

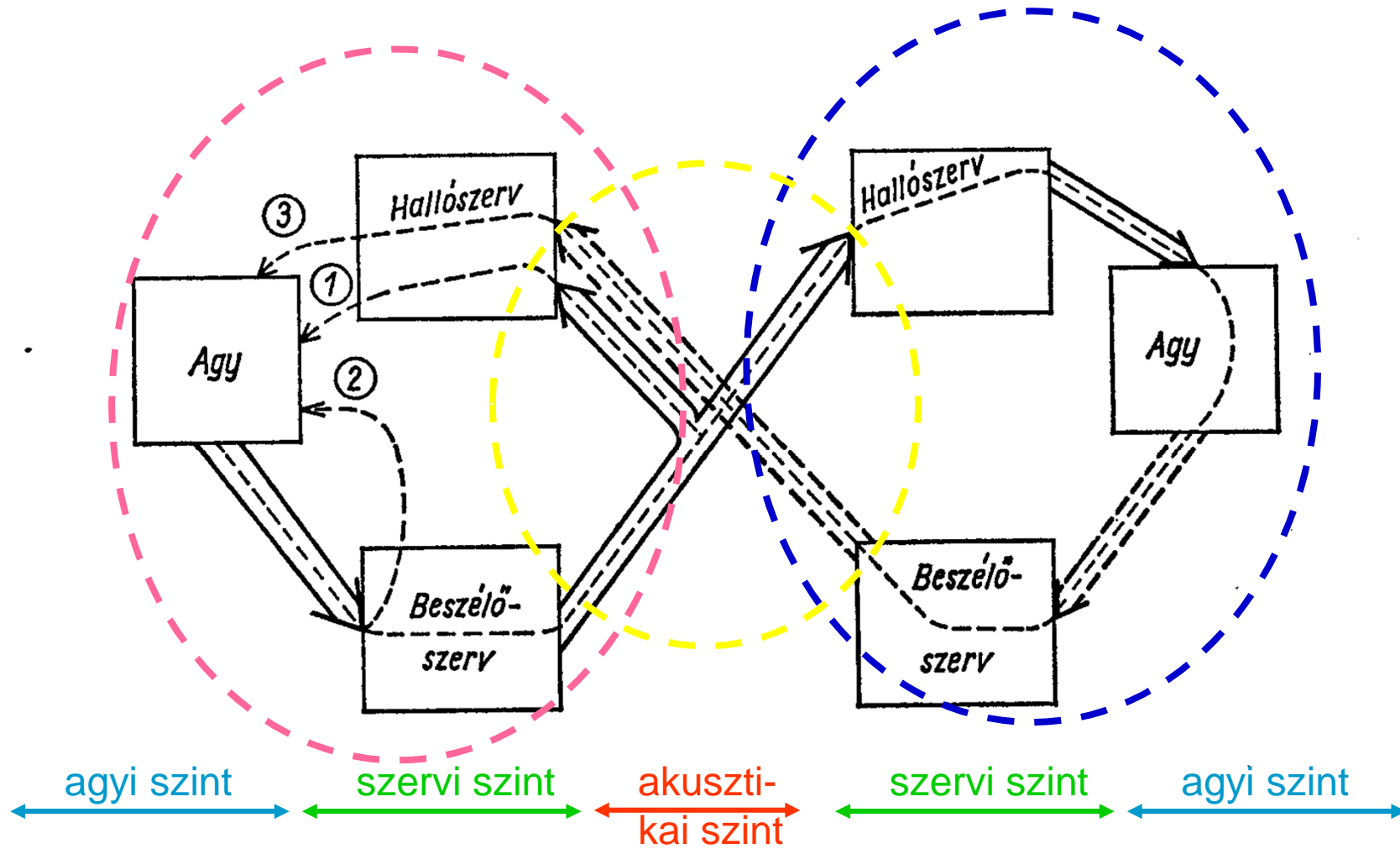
# **A beszéd**

**Segédlet a Kommunikáció-akusztika  
tanulásához**

# Bevezetés

- **Nyelv:**
  - az emberi társadalom egyedei közötti kommunikáció
  - az egyed gondolkodásának legfőbb eszköze
  
- **Beszéd:**
  - a nyelv elsődleges megnyilvánulása
  - az emberi kommunikáció **legelterjedtebb** formája
  - a **leghatékonyabb** emberi kommunikációs mód
  - kronológiai szempontból a **legelső**: jóval megelőzte az írást
  
- **Irodalom:**
  - Gordos-Takács: Digitális beszédfeldolgozás. Műszaki Könyvkiadó, Budapest, 1983.

