

MAGYAR **10** FONETIKAI FÜZETEK

HUNGARIAN PAPERS IN PHONETICS

A BESZÉD

AKUSZTIKAI ANALÍZISE

ÉS SZINTÉZISE

Kiadja az MTA
Nyelvtudományi Intézete
Budapest 1982



RZF 20/2003

MAGYAR FONETIKAI FÜZETEK
Hungarian Papers in Phonetics
10.

**A BESZÉD AKUSZTIKAI ANALÍZISE
ÉS SZINTÉZISE**

Szerkesztette:

BOLLA KÁLMÁN

Réger Zita
(1944 - 2001)
hagyatéka

A MAGYAR TUDOMÁNYOS AKADÉMIA NYELVTUDOMÁNYI INTÉZETE
BUDAPEST 1982

Technikai szerkesztő: GÓSY MÁRIA

Technikai munkatárs: NIKLÉCZY PÉTER
SZ. ZSIGÓCZKI IRÉN

HU ISSN 0134-1545
ISBN 963 8461 16 0

© MTA Nyelvtudományi Intézete, Budapest 1982 · Bolla Kálmán

Felelős kiadó: HAJDÚ PÉTER, az MTA Nyelvtudományi Intézetének igazgatója.

Készült 300 példányban, 12,8 (A/5) ív terjedelemben, térítésmentes terjesztésre.

8213180 MTA KESZ Sokszorosító, Budapest. F. v.: dr. Héczey Lászlóné.

THE ACOUSTIC ANALYSIS AND SYNTHESIS OF SPEECH-PRODUCTION

TARTALOM

CONTENTS

Előszó	5
Preface	5
BOLLA Kálmán: A magyar beszéd akusztikai szerkezetének analízise és szintézise. Kutatástörténeti áttekintés	7
The analysis and synthesis of the acoustic structure of Hungarian speech. A historical survey of the research	20
KISS Gábor—OLASZY Gábor: Interaktív beszéd-szintetizáló rendszer számítógéppel és OVE III szintetizátorral	21
An interactive speech synthesizing system with computer and OVE III synthesizer	45
OLASZY Gábor: A magyar mássalhangzók és a mássalhangzó—magánhangzó-kapcsolódások akusztikai szerkezetének analízise és szintézise	46
The analysis and synthesis of the acoustic structure of Hungarian consonants and consonant—vowel combination types	82
GÓSY Mária: A [b, d, g] mássalhangzók percepció vizsgálatára	84
The perceptual analysis of [b, d, g]	108
KISS Gábor—OLASZY Gábor: A magyar beszéd automatikus szintézisének első lépcsője	110
The first step in the automatic synthesis of Hungarian	116
BOLLA Kálmán: Folyamatos beszéd-szintetizáló rendszer magyar nyelven (VOXON)	118
VOXON: A system generating impersonal Hungarian speech by rule	128
FÖLDI Éva: A lengyel beszédintonáció elemzése szintézissel	129
The analysis of Polish intonation by synthesis	151
*	
A Magyar Fonetikai Füzetek első tíz kötetének bibliográfiája	153
The bibliography of the first ten issues of the Hungarian Papers in Phonetics	153
Új kiadványok	160
New publications	160

Balázs János: Magyar deákság. Anyanyelvünk és az európai nyelvi modell. Elvek és Utak. Budapest 1980. (<i>Vértes O. András</i>)	160
Fejezetek a magyar leíró hangtanból. Szerk. Bolla Kálmán. Budapest 1982. (<i>Gósy Mária</i>)	161

ELŐSZÓ PREFACE

Tanulmánykötetünkben a beszéd akusztikai szerkezetének az elemzésében és a beszédhangsorok mesterséges előállításában elért eredményeinkről számolunk be. Az MTA Nyelvtudományi Intézetének fonetikai osztályán már közel egy évtizede folyik olyan kutatómunka, amelyben a beszéd folyamat analizise és szintézise egymást kiegészítő és egységes metodológiát képez. A szintetikus beszéd számunkra nem végcél, hanem eszköz a hangtest akusztikai és fonetikai szerkezetének a jobb megismeréséhez, a hangszerkezeti elemek, az elemekből képződő fonetikai struktúrák nyelvi felhasználásának, funkciójának a feltárására. A szintézissel ellenőrizhetjük elemzéseink helyességét, pontosságát – ez már önmagában sem lekicsinylendő és elvetendő lehetőség –, de csak a szintézis révén tudunk a hangalak alkotórészeiben ellenőrizhető változtatásokat eszközölni és ezáltal a különböző hangszerkezeti összetevők szerepét „kitapogatni”, a lényeges és kevésbé lényeges, funkciós és funkcióatlan elemeket elkülöníteni egymástól az egyébként nagyon bonyolult felépítésű hangtestben. Akusztikai elemzésekkel és szintézissel foglalkozó írásokat korábbi füzetek is tartalmaznak. Ebben a füzetben a nyelvészeti-fonetikai eredmények mellett bemutatjuk a beszéd szintézis laboratóriumunkban használatos hardware-jét, technikai eszközeit, a szintézis általunk kidolgozott és alkalmazott technológiáját is.

A mesterséges beszéd korunkban, az automatizálás korában, nem marad a tudományos kíváncsiság keretei között, tehát nemcsak kutatólaboratóriumi termék, hanem a mikroelektronika rohamos fejlődése révén számos területen, termelési és társadalmi szférában nyer konkrét felhasználást az emberi munkaerő felszabadítása, az emberi tevékenység megkönnyítése és kiterjesztése érdekében is.

Korunkat okkal szokták a technika századának is nevezni. A gépek tömegestől törnek be mindennapi életünkbe. Az ember és a gép, a gép és az ember között új kapcsolatok alakulnak ki. Az EMBER elgépiesedésével szemben, e folyamat ellensúlyozására mindent meg kell tennünk a gépek „emberiesítése” (humanizálása) érdekében. E cél elérésében kapcsolatteremtő eszköz lehet a mesterséges beszéd, amellyel megvalósulhat a „párbeszéd” az ember és a gép között.

Bolla Kálmán

A MAGYAR BESZÉD AKUSZTIKAI SZERKEZETÉNEK ANALÍZISE ÉS SZINTÉZISE

Kutatástörténeti áttekintés

Bolla Kálmán

1. A hazai kísérleti fonetika történetében az utóbbi tizenkét év tudományunk fejlődését hosszabb távra meghatározó eredményeket hozott. Eredményen nemcsak könyvekben és cikkekben testet öltött tudományos teljesítményt értek, hanem tágabb összefüggésben a kutatásfejlesztés számos más területén történt előrehaladást is, mivel a kutatómunka színvonala és hatékonysága nagymértékben a tárgyi és személyi feltételek alakulásától függ. A munka minőségét a kutatás szervezésének, irányításának, összehangolásának a gyakorlata éppúgy befolyásolja, mint ahogy a szakmai információcsere állapota (pl. írásos publikációs lehetőségek, részvétel tudományos tanácskozáson, konferenciákon, szimpóziumokon; a hazai és külföldi kutatóhelyekkel való kapcsolatok, a kiadványcserék stb.) is érezhető hatást gyakorol a tudomány fejlődésére. A Fonetikai Füzetek korlátozott terjedelme miatt csak az MTA Nyelvtudományi Intézetének fonetikai osztályán – amely egyben országos bázisa is a magyar kísérleti fonetikának – végzett beszédakusztikai munkálatokról adhatok rövid áttekintést. A kutatástörténeti összefoglalás közlését a beszéd akusztikai analízise és szintézise témakörnek szentelt kötetünkben az alábbi körülmény teszi szükségessé:

Eddig még nem jelent meg szélesebb szakmai közönségnek szóló, átfogó írásos tájékoztatás az intézetünk fonetikai osztályán az utóbbi évtizedben végbement fejlődésről. A kötet egyes tanulmányai pedig nem vagy túlságosan egyoldalúan és hézagosan számolnak be az előzményekről. Szeretnénk, ha az olvasó több és teljesebb ismeretek birtokában, a kutatási folyamat egészének figyelembevételével ítélné meg munkánkat. E füzetben közölt eredmények egy hosszabb kutatási szakaszra épülnek, ugyanakkor az előmunkálatok nagy részéről írásos publikációban még nem számoltunk be. Az alábbiakban, érthetően, ezekről is szót ejtünk.

2. A hazai fonetika régebbi és újabb történetével, helyzetével, továbbfejlesztésének problémáival több önálló kiadvány és közlemény is foglalkozik. Az 1970 előtt megjelent írások adatait megtaláljuk A magyar hangtan válogatott bibliográfiájában (Szerk. Bolla–Molnár 1977, 27–74). Ehhez kiegészítésként hivatkozhatok még egy 1968-ban rendezett tanácskozáson elhangzott hozzászólásomra, amelyben a fonetikai kutatások állapotával és gondjaival foglalkoztam (Általános nyelvészetünk... 1969, 94–8). 1974-ben Büky Béla Beszédkutatás, halláskutatás és ezek rokonterületei Magyarországon 1960–1969 között címmel közölt bibliográfiája is tartalmaz a témába vágó munkákat (ÁltNyT X, 1974, 281–311). Vértes O. András az 1980-ban megjelent nagyszabású tudománytörténeti monográfiájában feldolgozta a magyar leíró hangtan történetét a kezdetektől az újgrammatikusokig terjedő korszakra (Vértes O. 1980). Folytatása az 1945-ig tartó időszakra ebben az évben jelent meg a Fejezetek a magyar leíró hangtanból című tanulmánykötetben (Szerk. Bolla 1982, 283–301). Ugyancsak a fenti kötet Bevezetésében is találunk a fonetikai kutatások történetéhez kapcsolódó tájékoztatást (uo. 9–12). Kiadványsorozatunk 6. füzetében, megemlékezve szaktudományunk két kiváló művelő-

jéről, Fónagy Ivánról és Tarnóczy Tamásról, közzétük műveik jegyzékét is, amely számos beszédakusztikai vonatkozású tanulmányt tartalmaz (Szende MFF 6. 172–8; Vicsi 168–72).

3. Visszatérve a Nyelvtudományi Intézet fonetikai osztályára és az ott folyó beszédakusztikai kutatómunkára, először a kutatás tervezésével, tárgyi és személyi feltételeinek alakulásával, azaz az érdemi kutatómunka megalapozásával összefüggő néhány – általunk lényegesnek tartott – körülményre szeretnék rámutatni.

a) 1969–1971 között az intézet vezetőségétől kapott megbízás és intenciók alapján kidolgoztunk egy olyan bővíthető és továbbfejleszhető kutatási programot, amellyel elérhettük a kísérleti fonetikában bekövetkezett lemaradásunk gyorsabb ütemű felszámolását és meggyorsíthattuk az európai élvonalhoz való felzárkózásunkat. E koncepció lényegesebb vonásai az alábbiak:

- komplex nyelvészeti-fonetikai tematikát fog át,
- strukturális–funkcionális metodológiával dolgozik; magában foglalja az objektív és a szubjektív módszerek együttes, egymást kiegészítő alkalmazását; összekapcsolja a műszeres és az elméleti kutatásokat;
- épít az interdiszciplináris együttműködés eredményeire,
- hagyományörző és nyitott a fonetika nemzetközi eredményeinek a befogadására; tapasztalatcserére és együttműködésre törekszik más országok fonetikusaival.

A fonetika tárgyának, valamint a nyelvészeti-fonetikai ismeretek terjedelmének, határainak a megítélését, s mindezzel összefüggésben a hangtan szakrendszerének a meghatározását illetően mind a mai napig nem alakult ki egységes álláspont tudományunk művelői között. Ma már egyre ritkábban találkozunk nyelvészkörökben olyan véleményekkel, miszerint a hangtan tulajdonképpen nem is igazán nyelvészeti diszciplína, csupán annak segédtudománya. Ugyanakkor számos nyelvelírásban legfeljebb bevezetésképpen szerepel a fonetika. Nálunk általánosnak tekinthető mind az elméleti igényű, mind pedig a gyakorlati célzatú (elsősorban a különböző szintű tankönyvekre gondolok) fonetikai munkákban az erősen leszűkített, többnyire csak a hangállomány bemutatására és képzési sajátságainak a leírására korlátozódó fonetika. (A szerzők gyakran az ortográfia szempontjából válogatva és súlyozva, a helyesírási problémák megoldásának szolgálatába állítva tárgyalják a fonetikai anyagot.) Már korábban kifejtettük, hogy a nyelvészeti fonetika megújulása (ezzel a címmel tartottam előadást 1975-ben a Magyar Nyelvtudományi Társaság Fonetikai és Beszédművelő Szakosztályának az ülésén) nem képzelhető el a szakismeretek differenciálódása, szakrendszerének a kibővítése nélkül, azaz nem hagyhatjuk figyelmen kívül sem a beszédképzés és beszédértés (a fonetikai kódolás és dekódolás) kortikális folyamatait, sem a beszédészlelés és beszédészlelési megjelenési formáját, sem a beszédészlelés, -értés auditorikus és perceptorikus vonatkozásait, sem pedig a fonológiai szempontokat (Bolla NyK LXXVIII, 1976, 292–9). Meggyőződésünk szerint a fonetika tárgya a nyelv hangalakja, a beszéd hangteste a maga teljességében, mind szegmentális, mind pedig szuprasegmentális megszerkesztettségét illetően strukturális és funkcionális aspektusban egyaránt. Ezért szerepel vizsgálódásainkban nagy súllyal a beszédintonáció témaköre is (erről bővebben: Bolla MFF 5. 1980, 40–69, továbbá Fejezetek... 1982, 21–2).

Nem foglalkozhatunk részletesen a fonetikai kutatásainkban alkalmazott módszer-

rekkel és eljárásokkal. Módszerünk egészében véve a strukturális–funkcionális metodológiával jellemezhető. Vizsgálatainkban az objektív elemzési módszereket a szubjektív megfigyelésekkel együtt, egymást kiegészítve alkalmazzuk. A műszeres mérések adatait csak a nyelvész szubjektív nyelvi mértékrendszerén átszűrve, a nyelv gondolatformáló és kommunikatív funkciója szempontjából értékelve és értelmezve tartjuk fonetikailag relevánsnak, valós nyelvi struktúrák és folyamatok egzakt kifejezőinek, amelyekből a szaktudomány számára elfogadható és új ismereteket nyerhetünk. A „mikroszkopikus” elemzés számára kiragadott hangtest-elemeket, az akusztikai szerkezet mikrometszeteit is mindig a nagyobb egészben, nyelvi funkciók szerveződésein, szélesebb összefüggéseikben értékeljük és minősítjük. Ez a metodológia és szaktudományos szemlélet vezetett el bennünket a komplex beszéd-szintézis megvalósításához, az imperszonális beszéd automatikus és azonos idejű mesterséges előállításához (ld. Voxon, Univoice).

Az interdiszciplinaritás a kísérleti fonetika egyik legjellemzőbb vonása. Sokan valljuk – fonetikusok, nyelvészek és a nyelvvel foglalkozó más ágazatok képviselői is –, hogy a tudomány mai fejlettségi szintjén a nyelv hangalakjának a vizsgálatában jelentősebb új eredmények elsősorban komplex kutatásból, nyelvészek, pszichológusok, fiziológusok, fizikusok alkotó együttműködéséből, matematikus és a számítástechnikában járatos szakemberek együttes erőfeszítéséből születhetnek. Ezért alakítottuk úgy a fonetika osztály külső kapcsolatait, hogy munkánkban támaszkodni tudjunk a társtudományokra, igényeljük közreműködésüket, és magunk is szívesen részt veszünk olyan kutatásaikban, ahol nyelvészeti-fonetikai szakismeretek felhasználhatók. Szilárd meggyőződésünk továbbá, hogy nincsen a nyelvtudománynak még egy olyan ágazata – a lexikográfiát is beleértve –, ahol a „csapatmunka” (team munka) annyira fontos, a munka minőségét és eredményességét annyira meghatározó volna, mint a kísérleti fonetikában. Ugyanakkor a szűk látókörű, önző egyénieskedésnek az intenzívebb továbbfejlődést fékező és bénító hatásait sem kivédeni, sem kikerülni nem tudjuk.

Ami pedig a kontinuitást illeti, a magyar fonetika sohasem volt irigylésre méltó helyzetben, s tegyük hozzá, ma sincs abban. Folyamatosságról csak nagyobb történelmi távlatban nézve beszélhetünk. Kempelen Farkas, Balassa József, Gombocz Zoltán, Csúry Bálint, Laziczius Gyula nevével fémjelvezhető az a tudománytörténeti folyamat, amelyben nagy megszakításokkal követik egymást a „csúcsok”. Gombocz Zoltán gyakran idézett megállapítása (ld. Molnár 1970a): „...hiszen talán egy tudományban sem szakad meg a fejlődés vonala annyiszor, mint a hangtanban, egy tudomány művelői sem kezdték a megfigyelés és rendszeralkotás munkáját annyiszor a legelejéről, abban a meggyőződésben, hogy töretlen úton járnak, mint a hangtan munkásai” (Gombocz Nytud. II, 241), úgy véljük, nem vesztette el érvényét a megjelenését követő hét évtizedre vonatkoztatva sem. Ezzel a sajátos helyzetű magyaríráshoz, hogy nálunk nem szerveződtek olyan rangos kísérleti-fonetikai laboratóriumok, nem alakultak ki olyan alkotóműhelyei a fonetikának, mint ahogyan a múlt század végétől kezdődően Kazányban, Moszkvában, Leningrádban, Helsinkiben, Kijevben, Prágában, Varsóban, Koppenhágában, Hamburgban és számos más európai országban sorra születtek fonetikai laboratóriumok, nálunk nem jött létre iskola ebben a tudományágazatban. Különösebb tévedés veszélye nélkül megkockáztatható az az állítás, hogy nem vagy alig volt olyan magyar nyelvész, aki fő profilban művelte volna a fonetikát. Igaz, hogy nálunk nem következett be a fonetika elkülönülése sem a nyelvészettől, átcsőszása más tudományterületekre (fizika, fiziológia,

pszichológia, logopédia), és nem szorult ki a tudományból sem, megmaradt mindvégig a nyelvészet „bölcsőjében”. Ez a körülmény azonban káros következményekkel is járt. A magyar fonetika elszigetelődött a továbbfejlődés számára oly fontos tudományközi hatásoktól, nem tudta feldolgozni, adaptálni a társtudományok által feltárt új ismereteket. Még Tarnóczy Tamás ismert fizikus beszédakusztikai munkái sem találtak kellő visszhangra a nyelvtudományban, nem integrálódtak szervesen a magyar fonetikába. A kedvezőtlen helyzetnek pozitív következménye az volt, hogy a hazai fonetika rákényszerült a külföldi műhelyekben való tájékozódásra, az új eredmények megismerésére és befogadására. (Ezt példázza Balassa, Gombocz, Laziczius, Hegedűs, Fónagy és Magdics munkássága is.) A nyitottság olyan örökségünk, amelyet megtartani és továbbvinni kötelességünk, mely továbbfejlődésünk fontos tényezője, sőt föltétele. Ennyit a koncepció magyarázatára. De térjünk vissza a jelenbe, illetőleg a közelmúlt fejlődésének a bemutatására.

b) Szóltunk a kutatási programunk néhány jellemző vonásáról, s most szóljunk a kutatói potenciál fejlesztéséről/fejlődéséről is röviden. Az 1971-ben 4 fős csoport a duplájára növekedett, s a külső együttműködéssel tovább erősödött. Új kutatókkal, villamosmérnökkel, technikussal, programozó matematikussal és laboránssal egészült ki a régi kutatócsoport. Ez a felállás minimálisan biztosítja tudományunk művelhetőségét (életben tartását és szerény mértékben való előbbrevitelét). Súlyosabb nehézségeket jelent számunkra elhelyezésünk szűkösége és megoldatlansága. A személyi állomány tudományos fejlődését jelzik a tárgyalt időszakban megvédett egyetemi doktori értekezések és kandidátusi disszertációk, a megjelentetett könyvek és tanulmányok.

c) 1970–1979 között megvalósítottuk a fonetikai laboratóriumunk rekonstrukcióját, amely valójában egy teljesen új laboratóriumi komplexum kiépítését jelentette. A beszerzett és a feladatnak megfelelően laboratóriumi rendszerbe állított kutatási eszközök és műszerek magas műszaki–technikai színvonalon képesek kielégíteni a nyelvészeti-fonetikai kutatás mai igényeit mind a beszédfiziológiai, mind az akusztikai, mind pedig a beszédészleléssel és a beszédértéssel kapcsolatos témakörökben. (Eszközeink ismertetésére a 4. pontban még visszatérünk.)

d) 1978-ban Magyar Fonetikai Füzetek (Hungarian Papers in Phonetics) címmel kiadványsorozatot indítottunk. Az eddig megjelent kötetekről és közleményekről bibliográfiai összeállítást készítettünk (ld. MFF 10. 153), így ennek részletesebb ismertetésétől ezúttal eltekinthetünk. A kiadvány a magyar fonetika fejlődése számára felbecsülhetetlen fontossággal és jelentőséggel bír, de ennek objektív megítélésére jobb, ha nem mi vállalkozunk. Száz körül mozog azoknak a külföldi kutatóhelyeknek és szakembereknek a száma Európában, Ázsiában és Amerikában, ahova eljut szerény kiadványunk. Az MFF jelentősen hozzájárul szakmai kapcsolataink fejlődéséhez, a kiadványcsere elősegíti szakmai kapcsolataink fejlődését, a kiadványcsere formájában hozzánk érkező „műhelymunkákból” (working papers) és kutatási beszámolókból pedig gyorsan tájékozódhatunk tudományunk legújabb eredményeiről. A hazai fonetika külföldi elismertségének a jeleként értékeljük, hogy újabb és újabb rangos külföldi műhelyek érdeklődnek munkánk iránt, igénylik kiadványunkat és figyelemre méltó szakmai információs forrásként tartják nyilván.

e) A fonetika kutatásában, oktatásában és alkalmazásában szerzett tapasztalatok kicserélésére és a problémák megvitatására 1977-től kezdődően évente tudományos ülés-

szakot szervezünk más szakmai szervezetekkel és bizottságokkal közösen. Szimpóziu-
mainkon az alábbi témakörök szerepeltek:

- Fonetika '77: Vizsgálatok a hangtan köréből (Hangadás, hangfejlődés, hangrend-
szer, intonáció, beszédjavítás). 1977. szeptember 28–29-én tartottuk a
Bárcsi Gusztáv Gyógypedagógiai Tanárképző Főiskolán.
- Fonetika '78: A beszédintonáció néhány elméleti, módszertani és gyakorlati prob-
lémája. Ideje: 1978. szeptember 26. Helye: az ELTE Bölcsészettudományi
Kara.
- Fonetika '79: Interlingvális hangtani egybevetések. Az MTA Roosevelti téri szék-
házában tartottuk 1979. szeptember 27-én.
- Fonetika '80: A nyelvészeti fonetika aktuális és vitás kérdései címmel a) a nyelv
artikulációs bázisa, b) a nyelv percepció bázisa, c) a supraszegmentális
hangszerkezet kérdései, d) a hangtani műszóhasználat problémái, valamint
e) a fonetika és a beszédzavarok témakörökben hangzottak el vitaindítók és
folytattunk élénk eszmecsere. A szimpóziomot a Művelődési Minisztérium
tanácsstermében tartottuk 1980. december 9-én.
- Fonetika '81: A mai magyar nyelv hangrendszere és helyesírásunk témában 1981.
december 8-án a Gyógypedagógiai Tanárképző Főiskola klubtermében tar-
tottuk meg nyilvános tanácskozásunkat.

f) Az alap kutatásban elért eredmények gyakorlati felhasználásának az elősegítése
céljából jó együttműködés alakult ki oktatási intézményekkel (egyetemekkel és főisko-
lákkal), kórházakkal, igazságügyi szervezetekkel és különböző területeken dolgozó gyar-
korlati szakemberekkel.

