

# Puheen tuotto ja havaitseminen I

## *Vokaalit*

Martti Vainio

Fonetiikan laitos, Helsingin yliopisto



# Puheentuoton lähde-suodin -malli



# Glottaalinen äänilähde

- Vokaalit saavat alkunsa *larynksissa*, missä tuotetaan kaikkien soinnillisten äänteiden perusta.
- Kurkunpään äänentuotto perustuu ns. *ääniraon* (glottis) nopeasti toistuvaan avautumiseen ja sulkeutumiseen.
- Tätä glottiksen toimintaa kutsutaan *fonaatioksi* (phonation).



# Fonaatio

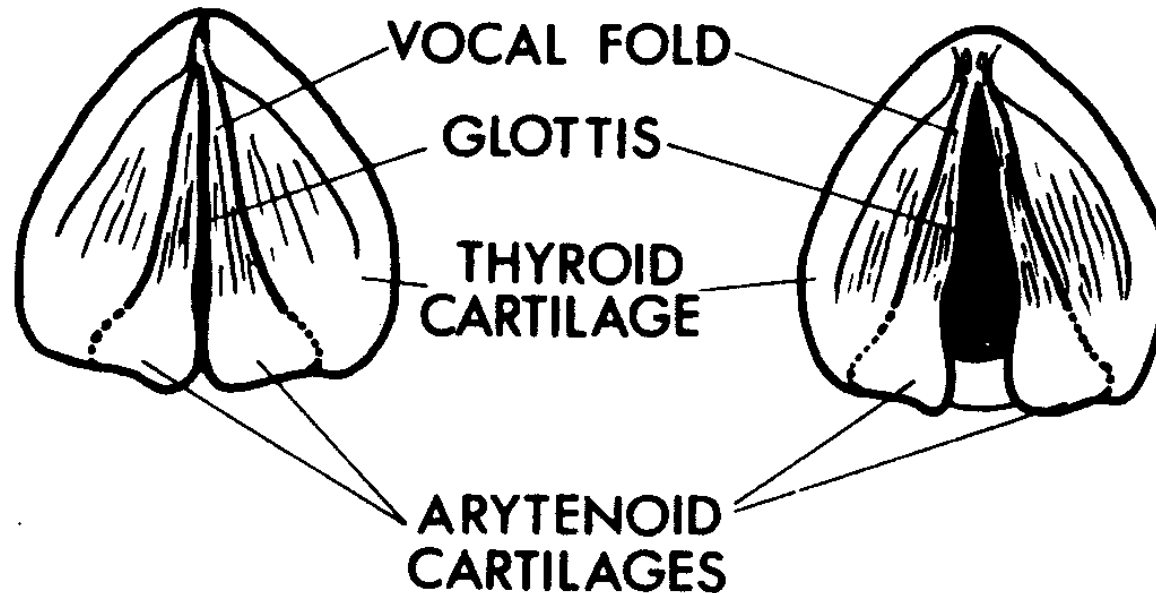
- Fonaatio perustuu siis *äänihuulten* (vocal folds) liikkeisiin. Äänihuulet ovat lihaksista, jänteistä ja kudoksesta koostuvat ulokkeet ilmaväylän molemmin puolin.
- Äänihuulten primitiivinen, alkuperäinen tarkoitus lienee ollut pitää vieraat esineet poissa keuhkoista, mutta niiden puheeseen liittyvä funktio on niin tärkeä, että niitä on alettu kutsua äänihuuliksi.



# Kurkunpää (larynx)

PHONATION

BREATHING



- Äänihuulten asentoa muutetaan ns. arytenoidiruston (arytenoid cartilage) välityksellä.

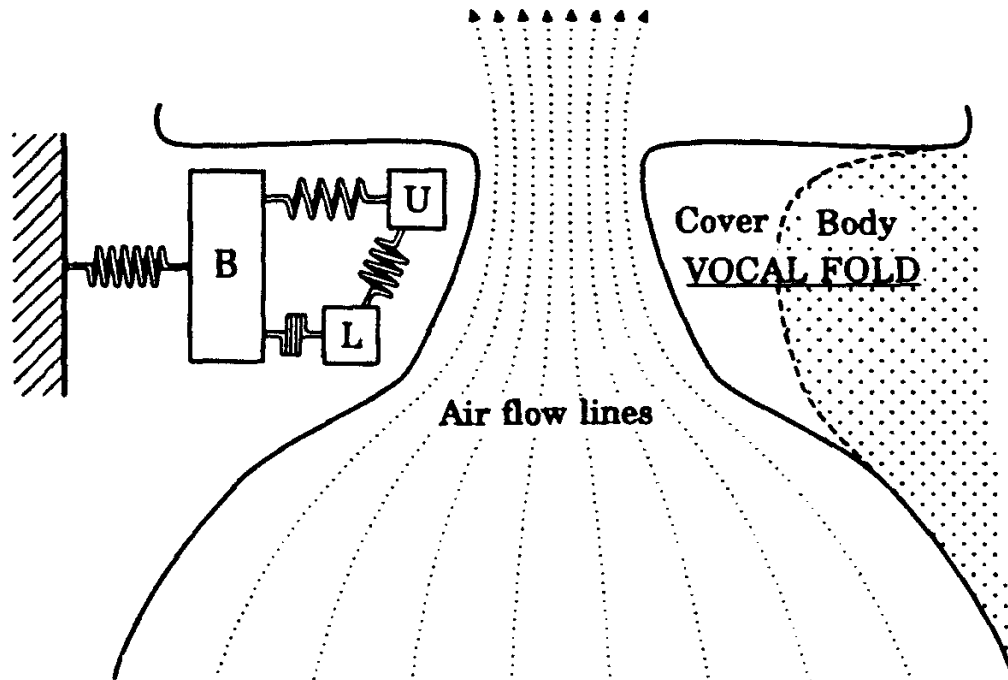


# Glottispulssi

- Fonaatiossa ääniraon toistuva sulkeutuminen ja avautuminen muodostaa sen läpi kulkevaan ilmavirtaan jonon pulsseja, jotka toistuvat jotakuinkin tasaisin välein.
- Pulssijonon spektraalinen muoto toimii pohjana kaikille soinnillisille ääniteille – etenkin vokaaleille. Näin ollen glottaalisen äänilähteen spektri on kuultavissa kaikissa soinnillisissa ääniteissä.
- Glottispulssin spektraaliset ominaisuudet lähes täysin riippuvaisia siitä, millä tavoin äänihuulet sitä tuottaessa värähtelevät.



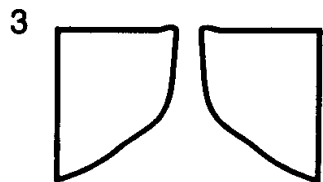
# Äänihuulten mekaaninen malli



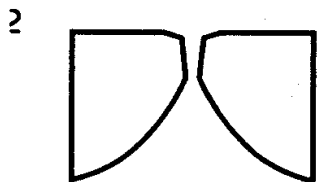
- Värähtely säilyy äänihuulissa, koska niiden osat liikkuvat toistensa suhteen yhtenä järjestelmänä, joka voidaan mekaanisesti kuvata edellisen kuvan kaltaisesti.



