

Beszédhangok és spektrális jellemzésük

Beszédinformációs rendszerek, 2. gyakorlat

Sztahó Dávid

sztaho@tmit.bme.hu

Tulics Miklós Gábrriel

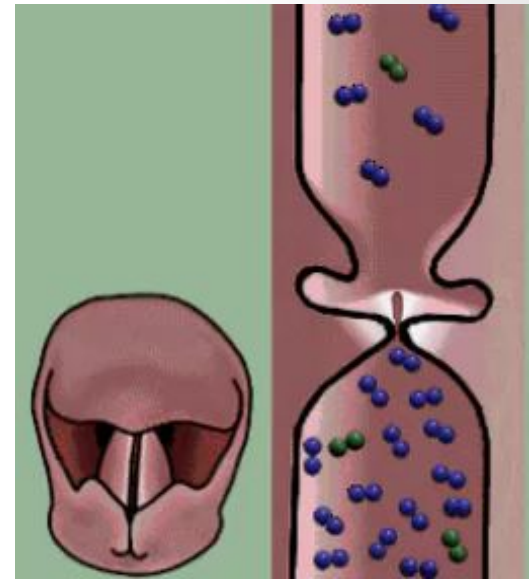
tulics@tmit.bme.hu

Tündik Máté Ákos

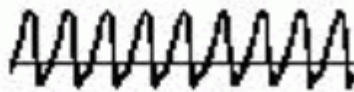
tundik@tmit.bme.hu

1. feladat

- Mi az alapfrekvencia definíciója?
 - A hangszalagoknál kialakult zöngéperiódusideje
- Hogyan állapítaná meg az alapfrekvenciát?
 - Időbeli megoldás
 - Periódusidő leolvasása
 - Frekvenciabeli megoldás
 - Vonalas spektrumon az alapfrekvencia helyének leolvasása
- Mi a formáns?
 - A hangképző üreg által felerősített felhangnyaláb.
- Hogyan állapítaná meg az F1 és F2 formánsokat?
 - Frekvenciabeli megoldás
 - Burkológörbe maximumhelyeinek leolvasása



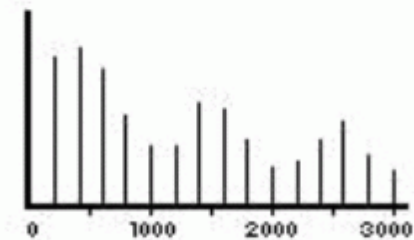
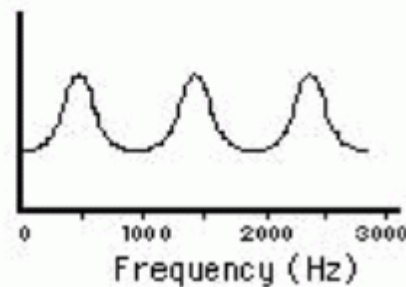
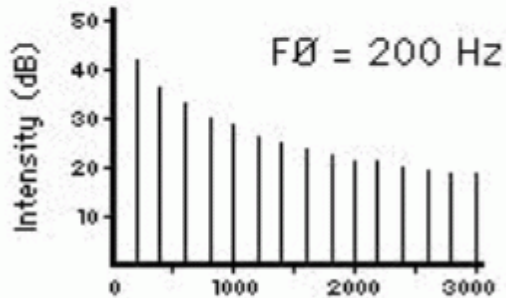
Glottal Pulses



Vocal Tract



Speech Signal



(a) Source Spectrum

(b) Filter Function

(c) Output Energy Spectrum

2. feladat

- A vonalas spektrumról tanultak alapján válaszoljon
 - a) Mit mutat meg egy vonalas spektrum a beszédre vonatkoztatva?
 - b) Milyen Fourier-eljárással határozható meg elméletileg a vonalas spektrum?
 - c) Rajzolja le a hangszalagok felett a gégében mérhető hangnyomás-idő függvény közelítő spektrumát, ha nő ejti ki az „á” hangot.
 - d) Hány felharmonikust tudunk megszámlálni a c) válasz ábráján a 4000 Hz-es pontig?
 - e) Rajzolja fel az ajkáról kisugárzott „á” hang hangnyomás-idő függvénye vonalas spektrumának közelítését (a magánhangzó háromszögről tanultakat is vegye alapul).
 - Sugárzási ellenállás
 - Magánhangzó háromszög

a) Mit mutat meg egy vonalas spektrum a beszédre vonatkoztatva?

- Alapfrekvenciával és felharmonikusokkal rendelkezik?
 - Igen: zöngés hang
- Vannak kiemelések a spektrumban?
 - Igen: formánsszerkezet
- Közvetett következtetések:
 - Beszélő neme (ha leolvasható az alaphang)
 - Hang megnevezése a formánsok leolvasásából (ha leolvashatóak)

b) Milyen Fourier-eljárással határozható meg elméletileg a vonalas spektrum?

- Fourier-sor, Fourier-transzformáció

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx \, dx \quad (k = 0, 1, 2, \dots)$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx \, dx \quad (k = 0, 1, 2, \dots)$$

Fűrészfog jel:

$$f(x) = 2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\sin kx}{k}$$

- c) Rajzolja le a hangszalagok felett a gégében mérhető hangnyomás-idő függvény közelítő spektrumát, ha nő ejti ki az „á” hangot.
- „á” hang: magánhangzó → vonalas spektrum
 - Női hang: alapprofrekvencia tartomány kb. 150-250 Hz
 - Felharmonikus szerkezet: az alaphang egész számú többszöröseinél
 - Amplitúdó csökkenés 12 dB/oktáv
 - Vonalas spektrum rajzolása 150-250 Hz-ig választható alapharmonikkal és egész számú többszörösöknél felharmonikkal
- d) Hány felharmonikust tudunk megszámolni a c) válasz ábráján a 4000 Hz-es pontig?
- $4000/\text{alaphang (legyen 200 Hz)} = 20$
- e) Rajzolja fel az ajkáról kisugárzott „á” hang hangnyomás-idő függvénye vonalas spektrumának közelítését (a magánhangzó háromszögről tanultakat is vegye alapul).
- Spektrum rajzolása a következő paraméterekkel:
 - Vonalas spektrum
 - 150-250 Hz alaphang és hozzá tartozó felharmonikusok
 - Formánsszerkezet kb. 900 és 1500 Hz-es két formánssal
 - Amplitúdó csökkenés: 6 dB/oktáv

3. feladat

- Vázolja egy 110Hz frekvenciájú szinusz-jel és egy ugyanilyen alapfrekvenciájú magánhangzó spektrumának jellemző tulajdonságait és ismertesse az ezzel kapcsolatban tanult fogalmakat!
 - Szinusz:
 - Vonalas spektrum
 - Egy frekvenciakomponens (110 Hz)
 - Magánhangzó:
 - Vonalas spektrum
 - Alapharmonikus + felharmonikusok
 - Formánsok
 - Rajz a paraméterek alapján
 - Fogalmak: alaphang, felharmonikus, formáns

4. feladat

- Írja le a magánhangzó akusztikai létrejöttének, valamint a végső spektrális kép előállításának folyamatát!
 - Lépések:
 - Levegő útja a tüdőből a hangszalagokon és a hangképző üregen át a szájon keresztül a szabad térbe
 - Hangszalagok: zöngé létrehozása
 - Hangképző üreg
 - Formánsok, sugárzási ellenállás
 - Száj: szinte akadálytalanul áthalad

(elmélet: beszéd, mint akusztikai produktum)

- Illusztrálja egy tetszőleges magyar magánhangzón keresztül!
Használjon ábrákat is!
 - Lásd korábbi ábra

5. feladat

- Egy beszédrészlet elemzésekor vonalas spektrumot látunk. Milyen adatokra tudunk következtetni?
 - zöngés hangról van szó
 - Vonalas-e a spektrum? Látható-e alaphang?
 - magánhangzóról van szó
 - Vonalas-e a spektrum? Van-e alaphang? Vannak-e formánsok?
 - Nehéz kérdés!
 - mássalhangzóról van szó
 - Zajszerű-e a spektrum?
 - hang hosszára
 - Látunk-e időbeli információt egy spektrumon?
 - beszélő korára
 - Miben befolyásolja a kor a beszédet?
 - kijelentő vagy kérdő mondatban van-e
 - Modalitást meghatározza: mondat dallama
 - Látszik-e ez?

6. feladat

- Egy telefonos rendszerben a következő felharmonikusokat sikerült kimérnünk: 400Hz, 800Hz, 1000Hz. Mennyi lehet az alapfrekvencia?
 - Fontos megjegyzés: CSAK ezeket mértük, vagy EZEKEN KÍVÜL is lehet még felharmonikus?
 - A felharmonikusok az alapfrekvencia egész számú többszörösei
 - Megoldási lehetőségek
 - 150 Hz
 - 100 Hz
 - 200 Hz
 - 400 Hz
 - 20 Hz
 - 80 Hz

7. feladat

- A Kossuth Rádió egyik férfi rádióbemondóját vizsgálják a gégeészeten. A száz szót kell kiejtenie, ezt veszik fel először egy jó minőségű mikrofonnal, másodsor egy vezetékes telefonon keresztül. A hangfelvételeket elemzik. Válaszoljon a feltett kérdésekre indoklással együtt.
 - a) Mennyi lehet a bemondó alapfrekvenciája?
 - Férfi bemondó: kb 100 Hz átlagos alapfrekvencia
 - b) Rajzolja fel a bemondó zöngéjének nyomás-idő függvényét. Jellemezze. 2 cm=10 ms
 - Ábra paraméterei:
 - Hullámforma rajzolása, kvázi fűrészfog jellegű jel, periódusidő: 0.01 s
 - c) Hány felharmonikust számolhatunk meg az „á” hangban a mikrofonos felvételen?
 - A felharmonikusok az alapfrekvencia egészszámú többszörösei
 - A mintavételi frekvenciától függ: $\#felh = \frac{mintavételi\ frekv./2}{alapfrekvencia}$
(a gyakorlatban kb. 4000 Hz-ig lehet leolvasni felharmonikusokat)