4. A beszédakusztikai analízis és szintézis sajátos hardware és software feltételek-
hez kötődik. Megteremtése a fonetikai laboratórium rekonstrukciójával egyidejűleg folyt.
Minthogy az eszközök megválasztása a kutatási feladatok függvényében történik, mi is
először a szakmai profilt és a kutatási célokat tisztáztuk. Ezután felmértük a műszerhasz-
nálat kooperációs lehetőségeit, majd kijelöltük azokat az eszközöket, illetőleg eszközfaj-
tákat, amelyek a fonetikai laboratórium alapfelszerelését, sine qua nonját képezik. Tá-
jékozódunk a vizsgálandó témák kutatására használható eszközök beszerezhetőségét il-
letően, és hozzáálltunk a műszerbeszerzési tervünk megvalósításához. Minthogy nincs
két teljesen egyforma kutatási tematika szerint működő kutatóhely, úgy nincsen két
egyforma fonetikai laboratórium se. Ezenkívül minden országban a számukra beszerez-
hető műszerekből, sokszor saját konstrukciójú eszközökből képezik ki laboratóriumai-
kat. Így a tervezésben nekünk is szükségszerűen saját tapasztalatainkra, ismereteinkre
és elképzeléseinkre kellett építenünk. Nem részletezem azokat a nehézségeket, amelyek-
kel szembe találtuk magunkat a munka során, de a történeti hűség kedvéért megemlí-
tem, hogy ez a feladat hosszú időn keresztül aránytalanul nagy munkabefektetést köve-
telt néhány embertől, míg többen arisztokratikus közönnyel szemlélték az „újrakezdés”
kínlóadásait, gyötrelmeit. Végül is 1979 elejére sikerült elképzeléseinket megvalósítanunk.
A továbbiakban, a témánál maradva, a beszédakusztikai kutatáshoz használatos labora-
tóriumi eszközöket és műszereket ismertetem.

1. Kép- és hangfelvevő stúdió

A beszéd folyamat akusztikai felépítésének a tüzetesebb megismerésére, valamint
a beszéddel kapcsolatos akusztikus jelenségeknek a műszeres vizsgálatára irányuló fone-

tikai kísérletek elengedhetlen követelménye a jó minőségű hangfelvétel. Mind a szubjektív elemzések, mind pedig a műszeres mérések megkívánják a vizsgálendő nyelvi anyag előzetes stúdiófelvételét. Stúdióink egy hangszigeteléssel ellátott, mikrofonokkal, videokamerával, vizuális és auditív jelzővel, továbbá világító berendezéssel felszerelt csendes szoba (camera silenta), amely vezetékes összeköttetésben van a mérőlaborban elhelyezett kép- és hangrögzítőkkal. Nem biztosít ugyan teljes hangnyelést és hangszigetelést, de a célnak megfelel. Hangfelvételre 1974-től egy svájci gyártmányú Studer A/80-as típusú **stúdiómagnetofon** és 1976-tól Sennheiser MKH 815 T típusú **kondenzátor mikrofont** használunk, míg a képfelvételekre Siemens gyártmányú **videokamera** áll rendelkezésünkre. Stúdióink kiegészítő és kisebb felszerelési tárgyait (pl. feszültségstabilizátor, előerősítő, végerősítő, hangszórók, videomonitor, fejhallgatók stb.) hazai gyártmányú eszközökből válogattuk össze. A stúdiómagnetofon üzembe helyezését és bemérését Zombori László hangmérnök és Kiss István elektrotechnikus, a Siemens Sirecord S típusú képmagnetofonét pedig a Híradástechnika KSZ végezte 1973-ban. A csendes szoba zajszint tesztjét 1979-ben Olasz Gábor villamosmérnök és Nikléczy Péter technikus készítette el.

2. A mérőlabor akusztikus elemzőeszközei

A beszéd akusztikus kifejeződési formáját tekintve így határozható meg: A beszéd hangalakja nem más, mint frekvenciáját, intenzitását időben folyton változtató összetett rezgéseknek szünetekkel váltakozó sora. Ebben a rendkívül gyors folyamatváltozásban kimutathatók ismétlődő rezgéssztereotípiák, kváziazonos rezgéskomplexumok, valamint nagyobb szekvenciákat átfogó szerkezeti formációk. Ez adja az akusztikus folyamat strukturáltságát, szerkezeti tagoltságát, meghatározható szabályosságok szerinti felépítettségét. Lényegében ez teszi alkalmassá jelzőrendszerként való felhasználásra. A beszéd folyamat akusztikai elemzésének a célja, hogy feltárja a hangtest szerkezetét, hogy kimutassa miből és hogyan épül fel a beszédhangsor, hogy megismerjük a különböző szerkezeti összetevőknek a nyelvi jelek realizálásában és elkülönítésében, illetve a nyelvi tartalmak kifejezésében játszott szerepét. Az akusztikai elemző- és mérőberendezések segítségével megvizsgálhatjuk a hangtestet akár fizikai teljességében, akár részenként is.

- A szovjet gyártmányú K 115 típusú 12 csatornás **oszcillográf**ot 1972 óta használjuk. Üzembe helyezését és beszabályozását Rappai Romuald technikus (BME) végezte. Az oszcillográffal a beszéd teljes akusztikus rezgésképét folyamatában vizsgálhatjuk, filmre, fotopapírra, ultraibolya sugárra érzékeny anyagra rögzíthetjük.
- Az EG 830 típusú **elektroglottográf** a dán F-J Electronics A/S által fonetikai kutatási célra kifejlesztett rendkívül hasznos műszer család tagja. 1974-től használjuk a gége szintjén végbemenő rezgések, azaz a zöngének a vizsgálatára. E műszer segítségével hallhatóvá tehető a zöngé, s megfigyelhető a hangfolyamnak a zöngé meglétéből és hiányából szerveződő tagoltsága.
- Az 1973-ban beszerzett **alaphangmagasság-mérőnk** (FFM 650 típ.) a már említett dán műszer családkhoz tartozik. A hanglejtésvizsgálatok nélkülözhetetlen eszköze. A műszerbe beépített elektronikus szűrők lehetővé teszik a beszélő alaphangmagasság-sávjának (hangterjedelmének) megfelelő beállítást. A hangszalagról vagy mikrofonról bevitt hangsorban, rezgéshalmazban ezáltal elkülöníthetjük a beszéd dallamát, hanglejtését, alaphangmagasság-változásait kifejező összetevőket, és egy kiíró berendezéssel tetszés szerinti felbontásban (25–1000 mm/s között) papírte-

kercsre vesszük. Az így nyert diagramról az idő és a frekvencia függvényében leolvassuk a beszéd folyamat magassági hullámzásait reprezentáló értékeket Hz-ben, az időtartamokat pedig ms-ban. Az alaprezgéssel párhuzamosan a műszer külön kimeneteléről megkapjuk a beszéd folyam teljes rezgésképét is ún. duplex oszcillogram formájában. A duplex oszcillogramon végezzük el a hangsor fonetikai egységekre (a kutatási célnak megfelelően beszédhangokra, szótagokra, beszédszakaszokra vagy szólamokra) tagolását, s ezután már könnyen megállapítható, hogy a folyamat mely pontjához milyen magassági érték tartozik. Az alaphangmagasság-mérő felhasználásával készült diagramokon gyorsan áttekinthető, globális és aprólékos feldolgozásra egyaránt alkalmas formában kapjuk meg a beszéd folyamat dallamszerkezetét.

- Szintén 1973-ban vettük használatba a beszéd dinamikai szerkezetének a vizsgálatára kifejlesztett dán gyártmányú **intenzitásmérőt** (IM 360 típus). A készülék többféle beállításban használható. Mérhetjük a komplex rezgéssor erősségi viszonyainak a változását, és beállítható egy vagy két tetszőlegesen kiválasztott frekvenciasávra (pl. az alaprezgésre, a magánhangzók vagy a mássalhangzók specifikus frekvenciasávjára). A készüléken több (5, 10, 20, 25 ms) integrálási idő állítható. Az erősségi viszonyoknak megfelelően változó, s a mingográf felvett elektromos impulzusokból rajzolódik ki a vizsgált hangsor intenzitásgörbéje. Leolvasása és adatolása a dallamgörbével egyező módon történik, az intenzitásértékeket dB-ben fejezzük ki.
- Az AFF 440 típusú **hangfrekvenciás szűrővel**, amelyet a fenti két műszerrel egy időben (1973-ban) szereztünk be Dániából, tovább bővíthetjük, szélesíthetjük a beszédakusztikai méréseinket és megfigyeléseinket.
- A fentebb ismertetett műszerből mérőláncot alakítottunk ki. Rögzítő berendezésként egy 4 csatornás EM 34 T típusú svéd gyártmányú **minográfot** (beszerzési ideje 1973) használunk. Tintaíróval működik, 120 mm szélességű papírekercsre egyidejűleg 4 görbét tud írni. A beszédintonációs vizsgálatainkat rendszerint duplex oszcillogramból, hanglejtésgörbéből és két intenzitásgörbéből álló intonogramokon (mingogramokon) végezzük. A mingográf regisztráló sebessége széles határok között (25–1000 mm/s) változtatható, s ezzel a diagramok felbontó képessége a kutatási célnak megfelelően növelhető vagy csökkenthető. A mérőlánc üzembe helyezésében Kontár Jenőné mérnök volt a segítségünkre, míg a kutatási technológiát magunk dolgoztuk ki.
- Később az ismertetett mérőlánchoz még két kiegészítő műszert készítettünk, **dinamikakompresszort** és **imitátort**. Ez utóbbival hallhatóvá tesszük az alaphangmagasság-mérővel és az intenzitásmérővel kielemezett hangsajátságokat. Az imitátorral, továbbá hangfelvevő és lejátszó berendezéssel kibővített lánc jól használható mind a beszédhallás-fejlesztésben, mind pedig a percepciósi kísérletekben. A két műszer Olasz Gábor mérnök és Nikléczy Péter technikus munkája. A hangszalagról betáplált beszéd felvételről az alábbi bemutatások végezhetőek el:
 - a) csak a szöveg alaphangmagasság-változásait halljuk változatlan időszerkezetben),
 - b) a szöveg erősségi változásait érzékeljük egy kiválasztott állandó vivőfrekvencián a természetes ejtéssel egyező időszerkezetben,

c) a szövegből a magassági és a dinamikai változások együtt szólalnak meg változatlan időszerkezettel, és végül

d) a hangsor zöngés szekvenciáit egy állandó frekvenciaértékkel, a zöngétlen részeket szünetekkel megvalósított demonstráció.

- A beszéd folyamat szubjektív megfigyeléssel történő elemzésére szolgáló eszközünk még az **elektronikus beszédlassító és -gyorsító** (Varispeech II SN-1103 típus.). Ezzel az 1976-ban az NSZK-ból beszerzett, a Lexicon Inc. által gyártott készülékkel a hangfolyam időszerkezetét tudjuk változtatni. A hangkazettáról lejátszott felvételt, a frekvenciamenet és dinamika jelentősebb torzítása nélkül, időben tág határok között nyújthatjuk vagy „zsugoríthatjuk”. A tempóváltozások nyelvi szerepének a vizsgálatára, a beszédsebesség-határok megállapítására, a beszédíram-sztereotípiák „kitapogatására”, valamint a szegmentális szerkezet építőelemeinek a megfigyelésére (hangminőség, hangátmenetek, hangkapcsolódások stb.) kiválóan alkalmas eszköz.
- Akusztikai elemzőberendezéseink külön csoportját képezik az ún. **hangszínképelemzők és dinamikus hangszínképírók** (hangspektrográfok). 1966-ban került a laboratóriumunkba egy Kay Electric (USA) gyártmányú 6061 A típusú Sona-Graph. A kutatás pontosságát és hitelességét garantáló műszaki beszabályozását, javítását Borsos Szabó Ferenc mérnök végezte el 1973-ban. Az akusztikai elemzésekhez ezen kívül használunk még egy 1972-ben beszerzett Brüel Kjaer gyártmányú frekvenciaanalizátort (Frequency Analyzer 2107). A dinamikus hangszínképíró az akusztikai vizsgálatokban semmi mással nem helyettesíthető alapműszer. 1974-től a szonográfokból az amerikai Voice Identification Inc. által továbbfejlesztett és tökéletesített 700-as szériaszámot viselő hangspektrográffal (Sound Spectrograph Series 700) dolgozunk. A hangszalagra felvett szöveg bármely 2,4 s-os (vagy ennél kisebb) szelvényéről részletes hangszínképet kapunk többféle, előre kiválasztható elemzésmódban (széles vagy keskeny sávú, magas vagy mély kiemelésű, egyenletes átvitelű, terepvonalas kiíratású, különböző skálanyújtásokkal készített elemzések, amplitúdómetszet). A speciális hőérzékeny papírra égetett hangszínképből kiolvashatók az elemzett hangszakasz idő-, frekvencia- és intenzitásadatai. A hangszínképelemző fonetikai felhasználhatóságát tovább növeli még a beépített szegmentátor, amellyel 0,01–2,4 s között tetszés szerint kiválasztott intervallumban meghallgatható a hangsor bármely pontjának a rezgése. Az elemzendő részek helyes kiválasztását, továbbá a vizsgálat pontosságát az auditív kontroll mellett egy, a közelmúltban beszerzett kétsugaras tárolós oszcilloszkóppal (5113 típus.) végezzük. Így a hangszínképelemzésre Olasz Gábor és Nikléczy Péter közreműködésével kialakítottunk egy megfelelő mérőláncot, valamint mérési technológiákat dolgoztunk ki a legkülönbözőbb fonetikai vizsgálatokra.

3. A beszéd szintézis hardware eszközei

A beszéd hangtestének mesterséges előállítására való törekvés nem újkeletű a fonetikai kutatásainkban, hanem – az előkészületeket is beleszámítva – több mint másfél évtizedes kísérletezés van már mögöttünk. A legutóbbi évek látványos eredményei sem úgy születtek meg, mint ahogyan a görög mitológia szerint Pallasz Athéné pattant ki teljes fegyverzetben Zeusz fejéből, hanem hosszabb időszak céltudatos kutatómunkájának a terméke. Nem egy ember „felfedezése és találmánya” csupán, hanem egy alkotó-

közösség együttműködésével, összehangolt munkájával elért eredmény. Alkotóközösségen nemcsak a fonetikai osztály dolgozóit értem, hanem azokat a külső munkatársakat is, akik szakmai tanácsokkal segítették munkánkat, vagy más módon működtek közre az eszközök kiválasztásában, beszerzésében és üzembe helyezésében, beleértve a hivatali dolgozókat is, intézetben belül és kívül egyaránt.

A szintézis számunkra sohasem volt öncél, hanem tágabb értelemben a nyelvi valóság megismerésének egyik hathatós eszköze, közelebből pedig az elemzés kontrollja és egy magasabb szintű elemzés lehetősége, az egész és a részek kölcsönviszonyainak a tanulmányozására legalkalmasabb eszköz és módszer. A beszéd dallamának és a magánhangzószerű hangoknak hanggenerátorral való előállításától az automatikus szövegszintézisig hosszú út vezetett. Minőségi változást jelentett kutatásainkban a svéd gyártmányú OVE III C típusú beszéd szintetizátor beszerzése és a kutatás szolgálatába állítása 1973–74-ben. Elektronikus vezérlés híján manuális működtetéssel használtuk 1979-ig, mikor is beérkezett és üzembe helyeztünk egy laboratóriumi kisszámítógépet, az amerikai DEC cégtől vásárolt PDP 11/34 típusú minicomputert. A manuális „vezérlés” időszakában 1977-ig elkészültünk a magyar magánhangzók mesterséges előállításával, beleértve az addig publikált adatok helyességének szintézissel való kontrollálását is és a magánhangzó-identitás felmérését a hangminőség alapján (Bolla MFF 1. 1978, 54–67; uo. Olasz 68–76). A hangzók terjedelmében tapasztalható folyamatváltozásokat 3–5 statikus metszettel, ún. hangszelettel próbáltuk megvalósítani a szintézis manuális korszakában. A beszéd szintézis eszközeinek a historikumához még csak annyit, hogy köszönettel tartozunk Tapolczai Lászlónak, a SZTAKI munkatársának, aki a számunkra legnehezebb szakaszban a számítógépes konfiguráció összeállításában és a software-fejlesztés elindításában nyújtott hathatós segítséget. A részletes hardware ismereteket az érdeklődők megtalálják e kötetben Kiss Gábor és Olasz Gábor írásában (Kiss–Olasz MFF 10. 1982, 21–46), megismétlenségük e helyütt szükségtelen.

5. Rátérünk utolsó kérdéskörünk ismertetésére, a magyar beszéd akusztikai szerkezetének a vizsgálatára irányuló fonetikai kísérletek összegezésére. Az 1970-ig terjedő időszakban a hazai beszédakusztikai kutatások eredményeiről elsősorban Tarnóczy Tamás (ld. Vicsi MFF 6. 1980, 168–72), Hegedűs Lajos (Hegedűs 1930; 1934; 1957), Magdics Klára (Magdics 1965; 1965a, 1966; 1964; 1970, valamint 1963), Fónagy Iván (ld. Szende MFF 6. 1980, 172–8) és Molnár József (Molnár 1970) munkái tájékoztatnak. Kutatócsoportunkban a hetvenes évek elején elvégeztük a beszédakusztikai szakirodalom kritikai feldolgozását, a saját mérési adatainkkal való egybevetését, beleértve a szintézissel való ellenőrzést is (ld. Bolla MFF 1. 1978, 53–67; uo. Olasz 68–76; Kassai 1979; 1982, 115–54 és Vértes O. 1982, 71–113).

A magyar beszéd akusztikai szerkezetének az elemzésére és szintézisére irányuló kutatásainkat az alábbi négy témacsoportra oszthatjuk: a) a beszéd folyamat hangelemei, b) a hangkapcsolatok, szóterjedelmű hangsorok, c) a beszédintonáció, azaz a szupraszegmentális hangszerkezetek és d) a beszédmű szövegfonetikai elemzése és szintézise; tehát a hangtest egészére kiterjedően folytattunk fonetikai vizsgálódásokat. Mindezen fölül még két kiegészítő témakörben végeztünk kísérleteket: a) foglalkoztunk a beszédfejlődés korai szakaszának hangjelenségeivel és b) a magyar beszéd folyamat egyes akusztikai sajátosságait egybevetettük az orosz, az angol és a lengyel nyelv fonetikai jellemzőivel.

Végül 1982-ben eljutottunk a komplex beszéd szintetizáló rendszerek (Voxon, Univoice, Minivoxon) kifejlesztéséhez.

