

Tartalomjegyzék

Előszó	xv
Szerkesztők – szerzők	xix
Rövidítések jegyzéke	xxv

EMBER, NYELV, BESZÉD

1. A beszéd és az információs társadalom	3
2. A beszéd komplex szerkezete	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
3. Fiziológiai, fizikai alapok	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	38
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, zörejgóc	50
3.4.	Pszichofizikai tényezők	56
3.4.1.	Hangosságérzékelés	58
3.4.1.1.	Tisztahangok hangosságérzékelése	58
3.4.1.2.	Összetett hangok hangosságérzékelése	60
3.4.1.3.	Hangosság és időtartam	64
3.4.2.	Hangmagasság-érzékelés	64
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	69
4.	A beszéd és az írás kapcsolata	73
4.1.	Írásrendszerek	73
4.2.	Hangjelölések	77
4.3.	Tagolási különbségek	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 szóstatisztika	86
4.6.1.	Betűstatisztika a hangalak figyelembevételével	86
4.6.2.	A magyar szavak eloszlásai	90
 A BESZÉD SZERKEZETI ELEMZÉSE		
5.	A beszéd szegmentális szerkezete	95
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	114
5.1.3.	A magyar mássalhangzók	116
5.1.3.1.	Zöngés zárhangok	118
5.1.3.2.	Zöngétlen zárhangok	119
5.1.3.3.	Zöngés réshangok	121
5.1.3.4.	Zöngétlen réshangok	122
5.1.3.5.	Zöngés zár-rés hangok	124
5.1.3.6.	Zöngétlen zár-rés hangok	125

5.1.3.7.	Közelítő hangok	126
5.1.3.8.	Pergőhang	126
5.1.3.9.	Nazális hangok	128
5.2.	A hangkapcsolódások típusai és szerkezeti sajátosságaik	129
5.2.1.	Magánhangzó-magánhangzó kapcsolódások	133
5.2.1.1.	A hiátustöltés jelensége	136
5.2.2.	Mássalhangzó-magánhangzó-mássalhangzó kapcsolódások ...	139
5.2.3.	Mássalhangzó-mássalhangzó kapcsolódások	142
5.2.3.1.	Kettős mássalhangzó-kapcsolódások	143
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.	Suttogás	166
5.3.3.	Irreguláris zöngképzés, glottalizáció, rekedtség	167
6.	A beszéd szuprasegmentális szerkezete	171
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	183
6.1.3.3.	Ellenőrző kérdés	188
6.1.3.4.	Választó kérdések	188
6.1.3.5.	Befejezetlen kérdések	189
6.1.4.	Más modalítások 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	196
6.4.	Időszerkezeti tényezők	199
6.4.1.	Artikulációs sebesség	199
6.4.2.	Beszédtempó	200
6.4.3.	Szünetek	201
6.4.4.	Ritmus	202
6.5.	A hangszínezet	204

BESZÉDTECHNOLÓGIA

7. A beszédtechnológia tudománya	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	232
7.1.4.1. Fonetikai alapú prozódiamódosítás	237
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
8. Adatbázisok a beszédtechnológia szolgálatában	261
8.1. Tanító adatbázisok gépi beszédfelismeréshez	268
8.1.1. Tanító adatbázisok a nyelvi tartalom gépi felismeréséhez	271
8.1.1.1. Beszédatadattárak az akusztikai-fonetikai modell betanítá- sához	272
8.1.1.2. Szövegadatbázisok a nyelvi modell betanításához	280
8.2. Beszédből készített elemzési adatbázisok beszéd-szintézishez	283
8.2.1. Hangelembázisok számok felolvasásához	284
8.2.1.1. Jó minőségű számfelolvasó hangelembázisának tervezése ..	286
8.2.2. Logatomalapú, diád-, triád-hangelembázis szövegfelolvasáshoz ..	292
8.2.2.1. Diád-hangelembázis	293
8.2.2.2. Triád a hangelembázisban	299
8.2.3. Nagyméretű beszédatadattárak szövegfelolvasókhoz	300

