

**A MAGYAR BESZÉD**  
**beszédakusztika, beszédtechnológia, beszédinformációs rendszerek**  
**Szrkesztette: Németh Géza, Olasz Gábor**  
**Áttekintő szerkesztő: Gordos Géza**  
 Akadémiai Kiadó  
**KIKNEK SZÓLNA A KÖNYV?**

A könyv a Budapesti Műszaki és Gazdaságtudományi Egyetem Távközlési és Médiainformatikai Tanszék Beszédtechnológiai és akusztikai laboratóriumainak kutatásai és fejlesztései köré épül, egyúttal kitekint a nemzetközi beszédkutatásra is (170 számozott fejezetet tartalmaz). Felöleli a beszéd fiziológiai, fonetikai, akusztikai és jelfeldolgozás témaköreit és a magyar beszéd jellemzőit mutatja be. A beszédtechnológiával foglalkozó rész bemutatja a beszéd számítógépes feldolgozásának minden területét (gépi felismerés és szintézis, tömörítés, beszédmódosítás, általános jelfeldolgozás. A beszédtechnológia gyakorlati alkalmazásaiból mindent bemutat, amit eddig elértek a kutatók, fejlesztők a távközlési alkalmazásoktól az egészségügyi fejlesztéseken keresztül a vakokat kiszolgáló rendszerekig.

A Magyar beszédet ilyen aspektusból tárgyaló könyv még nem jelent meg, tehát hiányt pótol.

Igénylik az egyetemek, főiskolák, valamint minden olyan oktatási hely, ahol informatikusokat képeznek. Jó támogatást adhat távközlési fejlesztőknek és döntéshozóknak, a beszédtechnológiai fejlesztések szakembereinek, új, multimédiás távközlési szolgáltatások tervezőinek.

Azonban ennél szélesebb rétegnek szeretnénk ajánlani. Segítheti a humán területek oktatását is (fonetika, beszédelemzés, nyelvészet és a beszéd kapcsolata, beszédpszichológia, egészségügyi betegségmegelőzés, tájékoztatás). Ajánljuk továbbá a középiskolások felső tagozatának is, valamint mindenkinek, akit érdekel a témakör (pl. fizikusok, nyelvészek, rádiósok, televíziósok, filmesek, tudományos media szakemberei). Nem folyamatos olvasmánynak szánjuk, hanem lexikon szerű kézikönyvnek. Átfogó tartalma miatt hasznos információkat találhatnak a fenti szakmák művelői, a mérnököktől a bölcsészekig.

A könyv másik célja megmutatni és rögzíteni a magyar beszéd akusztikai szerkezeti képét a 21. század elején, ismertetni a beszédtechnológia, mint új, interdiszciplináris tudományág eddig elért eredményeit, problémaköreit és alkalmazásait, főleg hazai vonatkozásban.

Szakszójegyzék segíti az olvasót a könyv végén. A könyvhöz rendelt honlap pedig sok olyan adatot tartalmaz, amit a könyvben a terjedelem korlátozottsága miatt nem lehet elhelyezni.

A 21. század új iparága a beszédtechnológia. Ez a könyv az első szisztematikus összefoglalás magyar nyelven a magyar beszédre vonatkozóan ebben az aspektusban.

További felvilágosítás: Olasz Gábor ([olaszy@tmit.bme.hu](mailto:olaszy@tmit.bme.hu)), 463-3883.