Hosszú lenne még a pusztán felsorolása is azoknak a kísérleteknek, amelyeket a tárgyalt időszakban végeztünk, és amelyekről nem jelent meg írásos közlemény. Saját kutatásaimban 140 hangfelvételt tartok nyilván, hetvenféle résztermében végeztem hangszínképelemzést, 130 körül van az intonációs felvételeim száma, 100-nál több kísérleti alánnyal dolgoztam, mesterségesen előállítottam az összes magyar magán- és más-salhangzót – a különböző személyek ejtésének megfelelő hangzásban és személytelen változatban is. Szintetizáltam a magyar beszédben előforduló szupraszegmentális szerkezet típusokat, dallamformákat, és 1982 márciusában létrehoztam az első magyar nyelvű folyamatos beszéd szintetizáló rendszert, a Voxont. A munkálatok főbb állomásait jelzik az alábbi kísérletek és a kísérletekről szóló híradások:

1968: Beszédhangszerű komplex hangok és dallamformák előállítása hanggenerátorral, valamint a beszéd akusztikus összetevőit elemzően bemutató demonstrációs hanganyag elkészítése. – Felhasználtam az egyetemi fonetika-oktatásban.

1969–76: A magyar magánhangzók akusztikai analízise és szintézise, továbbá kísérlet magyar beszédhangsorok mesterséges előállítására (*sás, sas, sós, süss, ás, is, ós, és, sí, se*). – Bemutattam az 1977. szeptember 29-én megtartott fonetikai szimpóziumon.

1974: A beszédintonáció kísérleti-fonetikai elemzésére kialakítandó módszerek és technológia beállítása. – Nem publikált műhelymunka.

1978: Első kísérlet függvénygenerátorral a dallamformák számítógépes szintézisére Olasz Gábornak, továbbá az ÉTI két munkatársának, Hartmann Tivadar villamosmérnöknek és Csúri Ottó programozónak a közreműködésével. – Nem publikált fonetikai kísérlet.

1977–79: Kísérlet a beszédintonáció (azaz a szupraszegmentális hangszerkezet) elemeinek a bemutatására elemzéssel és szintézissel. – Erről az MNYT Általános Nyelvészeti, valamint Fonetikai és Beszédművelési Szakosztályának ülésén adtam bemutatással egybekötött tájékoztatót 1979. május 29-én.

1978: A beszéd szupraszegmentális szerkezetének a fonetikai elemzése és leírása. Kísérlet az intonáció egzakt lejegyzésére, műszeres mérési eredmények alapján. – Ehhez kapcsolódik az MFF 5-ben (1980, 40–69) megjelent közleményem.

1979: Beszédünk zenei elemei, hanglejtésformák és zenei struktúrák. Kísérlet a beszédintonáció fonetikus és zenei lejegyzésének a megfeleltetésére. – Nem publikált műhelymunka.

1979: Magyar dallamformák természetes ejtésű hangsorokban, imitátoros felvételben és szintetizált formában. – A mérés technológia ellenőrzésére és továbbfejlesztésére szolgáló kísérletsorozat.

1979: Beszédsebesség és beszédérthetőség. A vizsgálat anyaga a Doktor Faustus várszínházbeli előadásáról készült hangfelvétel. – Nem publikált kísérlet.

1979: Kísérlet a beszéd ritmikai szerkezetének az elemzésére. A kísérlet anyaga a Magyar Rádióban Török Tamás rendezésében elhangzott Weöres Sándor költői munkáiból készült összeállítás. – Nem publikált kísérlet.

- 1979: Magyarországon az első mesterségesen előállított komplex beszédszöveg: „*Boldog Újévet! Sok sikert és boldogságot az új esztendőben.*” – 1979. december 18–28-a között szintetizáltam Kiss Gábor programozó közreműködésével. Bemutattam 1979 decemberében az intézet igazgatóságának, továbbá 1980. január 25-én a VI. Russzisztikai Konferencián, és elhangzott ugyanakkor a Nyíregyházi Rádió Észak-tiszántúli Krónikájában.
- 1980: Szintetizált akusztikai stimulusok az alpfrekvencia és az intenzitás perceptiójának a vizsgálatára. – Nem publikált intonációs kísérlet.
- 1980: A megszólítás intonációs eszközei (analízis és szintézis). – Ismertettem 1980 júniusában A számítógép felhasználása a beszéd folyamat akusztikai elemzésében és szintézisében címmel tartott intézeti előadásomban.
- 1980: 2 x 200 minta Török Annamáriával és Szalóczy Pállal a magyar beszédintonáció vizsgálatára. E terjedelmes nyelvi anyag lehetőséget adott az intonáció elemzésére a modalitás, a grammatikai megszerkesztettség, a közlés érzelmi töltésének a változása és a szöveggörnyezet függvényében. – Nem publikált műhelymunka a korábbi részvizsgálatok összegezéséhez.
- 1980: Magyar–oros hangátültetés. A kísérlet célja, hogy rávilágítson a két nyelv artikulációs mechanizmusainak a sajátosságaira a hangzásban érzékelhető összeegyeztethetetlenség (inkompatibilitások) alapján. – Bemutattam 1980. október 1-én a Budapesti Nemzetközi Beszédakusztikai Szimpóziumon.
- 1980: Párbeszéd ember és gép között címmel – munkatársaim, Földi Éva, Kiss Gábor, Nikléczy Péter és Olasz Gábor közreműködésével – kísérletet végeztem az emberi és gépi beszéd élőszovegbeni összeszerkeszthetőségének a tanulmányozására. – Bemutattam 1980. december 9-én a fonetikai szimpóziumon, elhangzott a Rádiólexikonban és 1981 januárjában a Debreceni Russzisztikai Napokon. Teljes szövege így hangzott:

Ember – *Jó napot!*

Gép – *Jó napot!*

E – *Ki vagy te?*

G – *Én beszélőgép vagyok.*

E – *Nem hiszem el! Számolj tízig!*

G – *Egy, kettő, három, négy, öt, hat, hét, nyolc, kilenc, tíz.*

E – *Jól van. Mondd utánam: Atyámat gyászolom.*

G – *Atyámat gyászolom. Jó?*

E – *Nem jó. Így mond: Atyámat gyászolom.*

G – *Atyámat gyászolom. Jó?*

E – *Igen. Jól van. Hallottad? Az űrhajósok visszatértek a Földre.*

G – *Igen? Az űrhajósok visszatértek a Földre?*

E – *Vissza ám, már tegnap.*

G – *Igen, az űrhajósok visszatértek a Földre.*

E – *Hogy vagy?*

G – *Köszönöm, jól.*

E – *Sokat dolgoztok?*

G – *Igen, sokat dolgozunk.*

E — *No, akkor nem zavarlak. Szervusz, viszontlátásra!*

G — *Viszontlátásra! Minden jót!*

1980–81: A szupraszegmentális hangszerkezet változása a közléstartalom és a szövegszerkezet függvényében, azonos szegmentális szerkezet mellett. Ez az analízis és szintézis együttes alkalmazásával végzett kísérlet, amelyre „*Az űrhajósok visszatértek a Földre.*” mondattal emlékezünk, nyitotta meg a szegmentális és a szupraszegmentális szerkezetek mesterséges szétválasztását és összeillesztését, továbbá a hangteststruktúrák modellezhetőségét vizsgáló kísérlet sorozatunkat. — Részleges bemutatására a már említett intézeti előadás keretében került sor.

1981: A „Magyar hangszellettár” összeállítása. A munka célja a magyar beszédfolyamat legkisebb akusztikai építőelemeinek a szintéziséhez szükséges adatok összegyűjtése és rendszerezése. Ez az adatbázis volt az alapja a Voxon beszédsszintetizáló rendszernek.

1982: Az első magyar nyelvű folyamatos imperszonális beszédsszintetizáló rendszer. — Bemutatása 1982 márciusában történt meg.

Ez idő alatt munkatársaim közül Olasz Gábor először a zöngé egyéni jellemzőit vizsgálta, majd 1980-ban a magyar mássalhangzók alapszintézisét végezte, 1981-ben a számnevek kimondatására és matematikai műveletek végzésére alkalmas számítógépes rendszert hozott létre, s ez év májusában Univoice néven mutatta be a Kiss Gáborral közösen kifejlesztett automatikus beszédsszintetizáló eljárást. Kassai Ilona a beszédhangok időtartamviszonyainak, majd ezt követően a hangkapcsolódások szabályszerűségeinek és akusztikus folyamatainak a feltárása terén végzett értékes munkájával járult hozzá a beszédakusztikai kutatásainkhoz (Kassai 1979; MFF 8. 1981, 63–86). Gósy Mária a beszédelemek percepciójának a vizsgálatában ért el figyelemre érdemes eredményeket (ld. Gósy MFF 4. 1979, 119–36; MFF 8. 1981, 87–103). Vértes O. András — a már korábban hivatkozott publikációi mellett — a Bevezetés a magyar hangstilisztikába címmel készülő munkájának a befejezéséhez jutott közel. Szende Tamás munkásságában, az elméleti vizsgálódások mellett, várhatóan nagyobb teret kap majd a kísérleti fonetika.

Kutatócsoportunkban Olasz Gábor és részben Kiss Gábor vállalkozásával 1980-tól új profil bontakozott ki a nyelvészeti-fonetikai alap kutatások eredményeinek a gyakorlati hasznosítására és szélesebb körben való megismertetésére.

E vázlatos kutatástörténeti áttekintést azzal a reménnyel zárjuk, hogy a hazai fonetika fejlődése hosszabb ideig nem szakad meg, mint ahogyan évszázados történetében ez már annyiszor bekövetkezett.

Irodalom

A magyar hangtan válogatott bibliográfiája (–1970). Szerk. BOLLA Kálmán és MOLNÁR József. Budapest 1977.

BOLLA Kálmán: A fonetikai szerkezetek interlingvális egybevetéséről (Problémavázlat). MFF 5. 1980, 40–69.

BOLLA Kálmán: A leíró hangtan vázlata. in: Fejezetek a magyar leíró hangtanból. Szerk. BOLLA Kálmán. Budapest 1982, 13–24.

BOLLA Kálmán: A magyar magánhangzók akusztikai analízise és szintézise. MFF 1. 1978, 53–67.

BOLLA Kálmán: A nyelvészeti fonetika szakágazatai. NyK LXXVIII, 1976, 292–9.

- BOLLA Kálmán: Az orosz magánhangzók analízise és szintézise. MFF 4. 1979, 33–79.
- BOLLA Kálmán: Bevezetés. in: Fejezetek a magyar leíró hangtanból. Szerk. BOLLA Kálmán. Budapest 1982, 9–12.
- BOLLA Kálmán: Gondolatok az egyetemi fonetikaoktatás továbbfejlesztéséről. in: A magyar nyelv kutatásának és oktatásának kérdései. ELTE NyelvtDolg 6. 1971, 137–46.
- BOLLA Kálmán: Hozzászólás a fonetikai kutatások helyzetéhez. in: Általános nyelvészetünk helyzete. Az alkalmazott nyelvészet helyzete Magyarországon. Budapest 1969, 94–8.
- BOLLA Kálmán–MOLNÁR József: A hazai fonetikai kutatások történeti áttekintése és a mai helyzet. Budapest 1971.
- BÜKY Béla: Beszédkutatás, halláskutatás és ezek rokonterületei Magyarországon 1960–1969 között. ÁltNyT X, 1974, 281–311.
- GOMBOCZ Zoltán: A fonetika történetéből. Nyelvtudomány II, 1908; III, 1911.
- GÓSY Mária: Akusztikai paraméterek és nyelvi funkció a beszéddallam és a nyomaték percepciójában. MFF 4. 1979, 119–36.
- GÓSY Mária: A szegmentális hangszerkezet percepciójáról. MFF 8. 1981, 87–103.
- HEGEDŰS Lajos: A szájüreg saját hangjai magánhangzó-állásoknál. MNy XXX, 1934, 96–9.
- HEGEDŰS Lajos: Beszédtempó-elemzések. Nyr LXXXI, 1957, 223–7.
- HEGEDŰS Lajos: Magyar hanglejtésformák grafikus ábrázolása. Kísérletfonetikai tanulmány. A Bécsi Collegium Hungaricum Füzetei V, 1930.
- KASSAI Iлона: A magyar beszédhangok időtartamviszonyai. in: Fejezetek a magyar leíró hangtanból. Szerk. BOLLA Kálmán. Budapest 1982, 115–54.
- KASSAI Iлона: A magyar beszéd hangsorépítési szabályszerűségei. MFF 8. 1981, 63–86.
- KASSAI Iлона: Időtartam és kvantitás a magyar nyelvben. Budapest 1979. NytudÉrt 102.
- MAGDICS Klára: A beszéddallamtól a zenei dallamig. NyK LXV, 1963, 341–60.
- MAGDICS Klára: A magyar beszédhangok akusztikai szerkezete. Budapest 1965. NytudÉrt 49.
- MAGDICS Klára: A magyar beszédhangok időtartama. NyK LXVIII, 1966, 125–39.
- MAGDICS Klára: A magyar nyelvjárások összehasonlító hanglejtésvizsgálatának első tanulságai. MNy LX, 1964, 446–62.
- MAGDICS Klára: Szintetizált anyaggal végzett magyar magánhangzófelismerési kísérletek. NyK LXVII, 1965a, 335–78.
- MOLNÁR József: A magyar beszédhangok atlasza. Budapest 1970.
- MOLNÁR József: A magyar fonetika története. in: Tanulmányok a magyar és a finnugor nyelvtudomány történetéből. 1850–1920. Szerk. SZATHMÁRI István. Budapest, 1970, 39–45.
- OLASZY Gábor: Szintetizált magyar magánhangzók formáns-intenzitás és formáns-sáv szélesség értékei. MFF 1. 1978, 68–76.
- SZENDE Tamás: A beszéd folyamat alaptényezői. Budapest 1976.
- SZENDE Tamás: Fónagy Iván köszöntése. MFF 6. 1980, 172–8.
- VÉRTES O. András: A magyar leíró hangtan története az újgrammatikusokig. Budapest 1980.
- VÉRTES O. András: A magyar leíró hangtan története az újgrammatikusoktól 1945-ig. in: Fejezetek a magyar leíró hangtanból. Szerk. BOLLA Kálmán. Budapest 1982, 283–301.
- VICSI Klára: Tarnóczy Tamás köszöntése. MFF 6. 1980, 168–72.
- WODARZ-MAGDICS, K.: Experiments in Hungarian vowel recognition. Linguistics LVI, 1970, 64–86.

THE ANALYSIS AND SYNTHESIS OF THE ACOUSTIC STRUCTURE OF HUNGARIAN SPEECH

A historical survey of the research

Kálmán Bolla

1. The last 12 years were definitive in the history of Hungarian experimental phonetics in the sense that the results achieved in that period will determine the development of the field in the long run. Here we only deal with the most important results we achieved at the Phonetic Department of the Institute of Linguistics of the Hungarian Academy of Sciences, which is the centre of experimental phonetic research in Hungary.

a) We worked out a research programme with an extendable and improvable scope which helped us do away with our backwardness and may make it possible for us to catch up with the European standard. The more important features of the programme are the following:

- it deals with a complex linguistic–phonetic range of topics;
- it has a structural–functional methodology; it includes the use of objective and subjective methods complementing each other, and it combines experimental and theoretical research;
- it takes the results of interdisciplinary research into account;
- it relies on the traditions, but is open to integrate the results of the international phonetic studies; it facilitates the exchange of ideas and cooperation with phoneticians in other countries.

b) The original team which consisted of 4 members in 1971 has increased in number: new experts arrived, an electrical engineer, a technician, a programmer, and a laboratory assistant.

c) Between 1970 and 1978 we managed to renew the phonetic laboratory: we, in fact, built up a new laboratory system. With its high technical level, it can meet the requirements of the current linguistic phonetic studies.

d) In 1978 we put out the publication *Magyar Fonetikai Füzetek* (Hungarian Papers in Phonetics) which presents the results of our research.

e) Since 1977 we have organized an annual meeting where other institutions and organizations also take part to exchange ideas concerning phonetic research, the teaching of phonetics and its practical application.

f) We are cooperating with different universities, schools, hospitals and experts in order to put the results of the basic research into practice.

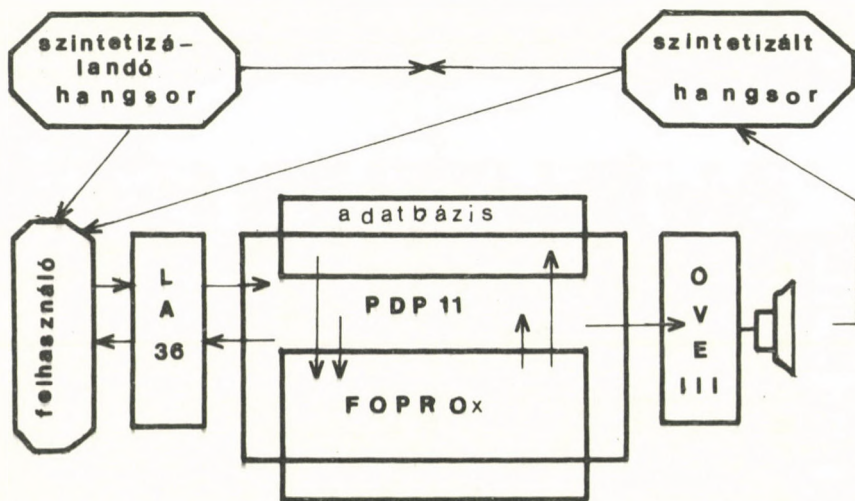
2. Two periods of the acoustic phonetic research are reviewed: a) the period before 1970, and b) the one beginning with 1970 and leading up to the present. Within the latter period we separately deal with the general development of linguistic phonetics (the scientific theory), the technical possibilities (hardware), computational technology (software), the methods of acoustic analysis and synthesis (technology), our results and the possibilities of future development.

INTERAKTÍV BESZÉDSZINTETIZÁLÓ RENDSZER SZÁMÍTÓGÉPPEL ÉS OVE III SZINTETIZÁTORRAL

Kiss Gábor—Olaszy Gábor

Dolgozatunkban bemutatjuk az MTA Nyelvtudományi Intézete fonetikai osztályának laboratóriumában kidolgozott INTERaktív BESzédszintetizáló Rendszert (INBERE). A rendszerrel jó minőségű (nemcsak magyar) hangsorokat lehet előállítani, elsősorban fonetikai alap kutatások számára, de a rendszerrel előállítható mesterséges hangsorok társadalmi és ipari alkalmazásának a lehetősége is meglehetősen széles.

A rendszer hardware és software részeit a következőkben ismertetjük (1. ábra).



1. ábra
Az INBERE architektúrája

Hardware részek

Rendszerünk a PDP 11/34-es számítógépből és OVE III (svéd gyártmányú) beszédszintetizátorból áll. A számítógép konfigurációjából felhasználjuk a beszédszintetizálás során a 64 Kbyte-os központi memóriát, a két floppy disk egységet és az LA-36-os konzol írógépet. A beszédszintetizátor kapcsolatát a számítógéppel 16 bites paralell interface biztosítja. A rendszer működésének megértéséhez és a mesterséges beszéd eredményes előállításához szükséges az OVE III alapos ismerete, ezért a következőkben először a beszédszintetizátort mutatjuk be.