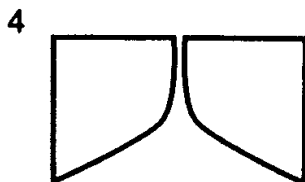
# Äänihuulten toiminta fonaation aikana



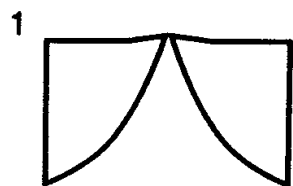
Mid-part of open phase



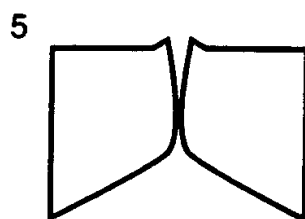
Early part of open phase



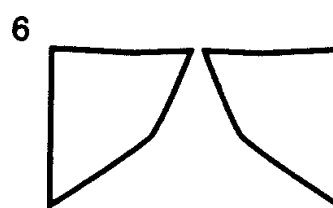
Last part of open phase



Vocal folds closed in upper part, about to be forced open



Vocal folds close in lower part of glottis



The closure is moving upward to the state at stage 1



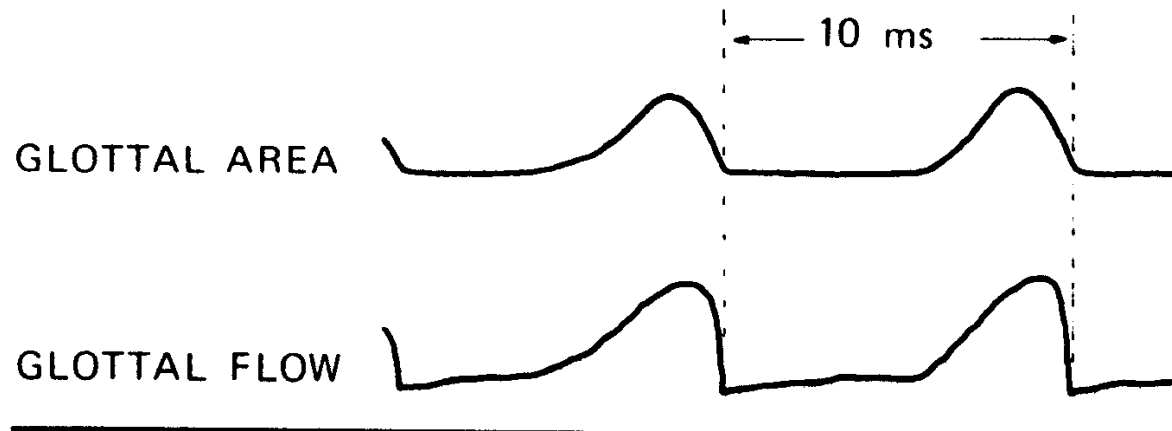


# Äänihuulten toiminta fonaation aikana

- Subglottaalinen paine pakottaa äänihuulet erilleen, jolloin niiden välisen ilman paine laskee ja kääntää niiden liikkeen suunnan sulkeutuvaksi. Ns. *Bernoullin efekti* aiheuttaa niiden sulkeutumisen äkkinäisesti.
- Tämän jälkeen subglottaalinen paine ja kudosten elastiset voimat aloittavat syklin uudelleen.



# Glottispulssin muoto



- Glottispulssi muistuttaa muodoltaan ns. *sahanteräaaltoa* (sawtooth wave).

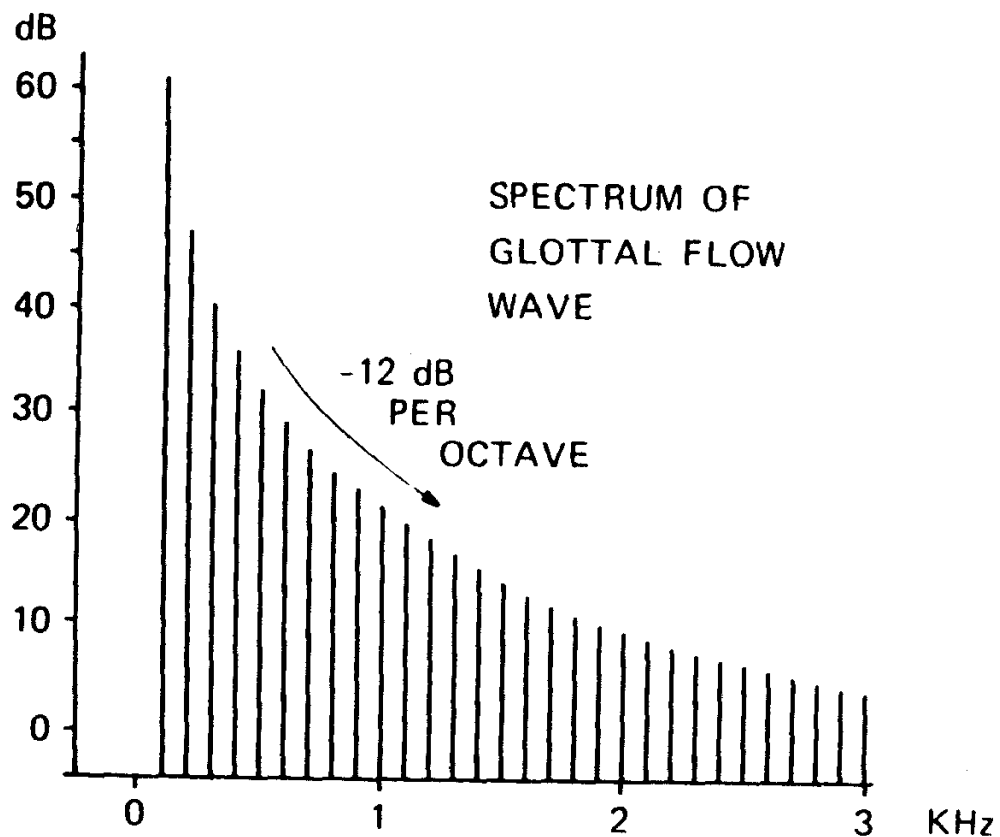


# Glottispulssin spektrin komponentit

- Glottispulssin spektrillä on kaksi erityisen tärkeää piirrettä:
  1. Spektraalisten komponenttien – siis perussävelen ja sen kerrannaistaajuuksien välimatka, joka on riippuvainen pulssin toistumisen taajuudesta.
  2. Komponenttien amplitudit taajuuksien suhteen, jotka ovat riippuvaisia perättäisten glottispulssien muodosta.



# Glottispulssin spektri



- Idealisoitu 100Hz:n taajuudella toistuvan glottispulssin spektri.



# Ääniväylän vaikutus fonaatioon

- Yleensä ääniväylän muoto ei vaikuta glottiksen toimintaan:
  1. Väylässä etenevien tasoaaltojen suhteen sen käyryydellä ei ole vaikutusta
  2. Väylässä olevat *sulkeumat* (constriction) vaikuttavat vain vähän pulssin muotoon.
- Glottispulssijono muistuttaa siis sahanteräaaltoa, joka voidaan analysoida tasavälein toistuviksi kerrannaistaajuuksiksi, joiden amplitudi laskee 12 dB jokaisella oktaavilla.

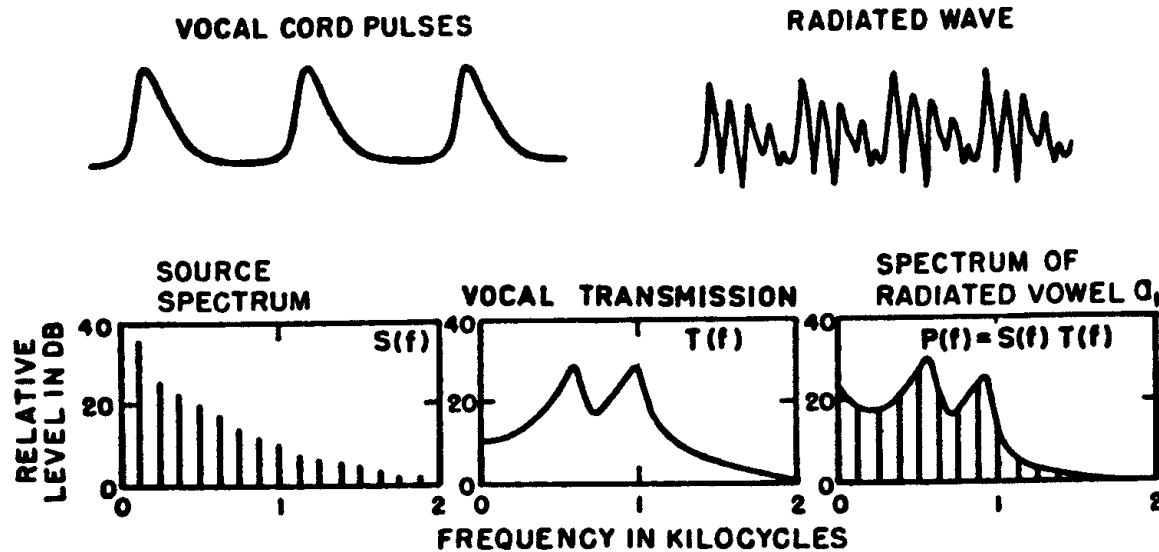


# Lähde-suodin -teoria

- Vokaalien ja voimakkaasti soinnillisten äänteiden suhteen glottis siis tuottaa niille niiden äänellisen perustan – voimme siis kuvitella, että glottiksen yläpuolinen ääniväylä toimii jonkinlaisena *suotimena* tai torvena, joka voimistaa lähteen tuottamista komponenteista niitä, jotka ovat lähellä väylän resonanssitaajuuksia.
- *Lähde-suodin -teorian* (source-filter theory) avulla voimme siis selvittää kuinka vokaalien spektrit muodostuvat glottaalisen lähteen ja ääniväylän suodinfunktion kombinaationa.