# A természetes beszédlánc



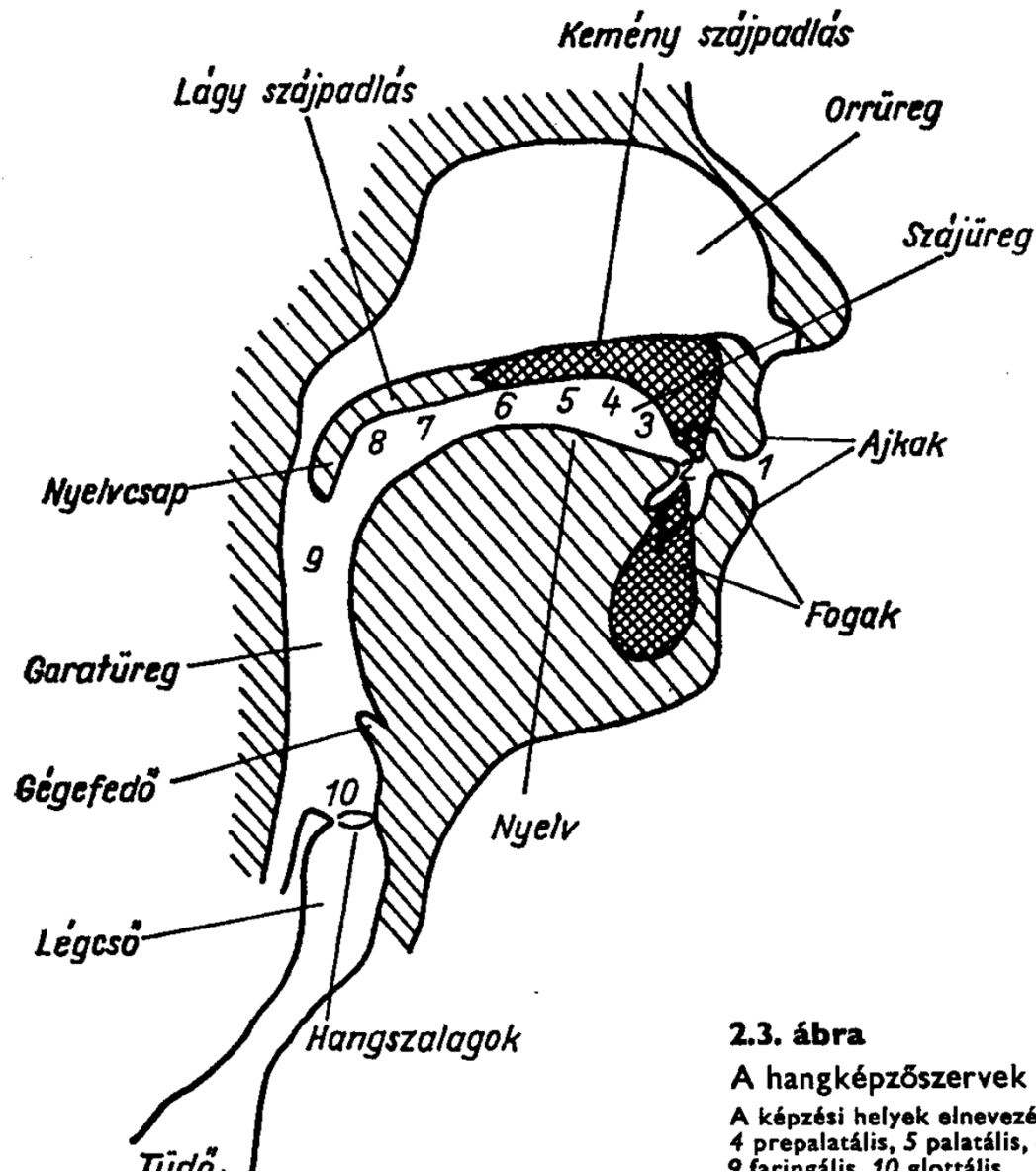
# A beszéd szerkezete

- Mondat > szó > szótag > beszédhang

- Beszédhang <> fonémák
  - Fonémakészlet: elemek olyan minimális számosságú halmaza, amelyből minden szó csak egyféleképpen állítható elő