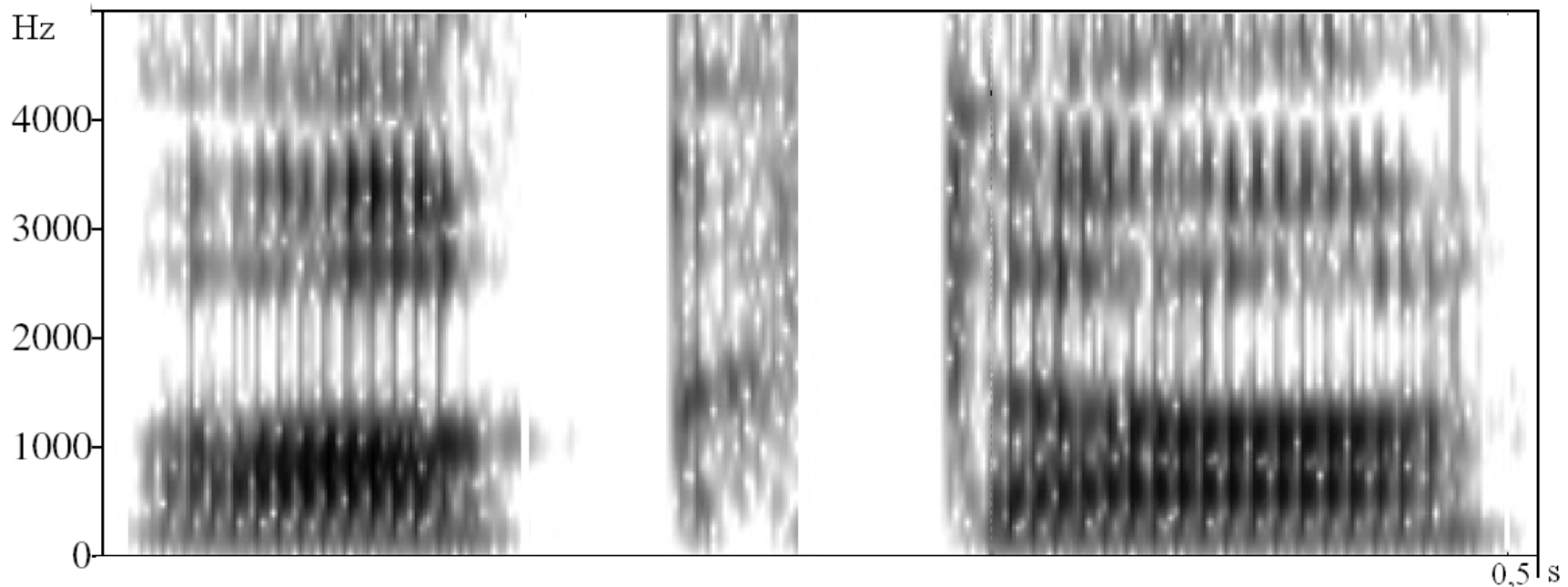
- d) Hány felharmonikust számolhatunk meg az á hangban a vezetékes telefonos felvételen?
- A telefon beszédátviteli frekvenciája: 300-3400 Hz
 - Ebben lehet, hogy nincs benne az első felharmonikus!
- e) Rajzolja fel a mikrofonnal felvett hangsor magánhangzójának közepén mérhető spektrumot. Adja meg a spektrum magánhangzóra jellemző pontjait.
- Lásd korábbi feladat
- f) Rajzolja fel a mikrofonnal felvett hangsor első hangjának közepén mérhető spektrumot. Adja meg a spektrum jellemző pontjait.
- „sz” hang spektruma:
 - Zajszerű spektrum, kiemelt frekvenciatartomány: kb. 2500 Hz-től
- g) Írja le hangszimbólumokkal, hogy mit hallhat az orvos a kétféle hangfelvételből kapott hangsor meghallgatásakor.
- telefonsávós felvétel levágja az energia nagy részét
 - Habár magát az alapfrekvenciát levágja a telefonsáv szűrője, a felharmonikusokból hallani fogjuk
 - Valószínűleg érteni fogjuk
 - Az agyunk „kiegészíti” az megtanult szótárunk alapján
 - Ha a szöveggörnyezet zavaró lenne, akkor minden olyan hang jó, aminek a képzése ugyanaz, csak a felső frekvenciákban térnek el (pl. „f”, „s”)

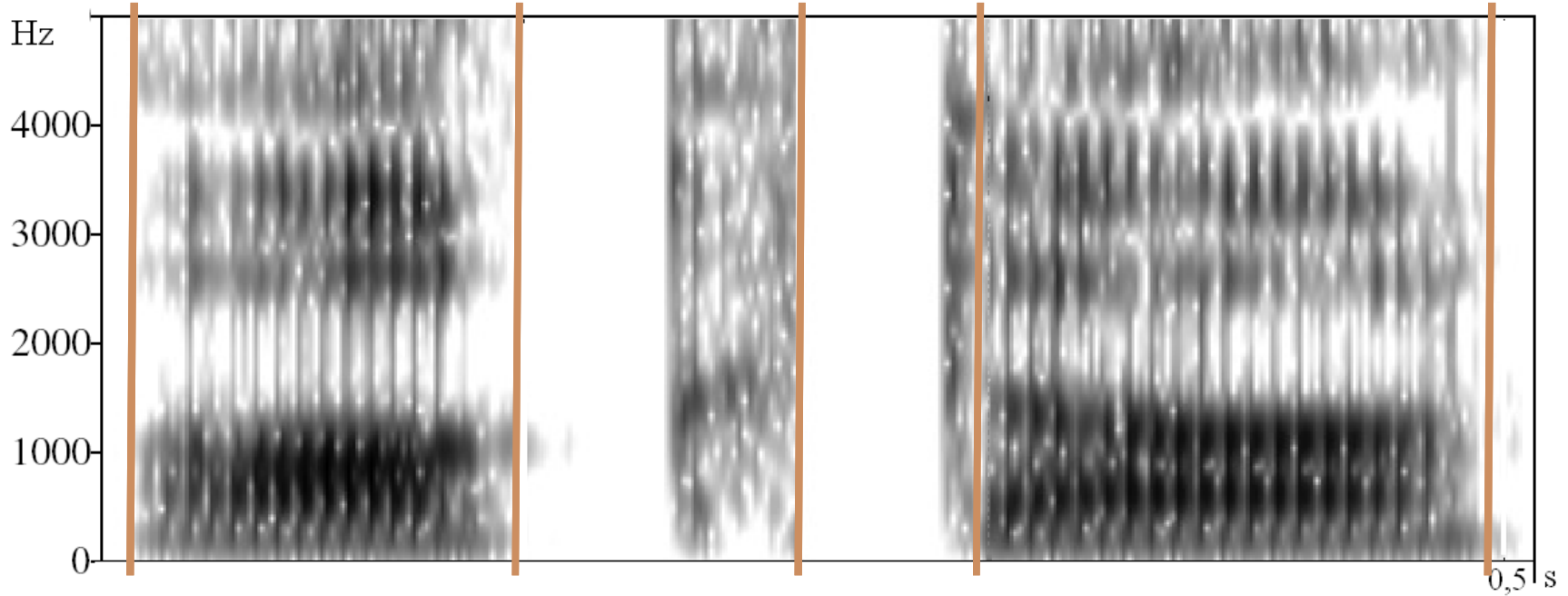
8. feladat

- A *kása* szó időfüggvényéből kell előállítania a *kácsa* szó időfüggvényét egy szokványos hangeditor programmal. Hogyan végzi el a feladatot?
 - „s” → „cs”
 - Zöngétlen marad
 - Réshang → zár-rés hang
 - Időtartambeli különbség
 - Súrlódási zörej időtartama szerint hallott hangok
 - $t > 100 \text{ ms}$ → réshang
 - $50 < t < 100 \text{ ms}$ → affrikáta
 - $t < 50 \text{ ms}$ → zárhang
 - DE: a képzési hely ugyanaz