## Tartalomjegyzék

<b>Előszó</b> .....	xv
<b>Szerkesztők – szerzők</b> .....	xix
<b>Rövidítések jegyzéke</b> .....	xxv
<b>EMBER, NYELV, BESZÉD</b>	
<b>1. A beszéd és az információs társadalom</b> .....	3
<b>2. A beszéd komplex szerkezete</b> .....	9
2.1. A beszéd körfolyamata, az emberi dialógus .....	10
2.2. A nyelvi tudás szintjei .....	12
2.3. Beszédformák .....	13
2.3.1. Spontán beszéd .....	14
2.3.2. Felolvasásos beszéd .....	17
<b>3. Fiziológiai, fizikai alapok</b> .....	19
3.1. A beszédképzés folyamata .....	19
3.1.1. Gégeszintű hangképzés .....	20
3.1.2. Az artikulációs csatorna .....	24
3.1.2.1. A koartikuláció .....	25
3.1.2.2. A gégeműködés és az artikulációs csatorna viszonya .....	26
3.2. A hallási folyamat .....	27
3.2.1. A fül szerkezete .....	28
3.2.2. Jelfeldolgozás a hallórendszerben .....	31
3.3. A beszéd fizikai jellemzése .....	38
3.3.1. A rezgőmozgás, a hang keletkezése .....	39
3.3.2. A hang terjedése a levegőben .....	40

3.3.3.	Kényszerrezgés, rezonancia .....	44
3.3.4.	Összetett rezgések .....	45
3.3.5.	A beszédjel elemzése .....	48
3.3.5.1.	Formáns, zörejjóc .....	50
3.4.	Pszichofizikai tényezők .....	56
3.4.1.	Hangosságérzékelés .....	58
3.4.1.1.	Tisztahangok hangosságérzékelése .....	59
3.4.1.2.	Összetett hangok hangosságérzékelése .....	61
3.4.1.3.	Hangosság és időtartam .....	64
3.4.2.	Hangmagasság-érzékelés .....	65
3.4.2.1.	Hangmagasságskálák .....	65
3.4.2.2.	Tisztahangok hangmagasságérzete .....	67
3.4.2.3.	Komplex hangok hangmagassága, virtuális hangmagasság ..	69
3.5.	Fizikai-nyelvi megfeleltetések .....	70
<b>4.</b>	<b>A beszéd és az írás kapcsolata .....</b>	<b>73</b>
4.1.	Írásrendszerek .....	73
4.2.	Hangjelölések .....	77
4.3.	Tagolási különbözőségek .....	79
4.4.	Az írott szöveg és a hangalak kapcsolata .....	81
4.5.	Hang- és szóhatárok kijelölése a beszéd hullámformáján .....	83
4.6.	Magyar hang-, betű- és szóstatistika .....	86
4.6.1.	Betűstatistika a hangalak figyelembevételével .....	86
4.6.2.	A magyar szavak eloszlásai .....	91
<b>A BESZÉD SZERKEZETI ELEMZÉSE</b>		
<b>5.</b>	<b>A beszéd szegmentális szerkezete .....</b>	<b>95</b>
5.1.	A magyar beszédhangok .....	99
5.1.1.	A beszédhangok osztályozása .....	99
5.1.1.1.	A beszédhangok specifikus időtartamai .....	101
5.1.1.2.	A beszédhangok specifikus intenzitásai, hangzósság .....	104
5.1.2.	A magyar magánhangzók .....	106
5.1.2.1.	A magyar magánhangzók időtartamadatai .....	113
5.1.3.	A magyar mássalhangzók .....	115
5.1.3.1.	Zöngés zárhangok .....	117
5.1.3.2.	Zöngétlen zárhangok .....	118
5.1.3.3.	Zöngés réshangok .....	120
5.1.3.4.	Zöngétlen réshangok .....	121
5.1.3.5.	Zöngés zár-rés hangok .....	123
5.1.3.6.	Zöngétlen zár-rés hangok .....	124