Az OVE III beszédszintetizátor ismertetése

Az OVE III manuálisan és digitálisan vezérelhető analóg beszédszintetizátor. Segítségével beszédhangok és hangsorok állíthatók elő. A szintetizátor statiku-

Ssz.	Paraméter	címbitek				adatbitek								tartomány	lépés		funkció		
		11	10	9	8	7	6	5	4	3	2	1	0						
1.	AO	0	0	0	0												0–31 dB	0.5 dB	zöngés hang
2.	AC	0	0	0	1												0–30 dB	2 dB	zörejh hang
3.	AH	0	0	1	0												0–30 dB	2 dB	hehezetes hang
4.	AN	0	0	1	1												0–30 dB	2 dB	nazális hang
5.	F ₀	0	1	0	0												50–308 Hz	0.8 %	alapfrekvencia
6.	F ₁	0	1	0	1												200–1234 Hz	0.8 %	első formáns
7.	F ₂	0	1	1	0												504–3109 Hz	0.8 %	második formáns
8.	F ₃	0	1	1	1												1008–6218 Hz	0.8 %	harmadik formáns
9.	N ₁	1	0	0	0												200–1234 Hz	3 %	nazális formáns
10.	AK	1	0	0	1												0–31 dB	0.5 dB	antiformáns
11.	K ₁	1	0	1	0												1008–6218 Hz	3 %	zörejjóc
12.	K ₂	1	0	1	1												2540–15668 Hz	3 %	zörejjóc
13.	B ₁	1	1	0	0												12–188 Hz	12 Hz	F ₁ sávszélessége
14.	B ₂	1	1	0	1												31–470 Hz	31 Hz	F ₂ sávszélessége
15.	B ₃	1	1	1	0												250–750 Hz	250 Hz	F ₃ sávszélessége
		1	1	1	1														

2. ábra
Az OVE III beszéd szintetizátor paramétertáblázata

adatbitek	frekvencia (Hz)				
	F ₀	F ₁ ,N ₁	F ₂	F ₃ ,K ₁	K ₂
0000 0000	50	200	504	1008	2540
0000 0100	51	206	519	1037	2614
0000 1000	53	212	534	1068	2691
0000 1100	55	218	550	1099	2770
0001 0000	56	224	566	1131	2851
0001 0100	58	231	582	1165	2934
0001 1000	59	238	599	1199	3020
0001 1100	61	245	617	1234	3109
0010 0000	63	252	635	1270	3200
0010 0100	65	259	654	1307	3294
0010 1000	67	267	673	1345	3390
0010 1100	69	275	692	1385	3490
0011 0000	71	283	713	1425	3592
0011 0100	73	291	734	1467	3697
0011 1000	75	300	755	1510	3805
0011 1100	77	308	777	1554	3917
0100 0000	79	317	800	1600	4032
0100 0100	82	327	823	1647	4150
0100 1000	84	336	848	1695	4271
0100 1100	87	346	872	1745	4397
0101 0000	89	356	898	1796	4525
0101 0100	92	367	924	1849	4658
0101 1000	94	378	951	1903	4795
0101 1100	97	389	979	1958	4935
0110 0000	100	400	1008	2016	5080
0110 0100	103	412	1037	2075	5229
0110 1000	106	424	1068	2136	5382
0110 1100	109	436	1099	2198	5539
0111 0000	112	449	1131	2263	5702
0111 0100	116	462	1165	2329	5869
0111 1000	119	476	1199	2397	6041
0111 1100	122	490	1234	2468	6218
1000 0000	126	504	1270	2540	6400
1000 0100	130	519	1307	2614	6588
1000 1000	133	534	1345	2691	6781
1000 1100	137	550	1385	2770	6979
1001 0000	141	566	1425	2851	7184
1001 0100	146	582	1467	2934	7394
1001 1000	150	599	1510	3020	7611
1001 1100	154	617	1554	3109	7834
1010 0000	159	635	1600	3200	8064
1010 0100	163	654	1647	3294	8300
1010 1000	168	673	1695	3390	8543
1010 1100	173	692	1745	3490	8793
1011 0000	178	713	1796	3592	9051
1011 0100	183	734	1849	3697	9316
1011 1000	189	755	1903	3805	9589
1011 1100	194	777	1958	3917	9870
1100 0000	200	800	2016	4032	10159
1100 0100	206	823	2075	4150	10457
1100 1000	212	848	2136	4271	10764
1100 1100	218	872	2198	4397	11079
1101 0000	224	898	2263	4525	11404
1101 0100	231	924	2329	4658	11738
1101 1000	238	951	2397	4795	12082
1101 1100	245	979	2468	4935	12436
1110 0000	252	1008	2540	5080	12800
1110 0100	259	1037	2614	5229	13175
1110 1000	267	1068	2691	5382	13561
1110 1100	275	1099	2770	5539	13959
1111 0000	283	1131	2851	5702	14368
1111 0100	291	1165	2934	5869	14789
1111 1000	300	1199	3020	6041	15222
1111 1100	308	1234	3109	6218	15668

3. ábra
Az OVE III formáns- és zörejfrekvencia-táblázata

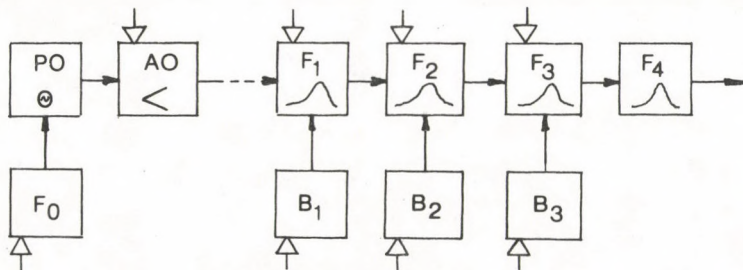
san külső kapcsolók, dinamikusan számítógép segítségével vezérelhető. (A készüléket Gunnar Fant és munkatársai tervezték, és a svéd FONEMA cég gyártja.)

Az OVE III a beszédhangok és hangsorok előállításához három alapáramkörrel: a zöngés (1), a nazális (2), és a zörej (3) áramkörrel, valamint a hehezetesítő (4) zörejforrással rendelkezik. A zöngés és nazális kört impulzusgenerátor (PO), a zörejkört zajgenerátor (CO) gerjeszti. A generátorokat, valamint az említett áramköröket és azok belső értékeit 16 paraméter segítségével – beleértve az időtartamot is – adhatjuk meg (2. ábra).

A frekvenciákat és az azoknak megfelelő bitkódokat a 3. ábra tartalmazza. A táblázatban az F_0, F_1, F_2, F_3 paraméterek esetében csak minden negyedik lehetséges frekvenciaérték van feltüntetve!

Ahhoz, hogy az OVE III beszéd szintetizátorral eredményesen és jó minőségben tudjunk hangsorokat szintetizálni, először is meg kell ismerkednünk az említett áramkörök működésével, gyakorlati méréseket kell végezni arra vonatkozóan, hogy az áramkörök milyen hatással vannak egymásra, az intenzitás-, a frekvencia- és az időviszonyok hogyan adhatók meg és hogyan kezelhetők digitálisan, számítógépprogram segítségével.

1. A zöngés hangot előállító négy pólusú, digitálisan vezérelhető szűrő. A szűrőt impulzusgenerátor jele gerjeszti. Az áramkör paramétereinek a meghatározására használt változók: AO (dB), F_0 (Hz), F_1 (Hz), F_2 (Hz), F_3 (Hz), B_1 (Hz), B_2 (Hz), B_3 (Hz) (4. ábra).



4. ábra

Az OVE III zöngés hangot előállító áramköre

A szűrő erősítését 30 dB-es átfogási sávban 0,5 dB-es lépésekben adhatjuk meg. Az alappfrekvencia F_0 -értéke, a szűrő három pólusának, vagyis az F_1, F_2, F_3 formáns helyének és sávzélességének értéke (B_1, B_2, B_3) változtatható. Ezeket a következők határok között választhatjuk meg:

$F_0 = 50 - 308$ Hz-ig 0,8 %-os lépésemelkedéssel 256 lépésben (alacsony fekvésű férfi- és női hangra);

vagy: $100 - 616$ Hz-ig 0,8 %-os lépésemelkedéssel 256 lépésben (magas fekvésű férfi-, női és gyermekhangra);

$F_1 = 200 - 1234$ Hz-ig 0,8 %-os lépésemelkedéssel 256 lépésben;

$F_2 = 504 - 3109$ Hz-ig 0,8 %-os lépésemelkedéssel 256 lépésben;

$F_3 = 1008 - 6218$ Hz-ig 0,8 %-os lépésemelkedéssel 256 lépésben

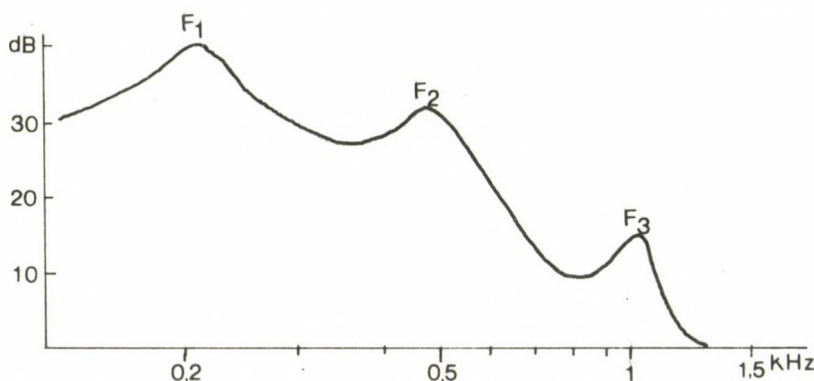
$B_1 = 12 - 188$ Hz-ig 12 Hz-es lépésemelkedéssel 16 lépésben;

$B_2 = 31\text{--}470\text{ Hz}$ -ig 31 Hz -es lépésemelkedéssel 16 lépésben;

$B_3 = 250\text{--}750\text{ Hz}$ -ig 250 Hz -es lépésemelkedéssel 4 lépésben.

Az F_4 formáns értéke állandó: 3500 Hz .

Ahhoz, hogy a magánhangzót előállító áramkör frekvencia- és intenzitásviszonyait jobban megértsük, meg kell ismerkednünk az említett, változtatható karakterisztikájú szűrő működésével. Ha az OVE III-mal magánhangzót vagy zöngés hangot akarunk előállítani, akkor az AO paramétert kell betáplálni. Ezzel az áramkör szűrőjének az alapkarakterisztikáját adjuk meg (5. ábra).



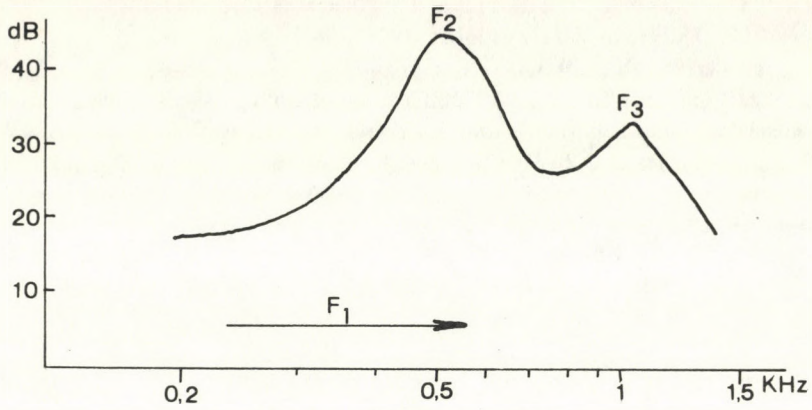
5. ábra

Az AO = 30 dB betáplálásával létrehozott szűrőkarakterisztika

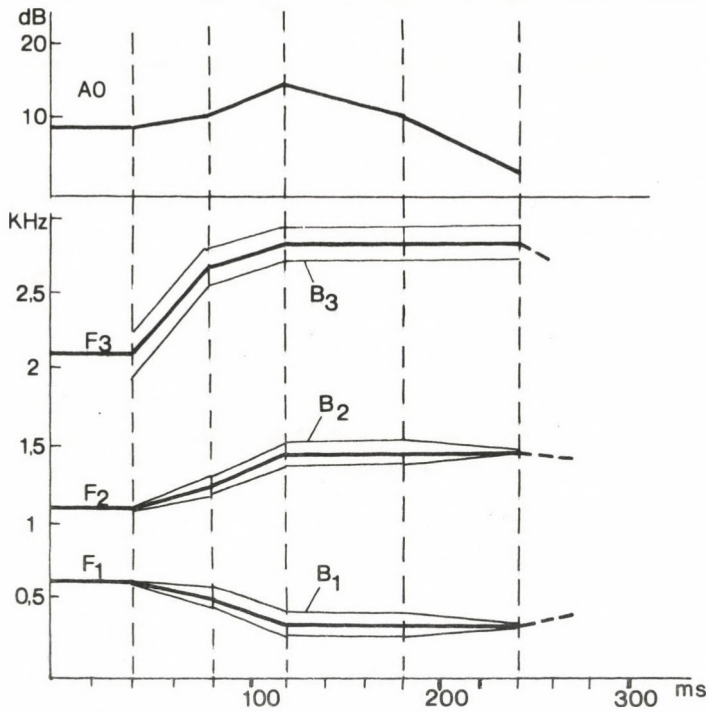
A görbén jelentkező csúcsok jelentik a formánsok legalacsonyabb frekvenciaértékeit. Az alapállásnál ezek a következők: $F_1 = 200\text{ Hz}$, $F_2 = 504\text{ Hz}$, $F_3 = 1008\text{ Hz}$. Ha az AO betáplálása után valamely zöngés hangra megállapított formánsértékeket is beadunk a szintetizátorba, akkor a beírt frekvenciaértékeknek megfelelően a görbe csúcsait eltoljuk a frekvenciatengelyen. Ha a tologatás folyamán a csúcsok közelebb kerülnek egymáshoz, akkor erősítik egymást és ellenkezőleg. Nézzük meg azt a példát, amikor az F_1 -et mozgatjuk a frekvenciatengelyen, az F_2 és az F_3 pedig az alapértékekre van beállítva. Az előbbiek szerint, ha az F_1 eléri az F_2 -t, akkor kettőjük együttes hatására intenzitáscsúcs keletkezik, majd, ha elhagyja az F_2 -t, az intenzitás fokozatosan csökken. Az F_3 viszonylatában hasonló jelenség zajlik le (6. ábra).

A beírt formánsok intenzitását külön-külön nem lehet szabályozni. A formánsintenzitásokat egy szűk határon belül (max. 12 dB) csökkenthetjük a sáv szélesség beadásával. Ha a formáns sáv szélessége nő, a mellette lévő formánshoz relatíve közelebb kerül, és így – a korábbiak szerint – intenzitásnövelő hatása van. A sáv szélességek használata a hangok finomabb árnyalatainak megvalósítására, az egyéni formánsamplitúdók kialakítására szolgál, de bizonyos hangoknál (pl. a [v]) a sáv szélesség a hang szerves építőköve is lehet. A formánsok és az intenzitások helyes csatolását az egyes hangszedetek között (vö. Olaszky MFF 8. 1981) jól meg lehet tervezni a 7. ábrán látható diagramtípus elkészítésével.

A fentiekből látható, hogy az OVE III beszéd szintetizátor magánhangzót előállító áramkörével létrehozott szintetizált hang vagy hangrész intenzitását nemcsak az AO



6. ábra
Az F_1 hatása az F_2 és F_3 intenzitására

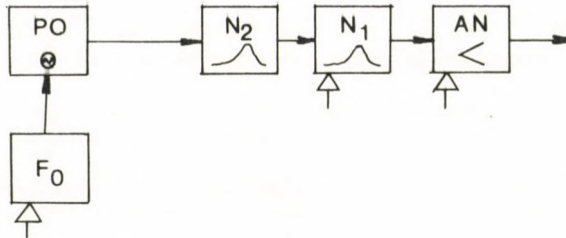


7. ábra
Az AO intenzitás, az F_1 , F_2 , F_3 és a B_1 , B_2 , B_3 mozgásának megtervezése a szintézishez

paraméter értéke határozza meg. A formánsok elhelyezkedése, egymáshoz való közelsége, valamint a sáv szélességek értéke is lényegesen befolyásolja a végleges intenzitást. A paramétereknek ezt a bonyolult egymásra hatását a szintézis során mindig hangsúlyo-

zattan figyelembe kell venni, ugyanis csak így sikerülhet a természetes ejtésű beszédhangok és hangsorok intenzitás szerkezetét a szintetizálás során jól megközelíteni.

2. A nazális áramkörrel az orrhangú beszédhangokat hozhatjuk létre. Felépítése olyan, mint a magánhangzó köré, csak két pólussal rendelkezik (8. ábra). A szűrőt ugyanazzal a PO impulzusgenerátorral gerjesztjük, mint a zöngés kört. Az áramkör paramétereinek meghatározására használt változók: AN (dB), N_1 (Hz).

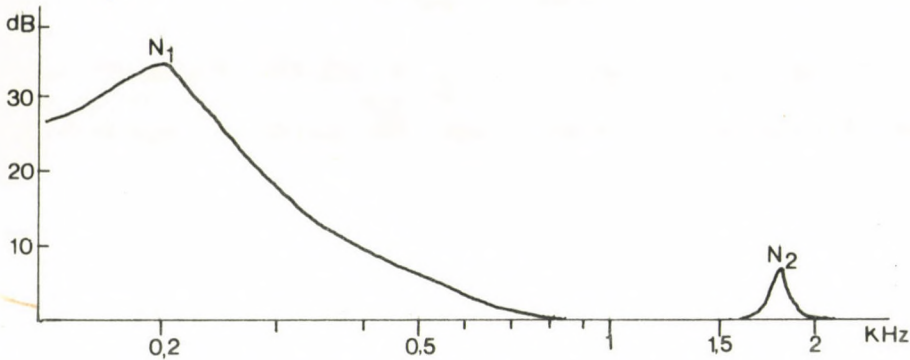


8. ábra

Az OVE III nazális hangot előállító áramköre

A nazális kör szűrőjének erősítését 30 dB-es átfogási sávban az AN-nel adhatjuk meg. Az erősítést 0,5 dB-es lépésekben lehet változtatni. A szűrő első pólusának a helye változtatható, értékeit a következő határok között választhatjuk meg.

$N_1 = 200-1234$ Hz-ig 3 %-os lépésemelkedéssel 64 lépésben. Az N_2 értéke állandó: 1800 Hz. Az AN változó betáplálásakor az AO-hoz hasonlóan a szűrő alapátviteli karakterisztikáját hozzuk létre (9. ábra).

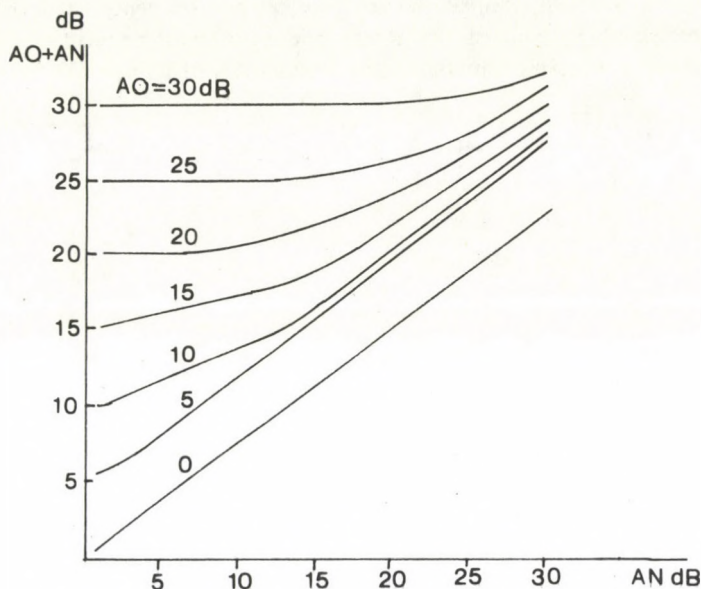


9. ábra

Az AN = 30 dB betáplálásával létrehozott szűrő karakterisztika

A karakterisztikából látható, hogy ha csak az AN paramétert adjuk be a szintetizátorba, akkor az N_1 alaphelyzetben (200 Hz) van. A karakterisztika 200 Hz-es intenzitásmaximumának értéke kb. 7 dB-lel kisebb, mint az AO körnél. Tehát az AN kör kisebb intenzitással vesz részt a beszédhangok megformálásában, mint az AO. Az N_1 formáns mozgatható ugyanazok a törvényszerűségek érvényesülnek a végleges hangintenzitás kialakulására vonatkozóan, mint a zöngés körnél. Ha az AN és az AO paramétereket együtt használjuk – ami gyakori eset nazalizált hangok létrehozásakor –, akkor figye-

lembe kell venni a két kör egymásra hatását az intenzitásviszonyok kialakításánál. Az AO és AN alapkarakterisztika esetén az eredő intenzitást a 10. ábra diagramjából olvashatjuk le.

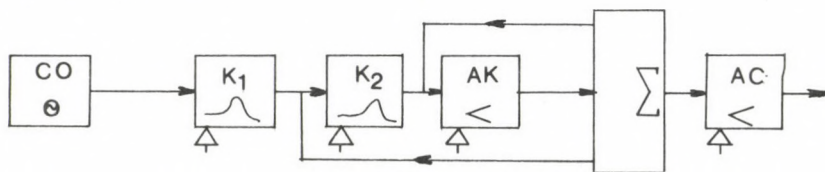


10. ábra

Az AN + AO eredő intenzitás kiszámítása

A fentiekből látható, hogy ha például egy vokális elemet és egy nazális elemet intenzitásban kapcsolni kell, pl. a *mama* szóban, akkor a folyamatos intenzitásmenetet olyan hangszekvenciákban valósítjuk meg, ahol az AO-t csökkentjük és ezzel egyidejűleg az AN-t növeljük.

3. A zörej áramkörrel a beszéd zörej- vagy zörejt is tartalmazó hangjait (komponenseit) hozhatjuk létre. Az áramkör alapját digitálisan vezérelhető, két pólussal és egy zérussal rendelkező szűrő alkotja. A szűrőt a CO zajgenerátor jele gerjeszti. Az áramkör paramétereinek a meghatározására használt változók: AC (dB), AK (dB), K_1 (Hz), K_2 (Hz) (11. ábra).



11. ábra

Az OVE III zörejhangot előállító áramköre

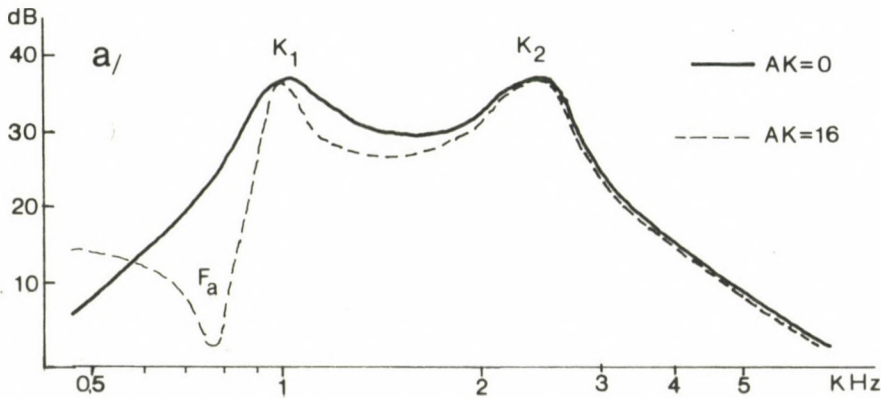
A szűrő erősítésének mértékét az AC változóval adhatjuk meg 30 dB-es átfogási sávban 2 dB-es lépésekben. A szűrő két pólusának helye változtatható, értéküket a következő határok között választhatjuk meg:

$K_1 = 1008\text{--}6218\text{ Hz}$ -ig 3 %-os lépésemelkedéssel 64 lépésben;

$K_2 = 2540\text{--}15668\text{ Hz}$ -ig 3 %-os lépésemelkedéssel 64 lépésben.