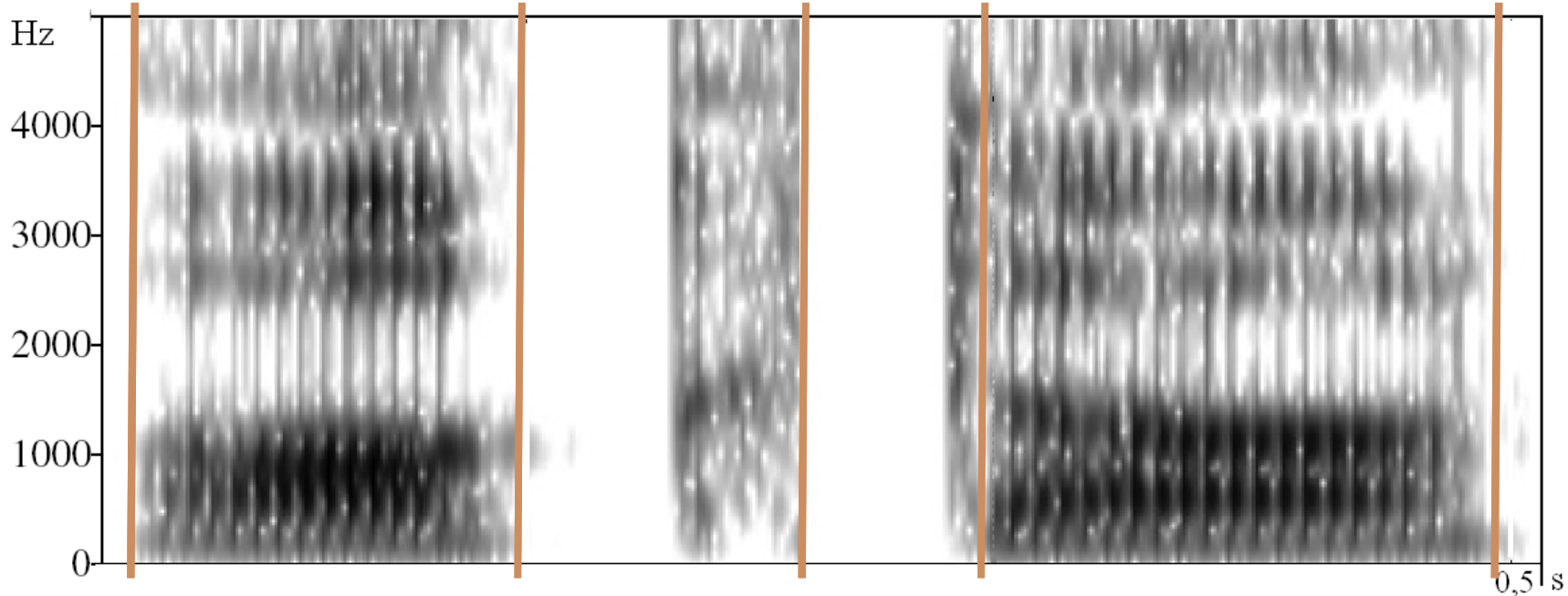
9. feladat

- Elemezze az alábbi spektrogramon látható értelmes magyar szót.
 - Húzza be a beszédhangok határait. Indokoljon!
 - Jellemezze a beszédhangok sajátosságait a tanultak alapján.
 - Próbálja meghatározni a hangsort. Indokoljon!
 - Minden hangra adja meg a gerjesztés típusát (részletesen) és helyét.
 - Női, férfi, vagy gyermek beszéde van a spektrogramon? Magyarázza meg a döntését.

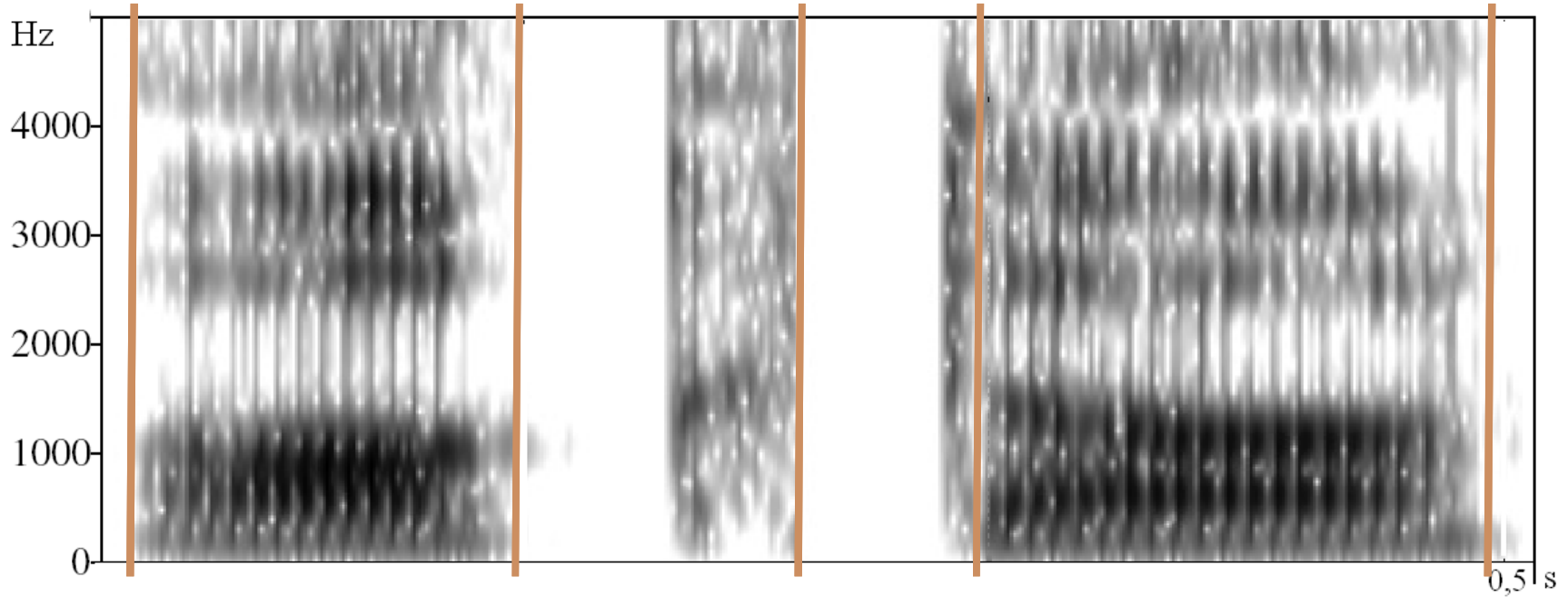




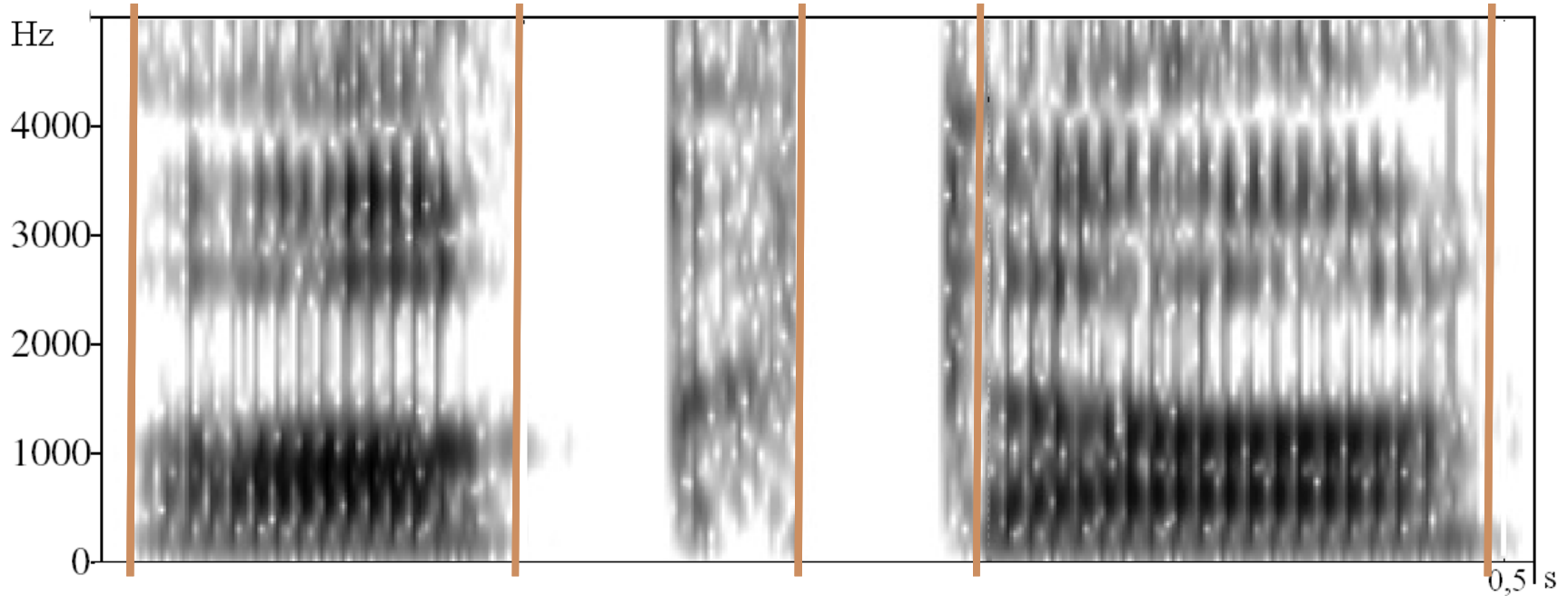
- Húzza be a beszédhangok határait!
 - Alapvető keresési elv: a hangok állandósult (tiszta fázis) és hangátmeneti szakaszokból állnak
 - Változásokat keresünk a spektrogramon
 - A zárhang időben összetett, így az nem két szomszédos átmenet között van
 - Átlagos hangidőtartam: 70 ms