8.2.3.1.	A szintézis fő építőelemei	302
8.2.3.2.	A beszédatadbázis 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 beszédatadbázis	319
8.4.3.	Elektronikus kiejtési szótár IPA-jelekkel és hangidőtartamokkal	320
8.4.4.	A magyar formánsadatbázis	324
8.5.	Spontánbeszéd-adatbázisok	330
9.	A beszéd gépi észlelése és felismerése	333
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 előállítása	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	350
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.	Környezetfüggő beszédhangmodellek	360
9.5.3.2.	A kényszerített illesztés	361
9.5.3.3.	Mintaillesztési példa a Viterbi-algoritmus használatára	362
9.6.	A beszéd-szöveg átalakítás alapvető tudásforrásai	364
9.6.1.	Az akusztikai modellek betanítása	364
9.6.2.	A nyelvi modell készítése	369
9.6.2.1.	Statisztikai N-gram modellek	369
9.6.2.2.	Környezetfüggetlen nyelvtanok	372
9.7.	Zajtűrő beszéd felismerés	375
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	380
9.7.4.	Beszéd-nem beszéd detektálás	381
9.8.	Beszélőadaptáció	381
9.8.1.	Az artikulációs csatorna normalizálása	383
9.8.2.	Akusztikai adaptáció	384

9.8.2.1.	Akusztikai adaptáció lineáris regresszióval	384
9.8.2.2.	Maximum a posteriori adaptáció	385
9.8.3.	Nyelvi adaptáció	386
9.9.	Beszélőfelismerés	387
9.10.	A prozódia szerepe a beszéd felismerésben	390
9.11.	Érzelemfelismerés	392
9.12.	Beszéd felismerés támogatása multimodális paraméterekkel	401
9.12.1.	A vizuális lényegkiemelés	403
9.12.2.	A vizuális és akusztikai modalitás integrálása	406
9.13.	Beszéd felismerők minősítése	407
10.	A beszéd gépi előállítása	411
10.1.	Kempelentől napjainkig	415
10.2.	Kötött szótáras beszéd szintetizátorok	421
10.2.1.	Hangminőségi skála	423
10.2.2.	Tervezési tanácsok a jó hangminőség elérésére	424
10.2.3.	A kötött szótáras rendszerek tervezési folyamata	426
10.2.4.	Fonetikai elvű modell szám-, dátum-, időpont-, pénzüsszeg- felolvasáshoz	427
10.3.	Automatikus szövegfelolvasás	429
10.3.1.	A beszéd modellezése szintézishez	430
10.3.1.1.	Hangsúlymeghatározás a szöveg alapján	430
10.3.1.2.	Az alaphékvencia változásának szabályalapú modellezése	442
10.3.1.3.	A beszéd dallam változatosságának statisztikai modellezése	446
10.3.1.4.	A beszéd időszerkezetének szabályalapú modellezése	449
10.3.1.5.	Komplex prozódiai modell	455
10.3.1.6.	Beszélő fej modellezése	458
10.3.1.7.	Érzelmi töltetű beszéd modellezése	466
10.3.2.	Az ortografikus magyar szöveg fonetikai átírásának gépi mód- szere	467
10.3.2.1.	A fonetikai átírás során kezelendő nyelvi jelenségek	467
10.3.2.2.	Eljárások a fonetikai átírás megállapítására	472
10.3.2.3.	Fonetikai átíró magyar nyelvre	480
10.3.3.	Ékezetek gépi helyreállítása	485
10.3.3.1.	Ékezetesítő eljárások	486
10.3.4.	A gépi szövegfelolvasók általános, elvi felépítése	488
10.3.5.	Formánsszintézis	491
10.3.5.1.	A MultiVox formánsszintetizátor szövegfelolvasáshoz	494
10.3.6.	Diád-, triádhullámformák összefűzésén alapuló technológia	497
10.3.6.1.	A ProfiVox szövegfelolvasó és fejlesztői rendszere	499
10.3.7.	Elemkiválasztás-alapú szövegfelolvasó	505

10.3.8. A rejtett Markov-modellen alapuló gépi szövegfelolvasás	512
10.3.9. Érzelmes szövegfelolvasás	518
10.4. Beszédszintetizátorok minősítése, szabványosítási javaslatok	520