5.1.3.7. Közelítő hangok .....	124
5.1.3.8. Pergőhang .....	126
5.1.3.9. Nazális hangok .....	127
5.2. A hangkapcsolódások típusai és szerkezeti sajátosságai .....	129
5.2.1. Magánhangzó-magánhangzó kapcsolódások .....	132
5.2.1.1. A hiátustöltés jelensége .....	135
5.2.2. Mássalhangzó-magánhangzó-mássalhangzó kapcsolódások ....	138
5.2.3. Mássalhangzó-mássalhangzó kapcsolódások .....	141
5.2.3.1. Kettős mássalhangzó-kapcsolódások .....	142
5.2.3.2. Három elemű mássalhangzó-kapcsolatok .....	161
5.2.3.3. Négyelemű mássalhangzó-kapcsolatok .....	162
5.2.3.4. A koartikulációs néma fázis jelensége .....	163
5.3. Szegmentális jelenségek a gége szintjén .....	165
5.3.1. Mikrointonáció .....	165
5.3.2. Irreguláris zöngéképzés, glottalizáció, rekedtség .....	166
5.3.3. Suttogás .....	169
<b>6. A beszéd szuprasegmentális szerkezete .....</b>	<b>171</b>
6.1. A beszéddallam .....	173
6.1.1. A mondatdallamok kapcsolódási rendszere .....	176
6.1.2. A kijelentés dallamszerkezetei .....	178
6.1.3. A kérdésformák dallamai .....	180
6.1.3.1. A kiegészítendő kérdés .....	180
6.1.3.2. Eldöntendő kérdések .....	184
6.1.3.3. Ellenőrző kérdés .....	188
6.1.3.4. Választó kérdések .....	189
6.1.3.5. Befejezetlen kérdések .....	189
6.1.4. Más modalitások dallamformái .....	190
6.1.4.1. A felszólítás dallama .....	190
6.1.4.2. A figyelmeztetés dallama .....	191
6.1.4.3. Az óhajtás dallamformája .....	191
6.2. A hangsúlyozás .....	192
6.3. Hangintenzitás mondatkeretben .....	197
6.4. Időszerkezeti tényezők .....	199
6.4.1. Artikulációs sebesség .....	200
6.4.2. Beszédtempó .....	201
6.4.3. Szünetek .....	201
6.4.4. Ritmus .....	203
6.5. A hangszínezet .....	205

## BESZÉDTECHNOLÓGIA

<b>7. A beszédtechnológia tudománya</b>	209
7.1. A beszéd számítógépes feldolgozása	209
7.1.1. Mintavételezés, kvantálás, visszaállítás	210
7.1.1.1. Mintavételezés	211
7.1.1.2. Kvantálás	215
7.1.2. Spektrális tulajdonságok meghatározása	220
7.1.2.1. Fourier-sor	221
7.1.2.2. Fourier-transzformáció	222
7.1.2.3. Teljesítménysűrűség-függvény	224
7.1.2.4. Ablakoló függvények	226
7.1.2.5. Idő- és frekvenciabeli felbontás	229
7.1.3. Zöngés-zöngétlen detekció	230
7.1.4. Jelfeldolgozás prozódiai módosításokhoz	233
7.1.4.1. Fonetikai alapú prozódiamódosítás	238
7.1.5. Kepsztrum	239
7.1.6. MFCC-paraméterek	240
7.1.7. Rejtett Markov-modellek	242
7.2. A beszéd tömörítése és átvitele	244
7.2.1. Kódolási alapelvek	245
7.2.1.1. A hullámforma kódolás	245
7.2.1.2. Parametrikus kódolás	245
7.2.1.3. Hibrid kódolás	246
7.2.2. Adaptív differenciális, predikciós kódoló	247
7.2.3. Nyílt hurkú predikciós kódoló	250
7.2.4. Zárt hurkú predikciós kódoló	255
<b>8. Adatbázisok a beszédtechnológia szolgálatában</b>	261
8.1. Tanító adatbázisok gépi beszédfelismeréshez	268
8.1.1. Tanító adatbázisok a nyelvi tartalom gépi felismeréséhez	271
8.1.1.1. Beszédatadatbázisok az akusztikai-fonetikai modell betanításához	272
8.1.1.2. Szövegadatbázisok a nyelvi modell betanításához	280
8.2. Beszédből készített elemzési adatbázisok beszéd-szintézishez	283
8.2.1. Hangelemzési adatbázisok felolvasásához	284
8.2.1.1. Jó minőségű számfelolvasó hangelemzési adatbázisának tervezése	286
8.2.2. Logatom alapú, diád, triád hangelemzési adatbázisok felolvasásához	292
8.2.2.1. Diád hangelemzési adatbázis	292
8.2.2.2. Triád a hangelemzési adatbázisban	299
8.2.3. Nagyméretű beszédatadatbázisok szövegfelolvasókhoz	300
8.2.3.1. A szintézis fő építőelemei	302
8.2.3.2. A beszédatadatbázis címkézése	305