Fonéma $x_i$	Kulcsszó	$P(x_i) \%$	$I(x_i)$	Fonéma $x_i$	Kulcsszó	$P(x_i) \%$	$I(x_i)$
(ε)	eke	11,22	0,3541	(l)	lap	5,26	0,2234
(e:)	élet	3,26	0,1610	(t:)	mellé	0,32	0,0265
(i)	igen	4,69	0,2070	(r)	ara	3,57	0,1616
(i:)	így	0,48	0,0370	(r:)	arra	0,01	0,0013
(ø)	öröm	0,84	0,0590	(n)	nap	6,88	0,2660
(ø:)	őrült	0,94	0,0633	(n:)	annál	0,12	0,0116
(y)	üres	0,39	0,0312	(p)	anya	0,43	0,0338
(y:)	űz	0,18	0,0264	(p:)	fonnyad	0,08	0,0082
(a)	akar	10,09	0,3340	(m)	malom	4,9	0,2132
(a:)	ádám	2,98	0,1510	(m:)	zümög	0,05	0,0055
(o)	ott	4,82	0,2109	(ts)	cápa	0,18	0,0164
(o:)	óra	0,95	0,0228		játszik		
(u)	ugat	0,96	0,0644	(ts:)	viccel	0,01	0,0013
(u:)	úr	0,26	0,0223	(dz)	brindza	0,002	0,0007
(p)	pipa	0,7	0,0501	(dz:)	edzz!	?	?
(p:)	nappal	0,05	0,0055	(dʒ)	lándzsa	?	?
(b)	baka	1,33	0,0829	(dʒ:)	briddzsel	?	?
(b:)	abba	0,3	0,0251	(tʃ)	csap	0,44	0,0344
(t)	túró	6,5	0,2563	(tʃ:)	fröccsen	?	?
(t:)	attól	0,64	0,0466	(c)	atya	0,01	0,0013
(d)	meder	2,37	0,1280	(c:)	hattyú	?	?
(d:)	addig	0,02	0,0024	(J)	agy	2,25	0,1232
(k)	eke	4,24	0,1933	(J:)	buggyan	0,001	0,0001
(k:)	akkor	0,07	0,0073	(v)	java	1,92	0,1095
(g)	egér	2,18	0,1203	(v:)	evvel	?	?
(g:)	reggel	0,03	0,0035	(s)	szó	1,56	0,0936
(f)	fa	0,77	0,0541	(s:)	lasszó	0,09	0,0091
(f:)	puffan	0,01	0,0013	(z)	záp	3,19	0,1586
(j)	java	1,68	0,0990	(z:)	bízz!	0,03	0,0035
	folyó			(ʃ)	sava	3,12	0,1561
(j:)	bajjal	0,001	0,0003	(ʃ:)	kossal	0,04	0,0045
(h)	hó	2,54	0,1346	(ʒ)	zsák	0,02	0,0024
(h:)	ahhoz	0,002	0,0005	(ʒ:)	rúzzsal	?	?

# Hangképző szerveink / 1



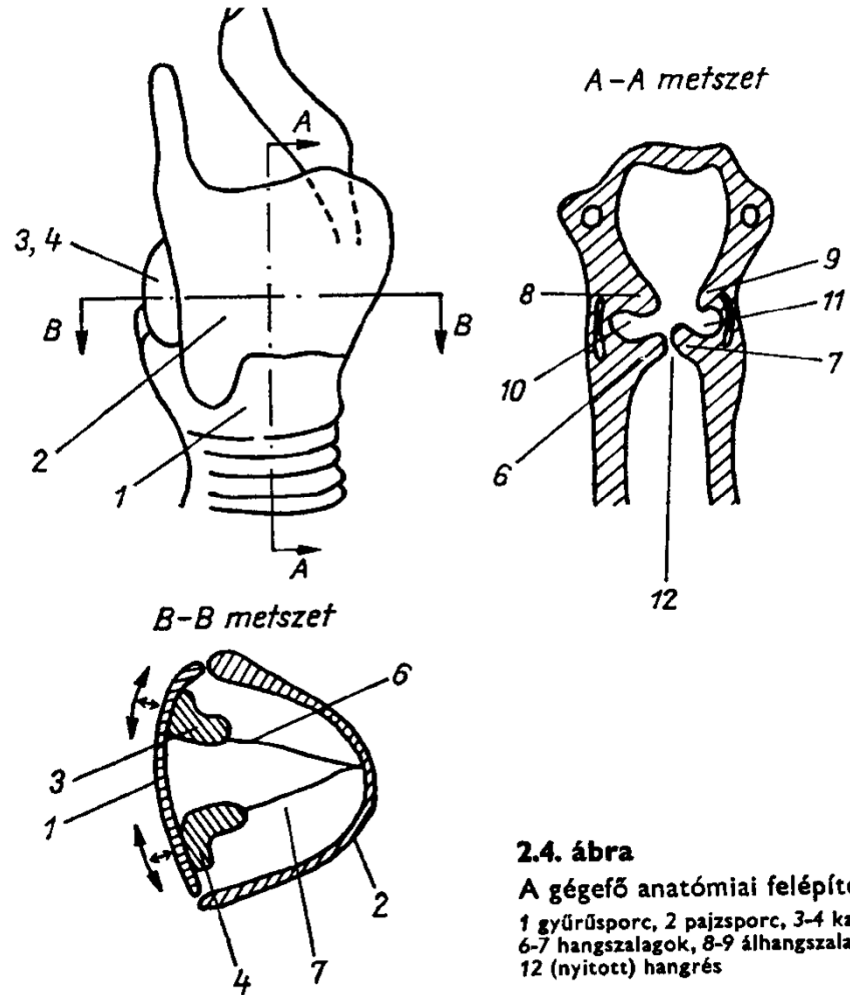
**2.3. ábra**

A hangképzőszervek vázlatos képe

A képzési helyek elnevezései: 1 labiális, 2 d  
4 prepalatális, 5 palatális, 6 postpalatális, 7  
8 faringális, 10 alveoláris

