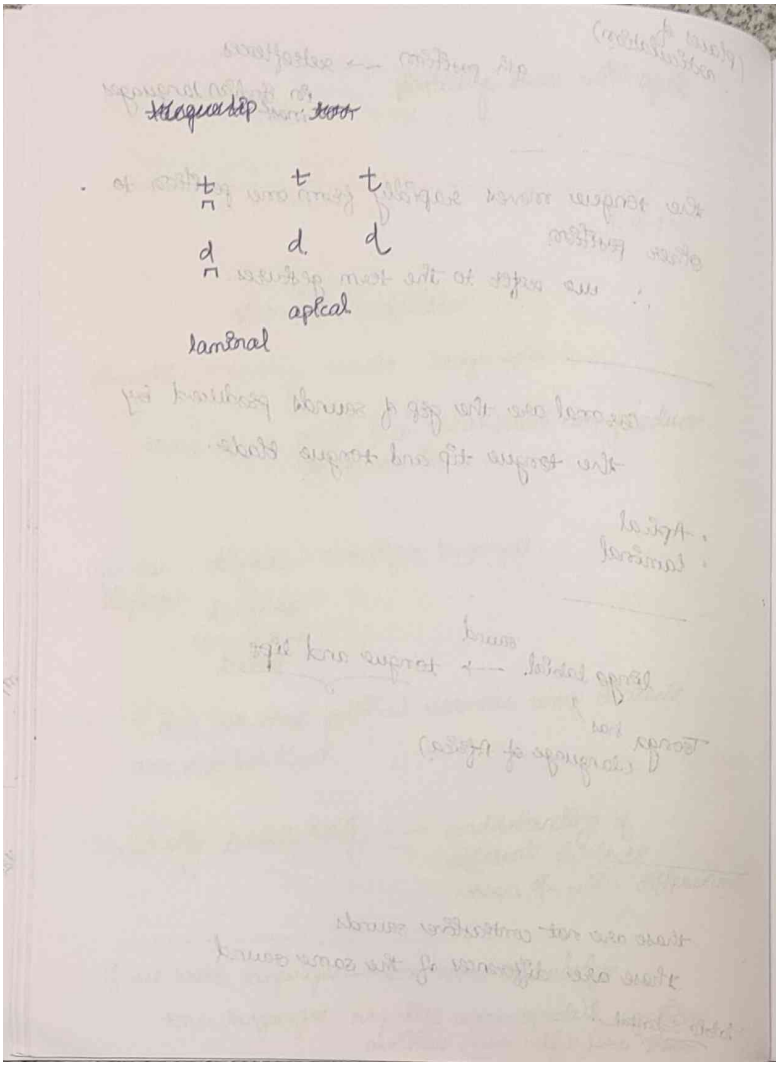
- Apical
- laminal

linguo labial sound → tongue and lips

Tsonga has? (language of Africa)

these are not contrastive sounds.
these are differences of the same sound

labio-dental.



6/11/24

Tamil and Telugu has robust retroflex sets

labiodental glides are more common in Indian languages

primary place of articulation ; secondary place of articulation.
manner of articulation
centrality vs laterality.
Nasality

(l)

extended: voiceless nasals
↳ (whispers)
voiced nasals.

ejectives → loud sounds

clicks
two closures
↳ so that it forms a pocket
then there is a release → resulting in a loud sound.

~~Det.~~

Dental Alveolar Post alveolar

t d

n

r

no distinction, everything is assumed to be alveolar

which are not showed on the IPA chart,

ex linguo labial

represented with diacritic

medal phonation

voice, voiceless

non modal phonation are also there

7/11/29

music/ opera from click sounds

phonation

breathy

b (bh) not just air release, but the glottis is ~~[illegible]~~

t th ← not the same as th

voiced

most open

(voiceless - Breathy ← Modal → creaky → glottal closure

↓

glottal stop

Gujarati has breathy voiced vowels, which contrasts with ~~normal~~ (modal) voicing.

forms minimal pairs

ghost

gost ← ex for English; pronunciation is

If Bharat becomes the name of our country, non Indians would not be able to produce /ɦ/, because only Indian languages has /ɦ/

In the US, the people put on think that (language news)

~~that~~ some people put on an accent for a creaky appearance at the end.
ex. yeah.

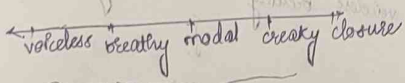
8/1/24

—— coronals.
sounds to be remembered.

Pulmonary ⟶ has only egressives
Bilabials, Velars } can languages can have ingressives & egressives

phonation. is not just voiced & voiceless
for expression we represent + & - voice.

languages have creaky & breathy, modal voice



minimal pairs

breathy ⟶ turbulence

wider paced repeating pattern

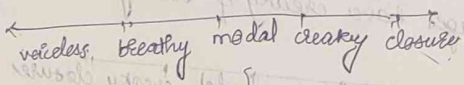
creaky spectrogram.

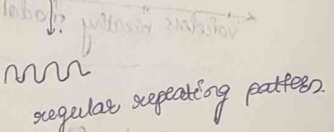
English → non contrastive glottal stops

Many languages → have contrastive glottal stops

If we record our own voice, towards the end we see creaky voice.

Spectral tilt → to distinguish btw breathy & creaky.





- fundamental freq.
- Harmonics → multiple of fundamental freq.
- Formants

amplitude & harmonics of higher freq helps us distinguish btw breathy & creaky.

Subtraction of amplitude & harmonics

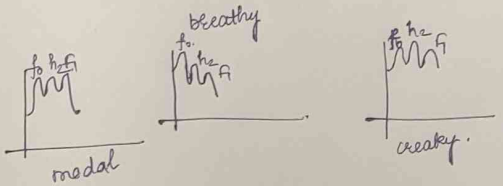
amplitude - harmonics

+ve value creaky

-ve value breathy.

second harmonic → closest to F_1

$(h_2 - f_0)$



lower f_0 → longer duration.