8.3.	Kiejtésikivétel-szótárak .....	310
8.4.	Oktatási, kutatási célú internetes adatbázisok .....	315
8.4.1.	A magyar hangkapcsolódások akusztikai bemutatása szavakban	315
8.4.2.	Mondatfajták beszédadatbázis .....	319
8.4.3.	Elektronikus kiejtési szótár IPA-jelekkel és hangidőtartamokkal	321
8.4.4.	A magyar formánsadatbázis .....	325
8.5.	Spontánbeszéd-adatbázisok .....	331
<b>9.</b>	<b>A beszéd gépi észlelése és felismerése .....</b>	<b>333</b>
9.1.	Gépi beszédészlelési feladatok .....	335
9.1.1.	A gépi beszéd felismerők osztályozása .....	336
9.2.	A beszéd gépi felismerésének alapjai .....	338
9.3.	Lényegkiemelési eljárások .....	340
9.3.1.	Normálás .....	342
9.3.2.	A tulajdonságvektorok előállítás .....	343
9.4.	Mintaillesztési eljárások .....	344
9.4.1.	Sablonbázisú mintaillesztés .....	344
9.4.1.1.	A dinamikus idővetemítés .....	346
9.4.2.	Statisztikai mintaillesztési módszerek .....	348
9.5.	A beszéd-szöveg átalakítás alapjai .....	351
9.5.1.	A beszéd felismerési feladat matematikai megfogalmazása .....	351
9.5.2.	Beszéd felismerés rejtett Markov-moddal .....	352
9.5.3.	Beszédhangalapú folyamatos beszéd felismerés .....	358
9.5.3.1.	Kontextusérzékeny beszédhangmodellek .....	360
9.5.3.2.	Mintaillesztési példa a Viterbi-algoritmus használatára .....	362
9.6.	A beszéd-szöveg átalakítás alapvető tudásforrásai .....	363
9.6.1.	Az akusztikai modellek betanítása .....	364
9.6.2.	A nyelvi modell készítése .....	368
9.6.2.1.	Statisztikai N-gram modellek .....	369
9.6.2.2.	Környezetfüggetlen nyelvtanok .....	371
9.7.	Zajtűrő beszéd felismerés .....	374
9.7.1.	Az átviteli csatorna hatását kompenzáló normalizációs eljárások	376
9.7.2.	Zajszűrő eljárások .....	377
9.7.3.	A beszélő személytől származó zajok kezelése .....	379
9.7.4.	Beszéd-nem beszéd detektálás .....	380
9.8.	Beszélő adaptáció .....	381
9.8.1.	Az artikulációs csatorna normalizálása .....	382
9.8.2.	Akusztikai adaptáció .....	382
9.8.2.1.	Akusztikai adaptáció lineáris transzformációval .....	383
9.8.2.2.	Maximum a posteriori adaptáció .....	384
9.8.3.	Nyelvi adaptáció .....	385