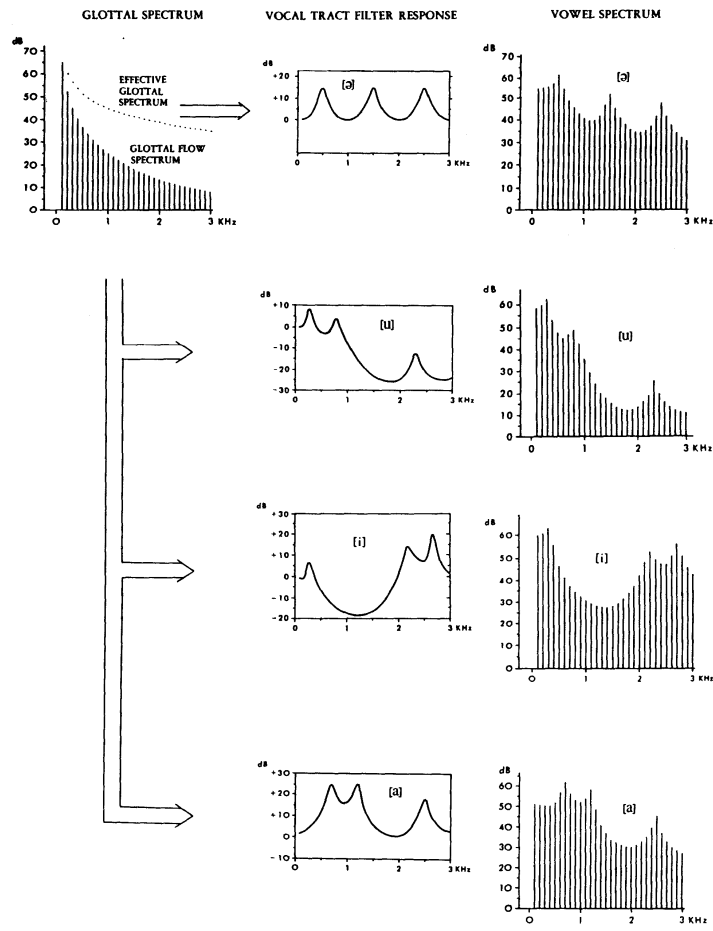
# Lähde-suodin -teoria (kuva)



- Gunnar Fant: The Acoustic Theory of Speech Production, 1960.
- Teorian perustana on olettamus, että ääniväylä on täysin riippumaton lähteestä.



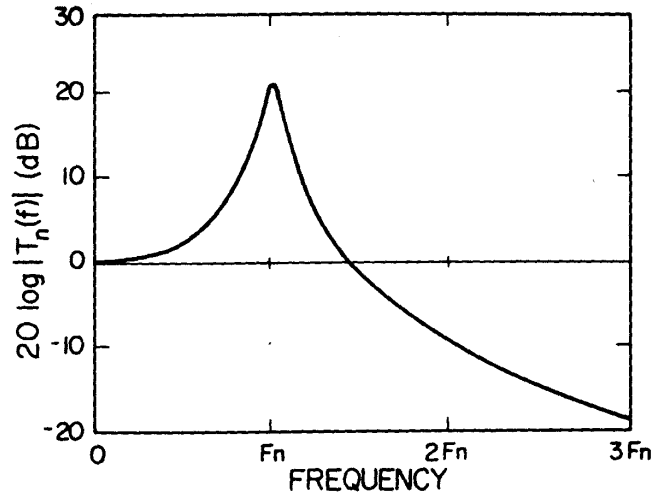
# Suotimen vaikutus lähteen spektriin





# Lopullisen spektrin kaltevuus

- Muiden kuin neutraalin keskivokaalin – [ə] – spektrin kaltevuus riippuu väylän resonanssien välisistä etäisyyksistä. Lähellä olevat resonanssit vahvistavat toisiaan ja vastaavasti resonanssit vaikuttavat negatiivisesti korkeammilla taajuuksilla oleviin resonansseihin.

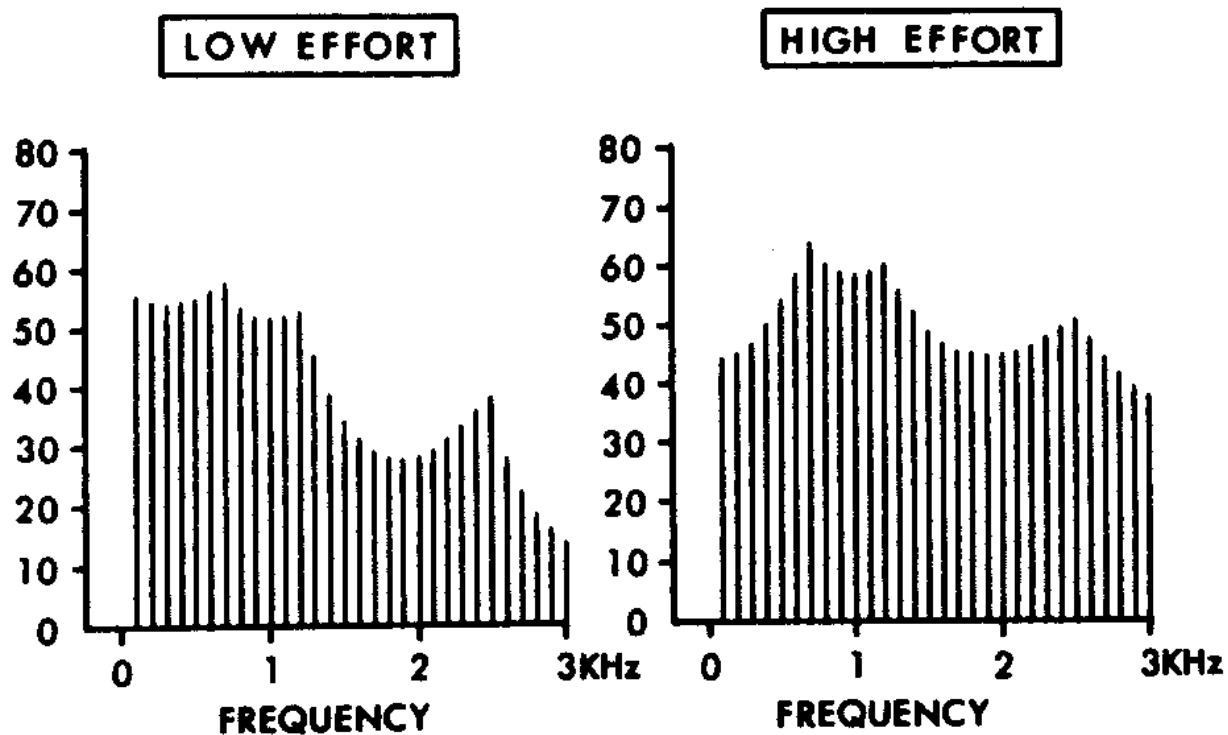


# Äänentuottotavan vaikutus spektriin

- Ääviväylän resonanssien vuorovaikutusten lisäksi vokaalien spektrin yleiseen muotoon (spectral envelope) vaikuttaa se miten glottispulssit tuotetaan.
- Toisin sanoen glottispulssin taajuuden lisäksi vokaalin lopulliseen spektriin vaikuttaa se, minkä muotoinen pulssi on. Muoto on taas riippuvainen pitkälti siitä, kuinka paljon sen tuottamiseen käytetään fyysistä toimintaa (vocal effort).
- Voimakkaasti tuotettu pulssi sisältää enemmän energiaa korkeilla taajuuksilla, jolloin spektri on vähemmän kalteva.



# Vocal effort



- Äänilähteen spektrin kaltevuus heikosti ja voimakkaasti tuotetuilla [ɑ] -vokaaleilla.