# Hangképző szerveink / 2

- A gégefő és a hangszalagok



**2.4. ábra**

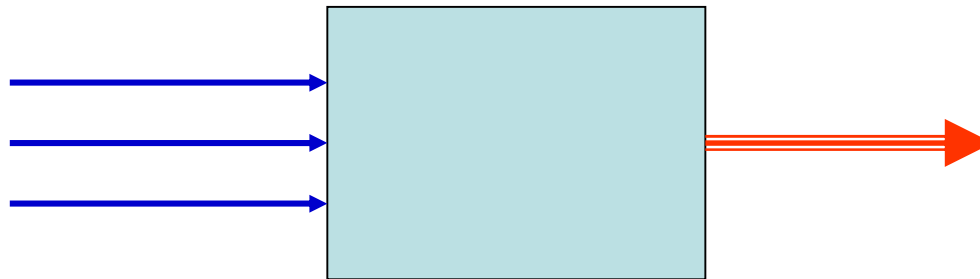
A gégefő anatómiai felépítése

1 gyűrűsporc, 2 pajzsporc, 3-4 kannaporcok, 5 gégefedő porc,  
6-7 hangszalagok, 8-9 álhangszalagok, 10-11 Morgani-féle üregek,  
12 (nyitott) hangrés



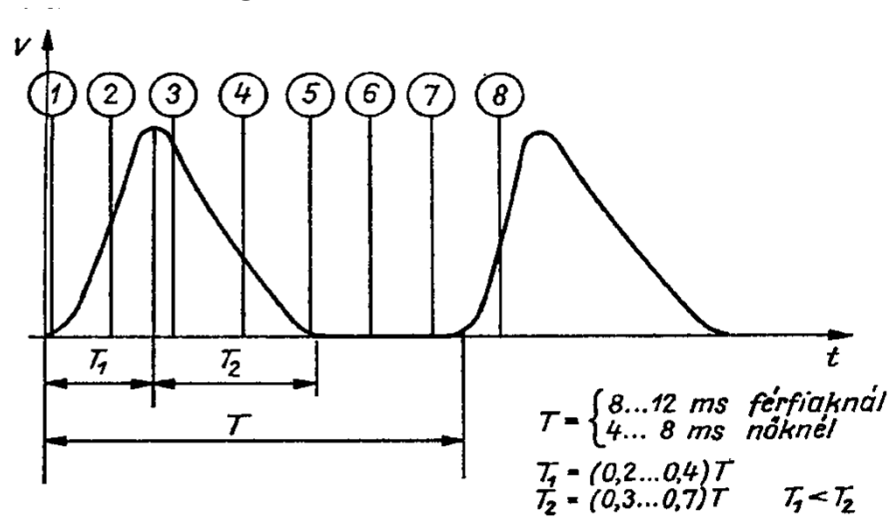
# A hangképzés módja

- Átviteli rendszer
  - rezonátor rendszer (toldalékcső, vokális traktus – vocal tract, artikulációs csatorna)
- Gerjesztések
  - zöngé
  - turbulencia
  - lökéshullám