9.9.	A prozódia szerepe a beszédfelismerésben .....	386
9.10.	Beszélőfelismerés .....	388
9.11.	Érzelemfelismerés .....	391
9.12.	Beszédfelismerés támogatása multimodális paraméterekkel .....	401
9.12.1.	A vizuális lényegkiemelés .....	401
9.12.2.	A vizuális és akusztikai modalitás integrálása .....	405
9.13.	Beszédfelismerők minősítése .....	406
<b>10.</b>	<b>A beszéd gépi előállítása .....</b>	<b>409</b>
10.1.	Kempelentől napjainkig .....	413
10.2.	Kötött szótáras beszéd szintetizátorok .....	419
10.2.1.	Hangminőségi skála .....	421
10.2.2.	Tervezési tanácsok a jó hangminőség elérésére .....	422
10.2.3.	A kötött szótáras rendszerek tervezési folyamata .....	423
10.2.4.	Fonetikai elvű modell szám-, dátum-, időpont-, pénzüsszeg- felolvasáshoz .....	425
10.3.	Automatikus szövegfelolvasás .....	427
10.3.1.	A beszéd modellezése szintézishez .....	428
10.3.1.1.	Hangsúlymeghatározás a szöveg alapján .....	428
10.3.1.2.	Az alapfrekvencia-változásának szabály alapú modellezése .	440
10.3.1.3.	A beszéd dallam változatosságának statisztikai modellezése .	443
10.3.1.4.	A beszéd időszerkezetének szabály alapú modellezése .....	447
10.3.1.5.	Komplex prozódiai modell .....	453
10.3.1.6.	Beszélő fej modellezése .....	456
10.3.1.7.	Érzelmi töltetű beszéd modellezése .....	464
10.3.2.	Az ortografikus magyar szöveg fonetikai átírásának gépi mód- szere .....	465
10.3.2.1.	A fonetikai átírás során kezelendő nyelvi jelenségek .....	465
10.3.2.2.	Eljárások a fonetikai átírás megállapítására .....	470
10.3.2.3.	Fonetikai átíró magyar nyelvre .....	478
10.3.3.	Ékezetek gépi helyreállítása .....	483
10.3.3.1.	Ékezetesítő eljárások. ....	484
10.3.4.	A gépi szövegfelolvasók általános, elvi felépítése .....	486
10.3.5.	Formánsszintézis .....	489
10.3.5.1.	A MultiVox formánsszintetizátor szövegfelolvasáshoz .....	492
10.3.6.	Diád, triád hullámformák összefűzésén alapuló technológia ...	495
10.3.6.1.	A ProfiVox több hangú, szövegfelolvasó és fejlesztői rendszere	497
10.3.7.	Elemkiválasztás-alapú szövegfelolvasó .....	503
10.3.8.	A rejtett Markov-modell alapuló gépi szövegfelolvasás .....	510
10.3.9.	Érzelmes szövegfelolvasás .....	516
10.4.	Beszéd szintetizátorok minősítése, szabványosítási javaslatok .....	517