# Ääniväylän vaikutus vokaalien spektreihin



# Ääniväylän mallit

- Akustisessa tutkimuksessa on huomattu, että yksinkertaisella – putkiin perustuvalla – ääniväylämallilla voidaan tuottaa luonnolliselta kuulostavia vokaaleita.
- Ääniväylää (pharyngeal-oral tract) voidaan mallintaa yhden tai usean putken kombinaatiolla.
- Koska ääni kulkee väylässä ns. *tasoaaltona*, ei väylän taipuvalla muodolla ole mainittavaa vaikutusta sen suodinfunktioon.
- Väylän poikkipinta-alan pienet poikkeamat eivät myöskään ole vaikutuksiltaan suuria.



# Formantti

- Tässä vaiheessa on hyvä määritellä, mitä tarkoitamme usein esiintyvällä termillä *formantti*.
- Teknisesti formantti on ääniväylän äänensiirtoon liittyvä resonanssi, jonka vaikutukset näkyvät mm. äänteiden spektreissä.
- On kuitenkin huomattava, että spektreissä ja spektrogrammeissa esiintyvät energian kasaumat eivät ole itse formantteja.
- Formantit ovat siis ääniväylään liittyviä akustisia piirteitä.



# Formanttien järjestys

- Puheessa formantit numeroidaan taajuuksiensa mukaan siten, että alin formantti on ensimmäinen (F1) ja sitä seuraa ylemmät formantit (F2 jne.).
- Kolmen ensimmäisen formantin taajuudet ovat suoraan riippuvaisia ääniväylän muodosta huulten, kielen ja leuan sekä kurkunpään muodostaessa eri konsonantteja ja vokaaleja.
- Ylemmät formantit pysyvät jotakuinkin vakiona puheen aikana.



# Ääniväylän vaikutus formantteihin

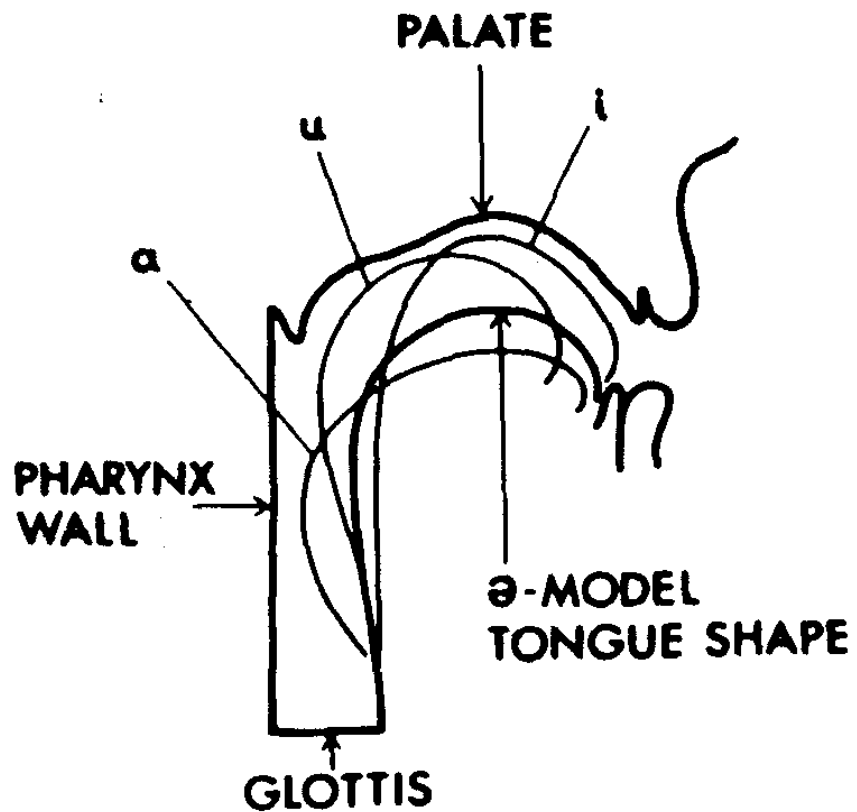
- Edellä olemme jo nähneet, että tärkein formanttien taajuuksiin vaikuttava tekijä on ääniväylän *pituus*.
- Toinen formanttitaajuuksiin vaikuttava tekijä on ääniväylään (yleensä) kielen avauella tuotettava *kapeuma*, jonka aste ja pituus vaikuttavat alempien formanttien taajuuksiin ja amplitudeihin.