A K_1 - és K_2 -vel két zörejgócot hozhatunk létre egyidejűleg. Az AK változó a zérus (antiformáns: F_a) helyének meghatározására szolgál. G. Fant kutatási eredményei szerint a beszéd zörej- vagy zörejes hangjainak előállítására olyan szűrőt használhatunk, amely nemcsak pólusokkal (maximumhelyek), hanem egy zérussal (minimum) is rendelkezik. Az antiformánsra egyrészt azért van szükség, mert a spektrum alacsonyfrekvenciás részén a zörejangoknak nincs komponensük, másrésztől segítségével lehet a zörejangok frekvenciaszerkezetére jellemző szűrőkarakterisztikát létrehozni. Az AK erősítést illetően három helyzetet különböztetünk meg.

a) Ha az AK értéke $< 20\text{ dB}$, akkor az F_a mindig a K_1 -nél alacsonyabb frekvencián van. Ezt az állapotot általában a [s] és a [ʃ] típusú hangok előállításánál használjuk. Ha zörejangot akarunk előállítani, akkor először mindig az AC változót kell betáplálnunk az OVE III-ba. Ezzel a zörej áramkör szűrőjének az alapkarakterisztikáját adjuk meg. (12.a ábra). Az AK értéke ekkor nulla.



12.a ábra

Ennél a helyzetnél a $K_1 = 1008\text{ Hz}$, a $K_2 = 2540\text{ Hz}$. Ha az AC után meghatározott zörejgócokat is beírunk a K_1 - és K_2 -be, akkor a pólusokat eltoljuk a frekvenciatengelyen. Ha a pólusok közelednek egymáshoz, akkor erősítik egymást és ellenkezőleg. A K_1 és K_2 egymásra hatásának számításában négy esetet különböztetünk meg:

– ha $0,5 K_2 < K_1 < K_2$, és a K_1 -et közelítjük a K_2 -höz, akkor az intenzitásnöve-

kedést az $I\text{ (dB)} = \frac{K_1}{K_2} \times 10 + AC\text{ (dB)}$ képlettel számíthatjuk,

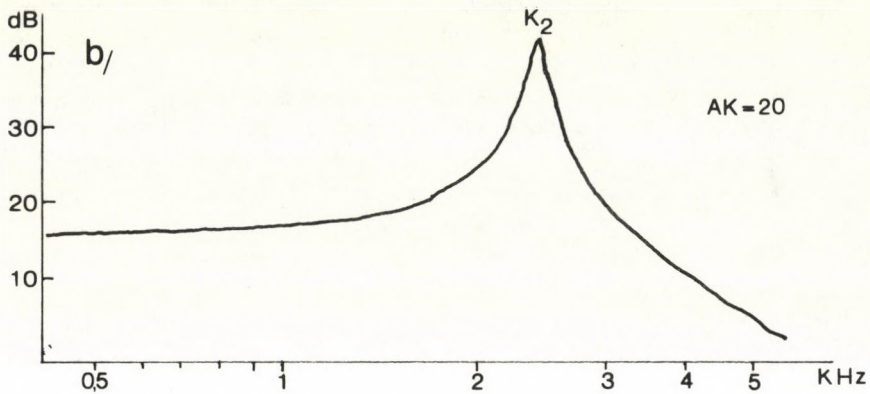
– ha a $K_1 = K_2$ -vel, akkor a kiemelés mértéke 10 dB ,

– ha a $K_1 > K_2$ -nél, akkor nagymérvű intenzitáscsökkenéssel kell számolni,

– ha a $K_1 \gg K_2$ -nél, akkor a csúcsok messze esnek egymástól, így az eredő intenzitás értéke egyenlő a beadott AC értékével.

Ha az AK változó dB-értékét nulláról növelni kezdjük, akkor az F_a egyre közelebb kerül a K_1 -hez (12.a ábra), és az $AK = 20\text{ dB}$ értéknél eléri azt.

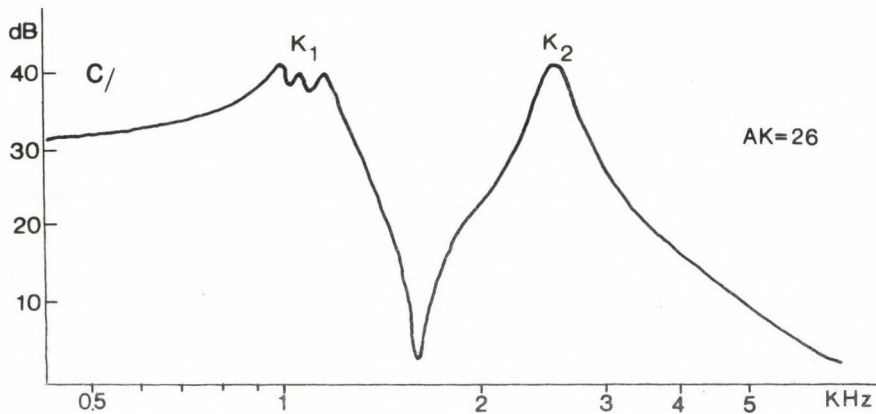
b) Ha az AK értéke egyenlő 20 dB-lel, akkor az F_a értéke megegyezik a K_1 értékével, így azt közömbösíti. Ekkor a szűrő karakterisztikájában csak a K_2 kiemelése játszik szerepet, az alsó sáv átvitele egyenletes (12.b ábra).



12.b ábra

Ezt az állapotot általában a [f] típusú hangok előállításánál használjuk.

c) Ha az $AK > 20$ dB, akkor az F_a a K_1 felett helyezkedik el a frekvenciatengelyen (12.c ábra).



12.c ábra

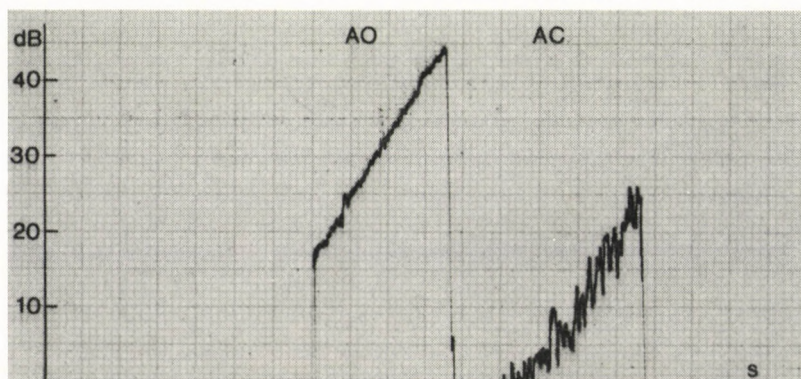
A zörejkör szűrőjének átviteli karakterisztikája különböző AK-értékekénél

Minél jobban növeljük az AK értékét, a 20 dB-től a 30 dB felé, annál messzebb lesz az antiformáns értéke a K_1 -től. Ezt a beállítást általában a [k] típusú hangok előállításánál használják.

Az antiformásra vonatkozó általános megállapítások a következők:

- az F_a frekvenciáját az $F_a = \frac{AK \times K_1}{20}$ (Hz) képlet adja, ha $30 > AK > 10$ dB (vö. 12.a ábra);
- ha a K_1 -et változtatjuk a frekvenciatengelyen, akkor az F_a is vele párhuzamosan fog mozogni, természetesen ugyanannál az állandó AK -értéknél;
- az antifórmásra tett megállapítások akkor is érvényesek, ha a K_2 frekvenciáját alacsonyabbra választjuk meg, mint a K_1 -ét.

Ha a zörejkört és a zöngés kört együtt használjuk – ami a zörejes zöngés réshangok szintetizálásánál elengedhetetlen –, ismerni kell a két áramkör egymásra hatását. Az AO erősebb intenzitást képvisel, mint az AC ugyanazon értéke (13. ábra).



13. ábra

A hangintenzitás, ha az AO-t és az AC-t 0-tól 30 dB-ig változtatjuk

Az AC zörejforrásának intenzitását csökkenteni kellett, hogy a magyarban előforduló hangsor végi, gyenge intenzitású zörejhangokat (pl. [t], [k]) megfelelő intenzitással tudjuk előállítani (az eredeti gyári beállítás túl magasnak bizonyult). Mivel a két áramkör szűrője (AO) egyrészt az alacsonyabb, másrészt (AC) a magasabb frekvenciákon engedi át a gerjesztő jelet, az eredő intenzitást többnyire a nagyobb intenzitással beadott szűrőkarakterisztika határozza meg.

4. A h e h e z e t e s í t ő á r a m k ö r alacsonyfrekvenciás zörejt (kb. 600 Hz-ig) állít elő, amit a korábbi áramkörökkel előállított bármely hanghoz hozzákeverhetünk. A zörej intenzitását az AH paraméterrel határozhatjuk meg. Értékét 30 dB-es sávon belül 2 dB-es lépésemelkedéssel adhatjuk meg.

Az OVE III fenti áramköreit és azok belső értékeit meghatározó paramétereket bármilyen kombinációban egyidejűleg is betáplálhatjuk a szintetizátorba. Így szinte végtelen variációs lehetőségünk van beszédhangtípusok előállítására.

Az OVE III digitális vezérlése

Mint már említettük, a szintetizátort 16 bites paralell „interface”-szel (DR-11K) kapcsoltuk össze a laboratóriumunkban működő PDP 11/34-es számítógéppel, vezérlése TTL logikai szintekkel történik. Az OVE III vezérlése 8 adatbit, 4 címbit, 1 „ready”,

1 „reply” és 1 törlőbit segítségével történik. Ezt a bitsorozatot a DR-11K kimenő tároló regiszteréből (output buffer register, OBR) kapja. A biteket az OBR-be a memóriából kell programutasítással beírni. Miután a program beírta a biteket az OBR-be, az interface kiadja a „kész” jelet, és az adatok a szintetizátor tároló regiszterébe (SZB) íródnak. A szintetizátor válaszzel jelzi, hogy az adatokat elfogadta. A bitsor beíródásával egy időben a szintetizátor megszólal. Ahhoz, hogy a DR 11K OBR-ébe a következő bitsorozatot be tudjuk írni, jelezni kell, hogy az előbbi bitsorozat továbbítása megtörtént. Erre szolgál az OBR 14. bitjének (interrupt enable bit) az 1-re állítása. Ha az OBR megkapta a 14. bite a logikai „igen” szintet (+5V), akkor képes új adatokat elfogadni, az előbbi folyamat kezdődhet előlről. A bitek kiosztása az OBR-ben a következő:

címbitek								adatbitek							
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

A hangokat és hangsorokat meghatározott időtartamú szeletekből építjük fel, amelyek mikrohangszeletekből állnak (vö. Olasz MFF 8. 1981). Egy ilyen mikrohangszelet akusztikai szerkezetének a felépítéséhez több paraméter megadására (pl. AO, F_0 , F_1 , F_2 , F_3 stb.) van szükség. Ezeket aktív paramétereknek hívjuk. Minden paramétert és annak belső értékét egy 16 bites szó reprezentál, amelyből csak 12 bitet használunk ki. A mikrohangszelet felépítéséhez szükséges paramétereket és értékeiket csak egymás után lehet az OVE III-ba „szavanként” betáplálni. Például egy [a:] hang építőköci a következők: AO, F_0 , F_1 , F_2 és F_3 . Ezeket egymást követő asszembler szintű programlépésekkel tápláljuk be a szintetizátorba. Egy ilyen lépés végrehajtása kb. 2 μ s ideig tart. Így egy mikrohangszeletet az építőköci számától függően kb. 50–60 μ s alatt lehet generálni. Ez az idő növekedhet, ha magasabb szintű programnyelven írt programmal működtetjük a szintetizátort. A következőkben nézzük meg, hogy az előbb említett kitarított hangzású [a:] hangot hogyan hozza létre a számítógép az input adatokból. Az input adatok negáltjai a következők:

frekvencia	digitális kód															
$F_0 = 79$ Hz	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1
$F_1 = 650$ Hz	1	1	1	1	1	0	1	0	0	1	0	1	1	0	1	1
$F_2 = 1250$ Hz	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1
$F_3 = 1950$ Hz	1	1	1	1	1	0	0	0	1	0	1	0	0	0	1	1

A programlépések, amelyekkel ezeket az adatokat sorba betápláljuk a szintetizátorba, a következők:

lépés	oktális kód	utasítás	az OBR tartalma	eredmény
1.	012737	mov # 177400, @ # 76774	AO	AO megszólal
2.	177400			
3.	76774			
4.	012737	mov # 040000, @ # 167770	14.	int. enab. bit
5.	040000		bit.	beírása
6.	167770			
7.	012737	mov # 175677, @ # 767774		
8.	175677		F ₀	AO+F ₀
9.	767774			szól
10.	012737	mov # 040000, @ # 167770	14.	int.
11.	040000		bit.	enab. bit
12.	167770			beírása
13.	012737	mov # 175133, @ # 767774		
14.	175133		F ₁	AO+F ₀ +F ₁
15.	767774			szól
16.	012737	mov # 040000, @ # 167770	14.	int.
17.	040000		bit.	enab. bit
18.	167770			beírása
19.	012737	mov # 174577, @ # 767774		
20.	174577		F ₂	AO+F ₀ + +F ₁ +F ₂
21.	767774			szól
22.	012737	mov # 040000, @ # 167770	14.	int.
23.	040000		bit.	enab. bit
24.	167770			beírása
25.	012737	mov # 174243, @ # 767774		
26.	174243		F ₃	AO+F ₀ +F ₁ + +F ₂ +F ₃
27.	767774			szól
28.	000000	halt		

Látható, hogy 28 programlépésre volt szükség az előbbi öt paraméter betáplálására. Ez egyetlen hangrészlet (mikrohangszelet) létrehozására volt példa. Mint már említettük, a hangokat és a hangsorokat is ilyen hangrészletek összekapcsolt sorozatával

kell létrehozunk. Ilyenkor a „halt” utasítás helyére kerül az órajel, ami a hangrészlet időtartamát határozza meg. Az órajel az időt méri, és csak a programban megadott időtartam eltelte után (pl. 10 ms) engedi meg, hogy az OBR-be a 14. bitet a program beírja, tehát a következő hangrészletre a bitsorozatot előkészítse. Hangsorok szintetizálása-
kor tehát az előbbi lépéssorozat zajlik le több százszor egymás után, mindig változó adatokkal.

Egyéb funkciók mellett a fenn leírt módon a szintetizátor digitális vezérlését a FOPROx számítógépprogram hajtja végre, amely az INBERE software magját képezi.

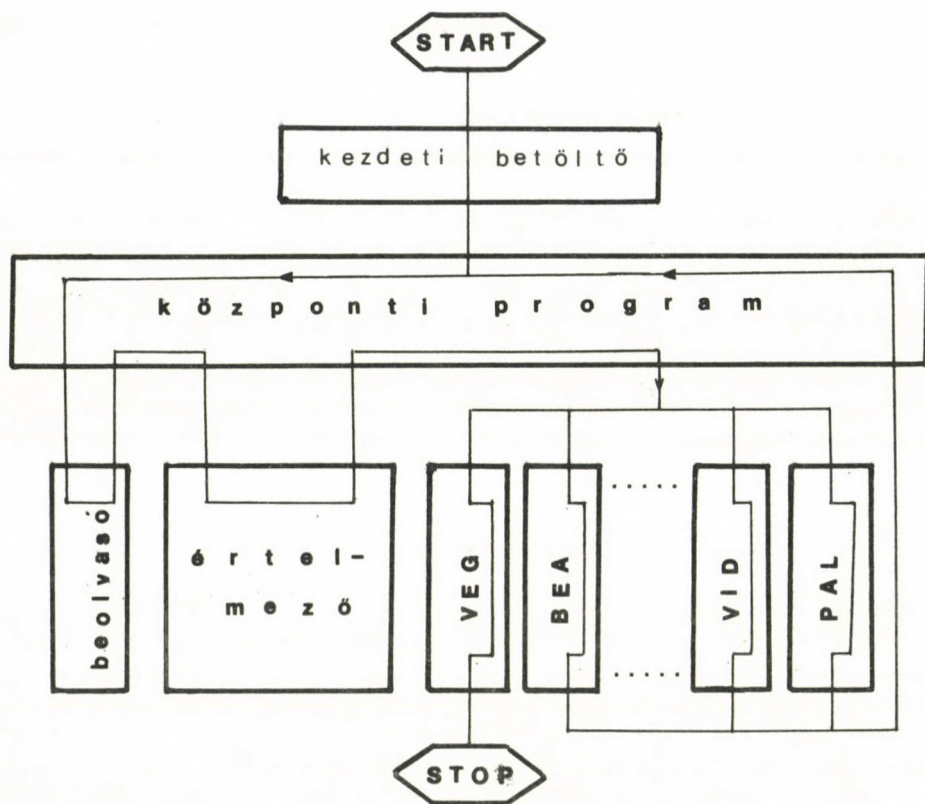
Software részek

A FOPROx program leírása

A beszédszintetizist a saját fejlesztésű FOPROx programmal végezzük. Az INBERE vezérlő részét, a FOPROx programot több lépcsőben fejlesztettük ki; egyrészt, mert a pontos felhasználói igények menet közben alakultak ki, másrészt magunk is a program építése során ismertünk fel újabb, jobb megoldási lehetőségeket. A FOPROx program – amelyet FORTRAN-IV PLUS nyelven, a FORTRAN EXTENSION felhasználásával írtunk meg – két nagyobb részből áll: 1. a parancsokat feldolgozó (interpretáló) és 2. a parancsokat végrehajtó részből. A programot a központi program mellett 32 subroutine alkotja, mintegy 1700 sor terjedelemben. Szűkös központi memória kapacitásunk miatt kénytelenek voltunk a programot OVERLAY technikával ellátni, amit elősegített annak strukturális felépítése (14. ábra).

Már a tervezés legelején célul tűztük ki azt, hogy a felhasználók részére könnyen elsajátítható és kényelmesen kezelhető rendszert adjunk. Még akkor is, ha a kényelmesebb rendszert nagy programozási pluszmunkával tudjuk létrehozni. Többek között ezért döntöttünk egy a felhasználó által – bár kötött formában –, de mégis tág határok között variálható utasítássorokkal vezérelhető rendszer kidolgozása mellett, egy „menü” rendszerben dolgozó programmal szemben (Nyman–Nyman 1978). Véleményünk szerint az a többletmunka, amely a programnak a parancssorokat értelmező, interpretáló részének a megírásához kellett – többszörösen megtérült a program kényelmes és gyors használata során.

A fent említett utasítássorokat egy ún. utasításnév vezeti be. Egy-egy utasításnév a felhasználó egy-egy kérésének szimbolikus jele. Jelenleg a rendszer 15 kérést fogad el, vagyis 15 fajta kéréssel fordulhatunk a beszédszintetizáló programhoz. Ezen kérések, utasítások segítségével lehet egyszeri futtatás során a max. 5 s-nyi mesterséges hangsor idő-, frekvencia- és intenzitásviszonyait hatékonyan változtatni, a floppy disken tárolt, a rendszer szerves részét alkotó adatbázist kezelni, karbantartani, és a mesterséges hangsorok előállítása során szükséges dokumentációs műveleteket végrehajtani. Az egyes kérések (utasítások) milyenségét a felhasználó a három karakteres utasításnevet követő paraméterlista révén határozza meg. Minden utasítás paraméterlistája meghatározott szintaktikai szabályok alapján épülhet fel. Amennyiben az utasítás nevében vagy annak paraméterlista részében a felhasználó szintaktikai hibát vét, ezt a rendszer hibaüzenet formájában közli. A hibaüzeneteket úgy határoztuk meg, hogy azok a rendszert kevésbé



14. ábra
A FOPROx program felépítése

ismerő felhasználó számára is didaktikus jelleggel bírjanak. A fent említett 15 utasítás alkotja a VEZER szimbolikus nyelvet.

A VEZER nyelv leírása

A VEZER nyelv bemutatása során először a leírásban felhasznált metanyelvi fogalmakat, majd a nyelvet alkotó utasításokat ismertetjük.

A leíráshoz felhasznált metanyelvi fogalmak

A karakter

A számítógép LA-36-os típusú – a szintetizáláskor használt – konzol írógépjén közvetlenül leüthető és megjeleníthető karakter, amely a következő lehet:

!, ", #, \$, %, &, ', (,)*, +, ,, -, ., /, 0, 1,
 2, 3, 4, 5, 6, 7, 8, 9, :, ;, <, =, >, ?, @, A, B,
 C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S,
 T, U, V, W, X, Y, Z, [, \,], ^, _ , ` , { , | , } , ~ .

A fenti 68 karaktert kiegészíti a szünet (space) karakter.

A parancssor

Maximum 130 karakterből álló karaktersorozat. A számítógép a ~ (tilde) karakter kinyomtatásával jelzi, hogy készen áll a következő parancssor beolvasására. A felhasználó ekkor kezdheti meg a parancssor begépelését, amelynek befejezését a „return” billentyű leütésével jelzi.

Az utasításnév

Három betűből álló karaktersorozat. Ez határozza meg a parancssorral igényelt műveletet. Az utasításnevet mindig szünet követi. Jelenleg a következő 15 utasításnév használatos:

BEA,EVA,ETI,LEO,MOR,JON,AGI,IDA,IMI,ZOL,TOM,VID,PAL,IVO,VEG.