# Gerjesztések

zöngé



turbulencia



lökéshullám

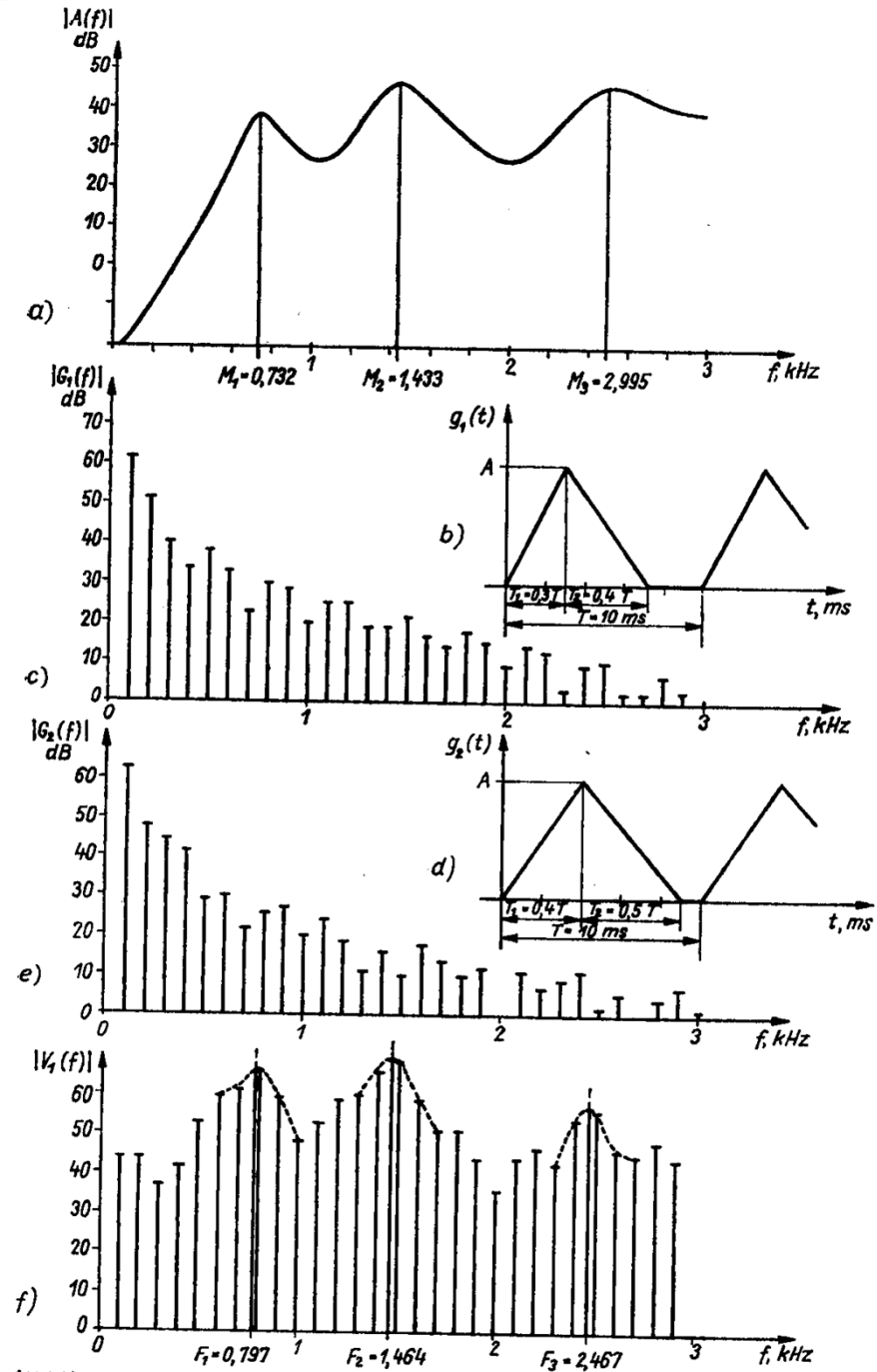




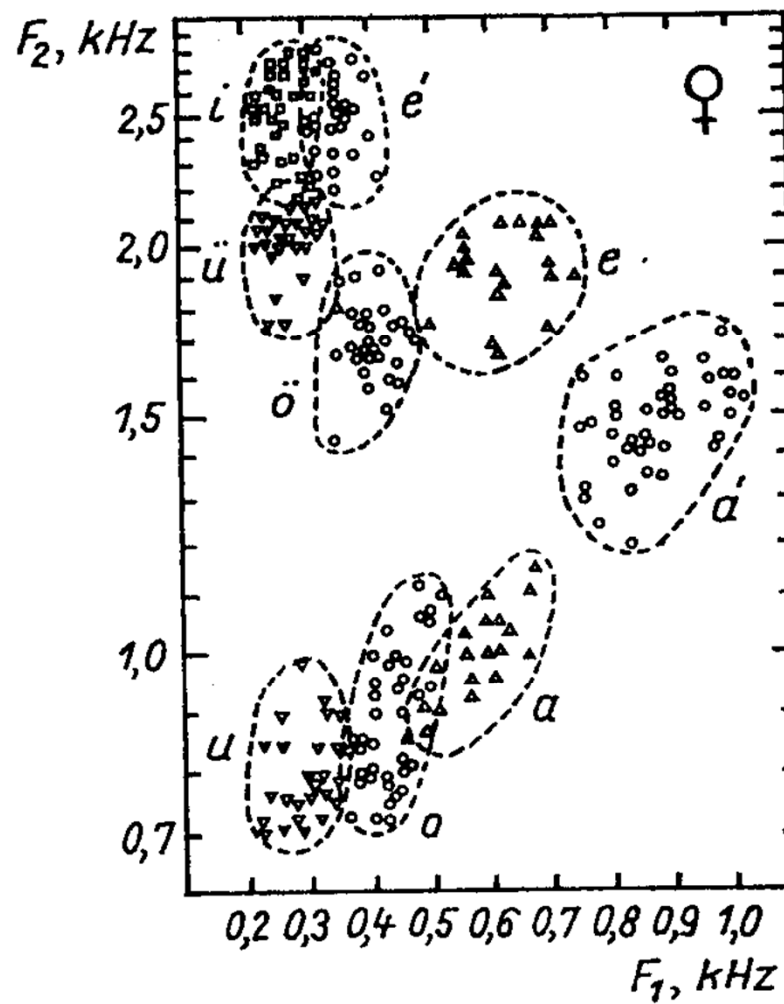
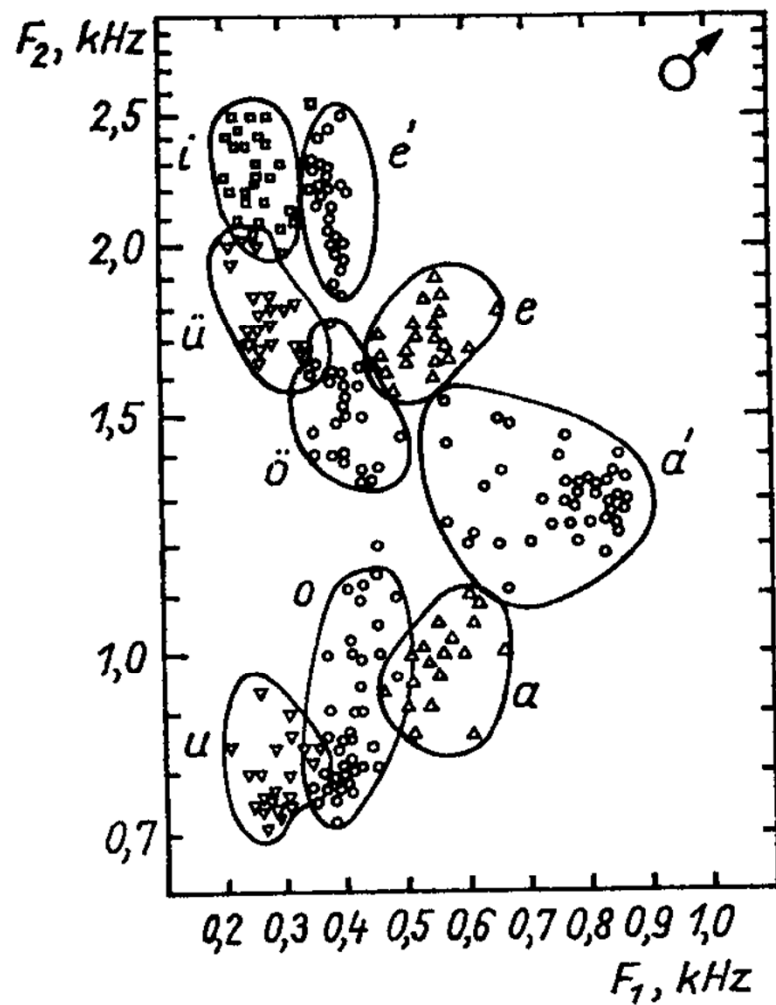
# A beszédhangok osztályozása