BESZÉDTECHNOLÓGIAI ALKALMAZÁSOK

11. Beszédinformációs rendszerek	525
11.1. A beszédinformációs rendszerek fő építőelemei	525
11.2. Emberi-gépi dialógus	527
11.3. A dialógus tervezése	528
11.4. Az akusztikai arculat	532
11.4.1. Az akusztikai arculat áttekintése	533
11.4.2. Infokommunikációs szolgáltatások és az akusztikai arculat . . .	535
11.4.2.1. Az akusztikai arculat összetevői infokommunikációs szol- gáltatásokban	535
11.4.3. Az akusztikai arculatot meghatározó néhány szolgáltatás vizs- gálata	537
12. Példák a beszédtechnológia felhasználásának területeiről	541
12.1. Beszédtömörítési megoldások a gyakorlatban	541
12.1.1. Kódoló ajánlások	543
12.1.2. A kódolók fejlődése	545
12.2. Gépi beszédminősítés távközlési rendszerekben	547
12.2.1. Hanganyag gyűjtése	547
12.2.2. Szubjektív beszédminősítés	548
12.2.3. Objektív beszédminősítő eljárások áttekintése	550
12.2.3.1. Az objektív minősítő eljárás lépései	551
12.2.3.2. Az objektív minősítő eljárások értékelése	552
12.3. Telefonos és mobilos alkalmazások	554
12.3.1. Telefonról elérhető e-levél felolvasó	555
12.3.2. SMS-felolvasó vezetékes telefonra	557
12.3.3. Mobiltelefonba épített SMS-felolvasó	560
12.3.4. Automatikus szám szerinti tudakozó	561
12.3.5. Gyógyszervonal, automatikus telefonos információs rendszer .	562
12.3.6. Automatikus, mobiltelefonos, helyfüggő kereső szolgáltatás . .	566
12.3.7. Automatikus áru- és árlista-felolvasó	569
12.3.8. Beszéddel vezérelt automatikus telefonközpontok	573
12.4. Internetes alkalmazások	574
12.4.1. Időjárás-előrejelzés írott szöveges és hangos modalitással	575
12.4.2. Híradókereső – internetes hang-videókeresés kulcsszavak alapján	576

12.4.3. Szövegfelolvasás a webfordítás színesítésére	578
12.5. Közlekedési alkalmazások	578
12.5.1. Vasútállomási utastájékoztató	579
12.6. Diktálórendszerek	579
12.6.1. Lelevező beszédfelismerő	580
12.7. Beszédtechnológia a vakok és gyengénlátók szolgálatában	582
12.7.1. Képernyőolvasás	583
12.7.2. Dramatizáló	586
12.7.3. Hangoskönyvek	588
12.7.4. Beszélő bankautomaták	590
12.7.5. NaviSpeech – beszélő navigátor látássérült gyalogosoknak ...	591
12.8. Beszédjel átalakítása mozgó száj képévé siketek kommunikációjá- nak segítésére	595
12.9. Beszédtanítás és beszédtechnológia	604
12.9.1. Beszédoktató varázsdoboz	607
12.9.1.1. Adatbázisok és modellezés	611
12.9.1.2. Képi megjelenítés	613
12.9.1.3. A kiejtés jószágának automatikus megítélése	617
12.9.1.4. Beszédoktatói módszertan a használathoz	618
12.10. Beszédkommunikátor beszédsérültek segítésére	620
12.11. Hallásmérés szintetikus beszéddel	624
12.11.1. A Mondom-2000 beszédhallást ellenőrző eljárás	624
13. Interfészek, szabványok, honlapok, programok	631
13.1. VXML	631
13.1.1. VoiceXML alkalmazásfejlesztés	632
13.1.2. VoiceXML alapú alkalmazások	634
13.2. Programozói interfész beszédtechnológiai alkalmazásokhoz (SAPI)	635
13.2.1. Microsoft Speech API	638
13.2.2. Java Speech API	641
13.3. MRCP	641
13.4. Intelligens beszédhang-időtartam mérő	643
13.5. Glottalizáló program	647
13.6. A könyvben szereplő honlapok beszédkutatáshoz, oktatáshoz, fej- lesztésekhez, döntéshozatalhoz	650
14. A beszédtechnológia jövője	653
Irodalomjegyzék	657

FÜGGELÉK

F. Hangkapcsolatok	691
F.1. CC hangkapcsolatok	691
F.2. CCC hangkapcsolatok	694
F.3. CCCC hangkapcsolatok	697
F.4. CVC hangkapcsolatok spektrogramjai	698
Tárgymutató	705