A hangazonosító

Egy szelet azonosítója, jele. A hangszelet (vö. Olasz MF 8. 1981) egy hang vagy hangrész, amelyben a paramétereknek nincs töréspontja, vagyis minden (a hangszeleteket alkotó aktív) paraméter értéke egy számmal vagy számpárral (től-ig) megadható. A felhasználó állapítja meg a hangszeleteket, és látja el azonosítóval azokat a szintetizálendő hangsor előkészítése során. Az egyes hangszeletekre a szintetizálás során ezekkel az azonosítókkal hivatkozhatunk. Az azonosító 1-5 karakterből állhat (ld. a felsoroltakat, de nem használható a szünet, a vessző és a kerek zárójelpár karakter). A számítógép központi memóriájában max. 150 szelet foglalhat helyet és a központi memóriában lévő szeletek összejeje nem haladhatja meg az 5000 ms-ot. Példák az azonosítók helyes megadásra: 120"1,A** ,1X,&,AZ%ZA

A # jel

Több hangszelet közös azonosítója, amely a # (hasch mark) karakterrel kezdődik és további két tetszőleges karakterből áll. A központi memóriában egyidejűleg max. 20 # jel helyezkedhet el. Példák a helyesen megadott # jelekre:

#11,#AB,#X%

A paraméter

A szintetizátor működését meghatározó 16 paraméternek a szimbolikus neve, amelyek két karakterből állnak. A paramétert mindig kettőspont követi. A nagytömegű adatbevitel megkönnyítése érdekében a számítógép elfogadja a paramétert követő karakternek a pontosvesszőt is. Ezt a módosítást azért tettük, mert a pontosvessző karakter beütéséhez nem szükséges a „shift” billentyű egyidejű lenyomása. Az érvényes paraméterek a következők:

TM,A0,F0,F1,F2,F3,B1,B2,B3,AN,N1,N2,AC,K1,K2,AK,AH.

Az érték

Pozitív egész szám vagy számpár lehet érték. Az értéként szereplő szám vagy számok nem lehetnek nagyobbak 32767-nél. Számpár esetén a két számot kötőjel (-) karakter választja el egymástól. Az érték mindig egymás felé nyitott csúcsos zárójelpárban: <> helyezkedik el. Példák helyesen megadott értékekre:

0-12,1222,896-1038

A szögletes zárójelpár ([])

A szögletes zárójelpárban lévő szintaktikai egység n-szer fakultatíven megismételhető: ([]n).

Az egyes utasítások

Az utasításokat a következő csoportokban ismertetjük:

- I. Adatbevivő utasítás: BEA
- II. Szeletszerkesztő utasítások: EVA, ETI
- III. Megszóllaltató utasítás: LEO
- IV. Szeletmozgató (adatbázis-kezelő) utasítások: MOR, JON, AGI
- V. Időmanipuláló utasítások: IDA, IMI
- VI. Dokumentáló utasítások: ZOL, TOM, PAL, VID, IVO
- VII. Megállító utasítás: VEG

Az egyes utasítások leírása az alábbi szempontok szerint történik:

- a) az utasítás szitaxisa,
- b) az utasítás szemantikája,
- c) példák az utasítás helyes használatára,
- d) az utasítás helytelen használata után megjelenő hibüzenetek.

I. Az adatbevivő utasítás

a) BEA azonosító (paraméter: <érték> [, paraméter: <érték>] n)
 b) Az utasítás segítségével lehet új hangszeletet definiálni, vagy egy meglévő hangszelet paramétereit változtatni, javítani. Az utasítás használatában figyelni kell arra, hogy új szelet definíciója esetén az első paraméternek mindig a TM-nek kell lennie, a már definiált szeletnek nem lehet a TM-értékét a BEA utasítással változtatni (ld. IDA utasítás). Az n értékének az utasítással megengedett hossza szab határt.

- c) ~BEA 1A1(TM:<50>,F0:<112-132>,F1:<654-800>,F2:<1131-1234>)
- ~BEA 1A1(F3:<2075-2263>)
- ~BEA 1A2(TM:<40>,A0:<12-16>,F0:<132>,F1:<800>,F2:<1234-1307>)
- ~BEA 1A2(F3:<2137>)
- ~BEA 1A1(A0:<1-12>)
- ~BEA PAUSA(TM:<250>)

- d) ~BEA 1A1(TM:<150>)
 **HIBA! **MAR DEF. SZELETNEK NEM LEHET UJ "TM:"JE.
 ~BEA 1A3(A0:<16-1>,F0:<132-122>,F1:<80-873>)
 **HIBA! ** UJ SZELET ESETEN ELSO A "TM:" !
 ~BEA 1A2(FQ:<132-125>)
 **HIBA! ** AZ FQ: KOMP. HIBAS !
 ~BEA 1A2(A0:<12-32>)
 **HIBA! ** A0 > 31 !
 ~BEA 1A2(F0:<45-132>)
 **HIBA! ** F0 < 50 !
 ~BEA 1A3**TM:<120>)
 **HIBA! ** A JEL MEGADASA NEM HELYES !
 ~BEA E5(TM:<120>0
 **HIBA! ** EGY VESSZO" HIANYZIK !

II. Szeletszerkesztő utasítások

1. Az EVA utasítás

- a) EVA (azonosító 1) azonosító 2
 b) Az azonosító 1-gyel jelölt szelet elnevezése azonosító 2-re. Az addig azonosító 1 nevű szeletre ezután azonosító 2 névvel hivatkozhatunk. Gyakran használatos utasítás, ugyanis sem az operatív, sem az adatbázisban nem lehet két egyforma azonosítójú szelet.

c) ~EVA (A3)1A3

d) ~EVA (S1) 1S1

**HIBA! ** A NEVVALTOZTATAS ROSSZ !

~EVA S1(1S1

**HIBA! ** A 4-5 POZICIO HIBAS AZ UT.-BAN!

2. Az ETI utasítás

- a) ETI (azonosító [, azonosító] n) #jel
 b) A kerek zárójelpár között felsorolt szeletek összekapcsolása, hogy ezután a #jellel hivatkozhatunk együttesen rájuk. Az n értékének az utasítás hossza szab határt. Megjegyzés: a felsorolt szeletek között egy szelet többször is előfordulhat!

c) ~ETI (1,2,15,24,24,33,39,40,41)#U3

~ETI (1,2,6,15,24,33,39,6,15,24,49)#U4

d) ~ETI (A1,B2,C3,D4,E5)XY

**HIBA! ** A JELEK OSSZEVONASA HELYTELEN !

III. Megszóltató utasítás

a) LEO (azonosító [, azonosító]_{n1})[I]

LEO (#jel [, #jel]_{n2})[I]

- b) A kijelölt szeletek egymás utáni megszóltatása, lejátszása. Egy utasítással egymás után maximum 300 szelet szóltatható meg, tehát: $1 \leq n_1 \leq 300$; az n_2 értéket pedig #jelekkel összevont szeletek száma határozza meg, itt sem lehet a lejátszandó sze-

letszám 300-nál több. A lejátszandó szeletek számának másrésről határt szab az a tény, hogy a parancssor nem lehet 130 karakternél hosszabb. Az I használata esetében a hangsorból csupán a szupraszegmentális szerkezet lejátszására kerül sor.

- c) ~LEO (1A1,1A2,1A3,1S1*,1S2*,1S3*,T1,T2,T3)
 ~LEO (#AT,#AT,#AT)
 ~LEO (#AA,#BB,#CC,#DD,#EE)I
- d) ~LEO (A11,A22,A33)
 HIBA! A11,A NEM DEF. SZELET!
 ~LEO (#AT,#AT,#AT,#AT,#AT,#AT,#AT,#AT)
 HIBA! EZ TOBB MINT 300 SZELET!

IV. Szeletmozgató, adatbázis-kezelő utasítások

1. A MOR utasítás

- a) MOR (azonosító [,azonosító] n_1)

MOR (#jel [, #jel] n_2)

b) Az operatív tárban lévő és az utasításban felsorolt szeletek eltávolítása a floppy disken elhelyezkedő adatbázisban. Az n_1 és n_2 lehetséges értéke megegyezik a LEO utasításnál megadottal. Ügyelnünk kell arra, hogy:

- a szeletek között nem lehetnek azonosak;
- az adatbázisba csak a már ott meglévő azonosítóktól különböző azonosítóval jelölt szeletet lehet elhelyezni;
- adott adatbázisban 2048 szelet helyezkedhet el, valamint az # jelek száma sem haladhatja meg az 512-t.

- c) ~MOR (A1*,A2*,A3*,1S11,2S12,3S13,T10)
 ~MOR (#PA)

- d) ~MOR (A1,A2,A3)

HIBA! SAJNOS >> A1 << SZELET MAR VAN A LEMEZEN!

2. A JON utasítás

- a) JON (azonosító [,azonosító] n_1)

JON (#jel [, #jel] n_2)

b) Az azonosítójukkal felsorolt szeletek felhozatala az operatív tárba az adatbázisból. Az operatív tárba csak a már ott lévő azonosítóktól különböző azonosítójával jelölt szeletet lehet felhozni az adatbázisból. Az n_1 , n_2 értéke megegyezik a LEO utasításnál megadott értékkel.

- c) ~JON (A3,S1,S2,S3)
 ~JON (#A1,#EE,#12)

- d) ~JON (T11,T12,T13)

HIBA! SAJNOS A >> T11 << SZELETET NEM TALALOM!

~JON.(A3,S1,S2,S3)

HIBA! SAJNOS >> A3 << SZELET MAR VAN A GEPBEN!

3. AGI utasítás

- a) AGI (azonosító [,azonosító] n)

b) Az azonosítójukkal felsorolt szeletek javított, módosított változatának ismételt

elhelyezése az adatbázisban. Az utasítás sikeres végrehajtásának feltétele, hogy az adatbázisban már legyen a megadott azonosítóval megnevezett szelet. Az $n \leq 300$.

c) \sim AGI (ELSO,MASOD,HARM)

\sim AGI (72,13)

d) \sim AGI(17,85)

****HIBA!**** A 4-5 POZICIO HIBAS!

\sim AGI (A1)

****HIBA!**** SAJNOS > A1 < SZ.NINCS A LEMEZEN!

V. Az időmanipuláló utasítások

1. Az IDA utasítás

a) IDA (azonosító [,azonosító] n_1) <érték>

IDA (#jel [, #jel] n_2) <érték>

b) A kijelölt szeletek idejének a változtatása. Az egyes szeletek idejét a következő képlet határozza meg: a szelet időtartama = $\frac{\text{a szelet TM-értéke}}{10}$ x a szelet IDA-értéke.

Ha az IDA utasításban az érték kisebb, mint 10, a szelet rövidebb ideig, ha az érték nagyobb, mint 10, a szelet hosszabb ideig hallható a BEA utasításban meghatározott értéknél. Az n_1, n_2 értékét lásd a LEO utasításnál. Minden szelet definiálásakor a szelet IDA-értékét a gép automatikusan 10-ben határozza meg.

c) \sim IDA (A1,A3,T30)<5>

\sim IDA (#PA)<50>

d) \sim IDA(#PA)45

****HIBA!**** 45 ERTEK ROSSZ !

\sim IDA(A1,A3)<5>

****HIBA!**** AZ IDA UT. HELYTELEN !

2. Az IMI utasítás

a) IMI <érték>

b) Az IMI utasítással lehet közvetett úton változtatni, alakítani a szelet, hang vagy hangsorok lejátszási időtartamát, mivel ennek az utasításnak a segítségével gyakorlatilag a számítógép órájának gyorsaságát változtatjuk. Az egyes szeletek lejátszási időtartamát (hosszúságát) a következő összefüggés fejezi ki:

$$\text{a szelet lejátszási időtartama} = \frac{\text{IMI-érték}}{100} \times \frac{\text{a szelet TM-értéke}}{10} \times \text{a szelet IDA-értéke}$$

A számítógépben mindaddig, amíg meg nem változtatják, az IMI-érték = 100-zal. Ebből látható, hogy ha az IMI utasításban az érték kisebb, mint 100, a lejátszási idő rövidül, ha pedig nagyobb az IMI-érték, mint 100, a lejátszási idő nő. Hangsorok megszólaltatása esetén a hangsort alkotó szeletek időtartama egyforma arányban változik.

c) \sim IMI <25>

\sim IMI <5000>

d) ~IMI<23>

****HIBA!**** AZ IMI UT. HELYTELEN!

VI. Dokumentáló utasítások

1. A ZOL utasítás

a) ZOL (azonosító [, azonosító] n_1) [@]

ZOL (#jel [, #jel] n_2) [@]

b) A felsorolt szeletek aktív paramétereinek táblázatos formában való kiírása. A @ használata esetében a táblázat fejléc nélkül készül. Az n_1 , n_2 értékére vonatkozó előírásokat lásd a LEO utasításnál.

c) ~ZOL (S11,2S12,3S13)

~ZOL (#AS)

~ZOL (A1,A2,A3)@

d) ~ZOL (Z,X,V)

****HIBA!**** Z,X,V NEM DEF. SZELET !

2. A TOM utasítás

a) TOM <érték>

b) Az adatbázis tartalomjegyzékének listaszerű kiírása. Az értékkel kell megadni, hogy az adatbázis hányadik elemétől hányadik eleméig történjen a tartalomjegyzék kiírása. Amennyiben az értékben a második szám értéke nagyobb, mint az adatbázis utolsó elemének a sorszáma, a listázás automatikusan megáll az utolsó elemnél.

c) ~TOM <25-34>

d) ~TON <4-6>

****HIBA!**** UT.-NEVE ROSSZ !

3. A PAL utasítás

a) PAL TETSZŐLEGES KARAKTERSOROZAT

b) A PAL utasítással megjegyzés, kommentár iktatható be a szintetizálásunk menétét dokumentáló lepreollóra. A negyedik pozíción az utasításorban itt is szünetnek kell állnia!

c) ~PAL MAGNETOFONFELVETEL AZ 1981/2-ES TEKERCSEN!

d) –

4. A VID utasítás

a) VID [@]

b) Az utasítás hatására a számítógép felsorolja az operatív tárban elhelyezkedő szeleteket, kiírja azok beadott TM-értékét, valamint a szeletek IDA-értékét. Ennek befejeztével az összevont szeletek, a # jelek csoportjainak kinyomtatása következik. A @ használata esetében csak az összevont szeletek csoportjait írja ki a számítógép.

c) ~VID

~VID @

d) –

5. Az IVO utasítás

a) IVO (azonosító [, azonosító] n_1) <érték> [N]

IVO (#jel [, #jel] n_2) (<érték> [N])

b) Az utasítás hatására a kijelölt szeletek AO és Fo értékeinek a menete kerül

kirajzolásra az LA-36-os konzol írógépen. Az érték itt az időtengely léptékét adja meg, míg az N használata során az F_0 frekvenciatengelyén a kalibráció nem 50–308 Hz-es, hanem 100–616 Hz-es intervallumban készül el. Az n_1, n_2 értéke megegyezik a LEO utasításban ismertetettel.

- c) \sim IVO (A \sim ,B \sim ,C \sim ,D \sim)<10>
 \sim IVO (A \sim ,B \sim ,C \sim ,D \sim)<20>N
 \sim IVO (#W%,#Q%)<10>
- d) \sim IVO (R,S,T)N
 **HIBA! ** N ERTEK ROSSZ !
 \sim IVO (R,S,T)
 **HIBA! ** ERTEK ROSSZ !
 \sim IVO (#RT,#RT,#RT,#TT)<12>
 **HIBA! ** #TT NEM DEF. ETI !

VII. Megállító utasítás

a) VEG

b) A VEG utasítás hatására a program futása befejeződik, és a következő üzenet jelenik meg:

REMELEM MUNKAJA EREDMENYES VOLT! VISZLAT:PDP11!

c) –

d) –

Mint már fentebb említettük, a FOPROx program strukturált architektúrájú. Ez lehetővé teszi számunkra, hogy akár utólag is kevés munkával módosítsuk; a felhasználó igényeinek megfelelően lássunk el kiegészítő funkcióval egy-egy utasítást. (Mint például a LEO utasítás esetében az I segédjel.) Sőt az ilyen felépítése a programnak megengedi, hogy könnyen bővíthessük, vagyis a FOPROx programhoz kapcsolhassunk újabb utasításokat feldolgozó alprogramokat. Jelenleg szupraszegmentális kutatást hatékonyan segítő programrészek fejlesztése folyik.

A következőkben bemutatunk egy felhasználói vezérnyelvi programlistát, amely az *És?* hangsor előállításának, lejátszásának, dokumentálásának és tárolásának menetét tartalmazza (15. ábra).

Eredmények, felhasználás

A dolgozatunkban ismertetett INBERE az első olyan számítógépes szintetizáló rendszer Magyarországon, amellyel mind a nyelvészeti-fonetikai kutatások, mind pedig a szintetizált beszéd egyéb felhasználására magyar (illetőleg bármilyen nyelvű) szavakat, mondatokat lehet létrehozni. A FOPROx programmal a hangsorok akusztikai szerkezetének bármely alkotóelemét rugalmasan lehet változtatni, így a rendszer ideálisan alkalmazható az analízis szintézissel (analysis by synthesis) eljárásra.

Az MTA Nyelvtudományi Intézetének fonetikai osztályán a magyar és az orosz beszédhangok, hangsorok akusztikai szerkezetének kutatásához a szintetizálási munkákat 1980 óta ezzel a rendszerrel végezzük.

.RUN DX1:FOPR01

*PAL AZ *E'S* HANGSOR ELO'ALLI'TA'SA

*PAL AZ ADATBA'ZISBO'LAZ E'1QQ,E'2QQ SZELETEK FELHOZATALA A KOIZPONTI TA'RBA.

*JON (E'1QQ,E'2QQ)

*PAL A KOIZPONTI TA'R TARTALMA'NAK KII'RATA'SA

*VID

JELEK	POZICIO	IDOALAP
E'1QQ	1- 6	10
E'2QQ	7- 13	10

OSSZEVONT JELEK: ---

*PAL A FENTI KE'T SZELET ATNEVEZE'SE

*EVA (E'1QQ)E'1

*EVA (E'2QQ)E'2

*PAL AZ *E'S* HANGSORT ALKOTO' TOVA'BBI SZELETEK LE'TREHOZA'SA

*BEA E'S(TM:<60>,A0:<14-6>,F0:<100>,F11<400-300>,F21<2000>,F31<2700>)

*BEA S1(TM:<80>,AC1:<15-20>,K1:<1900-1800>,K2:<3000>)

*BEA S2(TM:<90>,AC1:<20-5>,F\F\K11<1800>,K21<3000>)

*PAL AZ *E'S* HANGSORT ALKOTO' SZELETEK KO:ZOS #ES AZONOSI'TO' ALA' VALO' OISSZEVO'NA'SA

*ETI (E'1,E'2,E'S,S1,S2)#ES

*PAL SZUINET,NE'MA SZELET LE'TREHOZA'SA

*BEA PAUSA(TM:<300>)

*ETI (PAUSA,PAUSA)#PA

*VID

JELEK	POZICIO	IDOALAP
E'1	1- 6	10
E'2	7- 13	10
E'S	14- 19	10
S1	20- 27	10
S2	28- 36	10
PAUSA	37- 86	10

OSSZEVONT JELEK:

#ES 5 DB =E'1 ,E'2 ,E'S ,S1 ,S2 ,
#PA 2 DB =PAUSA,PAUSA,

*PAL AZ E'S HANGSORRA ME'ROD' DALLAM FELVITELE

*BEA E'1(F0:<110-118>)

*BEA E'2(F0:<118-125>)

*BEA E'9(F0:<125-135>)

*PAL AZ *E'S* HANGSOR HA'ROMSZORI LEJA'TSZA'SA,KO:ZBEN 1SEC SZUINETEKEL

*LEO (#ES,#PP,#ES,#PP,#ES)

HIBA! #PP NEM DEF. ETI !

*LEO (#ES,#PA,#ES,#PA,#ES)

*PAL AZ E'1,E'2 SZELETEK IDEJE'NEK MA'SFE'LSZERESRE NYU'JTA'SA

*IDA (E'1,E'2)<15>

*LEO (#ES,#PA,#ES,#PA,#ES)

*PAL AZ *E'S* HANGSOR EGE'SZE'NEK O:TSZO:RO:S IDO' ALATT VALO' LEJA'R\A'TSZA'SA

*IMI <500>

*LEO (#ES,#PA,#ES,#PA,#ES)

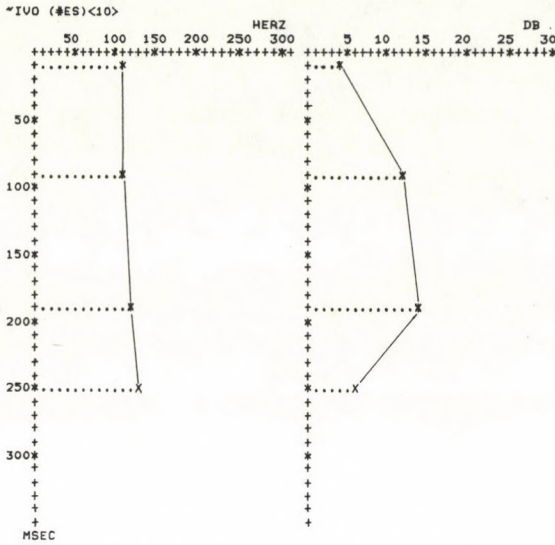
*IMI <100>

*PAL A HANGSOR ALAPHANG E'S INTENZITA'SMENETE'NEK KIRAJZOLTATA'SA

*IVO (#ES)10

HIBA! 10 ERTEK ROSSZ !