- Magánhangzók i, é, ü, ö, e á, a, o, u
- Mássalhangzók
  - zárhangok
    - nazálisok m, n, ny, ng
    - felpattanó zárhangok (explozívák)
      - zöngés b, d, g, gy
      - zöngétlen p, t, k, ty
    - pergőhang r
  - oldalrészhangok l, j
  - réshangok (frikatívák)
    - zöngés v, z, zs
    - zöngétlen f, s, sz, h
  - zár-réshangok (affrikáták)
    - zöngés dz, dzs
    - zöngétlen c, cs

# Magánhangzók jellemzői



# A magánhangzók formánsai



# Mássalhangzók jellemzői

- Réshangok
  - turbulens gerjesztés > vokális traktus rezonanciája
- Felpattanó zárhangok
  - nem tartható hosszan
  - nincs formánsstruktúra
  - lökeshullám > hangszalag holtidő > állandósult magánhangzó

# Korszerű beszédtechnológiák

- Gépi beszéd felismerés
  - izolált szavas felismerők
  - folyamatos beszéd felismerése
  - beszélő személy felismerése (biometrikus adatok)
  - .....
- Beszédtömörítés
- Gépi beszédkeltés
  - korlátozott szókészlet
  - korlátlan szókészlet
  - gépi felolvasás
  - .....

**Beszédérthetőség**

# A beszédérthetőséget befolyásoló tényezők

1. a beszélő artikulációs képessége
2. a hallgató hallóképessége
3. a tér, ill. terem tulajdonságai (teremakusztikai paraméterek)
4. a háttérzaj szintje

A 3. és 4. paraméter esetében ez konkrétan két arányt jelent:

3. a direkt-zengő tér arányt

és

4. a jel-háttérzaj arányt

# A beszédérthetőség mérése

- Szubjektív vizsgálatok
  - Szótagérthetőségi vizsgálat
    - értelmetlen szótagok megfigyeltetése számos kísérleti személy bevonásával
    - a helyesen megértett szótagok aránya az összes szótagokhoz képest (%-ban)
  - Mondatérthetőségi vizsgálat
- Objektív vizsgálatok
  - artikulációs index (AI): csak a zajt vette figyelembe
  - Speech Intelligibility Index (SSI): a visszhangosságot is figyelembe veszi
  - Speech Transmission Index (STI): objektív mérésekkel közelíti a szubjektív mennyiséget (nemzetközi szabvány)
    - a beszéd sokféle frekvenciájú komponenst tartalmaz, amelyek kismagasságú burkolóval amplitúdó moduláltak
    - az érthetőséghez ennek a modulációnak meg kell maradnia
    - a moduláció alakhű átvitelét a többutas terjedés (zengés) és a háttérzaj egyaránt rontja



## Zaj hatása: Artikulációs index (AI)

$$AI = \sum g \cdot \Delta L, \quad (2.4)$$

ahol  $g$  egy az oktávsvávtól függő súlyozó tényező és  $\Delta L$  a hallgató helyén mért beszéd-hangnyomásszint csúcsok és a beszédzavaró zaj oktáv-sávban mért hangnyomásszintjei között mért különbség a 250, 500, 1000, 2000 és 4000 Hz középfrekvenciájú oktávsvávokban. A  $g$  súlyzó értékek