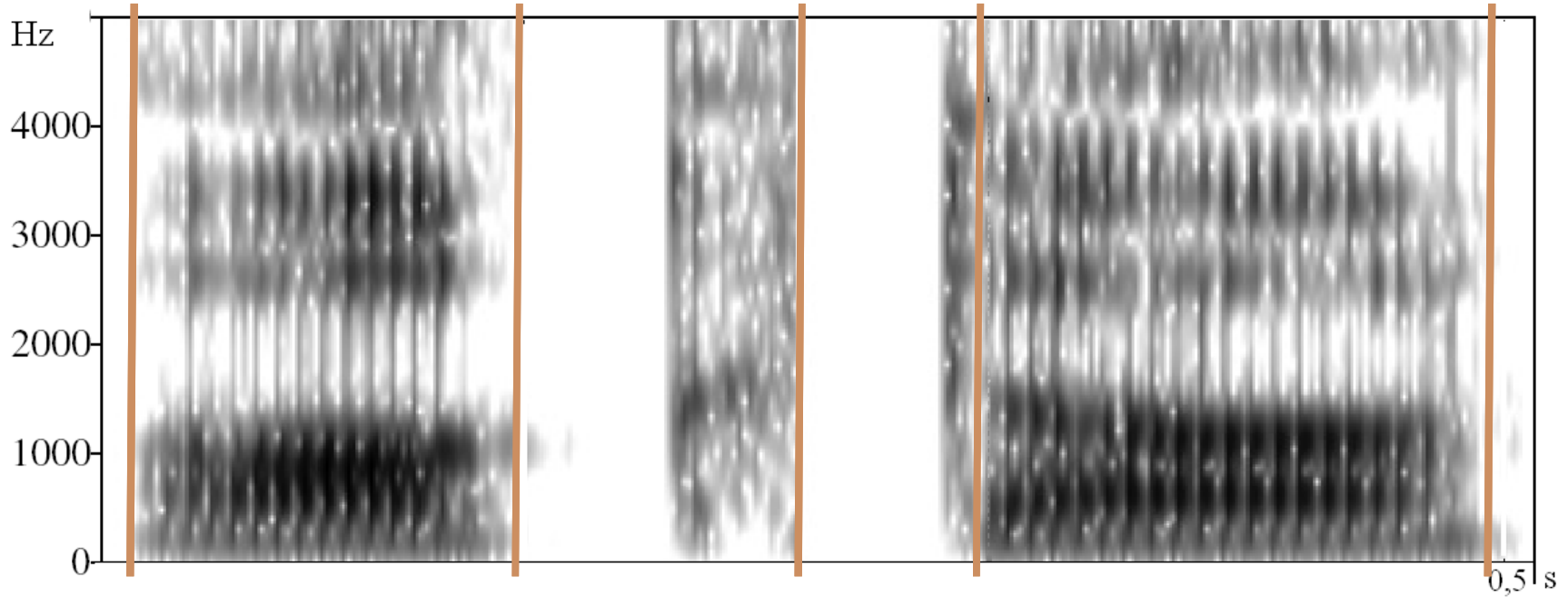
- Jellemezze a beszédhangok sajátosságait a tanultak alapján!
- Gerjesztés megállapítása
 - Van-e zöngé → zöngés gerjesztés
 - Zajszerű-e a spektrum → súrlódási zörej gerjeszti a hangot
 - Esetleg mindkettő jelen van...
- Rendelkezik-e a hang formánsszerkezettel?
 - A formánsok kiemelések a zöngés gerjesztésben → a spektrumon a kiemelt frekvenciák nagyobb amplitúdóval szerepelnek
 - Rendszerint csak a formánsok és a zöngé látszik, mivel ezek arányaiban sokkal nagyobb amplitúdóval rendelkeznek
- Időben egyszerű vagy összetett a hang?
 - Összetett: zárhang, zár-rés hang



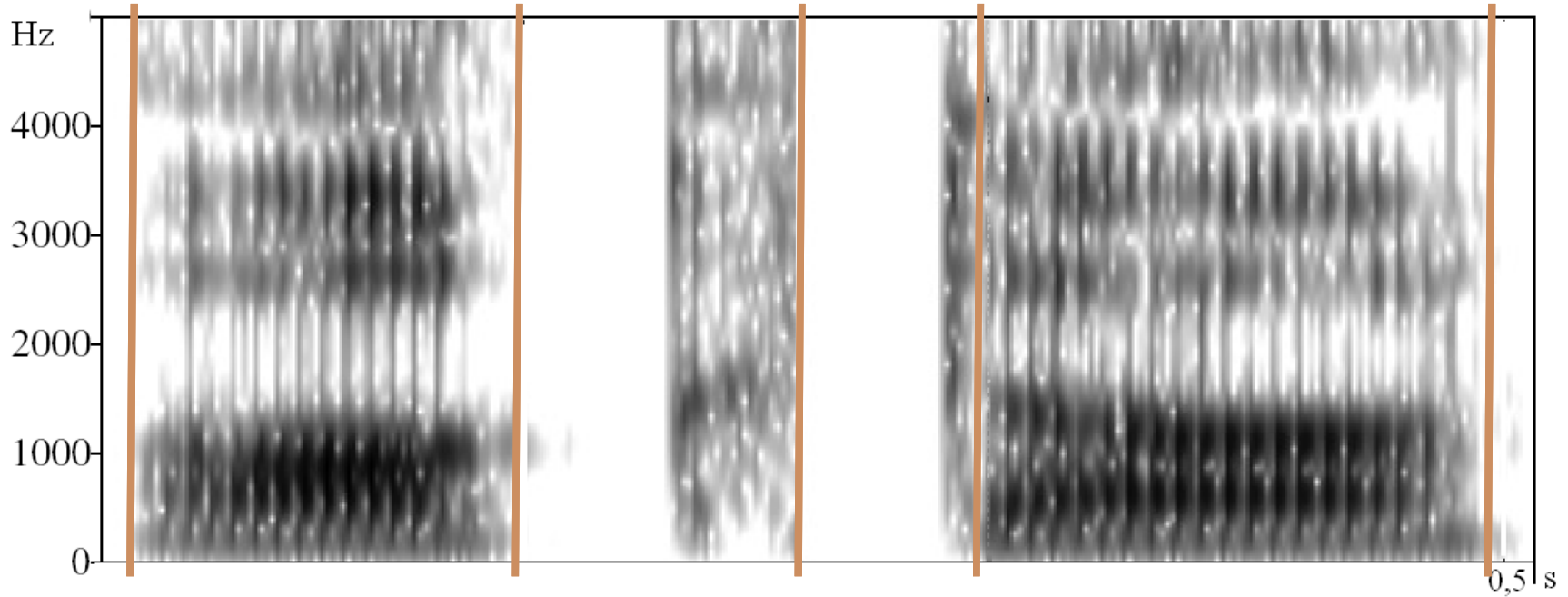
- 1. hang
 - Zöngés, formánsszerkezettel rendelkezik, nincs zörej, időben egyszerű
- 2. hang
 - Zöngétlen, nincs formánszerkezet, van zörej, időben összetett
- 3. hang
 - Zöngétlen, nincs formánszerkezet, van zörej, időben összetett
- 4. hang
 - Zöngés, formánsszerkezettel rendelkezik, nincs zörej, időben egyszerű



- 1. hang
 - Zöngés, formánsszerkezettel rendelkezik, nincs zörej, időben egyszerű
 - → magánhangzó
 - F1 kb. 600 Hz, F2 kb. 1000 Hz → „a”
- 2. hang
 - Zöngétlen, nincs formánszerkezet, van zörej, időben összetett
 - → zöngétlen zárfelpattanás
- 3. hang
 - Zöngétlen, nincs formánszerkezet, van zörej, időben összetett
 - → zöngétlen zárfelpattanás
- 4. hang
 - Zöngés, formánsszerkezettel rendelkezik, nincs zörej, időben egyszerű
 - → magánhangzó
 - F1 kb. 600 Hz, F2 kb. 1100 Hz → „a”
- Lehetséges hangsor: „a”, „k”, „t”, „a”



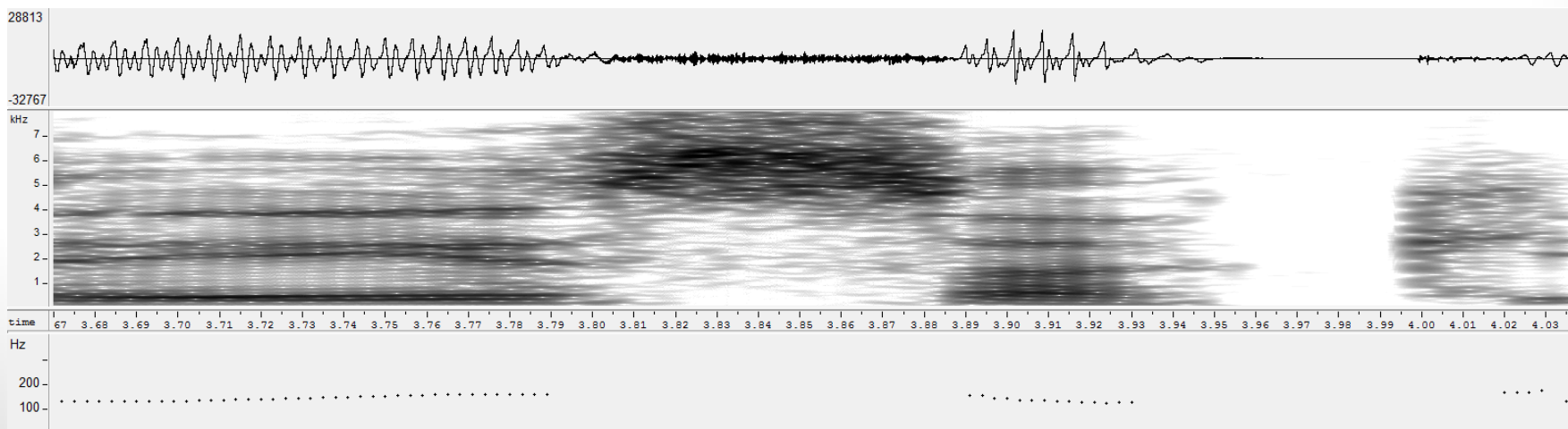
- Minden hangra adja meg a gerjesztés típusát (részletesen) és helyét!
- 1. hang
 - Zöngé, nincs zörej: hangszalagok, gége
- 2. hang
 - Nincs zöngé, van zörej: hangképző üreg
- 3. hang
 - Nincs zöngé, van zörej: hangképző üreg
- 4. hang
 - Zöngé, nincs zörej: hangszalagok, gége