15.a ábra Felhasználói vezényelvi programlista



*PAL AZ 'E'S?' HANGSORT ALKOTO' SZELETEK PARAMETEREINEK A KII'RATA'SA

*ZOL (#ES)

KODSZAM:.....

SZINTETIZALASI ADATLAP

DATUM:.....

* AZD	IDO*	IDA	+ AD	FD	AN	+ N1	F1	B1	F2	B2	F3	B3	+ AC	AH	AK	K1	K2
* E'1	60	15		4-	112-		400		2016		2691						
				-12	-119												
* E'2	70	15		12-	119-		400		2016		2691						
				-14	-126												
* E'S	60	10		14-	126-		400-		2016		2691						
				-6	-137		-300										
* S1	80	10											14-			1903-	3020
													-20			-1796	
* S2	90	10											20-			1796	3020
													-4				

*PAL A LE'TREHOZOTT HANGSOR ELHELVEZE'SE AZ ADATBA'ZISBAN

*MOR (#ES)

**HIBA! ** UT.-NEVE ROSSZ !

*MOR (#ES)

*PAL AZ ADATBA'ZIS TARTALOMJEGYZE'KE'NEK A KURRENS RE'SZE'NEK A KII'RATA'SA

*TOM <102-110>

SORSZ SZELET

102. E'100

103. E'200

104. E'1

105. E'2

106. E'S

107. S1

108. S2

#41 42- 61

#61 62- 81

#81 82- 101

#ES 104- 108

*PAL A FDPRQ1 PROGRAM FUTA'SA'NAK FELFU:GGESZTE'SE

*VEG

REMELEM MUNKAJA EREDMENYES VOLT I VISZLAT: PDP11

STOP --

15.b ábra
Felhasználói vezényelvi programlista

Irodalom

AB FONEMA: OVE III c Speech Synthesizer Manual. Type 21001.

FANT, G.: Speech sounds and features. Massachusetts 1973.

NYMAN, M.—NYMAN, G.: OVE III B-Puhesyntetisaattorin ohjausohjelman käyttöopas. Helsinki 1978.

OLASZY Gábor: Hangsorok számítógépes formánszintézisének előkészítése. MFF 8. 1981, 147–60.

AN INTERACTIVE SPEECH SYNTHESIZING SYSTEM WITH COMPUTER AND OVE III SYNTHESIZER

Gábor Kiss—Gábor Olasz

In the present paper the authors present the interactive speech synthesizing system developed at the Phonetics Department of the Institute of Linguistics of the Hungarian Academy of Sciences. The system is able to produce sound sequences of good quality mainly for the purpose of phonetic research.

The system's hardware consists of a PDP 11/34 computer and an OVE III speech synthesizer (made by FONEMA Sweden). The 64 K byte central memory, two floppy disk units and an La-36 console typewriter are used of the configuration of the computer in the production of synthetic speech.

The paper begins with the description of the OVE III speech synthesizer and its operation. The authors describe the three basic circuits of the speech synthesizer: the vowel (1), the nasal (2) and the noise (3) circuits; and the 16 parameters controlling the circuits. These parameters are the following: TM (ms), AO (dB), FO (Hz), F1 (Hz), F2 (Hz), F3 (Hz), B1 (Hz), B2 (Hz), B3 (Hz), AN (dB), N1 (Hz), AC (dB), AK (dB), K1 (Hz), K2 (Hz), AH (dB).

Then the authors present the controlling part of the system, their own FOPROx program, i.e. the software of the INBERE. The program, which was written in FORTRAN-IV PLUS, consist of two main parts:

a) the part which interprets the commands and b) the part which carries out the commands. The user can give commands to the system to perform different operations. The present paper gives the detailed syntactic and semantic description of the rule sequences and also presents the error signals the system gives to the user when he makes a syntactic error while typing the commands.

The present system accepts 15 commands. These commands together constitute the VEZER language. The authors begin their description of the VEZER language with metalinguistic notions. These are the following: „character”, „command sequence”, „command label”, „sound identifier”, „the signal” and „parameter”. Then they go on to describe the different commands in detail. The commands fall into the following groups:

1. Data feeding command: BEA
2. Section constructing command: EVA, ETI
3. Sounding command: LEO
4. Commands for handling the data basis: MOR, JON, AGI
5. Commands for temporal manipulation: IDA, IMI
6. Documentary commands: ZOL, TOM, PAL, VID, IVO
7. Stop command: VEG

With these commands it is possible to change the temporal, the frequency and the intensity characteristics of the sound sequence effectively during a single run which lasts for a maximum of 5 seconds. It is also possible to handle and maintain the data basis stored on a floppy disk and to carry out the necessary documentary operations.

A MAGYAR MÁSSALHANGZÓK ÉS A MÁSSALHANGZÓ—MAGÁNHANGZÓ-KAPCSOLÓDÁSOK AKUSZTIKAI SZERKEZETENEK ANALÍZISE ÉS SZINTÉZISE

Olaszy Gábor

Bevezetés

A magyar beszédben a leggyakrabban előforduló kételemű kapcsolat a mássalhangzó—magánhangzó-kapcsolódás CV és VC formája. Ezt statisztikai vizsgálatok is alátámasztják (vö. Kassai MFF 2. 1978). A számítógépes beszéd szintézissel készített hangsorok legfontosabb építőeleme tehát a # CV és a VC # kapcsolódási forma. E két kombinálásából adódik a VCV-elem.

A beszéd mesterséges előállítása számítógéppel megkívánja, hogy a beszédhangok és a hangkapcsolódások akusztikai szerkezetét az idő, az intenzitás és a frekvencia síkján egyaránt konkrét adatokkal adjuk meg, és meghatározzuk azokat az összefüggéseket, amelyek a három sík adatait összefogják.

A nyelvtudományi háttér

A magyar mássalhangzók és a mássalhangzó—magánhangzó-kapcsolódások akusztikai szerkezetének analizálásával már számos kutató foglalkozott. Tarnóczy Tamás úttörő munkáiban elsőként vizsgálta a 40-es évek végén a magyar laterális és a pergetett mássalhangzók, majd a zöngétlen spiránsok akusztikai szerkezetét. Majd Magdics Klára készített összefoglaló munkát hangspektrográf felhasználásával a teljes magyar hangállomány akusztikai vizsgálatáról. A munka 1965-ben jelent meg. Magdics külön fejezetet szentel a magyar mássalhangzók akusztikai szerkezetének, valamint a CV és VC hangkapcsolódások frekvenciaszerkezetének. 26 mássalhangzóra végzi el kísérleteit, és adja meg a mássalhangzók leírását, valamint a fenti hangkapcsolódások (26 msh és 8 mgh) összes CV- és VC-helyzetére a frekvenciaszerkezetet. Magdics munkája szintén úttörő jellegű, de nem törekszik teljességre, mivel nem foglalkozik a mássalhangzók teljes időtartamával, belső időszerkezetével, belső intenzitásviszonyaival, valamint a CV és VC kapcsolódásokban az intenzitásviszonyokkal. Nem közöl adatokat a mássalhangzó és a magánhangzó hangszórón belüli, egymáshoz viszonyított intenzitásértékére és a kapcsolódás során létrejövő intenzitásilleszkedésre. Magdics ezen kívül nem fordít figyelmet az F_1 és az F_3 mozgására sem, ami pedig a hangkapcsolódások leírásánál szintén szükséges. 1969-ben Fónagy Iván és Szende Tamás tollából jelent meg a „Zárhangok, réshangok, affrikáták hangszínképe” című tanulmány (NyK LXXI, 1969, 281–344). A szerzőpáros új módszerrel, megnövelt integrálási idejű hangspektrográffal vizsgálta a kérdéses mássalhangzók frekvenciaspektrumát az [a:], [i], [u] magánhangzók szomszédságában. A szerzők újszerű eredményeik ellenére megjegyzik, hogy „a spektrum integráció módszerével igyekeztünk további lépést tenni a magyar mássalhangzók elemzése terén. Az eredmények \gg integrálásától \ll , a magyar mássalhangzórendszer összefoglaló elemzésétől még igen sok választ el, mindenekelőtt a mássalhangzók szintézisén alapuló észlelés-lélektani vizsgálat” (Fónagy—Szende NyK LXXI, 1969, 282). Szende Tamás „A be-

széd folyamat alaptényezői” című könyvében a magyar mássalhangzók frekvenciaszerkezetének tárgyalása mellett már külön fejezetet szentel a hangok belső intenzitásviszonyainak, azonban csak általános következtetéseket von le konkrét szám adatok ismertetése nélkül (1976). Szende az összes magyar beszédhang átlagos időtartamát adatokkal is megadja, mint később részletesebb tanulmányában Kassai Ilona (1979), de egyikük sem szentel teret az egyes mássalhangzók belső időszerkezetének. Kassai, későbbi két tanulmányában (MFF 2. 1978; MFF 4. 1979) a hangkapcsolatokról és a magánhangzó–mássalhangzó-találkozásokról írva sok fontos adatot közöl, amelyek felhasználhatók lesznek a hangsorok szintézisében. Az alapszintézishez szükséges belső szabályok és törvényszerűségek feltárásával nem foglalkozik.

A fentiekből az tűnik ki, hogy az 1979-ig készült nyelvészeti munkák a mássalhangzók és a mássalhangzó–magánhangzó-kapcsolódási formák vizsgálatát más szempontok szerint közelítették meg; az alapszintézishez szükséges, konkrét adatokkal megadott részletes akusztikai hangleírásra még nem került sor.

A kitűzött cél

Jelen munkával olyan összefoglalót szándékoztam készíteni, amely elsőként tárgyalja a magyar mássalhangzó-állomány és a mássalhangzó–magánhangzó-kapcsolódási formák teljes akusztikai szerkezetét és egyfajta szintézisét. A tanulmányban a hozzáértő szakember (fonetikus, fizikus, mérnök, számítástechnikus, programozó, mérnök-nyelvész, logopédus, tanár, orvos stb.) megtalálja a magyar köznyelvi rövid mássalhangzók és a # CV, VCV, VC # hangkapcsolódások összes akusztikai adatának leírását szintetizálási példával illusztrálva. Ez tehát alapja lehet magyar hangsorok viszonylag gyors és kielégítő hangzású szintézisének, automatikusan működő, magyar nyelvű szöveget meghangosító szintetizáló rendszer kidolgozásának, valamint nyelvészeti vizsgálatoknak.

A probléma vázolója

Mivel a magyar hangsorok leggyakoribb eleme a mássalhangzó és a magánhangzó kapcsolódása, szavak, mondatok vagy nagyobb szövegegységek szintéziséhez szükséges, hogy elvégezzük a mássalhangzók, valamint a # CV, VCV és VC # hangkapcsolódások teljes akusztikai szerkezetének feltárását, majd szintézisét. A # CV hangkapcsolódást mindig abszolút hangsor elejként, a VC # hangkapcsolódást pedig hangsor végiként vizsgáltam. E háromféle szekvenciát tekintettem a szintézissel felépítendő hangsorok építőelemeinek. A három hanghelyzetre azért kellett a mássalhangzók vizsgálatát elvégezni, mert azok akusztikai szerkezete sok esetben más a hangsor eleji, más a hangsor belseji és más a hangsor végi helyzetben. A mássalhangzók és a fenti három építőelem teljes akusztikai szerkezetének a feltárásához – az eddigi eredmények felhasználása mellett – fel kellett tárni azokat az akusztikai paramétereket is, amelyeket a korábbi kutatások nem érintettek. Ezek a következők:

a) A mássalhangzóknál:

- milyen a frekvenciaszerkezet a belső időszerkezet függvényében,
- milyen az intenzitás szerkezet a belső időszerkezet függvényében.

A mássalhangzók többségének frekvenciaszerkezete és intenzitás szerkezete is összetett, a hangon belül is változik. Például a zöngés zárhangok esetében a zöngé intenzitásának alakulása az idő függvényében, a zárlejtés kialakulása a zöngéből, egymáshoz viszo-

nyitott intenzitásuk, a zöngés zárfelepattanás időszerkezetének és frekvenciaösszetevőinek az adatai mind jellemzőek az adott mássalhangzóra. Ezek az értékek szigorúan összefüggnek egymással, egyik vagy másik megváltoztatása a hang minőségének megváltozását jelentheti (vö. Gordos 1979; Gósy MFF 10. 1982).

b) A mássalhangzó–magánhangzó-kapcsolódásoknál (# CV, VCV, VC #):

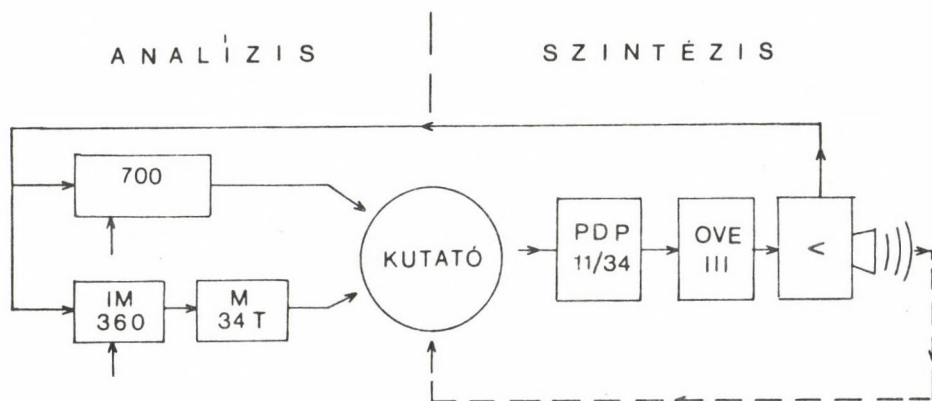
- milyen a mássalhangzó és a vele csatlakozó magánhangzó(k) egymás frekvenciaszerkezetére gyakorolt hatása, vagyis a frekvenciailleszkedés,
- milyen a mássalhangzó és a vele csatlakozó magánhangzó(k) egymáshoz viszonyított intenzitása,
- milyen a mássalhangzó amplitúdókapcsolódása a magánhangzó(k)hoz, vagyis az intenzitásilleszkedés (vö. Olasz MFF 8. 1981).

A fentiek szerint tehát a # CV, VCV és VC # kapcsolódásokra meg kellett határozni a teljes akusztikai szerkezet képét és azt, hogy hogyan lehet az így kapott adatokat olyan szabályrendszerbe foglalni, amelyet később a számítógépes beszéd-szintézisnél fel lehet használni.

A munka tárgya és módszere

A magyar mássalhangzók és a mássalhangzó–magánhangzó-kapcsolódások akusztikai szerkezetének vizsgálatát és szintézisét a [b], [p], [d], [t], [ʃ], [c], [g], [k], [m], [n], [ɲ], [ŋ], [v], [f], [z], [s], [ʒ], [ʃ], [j], [ç], [h̃], [h], [x], [b̃v], [p̃f], [d̃z], [t̃s], [d̃ʒ], [t̃ʃ], [l], [r] rövid magyar mássalhangzókra, valamint azoknak az [a:], [ɔ], [o], [u], [y], [ø], [i], [ɛ:], [ɛ] magánhangzókkal való összes kapcsolódási előfordulására, mintegy 864 hangkapcsolatra végeztem el 1980–81-ben.

A mássalhangzók és a fenti hangkapcsolódások bonyolult akusztikai szerkezete miatt az analízist nem lehetett a szintézis alkalmazása nélkül elvégezni, ugyanis kiderült, hogy sok esetben a tárgyalt beszédhangok és kapcsolódásaik akusztikai szerkezetében sok olyan – egyébként jelentéktelennek tűnő – adat játszik lényeges szerepet, amelyek



1. ábra

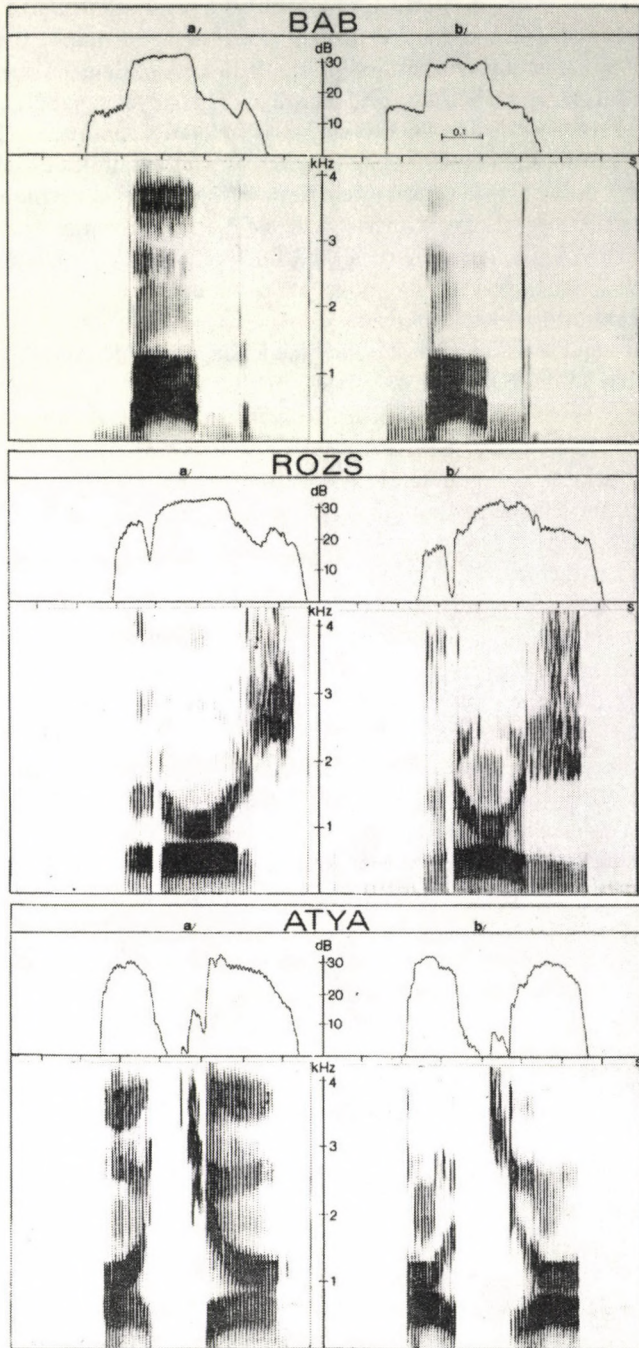
Az analízis–szintézis rendszere

hangformáló szerepét csak a mesterségesen előállított hangsorokban lehet megismerni (pl. a [l] hang intenzitása a V előtti szakaszban). Ezért az egész munka során az „analízist szintézissel” (analysis by synthesis) eljárást alkalmaztam. Ez a módszer egy zárt láncú vizsgálati forma, amelyben a kutató játssza a központi szerepet (1. ábra). A munka során az elemzési adatokat (időszerkezet, frekvenciaszerkezet és intenzitás- struktúra) laboratóriumunkban 700-as hangspektrográffal készített dinamikus hangspektrogramokról és IM 360-as típusú intenzitásmérő segítségével készített intenzitás-görbékről kaphatjuk meg. Ezeket a kutató, megfelelő átkódolás után, számítógépbe táplálja, és előállítja a vizsgált elem szintetizált megfelelőjét. Ezt percepció tesztnek és újbóli műszeres elemzésnek veti alá. Az így kapott új adatokat értékeli, és ezek ismeretében finomítja a számítógépnek korábban beadott adatokat. A folyamat többszöri megismétlésével végül kialakul a vizsgált hang vagy hangkapcsolódás akusztikai szerkezetére ténylegesen jellemző adatsorozat, amely már tartalmazza a vizsgált beszédhangra vagy hangsorra, az adott időpillanatonkénti frekvenciaszerkezet és intenzitásstruktúra Hz- és dB-értékeket (2.a,b ábra). A beszéd szintézist a saját fejlesztésű FOPROx programmal végezzük. A szintézis programjáról és hardware eszközeiről részletesen ugyanennek a kötetnek az „Interaktív beszéd szintetizáló rendszer számítógéppel és OVE III szintetizátorral” című tanulmányában (Kiss–Olaszy), valamint Olaszy Gábornak az MFF 8-ban megjelent munkájában található leírás.

A szintézis során a mássalhangzókat és a kapcsolódó magánhangzókat többnyire három hangszeléből állítottam össze. A hangszelettel egy beszédhang egy részét állítjuk elő. A hangszelet időtartamát a szintézist végző személy határozza meg. A hangszeletbe kell beírni az adott hangrész akusztikai (frekvencia-, intenzitás-) szerkezetének az adatait (részletesebben vö. Olaszy MFF 8. 1981). A magánhangzók esetében a középső szelet képviseli a tiszta fázist, vagyis az állandósult állapotot, a két szélső pedig a csatoló szelet szerepét tölti be. A csatoló szelettel valósítjuk meg a mássalhangzóhoz való formánssimulást, vagyis a frekvenciailleszkedést a V-ben, valamint a kapcsolódási szakaszra jellemző amplitúdómenetet, azaz az intenzitásilleszkedést. A mássalhangzóknál annak ellenére, hogy azokat is többnyire három szeletből állítottam elő, nem lehet ilyen egyértelműen meghatározni a szeletek szerepét a hang bonyolult akusztikai felépítéséből adódóan. Ezért a mássalhangzóknál az egyes szeletek szerepét minden esetben külön állapítottam meg. A hangátmenet időtartamát – mivel erre irodalmi adatot nem találtam – a magánhangzóknál többnyire a teljes hangidőtartam 30 %-ának vettem, a mássalhangzóknál egyedileg állapítottam meg. A munka során az első elemzéshez saját ejtésű mássalhangzóimat használtam # CV és VC # hangkapcsolatokban. A VCV kapcsolatokot szavakban vizsgáltam. Ezzel a módszerrel ugyanazon a példaanyagon sikerült elemezni maguknak a mássalhangzóknak az akusztikai szerkezetét és a mássalhangzó–magánhangzó-kapcsolódások során megvalósuló frekvencia-, idő- és intenzitásjellemzőket.