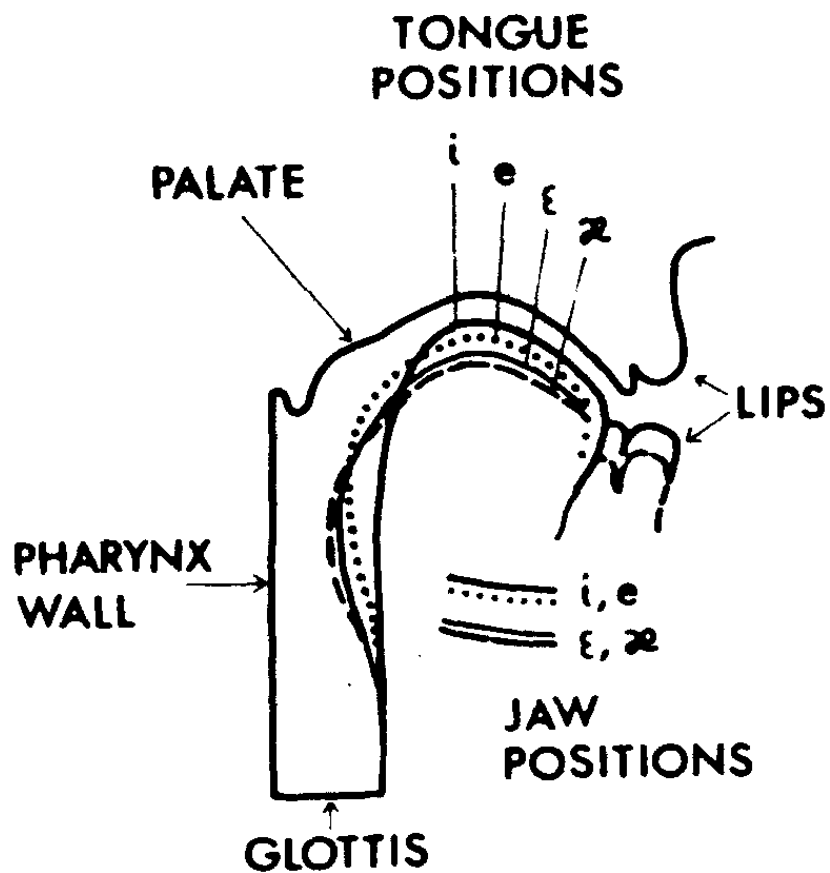
# Kardinaalivokaalit

## BASIC VOWEL SHAPES



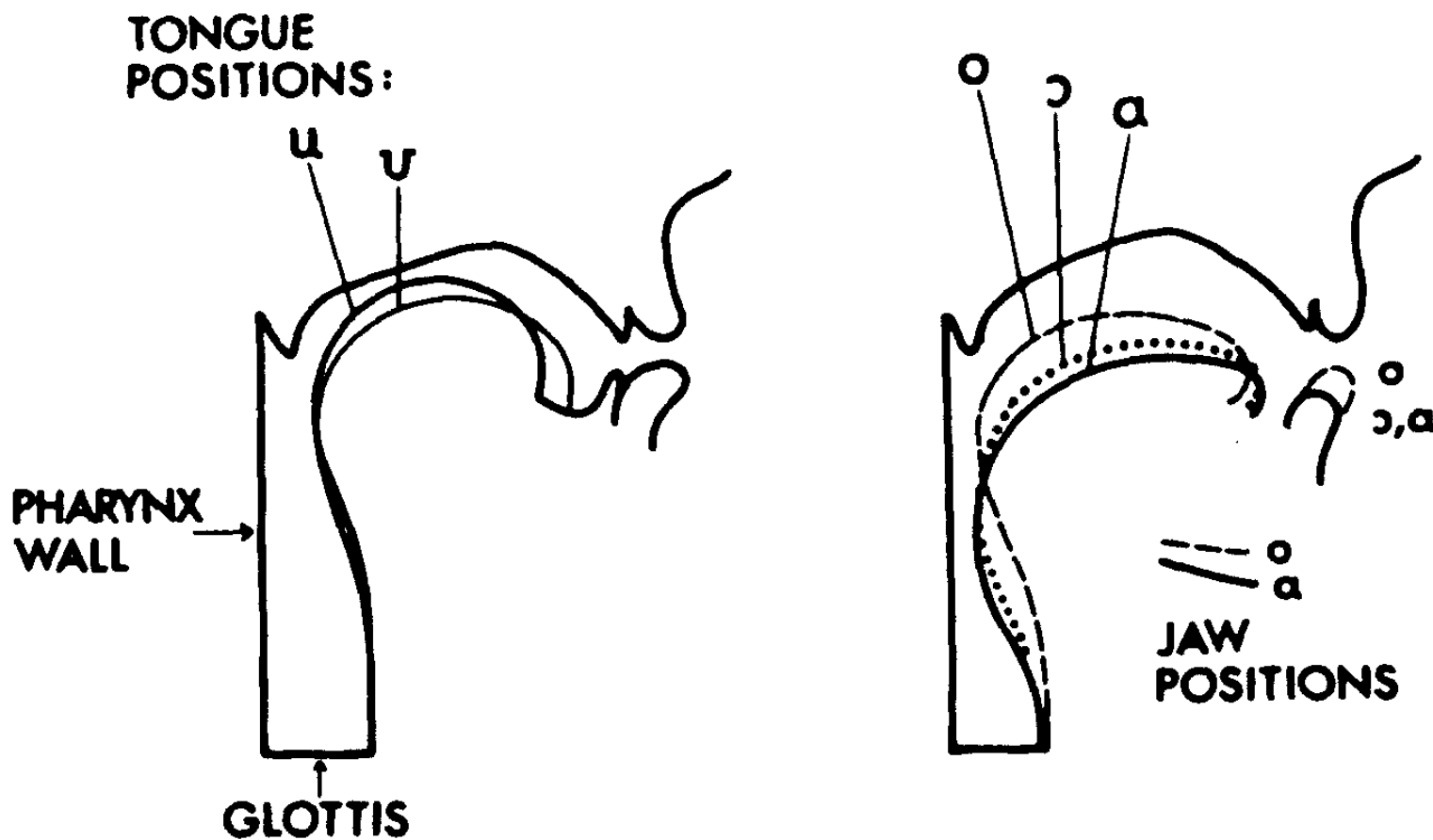
# Etuvokaalit

## FRONT VOWEL SHAPES

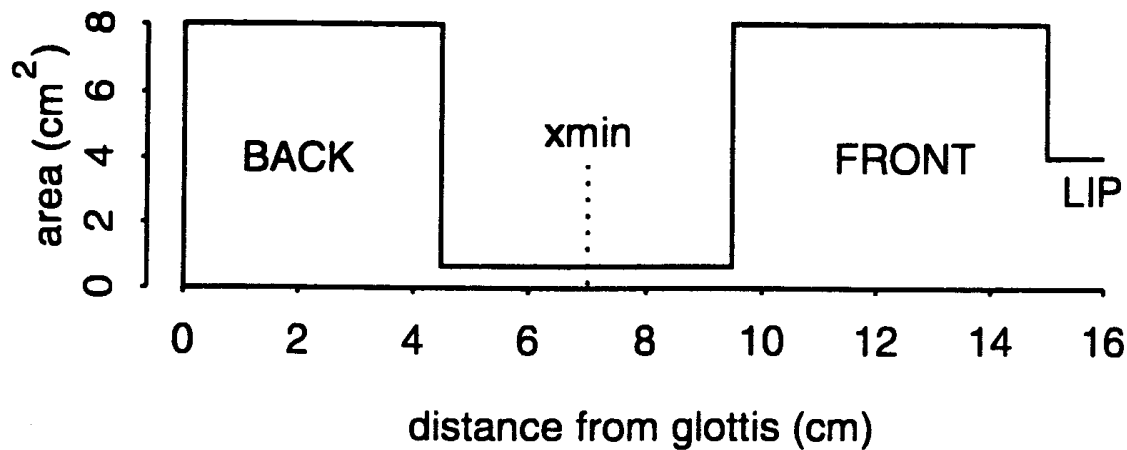


# Takavokaalit

## BACK VOWEL SHAPES



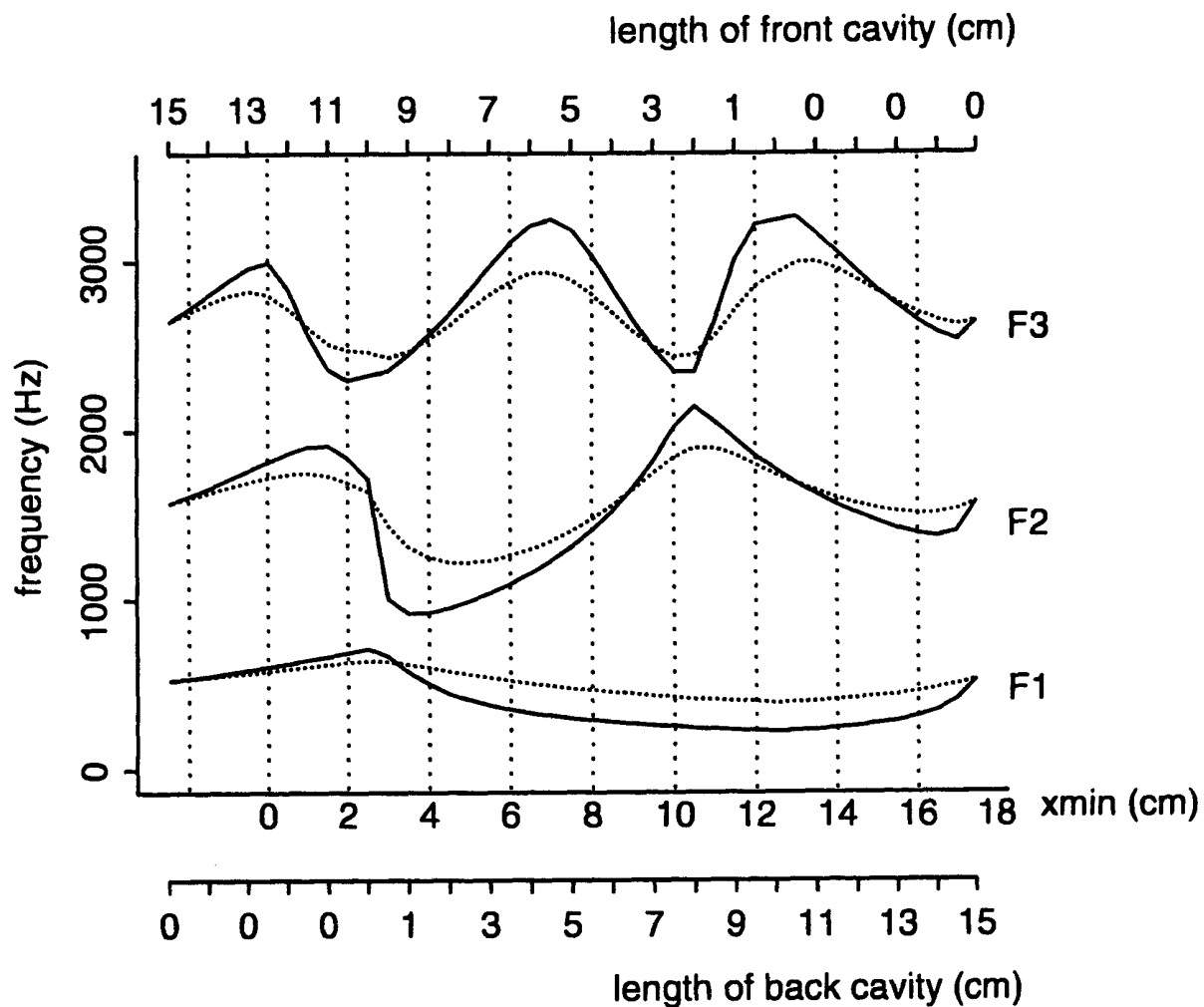
# Neljän putken ja kolmen parametrin malli