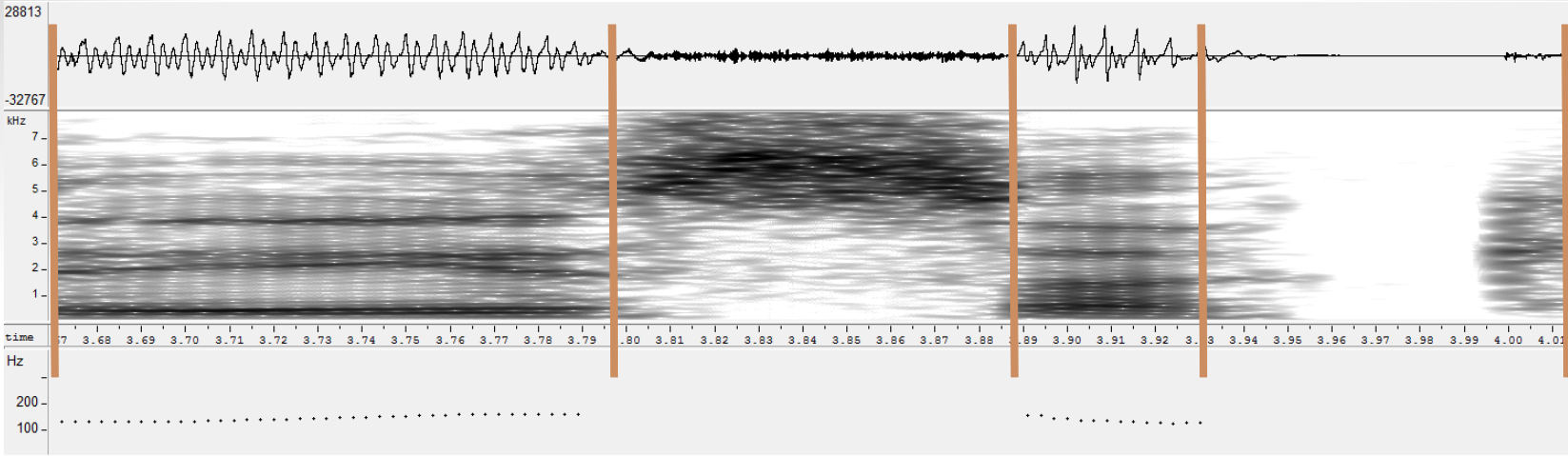


- Női, férfi, vagy gyermek beszéde van a spektrogrammon?
- Nem eldöntése: átlagos alaphang
 - Kb. 100 Hz (80-150 Hz): férfi
 - Kb. 200 Hz (150-250 Hz): nő
 - Kb. 400 Hz (250-500 Hz): gyerek
- Az ábrán kb. 110 Hz az első hangon mért átlagos alaphang
 - Férfi

10. feladat

- Ráilleszthető-e az ábrán látott hullámforma (és spektrogramm) a következő elhangzott beszéd valamely részére: „az északi szél”? Ha igen, akkor mely részére, ha nem, akkor miért nem?
- Férfi vagy nő a hullámformához tartozó beszélő?





- Milyen hangok láthatóak az ábrán?
- 1. hang
 - Zöngés, formánsszerkezettel rendelkezik, egyszerű
 - Valószínűleg magánhangzó
- 2. hang
 - Zörejes, zöngé ninc, egyszerű
 - Valószínűleg réshang
- 3. hang
 - Zöngés, formánsszerkezettel rendelkezik, egyszerű
 - Valószínűleg magánhangzó
- 4. hang
 - Zörejes, zöngé ninc, időben egyszerű
 - Valószínűleg felpattanó

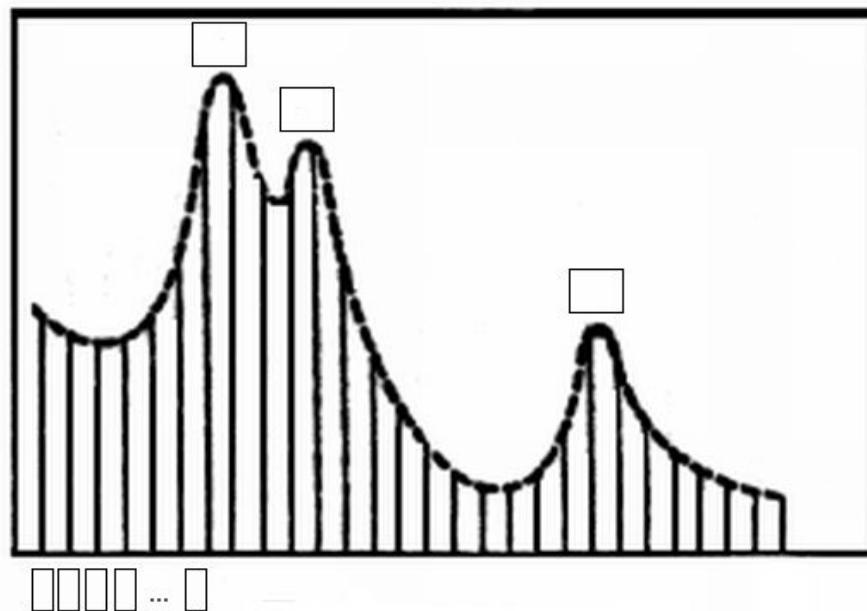
Keressük meg, milyen hangsorozatra illik rá az elhangzott „az északi szél”-ből!

11. feladat

- Rajzolja fel a férfi ejtésű „ő” magánhangzó képzésekor keletkezett zöngéhang, valamint a végleges beszédhang spektrumát 100-4000 Hz-es frekvenciatartományra
 - Rajzolás szempontjai
 - Zöngéhang
 - Alapharmonikus
 - Felharmonikusok, 12 dB/oktáv amplitúdó esés
 - Férfi: 80-150 Hz alaphang
 - Végleges hang
 - Formánsok: 500Hz, 1500Hz
 - 6 dB/oktáv amplitúdó esés
 - A végleges hang gerjesztése a zöngé, tehát a különbség ott van, amit a hangszalagok utáni hangképző üreg módosít a zöngén
 - miben egyezik a két spektrum
 - Vonalas, alaphang és felharmonikusok megegyeznek
 - miben különbözik a két spektrum
 - A felharmonikusok amplitúdó esése eltér
 - A végső hangnál létrejönnek a formánsok → frekvenciabeli kiemelések

12. feladat

- Egészítse ki az ábrát, beírva a keretekbe a megfelelő jelöléseket. Ismertesse az ezzel kapcsolatban tanult fogalmakat!
- Mely beszédhang típushoz tartozik az ábrán látható spektrum?
 - Vonalas spektrum
 - Alapfrekvencia: zöngés hang
 - Jól látható formánsszerkezet
 - magánhangzó



13 feladat

„házi feladat”

- Jellemezze a saját nevének kiejtett formáját beszédképzési és akusztikai szempontból.
 - a) Osztályozza nevének hangjait külön-külön legalább 3 osztályozási forma szerint.
 - Magánhangzó-mássalhangzó, rövid-hosszú hang, orális-nazális, gerjesztési hely, gerjesztési forma, egyszerű-összetett
 - b) Rajzolja fel a kiejtett név beszédjelére jellemző közelítő intenzitás-idő függvényt és jelölje benne a hanghatárokat is. Időosztás: 100 ms = kb. 2 cm.
 - c) Helyezze el a magánhangzókat az $x=F1$, $y = F2$ koordináta rendszerben a magánhangzó háromszög sarokpontjainak figyelembevételével. Frekvenciaosztás: 500 Hz= kb. 1 cm.
 - d) Adja meg a mássalhangzókat jellemző 3 paramétert mindegyik mássalhangzóra.
 - e) Jellemezze a hangkapcsolódások során végbemenő akusztikai változásokat.

14. feladat

- Milyen frekvenciatartományban legintenzívebb az 'sz' hang spektruma?
 - kb. 30Hz-100Hz
 - kb. 100Hz-300Hz
 - kb. 300Hz-1000Hz
 - kb. 1000Hz-3kHz
 - kb. 3kHz-10kHz
 - kb. 10kHz-20kHz

Appendix

- A magyar beszéd (19.o. – 26.o, 3. fejezet: Fiziológiai, fizikai alapok)

3.1 A beszédképzés folyamata

- *Beszédkeltés alapvető fiziológiai szervei:*

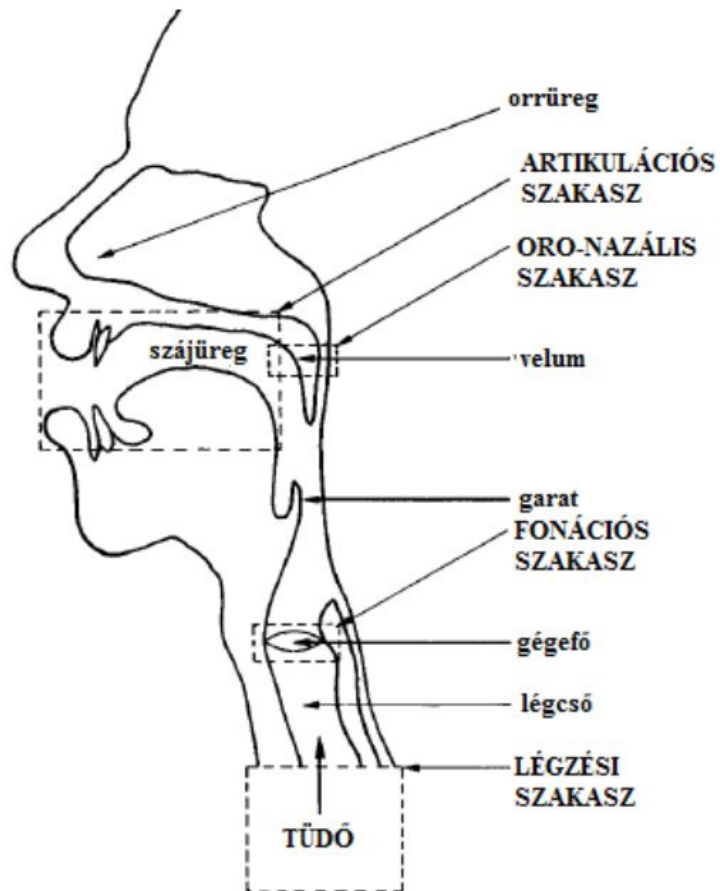
1. *Tüdő, légcső – szubglottális tér : levegőáram biztosítása*
2. *Gége – glottális tér: zöngés/zöngétlen hangképzés*
3. *Garat, száj- és orrüreg – szupraglottális tér :*

a beszédhangok az artikulációs csatornában megkapják végleges hullámformájukat

→ 1 + 2: hangkeltés, 3 : hangformálás



Appendix



Artikulációs (hangképzési) szakasz: toldalékcső

Oro-nazális szakasz: lágy szájpadlás (velum)