$f_{\text{közép}}$	250	500	1000	2000	4000
$g$	0,0018	0,0050	0,0075	0,0107	0,0083

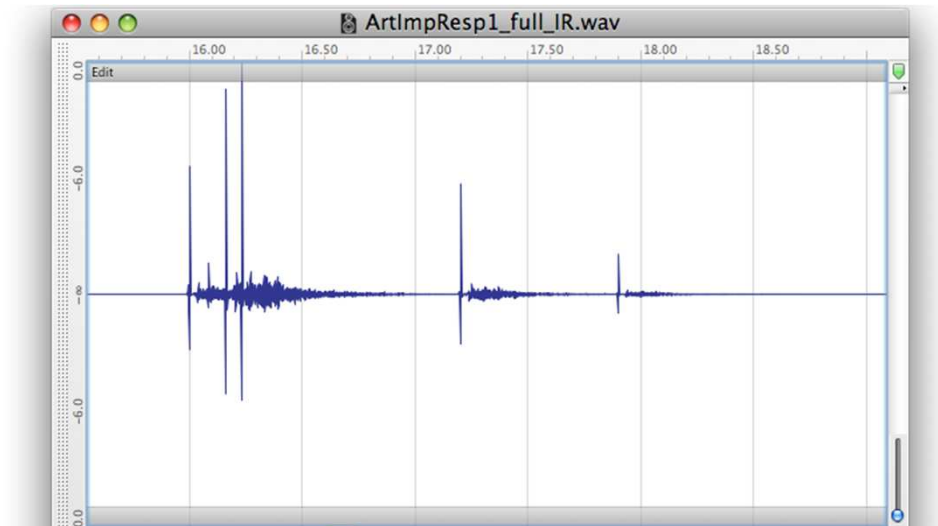
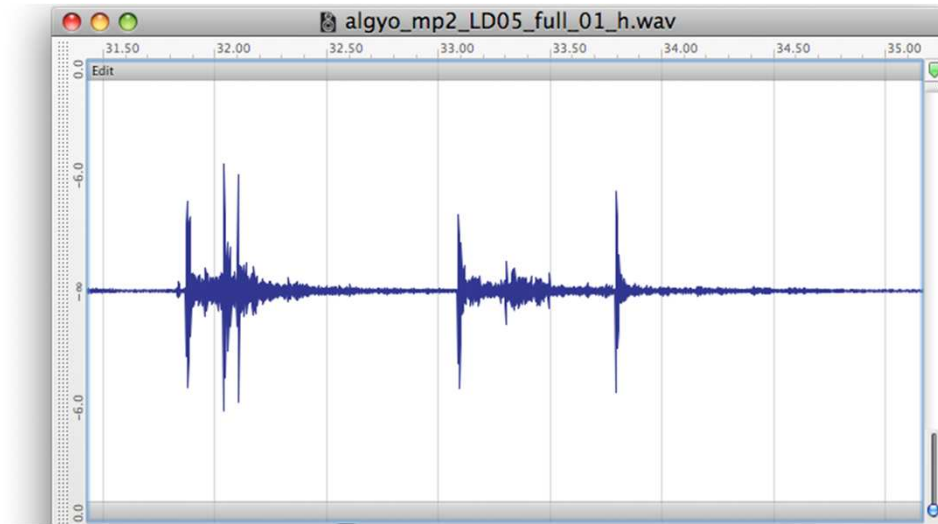
A beszéd érthetősége és az AI artikulációs index közötti összefüggés a következő:

AI	Beszédérthetőség
0,1	igen rossz
0,3	nem megfelelő
0,3 ÷ 0,5	megfelelő
0,5 ÷ 0,7	jó
> 0,7	igen jó

**Helyettesítő ökölszabály:** legyen a jel-zaj viszony a lehető legnagyobb, **min. 5-10 dBA**

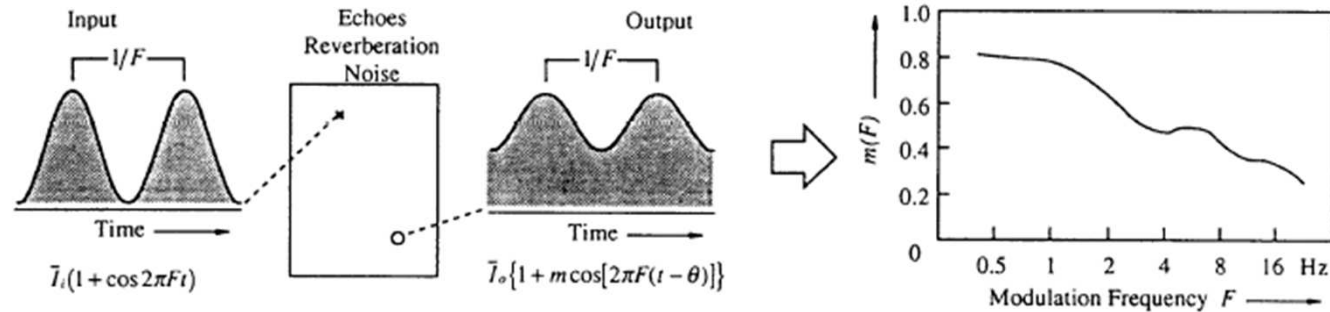
# Zengés hatása: visszhanggal torzított felvétel