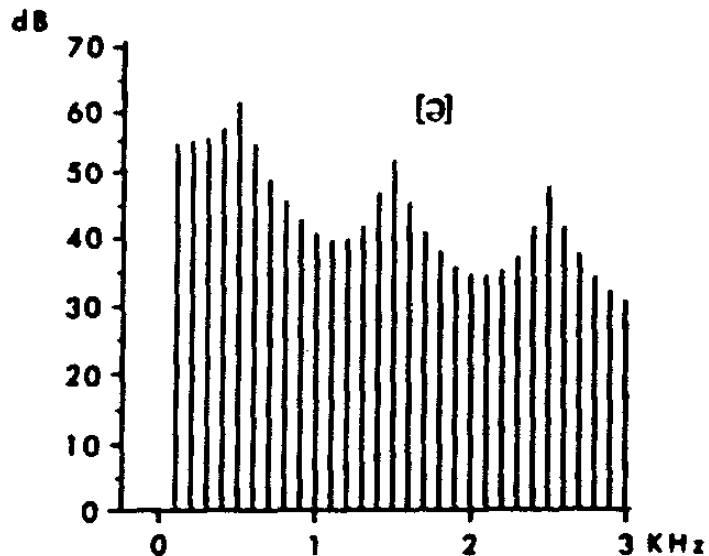
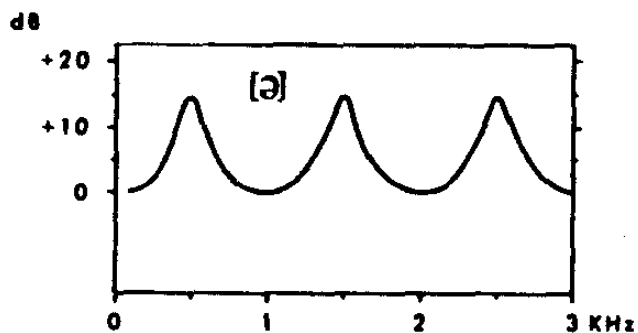
- Parametrit ovat *kapeuman horisontaalinen etäisyys, kapeuman pinta-ala ja huuliaukon pituus.*



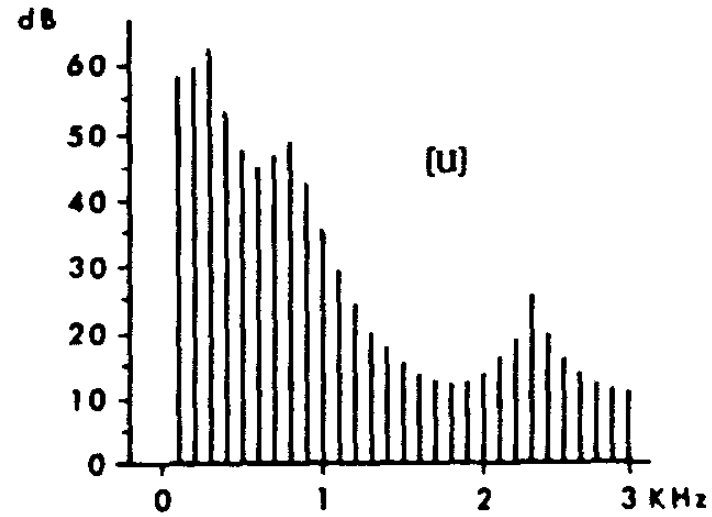
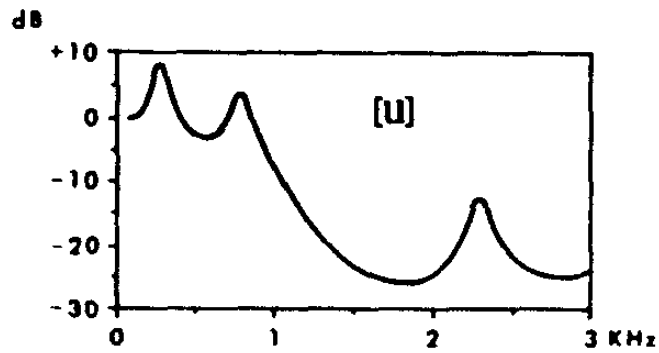
# Parametrien vaikutus formanttitaajuuksiin



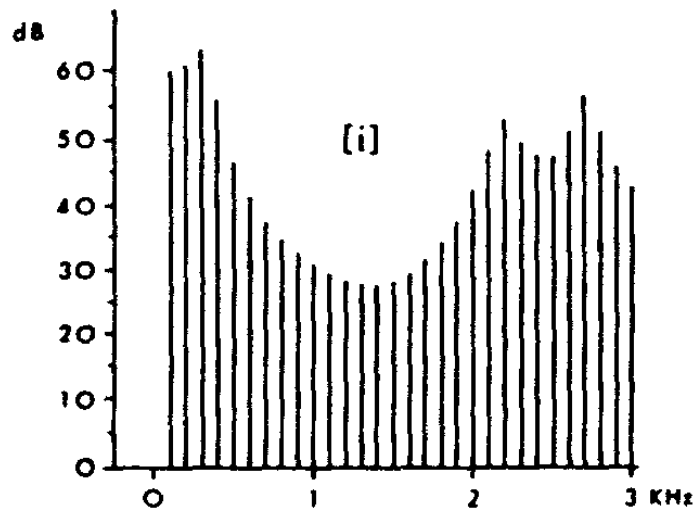
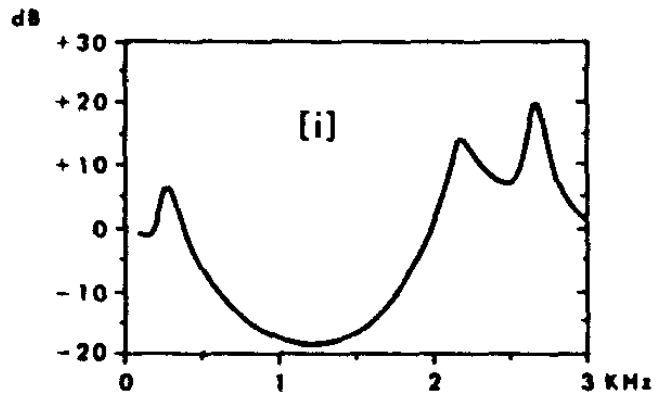
# Neutraali keskivokaali



[u]

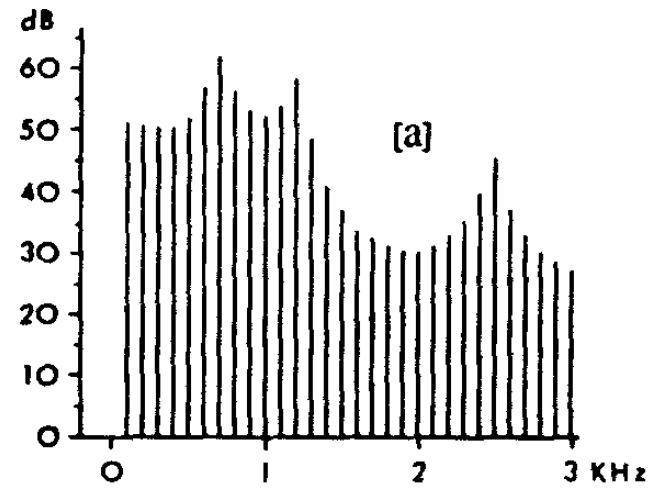
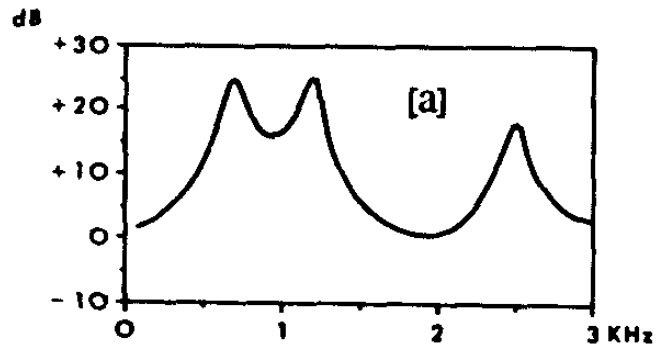


[i]





[a]