Fonációs (zöngéképző) szakasz: gégefő (larynx), hangszalagok

Respirációs (légzési) szakasz: tüdő

Forrás: Balogné Bérces Katalin : Fonetika és fonológia (előadás)

PPKE BTK, Budapest/Piliscsaba, 2015. október

<https://btk.ppke.hu/uploads/articles/463195/file/BBK-Kolozsv%C3%A1r-2-fonetika-mgh.pdf>

Appendix

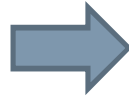
3.1.1 Gégeszintű hangképzés (glottális szakasz)

- A tüdőből kipréselt levegőáram a légcsőből a gégébe jut
 - A hangszalagok állása (kannaporcok közreműködésével) meghatározza, hogy a levegőáram hogyan jut tovább a száj- és orrüregbe
 - ❖ Tág légző állás : kezdeti
 - ❖ Suttogó állás
 - ❖ Zárállás
 - ❖ H-állás
 - ❖ **Fúvó állás**
 - ❖ **Zöngéállás**
- ← Ezek váltakoznak – zöngétlen/zöngés beszédhangok
- Zöngéállásnál a **hangszalagok kváziperiodikusan szétnyílnak-összezáródnak**, a levegő nyomásviszony-változásának következtében (tüdőből préseljük) → létrejön a **zöngéshang**

Appendix

3.1.1 Gégeszintű hangképzés – Zöngés hang

- A zöng (fonáció) kialakulása tehát **két fő lépésre**, egy nyitódás és egy záródás folyamatra osztható
- Ez a **zöngés hang ejtésének ideje alatt folyamatosan ismétlődik**, melyet egy **periódusidővel** jellemzünk
 - ❖ Férfiaknál : 8-12 ms
 - ❖ Nőknél : 4-6 ms
 - ❖ Gyermeknél még rövidebb
- Élettani tényező befolyásolják -> „kvázi” periodikus → fűrészfog jellegű
- A zöng amplitúdója a hangerősséget felelős
- **Átlagos alaphangfrekvencia** a beszélő nemére enged következtetni:
 - ❖ Férfi: 100 Hz, Nők : 180 Hz, gyermekek: több száz
- De a beszélő is széles skálán képes változtatni : 50-500 Hz között



Alaphang(F0): beszéd alaphangfrekvenciája
Reciprok kapcsolat a periódusidővel(1/s,Hz)

Appendix

3.1.1 Gégeszintű hangképzés – Zöngés hang

- Az egészséges zöngéhang az alapja a tiszta, karakteres beszédnek
- A zöngéképzést számos tényező zavarhatja
 - ❖ Pl. betegség: időskorban zöngétlen hangok zöngéssé válnak
 - ❖ Irreguláris zöngé: Amplitúdóingadozás lép fel a beszédben -> glottalizáció (fogyó levegő)
- Speciális zöngéképzés: **Fojtott zöngé**
 - Az orális és nazális csatorna le van zárva, nem áramlik ki levegő az artikulációs csatornából → pl. baba szóban a „b”-nél
 - Rövid időtartamú, amíg kisebb a szájüregben lévő nyomás a légcsőből érkezőnél.
 - Az arcüreg megrezegteti a külső levegőt

Appendix

- **Zöngétlen hangok esetén**

- ❖ Nincs fonáció

- ❖ Fúvó állásban a gége nem jelent akadályt a tüdőből kipréselt levegőnek -> turbulens áramlás → réshangok

- ❖ Az akadályok az artikulációs csatornában hozhatók létre:

1. Zárhangok: Lökéshullámok eredményeképpen, néma szakasz
2. Zár-rés hangok

- ❖ Kivétel: H-állás

Appendix

3.1.2 Az artikulációs csatorna, toldalékcső (szupraglottális)

- A hangszalagok és az ajkak közötti hangképzési terület
- Három üregrész: garat, szájüreg, orrüreg
- Itt formálódnak ki a beszédhangok, **artikulációs szervek** segítségével

❖ **Aktív:**

1. **Nyelv:** előre-hátra ill. fel-le mozgás : magánhangzó képzés \leftrightarrow ha akadályt is képez: mássalhangzó (zárhang)
2. **Ajak:** kerekítés, széthúzás: magánhangzó \leftrightarrow összezárás + fogakkal interakció : mássalhangzó

❖ **Passzív:** orr (orrüreg)

- Így adódik a mássalhangzók csoportosítása:
 - **Képzés helye:** bilabiális | labiodentális | dentialveorális | alveorális | palatális | veláris | faringális | nazális
 - **Képzés módja:** zárhang | réshang | zár –rés hang | pergőhang | közelítőhang

Appendix

3.1.2.2 A gégeműködés és az artikulációs csatorna viszonya

- A **gégében**, levegőráamlás által **létrejött hang** a beszéd alapkomponense → **Az artikulációs csatorna erre a hangra épít, a beszéd itt nyeri el végleges rezgésformáját.**
- A kisugárzott beszédhangok jellemző frekvenciakomponenseiket
 - A toldalékcsőben kialakult rezonanciafrekvenciáknak (~ magánhangzók)
 - Létrehozott szűkületek helyének és alakjának (~ mássalhangzók)
köszönhetik.
- A toldalékcső rezonátorüregként vesz részt a hangképzésben -> bonyolult rezgésformák jönnek létre, többféle rezonanciafrekvencia van a rezonátorüregben

Appendix

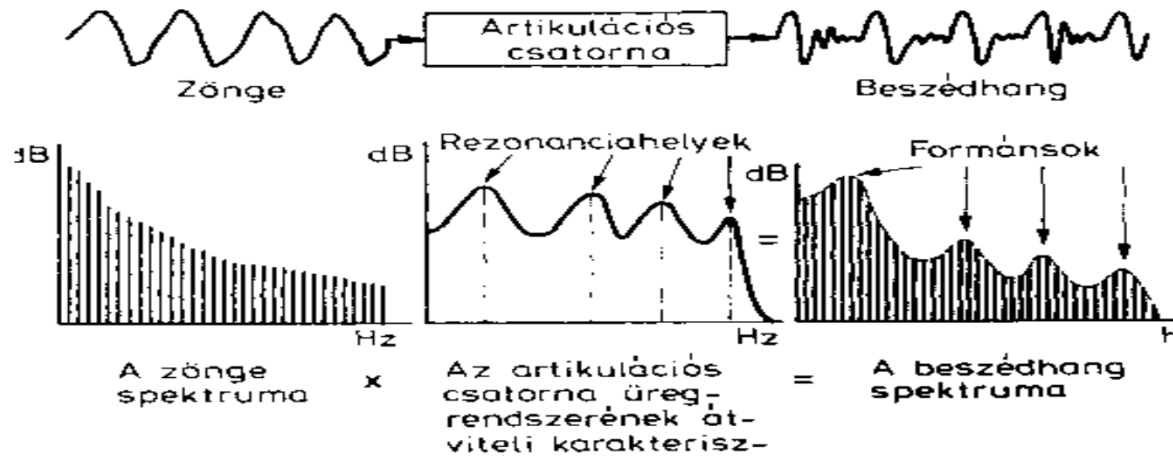
3.3.5.1. Formáns, zörejgóc

- A zöngés elemek frekvenciaszerkezete a gégeszintű hangforrás és a toldalékcső (mint rezonátorüreg) együttes tulajdonságaiból adódik, pillanatnyi spektrummal jellemezhető -> akusztikus vetület
- Fonáció esetén a zöngé rezgésformája gerjeszti az artikulációs csatornát
- A **zöngé: az alaphang + felharmonikusok**, kb. 5000 Hz-ig (gyakorlatban 4000 Hz-ig tudjuk leolvasni)
- **A vonalas spektrumban** a felhangok amplitúdó 12 db/oktáv csökkenéssel vannak jelen, legnagyobb amplitúdójú az alaphang
- **Felharmonikusok** egyes csoportjai **a pillanatnyi rezonanciafrekvenciákon és környékükön felerősödnek: Formáns**
- A vonalakra helyezett burkológörbe csúcsainál vannak
- A nagyobb sorszámú távolabb van az alaphangtól, nagyobb a sáv szélessége, az amplitúdó esése: 6 db/oktáv

Appendix

3.3.5.1. Formáns, zörejgóc

- A magánhangzókra jellemző az F1-F2 + azok sávszélessége
- Zárhangok zárszakaszában nem mérhető
- Zöngétlen hangoknál turbulens áramlások, zárfelpattanási zörejek (lökéshullámok) gerjesztik a csatornát, képzés helyének megfelelő **zörejgócok** alakulnak ki