- Eredeti hangminta 
- Helyszíni felvétel 
- Mért impulzusválasz 
- Zengetett hangminta 



# STI számítása

- modulációs transzfer függvényen keresztül:



- mérése régebben kisfrekvenciásan modulált sávkorlátozott zajjal, újabban impulzusválasszal

- ebből látszólagos jel-zaj értéket képezünk:

$$SNR_{kf} = 10 \log_{10} \frac{m_{kf}}{1 - m_{kf}}$$

- majd több lépés után:

$$STI = \sum_{k=1}^7 W_k MTI_k$$

$w_k$	0.13	0.14	0.11	0.12	0.19	0.17	0.14
-------	------	------	------	------	------	------	------

# STI jellemzői

- tehát függ
  - a jel-zaj viszonytól
  - a visszhangosságtól
  - a sávszélességtől
  - nemlineáris torzításoktól
  - stb.

Sávszélesség	CIS
350-3000 Hz	0,62
350-5700 Hz	0,77
180-3000 Hz (nagyjából ez felel meg a végpont elektroakusztikai elemei sávszélességének)	0,75
180-5700 Hz	0,86
85-11000 Hz	1

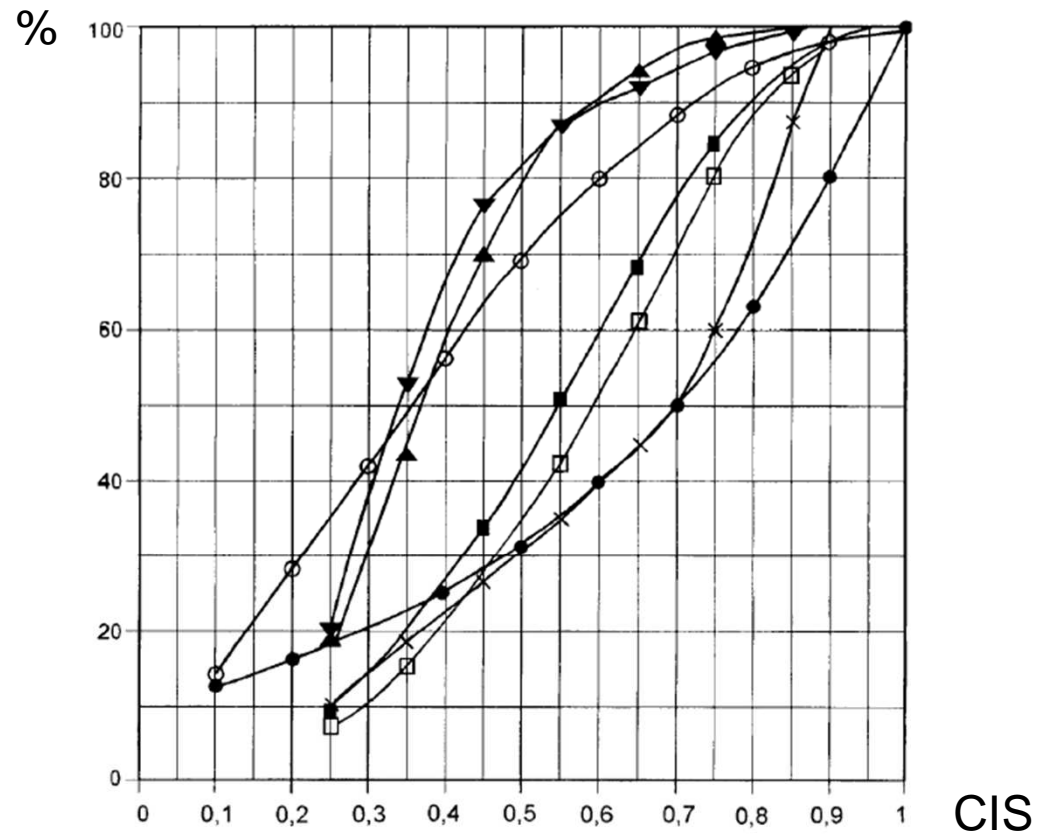
# STI értékelése

STI értéke	Beszédérthetőség minősítése
0 – 0.3	Elfogadhatatlan (érthetetlen)
0.3 – 0.45	Rossz
0.45 – 0.6	Elfogadható
0.6 – 0.75	Jó
0.75 – 1	Kiváló

A sokfajta beszédérthetőségi mérték összehasonlítható a CIS érték (**Combined Intelligibility Scale**) segítségével:

$$\text{CIS} = 1 + \lg (\text{STI}).$$

# Mérhető jellemzők és a CIS érték kapcsolata



- ▼ Fonetikailag kiegyenlített szófolyamat (256 szó)
- ▲ Rövid mondatok
- Másállhangzók artikulációs százalékaránya (100-(%AL<sub>cons</sub>))
- Fonetikailag kiegyenlített szófolyamat (1000 szó)
- 1000 szótag
- × Artikulációs mutató (AI)
- Beszédátviteli mutató (STI x 100)