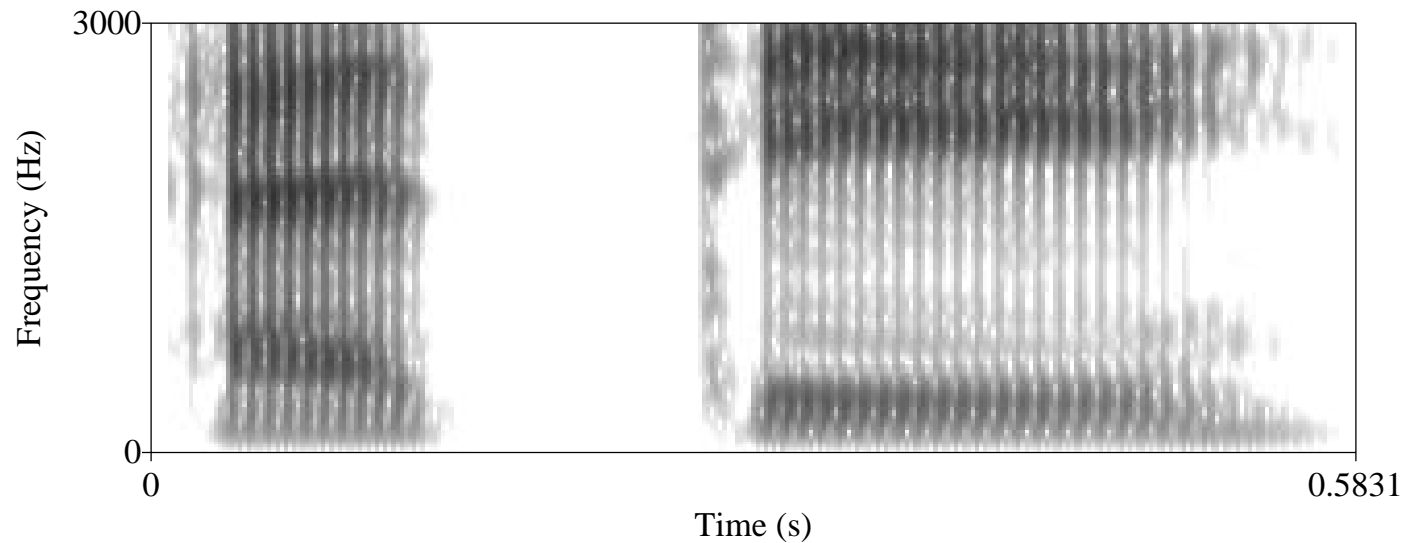


# Vokaalien havaitsemisesta



# Vokaalit

- Vokaalit tunnustetaan formanttirakenteestaan, joka näkyy selvästi leveäkaistaisessa spektrogrammissa:



# Vokaalien tavoitearvot

- Yleensä vokaaleja tutkittaessa käytetään arvoja, jotka toteutuvat vokaalin keskiosassa ja edustavat ns. *tavoitearvoja* (targets).
- Tavoitearvo määritellään usein niin, että se on sellainen intervalli vokaalin sisällä, jolloin formanttien taajuudet ovat käytännöllisesti katsoen muuttumattomia. Joskus käytetään myös yksittäisiä arvoja vokaalin keskeltä mitattuina. Esim. lasketaan sellainen neljännes vokaalin kestosta, jonka aikana perättäisten formanttiarvojen euklidinen etäisyys on pienin ja keskiarvoistetaan siitä.
- *Steady state* -periaatteen ongelmana ovat mm. se ettei vokaaleilla usein ole sellaista ja eri formanttien muuttumattoman alueet eivät välttämättä ole samoja.

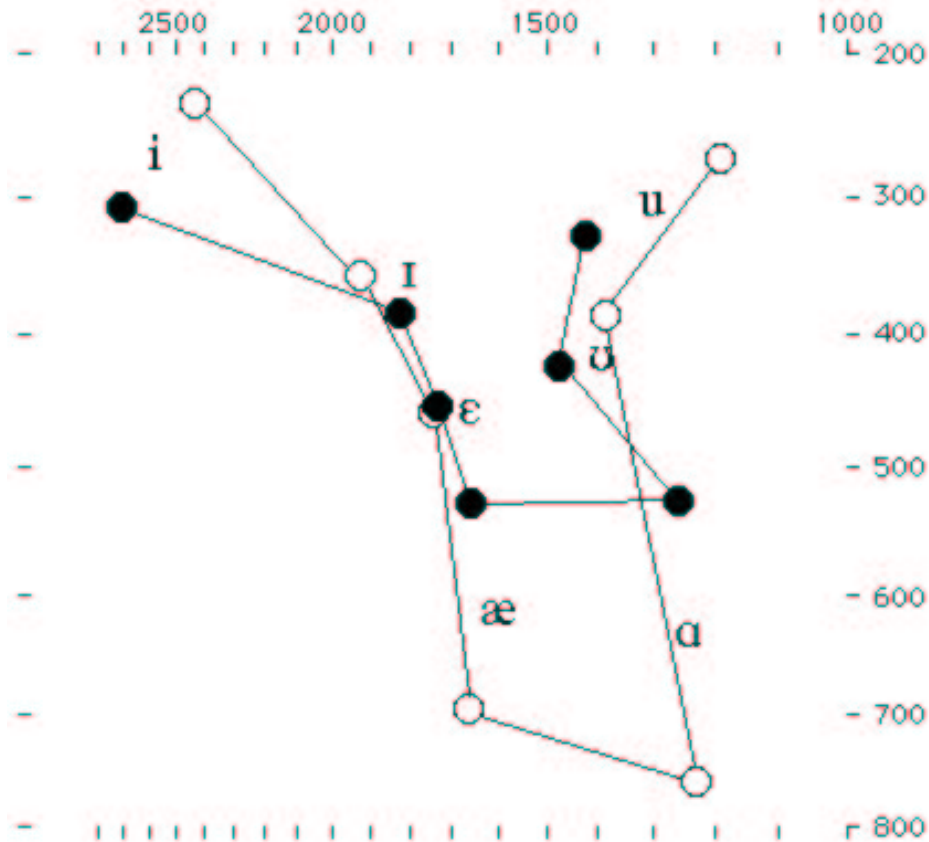


## F1/F2 avaruus

- Vokaaleja kuvataan usein pisteenä kahden ensimmäisen formantin määräämässä kaksiulotteisessa avaruudessa.
- F1/F2 avaruuden psykologinen todellisuus perustuu pitkäaikaisiin havaintoihiin ja tutkimuksiin, joissa on havaittu, että vokaalien kahden ensimmäisen formantin arvot hyvin pitkälle määräävät sen laadun (mm. Willis 1829, Helmholtz 1863, Bell 1879).
- Myös havaintokokeiden avulla on osoitettu, että kahden ensimmäisen formantin arvot ovat havainnon kannalta tärkeimmät vokaaleja erottelevat tekijät.



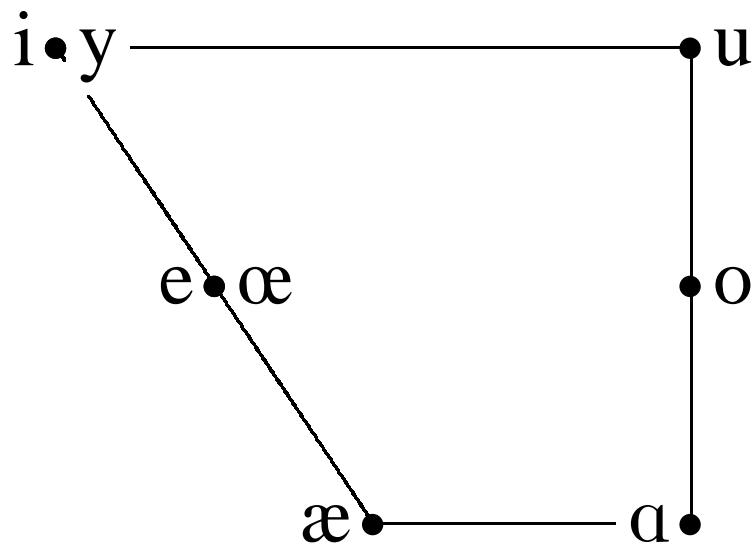
# Formanttikartta



- Formanttikartta piirretään usein käännteisin akseleihin siten, että formanttien sijainti muistuttaa niiden artikulatorista "sijaintia".



# Suomen vokaalit vokaalikulmiossa



## Ylemmät formantit ja f0

- Ylempien formanttien (etenkin F3) vaikutus on selvä tiettyjen vokaalien suhteen, mutta niiden yleisestä vaikutuksesta ollaan montaa eri mieltä. On ehdotettu erilaisia – enemmän tai vähemmän – abstrakteja ylempiä formantteja: F2' on F2 - F4 yhteisvaikutus.
- Joskus kuitenkin perustaajuus (siitä huolimatta, että lähde ja suodin ovat erotettavissa toisistaan) vaikuttaa vokaalin havaitsemiseen (F1 - f0 (barkeissa)). Ilmiö on tuskin merkittävä, sillä jatkuvassa puheessa perustaajuuden arvoja määräävät prosodiset tekijät.