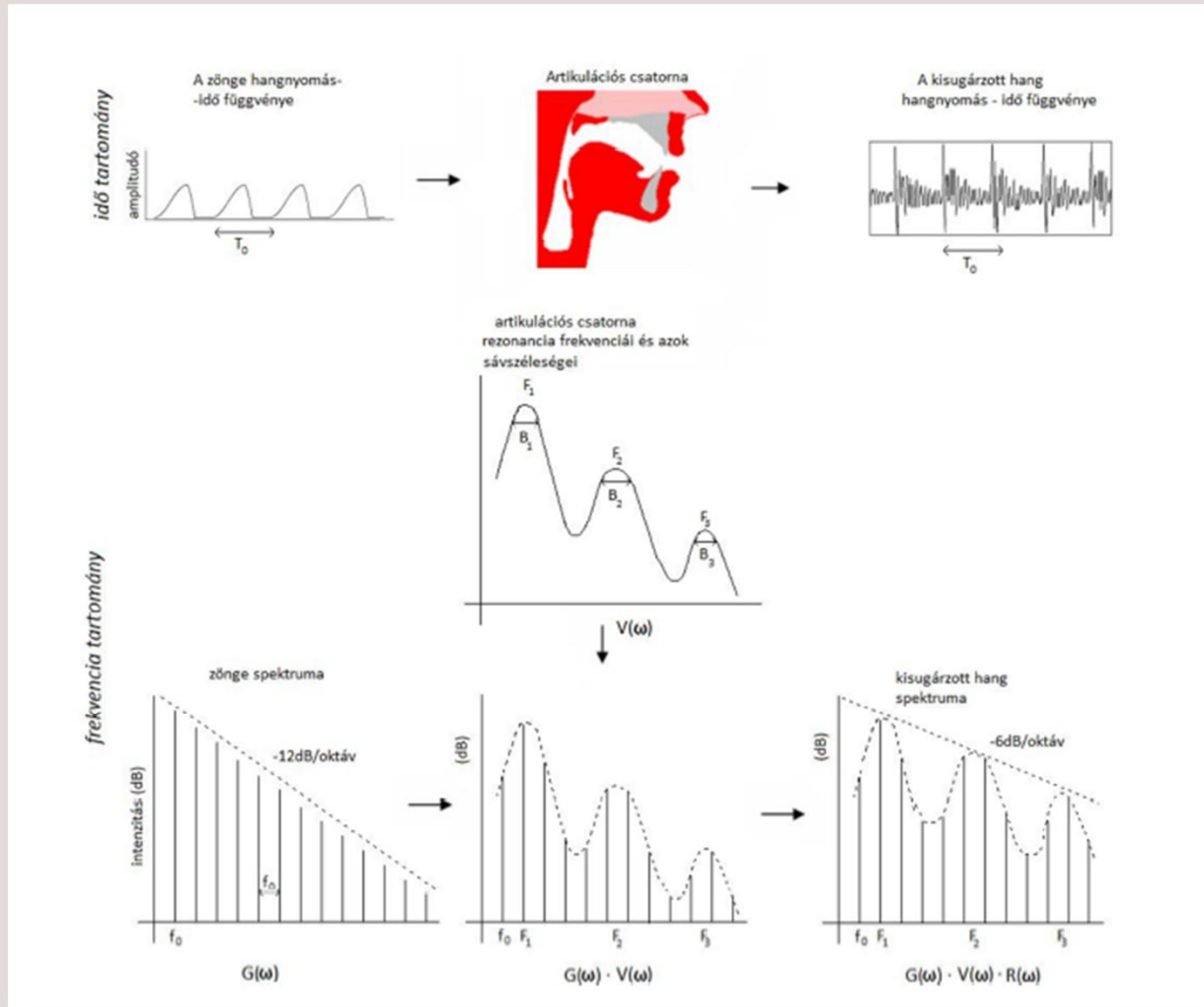


Appendix

3.3.5.1. Formáns, zörejtől – *A gerjesztett szűrő modellje*

- **A hangképző szervek komplex akusztikai rendszerét egyszerűbb modellekre képezzük le**, pl. az artikulációs csatornát is modellezük, hogyan képezzük a beszédhangokat
- $G(\omega)$: az alaphang + felhangok együttesét leírja, -12 dB/oktáv intenzitáscsökkenés. A hangszalag rezgésével létrejött hangforrás színeképe (Glottal)
- $V(\omega)$: az artikulációs csatorna befolyásolja ezt a színeképet, üregrendszer több rezonanciafrekvenciával -> kialakulnak a formánsok. (Vocal Tract – átviteli függvénye befolyásol)
- $R(\omega)$: a hangtérbe jutó hangszín képét befolyásoló átviteli függvény. Magasabb frekvenciákon kisebb a sugárzási ellenállás, így a felső frekvenciatartomány intenzitását megemeli 6 db/oktávval.
- mindig van környezeti hatás: $N(\omega)$, additív komponens
- $\rightarrow G(\omega) * V(\omega) * R(\omega) + N(\omega) = S(\omega)$ –a beszéd eredő színeképe

Appendix



Forrás: Vicsi Klára: A beszéd akusztikai-fonetikai leírása

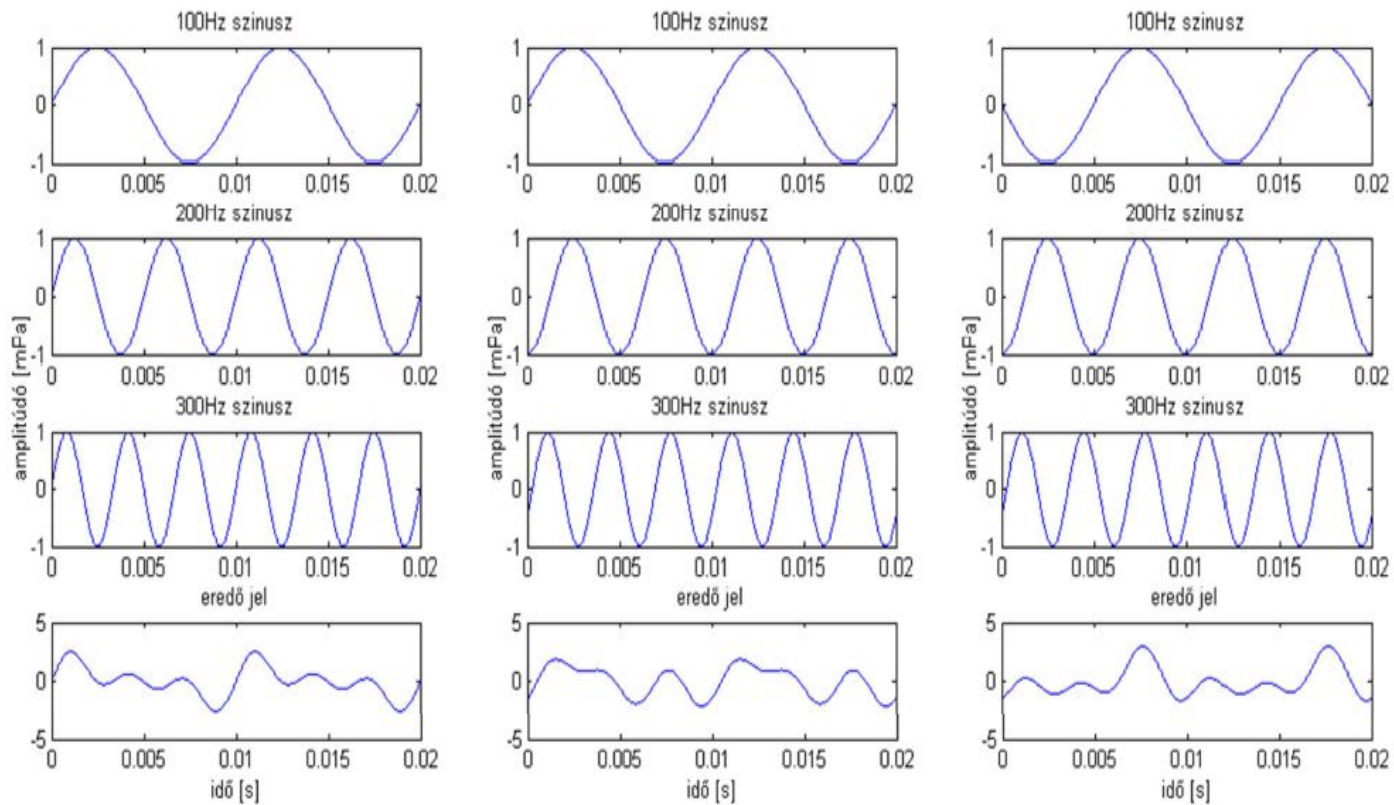
Appendix

3.3.A beszéd fizikai jellemzése

3.3.4. Összetett rezgések

- A szinuszos formájú harmonikus rezgés által keltett tisztahang ritka jelenség (hangvilla)
- Következésképp: többnyire összetett rezgésekkel találkozunk
- Ha lineáris szuperpozícióval összegzünk tiszta szinuszos jeleket, kialakulhatnak nem tiszta szinuszos periódusos rezgések is
- Az összetevő rezgések frekvenciájának aránya egész szám a legalacsonyabb frekvenciájú rezgéshez viszonyítva
- Így ismét periodikus rezgéseket kapunk, a frekvencia a legkisebb összetevője lesz, de alakja nem szinuszos, függ az összetevő frekvenciáktól , és kezdőfázistól (nem befolyásolja az érzeti hangzást)

Appendix



Forrás: Vicsi Klára: A beszéd akusztikai-fonetikai leírása

Appendix

3.3.4. Összetett rezgések frekvenciaelemzése

- Összetett rezgéseket veszünk fel a mikrofonnal is
- Szuperpozíció elve: egymással egész számú viszonyban álló frekvenciájú szinuszos rezgések párhuzamos összetétele -> periodikus rezgés
- Ennek megfordítása is igaz: A periodikus rezgések részelemei szinuszos rezgések
- Fontos megállapítani a frekvenciakomponenseket -> frekvenciaelemzés
- **Fourier szerint: lineáris rendszerekben bármely összetett rezgés időfüggvénye felbontható különböző frekvenciájú, amplitúdójú és fázisú harmonikus komponenseire – szinusz/koszinusz jelek sokaságára**
- Az összetett periodikus rezgések : Fourier-féle rezgések

Appendix

Összetett rezgések frekvenciaelemzése

Fourier-féle elemzési módszer

Fourier tétel – bármely periodikus, de nem szinuszos rezgés felbontható olyan szinuszos rezgésekre, amelynek frekvenciái az alaphang egészszámú többszörösei.

hallószervünk fázisérzéken!