## BESZÉDTECHNOLÓGIAI ALKALMAZÁSOK

<b>11. Beszédinformációs rendszerek</b> .....	523
11.1. A beszédinformációs rendszerek fő építőelemei .....	523
11.2. Emberi–gépi dialógus .....	525
11.3. A dialógus tervezése .....	526
11.4. Az akusztikai arculat .....	530
11.4.1. Az akusztikai arculat áttekintése .....	531
11.4.2. Infokommunikációs szolgáltatások és az akusztikai arculat ....	533
11.4.2.1. Az akusztikai arculat összetevői infokommunikációs szol- gáltatásokban .....	533
11.4.3. Az akusztikai arculatot meghatározó néhány szolgáltatás vizs- gálata .....	535
<b>12. Példák a beszédtechnológia felhasználásának területeiről</b> .....	539
12.1. Beszédtömörítési megoldások a gyakorlatban .....	539
12.1.1. Kódoló ajánlások .....	541
12.1.2. A kódolók fejlődése .....	543
12.2. Gépi beszédminősítés távközlési rendszerekben .....	545
12.2.1. Hanganyag gyűjtése .....	545
12.2.2. Szubjektív beszédminősítés .....	545
12.2.3. Objektív beszédminősítő eljárások áttekintése .....	548
12.2.3.1. Az objektív minősítő eljárás lépései .....	549
12.2.3.2. Az objektív minősítő eljárások értékelése .....	550
12.3. Telefonos és mobilos alkalmazások .....	553
12.3.1. Telefonról elérhető e-levél felolvasó .....	553
12.3.2. SMS-felolvasó vezetékes telefonra .....	555
12.3.3. Mobiltelefonba épített SMS-felolvasó .....	558
12.3.4. Automatikus szám szerinti tudakozó .....	559
12.3.5. Gyógyszervonal, automatikus telefonos információs rendszer ..	561
12.3.6. Automatikus, mobiltelefonos, helyfüggő kereső szolgáltatás ...	564
12.3.7. Automatikus áru- és árlista-felolvasó .....	567
12.3.8. Beszéddel vezérelt automatikus telefonközpontok .....	571
12.4. Internetes alkalmazások .....	572
12.4.1. Időjárás-előrejelzés írott szöveges és hangos modalitással ....	573
12.4.2. Híradókereső – internetes hang-videókeresés kulcsszavak alap- ján .....	574
12.4.3. Szövegfelolvasás a webfordítás színesítésére .....	576
12.5. Közlekedési alkalmazások .....	576
12.5.1. Vasútállomási utastájékoztató .....	577
12.6. Diktálórendszerek .....	577



12.6.1. Leletező beszédfelismerő . . . . .	578
12.7. Beszédtechnológia a vakok és gyengénlátók szolgálatában . . . . .	580
12.7.1. Képernyőolvasás . . . . .	581
12.7.2. Dramatizáló . . . . .	584
12.7.3. Hangoskönyvek . . . . .	585
12.7.4. Beszélő bankautomaták . . . . .	588
12.7.5. Navi-Speech beszélő navigátor látássérült gyalogosoknak . . . . .	588
12.8. Hallássérültek segítése beszédtechnológiával . . . . .	593
12.9. Beszédtanítás és beszédtechnológia . . . . .	602
12.9.1. Beszédoktató varázsdoboz . . . . .	605
12.9.1.1. Adatbázisok és modellezés . . . . .	608
12.9.1.2. Beszédoktatói módszertan a használathoz . . . . .	616
12.10. Beszédkommunikátor beszédsérültek segítésére . . . . .	619
12.11. Hallásmérés szintetikus beszéddel . . . . .	622
12.11.1. A Mondom-2000 beszédhallást ellenőrző eljárás . . . . .	623
<b>13. Interfészek, szabványok, honlapok, programok . . . . .</b>	<b>629</b>
13.1. VXML . . . . .	629
13.1.1. VoiceXML alkalmazásfejlesztés . . . . .	630
13.1.2. VoiceXML alapú alkalmazások . . . . .	632
13.2. Programozói interfész beszédtechnológiai alkalmazásokhoz (SAPI) . . . . .	633
13.2.1. Microsoft Speech API . . . . .	636
13.2.2. Java Speech API . . . . .	638
13.3. MRCP . . . . .	639
13.4. Intelligens beszédhang-időtartam mérő . . . . .	641
13.5. Glottalizáló program . . . . .	645
13.6. A könyvben szereplő honlapok beszédkutatáshoz, oktatáshoz, fejlesztésekhez, döntéshozatalhoz . . . . .	648
<b>14. A beszédtechnológia jövője . . . . .</b>	<b>651</b>
<b>FÜGGELEK</b>	
<b>A. Hangkapcsolatok . . . . .</b>	<b>655</b>
A.1. CC hangkapcsolatok . . . . .	655
A.2. CCC hangkapcsolatok . . . . .	658
A.3. CCCC hangkapcsolatok . . . . .	663
A.4. CVC hangkapcsolatok spektrogramjai . . . . .	664
<b>Hivatkozások . . . . .</b>	<b>673</b>

Tartalomjegyzék

---

xiii

**Tárgymutató**..... 691