# Vokaalien pituus

- Monissa kielissä (esim. englanti) vokaalien erot tuotetaan usein niiden *kesto*a vaihtelemalla – näin tapahtuu etenkin silloin kun vokaalit ovat spektraalisesti hyvin samankaltaisia. Esim. englannin *bit* vs. *beat* ([ɪ] ja [iː]). Ns. tense – lax -oppositio.
- Kysymyksessä ei kuitenkaan (todennäköisesti) ole pelkästään kestoerot, vaan myös formanttien ns. *lopputransition* (offglide) suhteellinen osuus vokaalin keustosta vaihtelee; lax-vokaalien lopputransitio on yleensä pidempi.

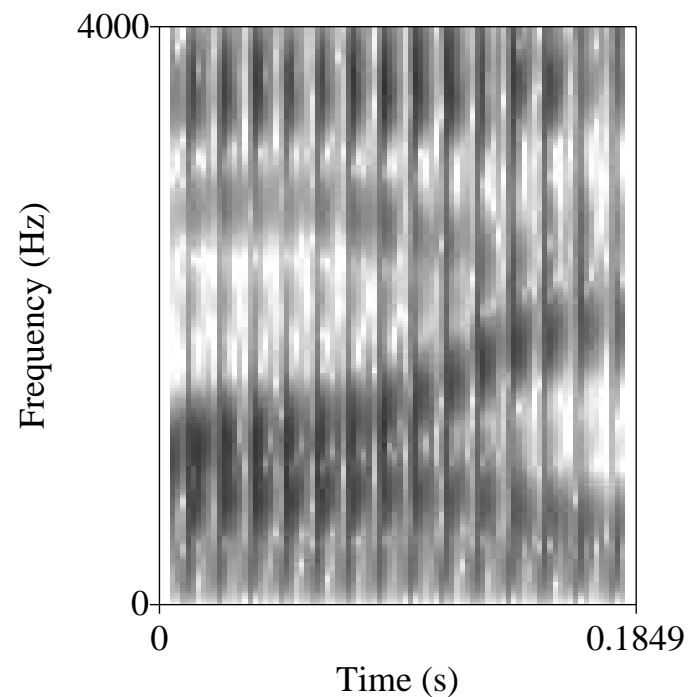
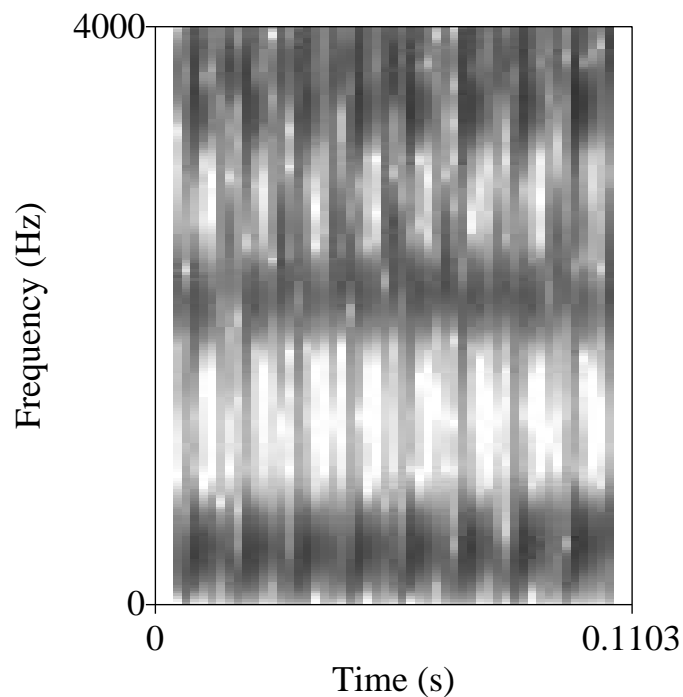


# Monoftongit ja diftongit

- Auditorisesti *monoftongien* ja *diftongien* ero on yleensä selvä.
- Fonologisen teorian suhteen diftongit tuottavat kuitenkin päänvaivaa akustinen moniarvoisuus tulisi saada *piirrespesifikaatio*ltaan mahdollisimman kompaktiksi eikä se aina onnistu.
- Tutkimuksen mukaan ensimmäinen tavoite (target) ja muutoksen suunta ovat tärkeimmät diftongien tärkeimmät havaintovihjeet.



# Monoftongit ja diftongit (jatkoa)



- Diftongilla on selvästi kaksi tavoitetta, joista toinen jää usein toteutumatta.



# Dynaamiset spektraaliset vihjeet

- Vokaalien formanttien *siirtymät* (transitiot (onglide, offglide)) vaikuttavat (joidenkin tutkimusten mukaan merkittävästi) vokaalien erottamiseen: Vokaalit tunnistetaan yleensä paremmin konsonanttiympäristössä (CVC) kuin irrallaan.
- Toisaalta jos korvaamme tavoitteen hiljaisuudella CVC-kontekstissa, vokaalit tunnistetaan yhtä hyvin kuin ehjissä ärsykkeissä. Havainnon kannalta tärkeät akustiset vihjeet ovat siis jakaantuneet koko vokaalin alueelle. Kysymyksessä on eräs puheen perusilmiöistä; *artikulatorisista eleistä*, jotka muuttuvat ajassa.



# Vokaalien reduktio: fonologinen

- *Reduktiota* on kahdentyyppistä: fonologista ja foneettista.
- Fonologinen reduktio tarkoittaa sellaista prosessia, jossa vokaali ns. neutralisoituu menettäen painon; esim. englannin **major** vs. **majority** = /ei/ vs. /ə/.



# Vokaalien reduktio: foneettinen

- Foneettisessa reduktiossa vokaalin laatu muuttuu *segmentaalisesta* ja *prosodisesta* kontekstista johtuen. Artikulatorisella tasolla puhutaan ns. *target undershoot* -ilmiöstä, jossa tavoite on piste spatiaalisessa tai auditorisessa avaruudessa.
- Yleensä reduktio myös lyhentää vokaalien kestoja.
- Spontaanissa puheessa reduktio lisääntyy, mutta vaihtelu vastaavasti kasvaa. Vokaalireduktion spektraalisia muotoja ovat *sentralisaatio* ja *kontekstuaalinen assimilaatio*.



# Tempo

- Jatkuva puhe on nopeudeltaan ns. *sitaatiopuhetta* suurempaa – tämä johtaa luonnollisesti äännekestojen lyhenemiseen. Samalla artikulaatioelimille jää vähemmän aikaa tavoittaa maalinsa, joka taas johtaa reduktioon.
- Kuulijat kuitenkin osaavat perkeptuaalisesti kompensoida reduktiosta johtuvia puutteita joten nopeampi ja redusoidumpi puhe ei vaikuta vokaalien havaitsemiseen. Toisaalta puhujat osaavat myös kompensoida ja välttää nopeamman tempon aiheuttamaan reduktiota esim. nopeuttamalla artikulointia ja organisoimalla koko artikuloinnin uudella tavalla.



# Paino ja kontekstin vaikutus

- Periaatteessa *lausepainolliset* vokaalit ovat painottomia *perifeerisempiä* (siis vähemmän redusoituneita) F2/F2-avaruudessa.
- Vokaalien konsonanttikonteksti vaikuttaa myös niiden havaitsemiseen – esim. formanttiarvoiltaan [u]:n ja [i]:n välimaastoon sijoittuva ärsyke havaitaan [i]:nä jos se on [wVw] kontekstissa. u-väritteisyys todetaan labiaalisesta kontekstista johtuvaksi.





# Selkeä puhe ja redundanssi

- Neuvottaessa puhujia puhumaan maksimaalisen selkeästi, tuottavat he vokaalit normaalia pidempinä ja vieläkin perifeerisempinä kuin he tuottaisivat ne ns. sitaattimuotoisina.
- Edellinen on relevanttia siinä mielessä, että tiedetään puheen selkeyden vaihtelevan sen mukaan, miten hyvin lähinnä sanat ovat ennustettavissa ympäristönsä avulla. Puhutaan ns. vanhasta vs. uudesta informaatiosta (given vs. new). On myös todennäköistä että ilmiö toimii myös sanojen sisällä (ainakin kielissä, joissa sanat voivat sisältää sekä uutta että vanhaa informaatiota – suomen *kongruenssi* "vanhaa informaatiota).