Fourier elemzés során az időfüggvényeket frekvencia függvényekké alakítjuk át.

$$f(t) = \sum_{k=0}^{\infty} a_k \cos(k \cdot 2\pi f_0 \cdot t) + \sum_{k=1}^{\infty} b_k \sin(k \cdot 2\pi f_0 \cdot t)$$

$$f_0 = \frac{1}{T}$$

diszkrét rezgések a_k b_k amplitúdóit és fázisát a Fourier együtthatók adják meg

$$a_k = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} f(t) \cdot \cos(k \cdot 2\pi \cdot f_0 \cdot t) dt$$

$$b_k = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} f(t) \cdot \sin(k \cdot 2\pi \cdot f_0 \cdot t) dt$$

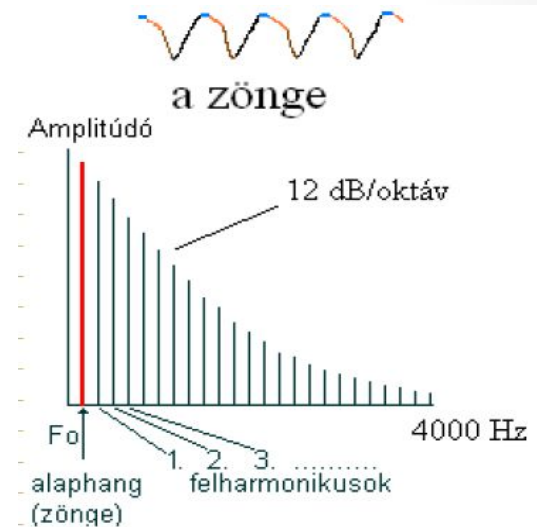
$$\varphi_k = \arctan \frac{b_k}{a_k}$$

Ezt a frekvenciafüggvényt (frekvencia, amplitúdó és fázis adatok összességét) nevezzük *spektrum*nak vagy színeképnek.

Appendix

3.3.4. Összetett rezgések frekvenciaelemzése

- Frekvenciakomponensekre bontás → összetett rezgés hangnyomás-idő függvényét felbontjuk frekvenciatartományban értelmezhető függvénné
- Többféle spektrum: nyomásamplitúdó, teljesítmény, energia: az adott időpontban a frekvencia-összetevők melyik eloszlást adják meg.
- Periodikus, összetett rezgések jellemző alapperiódusa az **alaphang**: az összetett hangot felépítő összes frekvencia-összetevő közül a legalacsonyabb frekvenciájú (F_0)
- Többi összetevő - felharmonikusok: a legalacsonyabb frekvenciájú alaphang egész számú többszörösei → vonalas spektrum jellemző a periodikus rezgésekre, hiszen csak a felhangok frekvenciáin vannak spektrumösszetevők
- Hallási tulajdonságunk révén akkor is meg tudjuk ítélni az alaphangot, ha nincsen benne az F_0 a beszédjelben (pl. telefon)



Appendix

3.3.4. *Összetett rezgések frekvenciaelemzése*

- A nemperiodikus jelekre folytonos spektrum jellemző
 - sok frekvenciájú szinuszos összetevőből állnak, a frekvenciatartomány bármely pontján lehetnek, az összetevők nem meghatározott frekvenciáknál koncentrálnak
- Ábrázolások a frekvenciatartományban
 - ❖ Spektrum: 2D – vízszintes tengely a frekvencia, függ.: intenzitás
 - ❖ Spektogram: 3D – x: idő, y: frekvencia, z: amplitúdó(szürkeárny.)

Appendix

3.3.5. A beszédjel elemzése

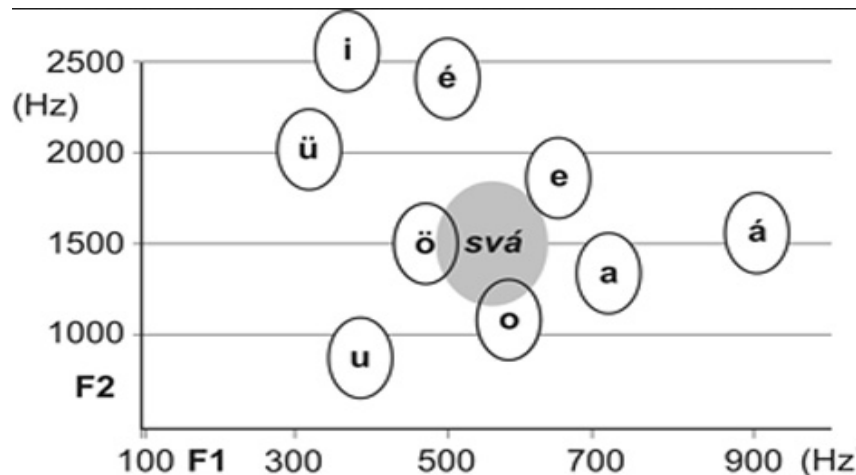
- Fontos, mekkora időintervallumot választunk
 - Rövid ablakszélesség : Időben gyors változásokhoz: (pl.zárfelpattanás), zöngés hangperiódusokból az alapfrekvencia kiszámítható
 - Hosszú időablak: Finom frekvenciafelbontás, zöngés hangok felhangjai látszanak
 - Minél feketébb a kép, annál nagyobb a hangrész frekvenciaösszetevőjének az intenzitása
 - Zöngés-zöngétlen szakaszok megállapítására is alkalmas
 - Kisebb-nagyobb energiájú szakaszok váltogatják egymást
 - A TF (mintavételi ablak) $\cdot \Delta\omega$ (frekvenciafelbontás, frekvenciaösszetevők távolsága DFT-ben) = állandó érték.

Appendix

- A magyar beszéd (106.o, 5. fejezet: A beszéd szegmentális szerkezete)

5.1.2 A magyar magánhangzók

- 9-féle + vannak hosszú párjaik + svá
- Három fő artikulációs alapparaméter: nyelv, ajkak, állkapocs helyzete
- Egyszerű szerkezet, kitartva ejthető, zöngés gerjesztésűek → a zöngé kváziperiodikus jelként gerjeszti a gége feletti artikulációs csatornát
- A magánhangzók formánsai adott sávon belül mozognak → az artikulációs csatorna hosszabb: lefelé tolja a formánsokat + a koartikulációval is mozoghatnak



Appendix

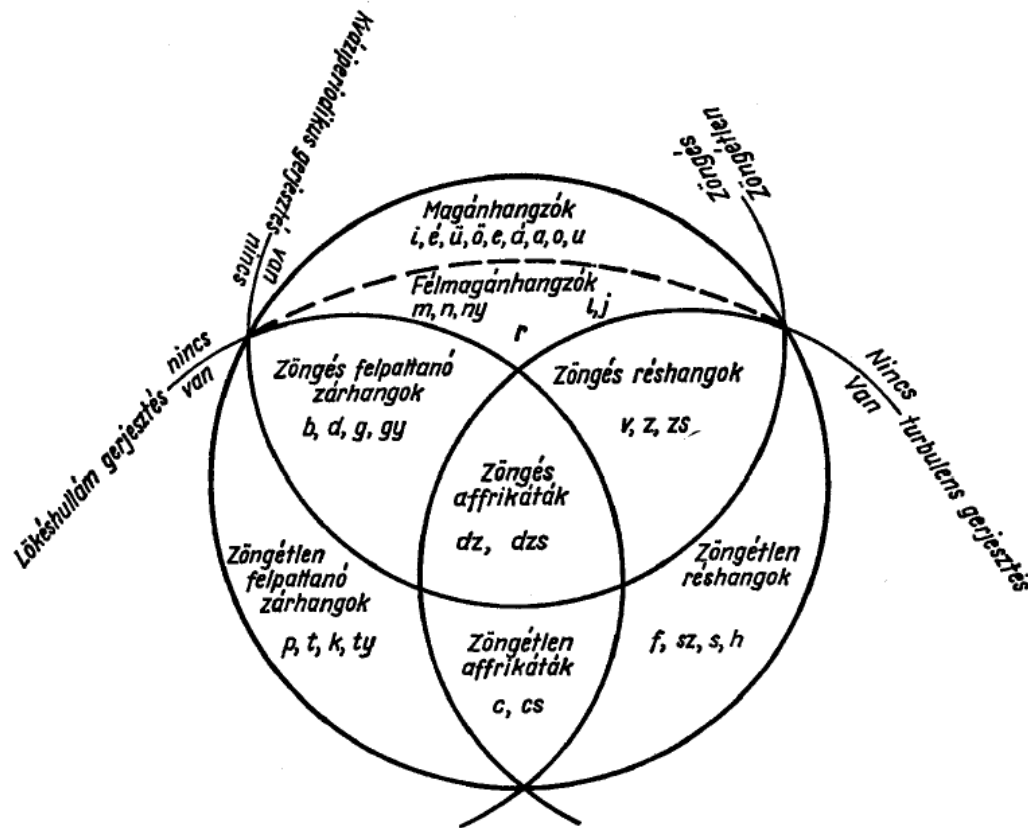
A magyar mássalhangzók akusztikai jellemzése (átlagos, az adott csoportra jellemző tájékoztató értékekkel)

Mássalhangzók	Szerkezet	Gerjesztés	F1 (Hz)	F2 (Hz)	Z1 (Hz)	Z2 (Hz)
felpattanó zöngés zárhangok	összetett	zöngés	300–500		500–2200	
felpattanó zöngétlen zárhangok	összetett	zörejes			500–2200	
nazálisok	egyszerű	zöngés	300	800–2200		
zöngés réshangok (kivéve [v])	egyszerű	kevert	300	1400–1800	1600–4600	3200–8000
[v]	egyszerű	kevert	300	1200	1500–7000	
zöngétlen réshangok	egyszerű	zörejes			1600–4600	3200–8000
zöngés affrikáták	összetett	kevert	300	1400–1800	1600–4800	3500–7000
zöngétlen affrikáták	összetett	zörejes			1600–4800	3500–7000
pergőhangok	egyszerű	zöngés	300–500	1000–1600		
közelítőhangok	egyszerű	zöngés	300–400	1600–2500		

Forrás: Gósy Mária: Fonetika, a beszéd tudománya c. könyv – A beszédhangok akusztikai szerkezete c. fejezet

Appendix

Gordos-körök a beszédhangok osztályzására



Forrás: Gordos-Takács: Digitális beszédfeldolgozás (1983)