A szintézis egyéb szempontjai

A beszéd akusztikai szerkezete és az artikuláció között szoros összefüggés van. Az általam végzett fenti vizsgálatok csak a mássalhangzók és a # CV, VCV és VC # kapcsolódások akusztikai szerkezetét érintették. A magyar beszédhangok, így a mássalhangzók artikulációs sajátosságairól Bolla Kálmán „Magyar hangalbum” című munkájában található röntgenogramok, labiogramok, lingvogramok és palatogramok (Bolla MFF 6.



2.a ábra

A *bab*, *rozs*, *atya* mintaszavakról készült intenzitásgörbék és hangszíneképek
(a/ természetes, b/ szintetizált)

* BAB	TM	IDA	AO	FO	AN	N1	F1	B1	F2	B2	F3	B3	AC	AK	K1	K2
B1	40	10	0-	116												
			-3													
B2	60	10	3-	116												
			-2													
BA3	40	1	10-	116			400-		1008		2614					
			-14				-519									
BA	40	7	14	116			519-		1008		2614					
							-599									
A1	40	3	14	116			599		1008		2614					
AB	70	8	14	116			582-		1008		2614					
							-449									
BU1	70	10	6-	116												
			-2													
BU2	30	2	2	116			300-	96	1199	155	2329	500				
							-400									
BU3	30	2	2	116			400-	96	1199	155	2329	500				
							-300									

* ROZS	TM	IDA	AO	FO	AN	N1	F1	B1	F2	B2	F3	B3	AC	AK	K1	K2
RE1	40	6	0-	116			224-	96	1307	93-	2851	500				
			-2				-267			-155						
RE2	40	6	2-	116			291-	96	1307	155	2851	500				
			-4				-356									
RE3	40	5	4-	116			356-	96-	1037	155-	2851	500-				
			-0				-400	-36		-62		-250				
RO	60	8	4-	116			449		1199-		2468-					
			-10						-898		-2770					
O1	50	5	10	116			462		898		2397					
OZS	70	8	12	116			436-		951-		2329-					
							-300		-1600		-2016					
OZSV1	50	10	10-	116			300-		1647-		2016-		10-	16	1796	4032
			-4				-252		-1796		-2329		-20			
ZSV2	50	10	4-	116			252-		1796		2329		20	16	1849	4032
			-2				-212									
ZSV3	40	10	4-	116			212		1796		2329		20-		1849	4032
			-0										-2			

* ATYA	TM	IDA	AO	FO	AN	N1	F1	B1	F2	B2	F3	B3	AC	AK	K1	K2
A1	60	4	2-	116			582		1008		2614					
			-12													
A9	40	5	14	116			582		1008		2614					
ATY	60	10	12-	116			582-		1037-		2614-					
			-8				-400		-1600		-2468					
TYN	40	25														
TY1	30	5											10		3490-	4658-
															-3020	-4032
TY2	30	10											10-		3020-	4032
													-4		-2691	
TYA	60	15	2-	116			356-		1903-		2540-					
			-12				-550		-1037		-2397					
A10	40	2	12-	116			582		1008		2329-					
			-10								-2397					
AV	70	10	10-	116			582-		1008-		2397-					
			-4				-534		-1068		-2468					

2.b ábra

A bab, rozs, atya mintaszavak szintézisének adatai

1980). A mássalhangzó–magánhangzó-kapcsolódások – mint hangsorépítési elemek – artikulációs folyamatáról vizsgálati adatokkal még nem rendelkezünk.

A hangleírásokban a V helyébe bármelyik magyar magánhangzó behelyettesíthető a tiszta fázisára megadott formáns- és intenzitásértékével. Így bármilyen # CV, VCV és VC # alakú hangsorépítő elem előállítható, tehát szavakat, hosszabb szekvenciákat is szintetizálhatunk.

A szintetizált mintaszavak nem tartalmaznak szupraszegmentális hangszerkezetet, az alapfrekvencia értéke állandó, 116 Hz. A magánhangzók tiszta fázisának betáplált intenzitásértéke: AO = 14 dB, néhány kivételtől eltekintve. A szintézis leírásában szereplő hangszeltek AO, AN, AC intenzitásadatai csak akkor érvényesek, ha a V-ben az intenzitás. AO = 14 dB. Más V-intenzitásnál az adatokat arányosan kell kiszámítani. A V intenzitásán minden esetben a tiszta fázisra jellemző értéket értem.

A vizsgálatok eredményeit a hosszú mássalhangzókra is felhasználhatjuk. A rövid és hosszú mássalhangzók frekvenciaszerkezetében ugyanis nincs olyan lényeges különbség, hogy azokat külön kelljen tárgyalni. Bolla Kálmán a magyar mássalhangzók röntgenografikus vizsgálatának eredményeként megállapítja, hogy „a zárhangoknál a zárfelpattanás előtti szakasz megnyúlásával fejeződik ki a hosszúság”. Vizsgálatai azt is kimutatták, hogy „a tiszta fázis és az átmeneti szakaszok arányát tekintve megállapítható a tiszta fázis tartamának megnövekedése, míg az átmeneti szakaszok változatlanok, vagy csak igen kis mértékben változnak a hosszúságtól függően” (Bolla MFF 7. 1981, 13). Ezt támasztják alá a szintézissel végzett kísérleteim is. Hasonló megfontolások alapján a rövid magánhangzók hangsor belseji illeszkedési és kapcsolási adatait felhasználhatjuk a hosszú magánhangzók tartalmazó hangsorok szintéziséhez is.

A zöngés zárhangok zárfelpattanás előtti szakaszára a fojtott zöngé meghatározást használom a dolgozatban. Ezzel az elnevezéssel a hangképzésnek azt a speciális fajtáját akarom érzékeltetni, amikor is nincs levegőkiáramlás sem a szájon, sem az orron; a hangszalagok rezgését az izmok és a bőr veszi át, és így válik hallhatóvá ezen hangok zöngéje.

A magánhangzó után következő zöngétlen zárhangoknál a néma fázis elején kialakul egy kis intenzitású zörejelem is, amit az eddigi kutatások is kimutattak (vö. Tarnóczy–Vicsi *Acustica* 43. 1979). Ez a zörejelem a hang hangsorbéli hangzása szempontjából nem játszik lényeges szerepet, ezért a szintézis folyamán ennek a zörejnek a megvalósítását elhagytam. Ugyanezen megfontolásokból nem foglalkoztam azzal a gyenge zörejelemmel sem, amelyik a magánhangzó utáni [ʃ] és [ʧ] hangoknál a néma fázis kezdeti szakaszában figyelhető néha meg.

A dolgozatban a mássalhangzók intenzitásértékeinél és időszerkezeti adatainál tudatosan a körülbelül kifejezést használom. Ezek az intenzitás- és időtartamértékek ugyanis feszítettebb vagy lazább, pongyolább ejtéseknel szórhatnak.

A vizsgálatok során minden mássalhangzót előállítottam egy vagy két szintetizált mintaszóban is. Ezek a következők:

[b] – <i>bab, bibe</i>	[g] – <i>giga, tag</i>	[ŋ] – <i>hangos, ing</i>
[p] – <i>pap, pipa</i>	[k] – <i>kap, akik</i>	[v] – <i>vívás, öv</i>
[d] – <i>dada, ad</i>	[m] – <i>mami, vám</i>	[f] – <i>FIFA, döf</i>
[t] – <i>Tata, köt</i>	[m̥] – <i>hamvas</i>	[z] – <i>zuzá, húz</i>
[ʃ] – <i>gyagya, vagy</i>	[n] – <i>nana, ön</i>	[s] – <i>Szusza, húsz</i>
[c] – <i>tyúk, atya</i>	[ɲ] – <i>nyanya, lány</i>	[ʒ] – <i>Zsuzsi, rozsz</i>

[ʃ] – <i>sisak, sas</i>	[x] – <i>ihlet, pech</i>	[d̂3] – <i>dzsámi, lándzsa</i>
[j] – <i>jaj, hajít</i>	[b̂v] – <i>habverő</i>	[t̂f] – <i>csacsi, ács</i>
[ç] – <i>lépj</i>	[p̂f] – <i>copf, copfos</i>	[l] – <i>Lali, fél</i>
[ĥ] – <i>nahát</i>	[d̂z] – <i>madzag, edz</i>	[r] – <i>Ráró, ár</i>
[h] – <i>hal</i>	[t̂s] – <i>cica, Vác</i>	

A fenti mintaszavak intenzitásgörbéit, hangszínképeit és szintetizálási adatlapjait tartalmazó tablók (vö. a mintákat a 2.a, b ábrán.) az MTA Nyelvtudományi Intézetének fonetikai laboratóriumában megtalálhatók.

A vizsgálatok eredménye

A munka végleges eredményét a következő hangleírások tartalmazzák. Ezek az első olyan hangleírások, amelyek egységesen tárgyalják nemcsak a magyar mássalhangzók, de a # CV, VCV, VC # hangkapcsolódások – mint hangsorépítő elemek – teljes akusztikai szerkezetét is. Minden mássalhangzót külön hangleírás tárgyal, amely a következőket tartalmazza.

Az akusztikai szerkezet jellemzése

- A mássalhangzó időszerkezeti adatai # CV-, VCV- és VC #-helyzetben.
- A mássalhangzó intenzitás- és frekvenciaszerkezetének adatai az előbbi három helyzetben.
- A csatlakozó magánhangzó intenzitás- és frekvenciaszerkezetének adatai a CV és VC hangkapcsolódások átmeneti fázisában. A csatlakozási pontot a „cs” indexszel jelöljük.

A hang szintézise: a mássalhangzó szintézisének adatai az OVE III szintetizátorhoz.

Bilabiális, zöngés zárhang

[b]

Az akusztikai szerkezet jellemzése

- A hang időtartama hangsor eleji helyzetben kb. 110/15 ms, VCV helyzetben 80/15 ms, hangsor záró helyzetben pedig kb. 130/30 ms.
- Hangsor elején a zöngé nulla intenzitásról indul, és a képzési idő harmadánál éri el a maximumát, amely kb. 12 dB-lel alacsonyabb, mint a V intenzitása. A zárfelpattanásig a zöngé intenzitása 2–5 dB-nyit csökken. Magánhangzó utáni helyzetben a zöngé a V erősségénél kb. 10 dB-lel kisebb intenzitásról indul, és a zárfelpattanásig 4–6 dB-t csökken. A zöngés zárfelpattanás hangsor eleji és VCV-helyzetben ugyanaz. Intenzitása robbanásszerűen nő a V-ig, frekvenciaszerkezete: $F_1 = 300$ Hz, az F_2 értékét a V típusa határozza meg, $F_3 = 2300$ Hz. A C és V F_2 -je között az alábbi kölcsönhatás mutatható ki:

$$\begin{aligned} - \text{ha } 600 \text{ Hz} \leq VF_2 \leq 1600 \text{ Hz} &\rightarrow CF_2 = VF_2 \text{ Hz} \\ - \text{ha } VF_2 > 1600 \text{ Hz} &\rightarrow CF_2 = 1600 \text{ Hz} \end{aligned}$$

A zárfelpattanás szerkezete hangsor záró helyzetben megváltozik. Intenzitása 15 ms-on keresztül nő, majd a képzés végére 0-ra csökken. A maximumnál az intenzitás kb. 10 dB-lel alacsonyabb, mint a V erőssége. Formánsszerkezete: $F_1 = 300$ Hz, $F_2 = 1200$ Hz, $F_3 = 2300$ Hz.

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban:
CV-helyzetben $F_{1cs} = 300$ Hz

$$F_{2cs} = VF_2 \text{ Hz} \quad (\text{ha } 600 \text{ Hz} \leq VF_2 \leq 1600 \text{ Hz})$$

$$F_{2cs} = \frac{VF_2 + 1600}{2} \text{ Hz} \quad (\text{ha } VF_2 > 1600 \text{ Hz})$$

VC-helyzetben $F_{1cs} = 350$ Hz

$$F_{2cs} = VF_2 \text{ Hz} \quad (\text{ha } 600 \text{ Hz} \leq VF_2 \leq 1500 \text{ Hz})$$

$$F_{2cs} = \frac{2VF_2 + 1500}{3} \text{ Hz} \quad (\text{ha } VF_2 > 1500 \text{ Hz})$$

A V fokozatosan csökkenti intenzitását a csatlakozási szakaszban, majd beleolvad a zöngébe.

A [b] hang szintézise CV-helyzetben

1. szelet: $T = 35\text{--}40$ ms, $AO = 0 \rightarrow 3$ dB

2. szelet: $T = 65\text{--}70$ ms, $AO = 3 \rightarrow 2$ dB

3. szelet: $T = 10\text{--}15$ ms, $AO = 3 \rightarrow 14$ dB, $F_1 = 300$ Hz, F_2 értéke a V-től függ,
 $F_3 = 2300$ Hz.

A V csatoló szeletében a formánsokat a tiszta fázishoz hajlítjuk, az $AO = 14$ dB.

Bilabiális, zöngétlen zárhang

[p]

Az akusztikai szerkezet jellemzése

a) A hang néma fázisát hangsor eleji helyzetben nem érzékeljük, VCV- és VC #-helyzetben a néma fázis időtartama kb. 100 ms. A zárfelpattanási zöreje – amely két részből tevődik össze – # CV- és VCV-helyzetben kb. 10 ms-os. Az első rész – az intenzív felpattanás – kb. 4 ms, az utána következő lecsengő rész kb. 6 ms. Hangsor végi helyzetben a zárfelpattanás kb. 10 ms, a lecsengő zörejelem kb. 50 ms.

b) Hangsor eleji és VCV-helyzetben az intenzív zárfelpattanási zöreje 10–14 dB-lel, hangsor záró helyzetben 20–25 dB-lel gyengébb, mint a V intenzitása. A lecsengő zörejelem intenzitása fokozatosan csökken, a képzés végére szinte nulla. A zárfelpattanás frekvenciaspektrumában 400–5000 Hz-ig található zörejkomponensek egyenletes eloszlásban.

c) A V formánsainak csatlakozási értékei a C-hez ugyanazok, mint a [b] hangnál.

A [p] hang szintézise # CV-helyzetben

1. szelet: $T = 4$ ms, $AC = 16$ dB, $K_1 = 1600$ Hz, $K_2 = 2800$ Hz. $AO = 2$ dB, $F_1 = 400$ Hz, $F_2 = 1500$ Hz, $F_3 = 2500$ Hz. Az átviteli karakterisztika simítására a sáv szélességeket is megadhatjuk.

2. szelet: $T = 6$ ms, $AC = 10 \rightarrow 2$ dB, $K_1 = 1600$ Hz, $K_2 = 2800$ Hz.

A V csatoló szeletében az intenzitást hirtelen kell felvezérelni, $AO = 0 \rightarrow 14$ dB.

Dentális, dentalveoláris, zöngés zárhang

[d]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben 100/10 ms, VCV-helyzetben kb. 70/10 ms, hangsor záró helyzetben kb. 150/50 ms.

b) Hangsor elején a zöngé intenzíven indul, már kb. 10 ms után eléri a teljes intenzitását, amely kb. 15 dB-lel kisebb, mint a V-é. A zárfelpattanásig a zöngé szintje állandó, feszített ejtésnél azonban a zárfelpattanás előtt lecsökkenhet és így intenzitásminimum keletkezhet. Magánhangzó utáni helyzetben a zöngé a V-hez simulva kb. 4 dB-lel magasabb pontról indul és VCV-helyzetben kb. 20 ms, VC-helyzetben kb. 40 ms után éri el a # CV-helyzetben megadott szintet. A zöngés zárfelpattanás hangsor eleji és VCV-helyzetben megegyezik egymással. Intenzitása csak 2–4 dB-nyit növekszik, frekvenciaszerkezete: $F_1 = 400$ Hz, F_2 a V típusától függ, $F_3 = 2600$ Hz. A C és a VF_2 -je között az alábbi kölcsönhatás mutatható ki:

$$- \text{ha } VF_2 < 1300 \text{ Hz} \quad \rightarrow CF_2 = 1300 \text{ Hz}$$

$$- \text{ha } 1300 \text{ Hz} \leq VF_2 \leq 1800 \text{ Hz} \rightarrow CF_2 = VF_2 \text{ Hz}$$

$$- \text{ha } VF_2 > 1800 \text{ Hz} \quad \rightarrow CF_2 = 1800 \text{ Hz}$$

Hangsor záró helyzetben a zárfelpattanás zöngés eleme zörejjel párosul. Intenzitása erősen függ az ejtéstől, a zöngé szintjéből kb. 5 dB-nyit emelkedik ki, majd a képzés utolsó 25 ms-ában a hangerő fokozatosan nullára csökken. Frekvenciaszerkezete: $F_1 = 350$ Hz, $F_2 = 1600$ Hz, $F_3 = 2600$ Hz. A zörejjelem két intenzív komponense kb. 1660 Hz-en és 4000 Hz-en található.

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban: CV-helyzetben a V formánsai a C zárfelpattanási formánsairól indulva hajlanak a tiszta fázishoz.

$$\text{VC-helyzetben } F_{1cs} = 300 \text{ Hz}$$

$$F_{2cs} = \frac{2VF_2 + 1800}{3} \text{ Hz (ha } VF_2 > 1800 \text{ Hz)}$$

$$F_{2cs} = VF_2 \text{ Hz (ha } 1300 \text{ Hz} \leq VF_2 \leq 1800 \text{ Hz)}$$

$$F_{2cs} = \frac{2VF_2 + 1300}{3} \text{ Hz (ha } VF_2 < 1300 \text{ Hz)}$$

A [d] hang szintézise # CV-helyzetben

1. szelet: $T = 20$ ms, $AO = 4$ dB, $F_1 = 245 \rightarrow 206$ Hz,

2. szelet: $T = 70$ ms, $AO = 4 \rightarrow 2$ dB,

3. szelet: $T = 10$ ms, $AO = 0 \rightarrow 6$ dB, $F_1 = 400$ Hz, $F_2 = a$ V szerint, $F_3 = 2600$ Hz, (fontos az F_1 400 Hz-ről való indítása, mivel ez adja meg a [d] hang jellegzetes hangzását [vö. Gósy MFF 10. 1982]).

Az akusztikai szerkezet jellemzése

a) A hang néma fázisát hangsor eleji helyzetben nem érzékeljük, VCV- és VC # helyzetben a néma fázis kb. 100 ms, illetve kb. 130 ms. A zárfelpattanási zörejt – amely két részből tevődik össze – # CV- és VCV-helyzetben kb. 25 ms-os. Ebből az intenzív felpattanás kb. 10 ms, az azt követő lecsengő zörejelem kb. 15 ms. Hangsor záró helyzetben a zárfelpattanás szinte egybeolvad a lecsengő zörejjel, kettőjük együttes időtartama erősen függ az ejtéstől, elérheti a 100 ms-ot is.

b) Hangsor eleji és VCV-helyzetben az intenzív zárfelpattanás 15–25 dB-lel, hangsor záró helyzetben 25–30 dB-lel gyengébb, mint a V. A lecsengő zörejelem intenzitása fokozatosan nullára csökken a képzés végére. A [t] zárfelpattanása az 1600 Hz – 2000 Hz-es és a 3000 Hz–4000 Hz-es frekvenciasávban tartalmaz zörejjösszetevőket. A C és a V F_2 között a következő kölcsönhatás mutatható ki:

- ha $VF_2 < 1600$ Hz $\rightarrow CK_1 = 1600$ Hz
- ha $1600 \text{ Hz} \leq VF_2 \leq 2000$ Hz $\rightarrow CK_1 = VF_2$ Hz
- ha $VF_2 > 2000$ Hz $\rightarrow CK_1 = 2000$ Hz

Hangsor záró helyzetben a [t] zárfelpattanásának frekvenciaszerkezetére a V hatással van:

- ha $VF_2 < 1700$ Hz $\rightarrow CK_1 = 1700$ Hz
- ha $1700 \text{ Hz} \leq VF_2 \leq 1900$ Hz $\rightarrow CK_1 = VF_2$ Hz
- ha $VF_2 > 1900$ Hz $\rightarrow CK_1 = 1900$ Hz

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban:

CV-helyzetben $F_{1cs} = 300$ Hz

$$F_{2cs} = \frac{VF_2 + 1600}{2} \text{ Hz} \quad (\text{ha } VF_2 \ll CK_1)$$

$$F_{2cs} = VF_2 \quad (\text{ha } 1600 \text{ Hz} \leq VF_2 \leq 2000 \text{ Hz})$$

$$F_{2cs} = \frac{VF_2 + 2000}{2} \text{ Hz} \quad (\text{ha } VF_2 > CK_1)$$

A V intenzíven indul a [t] után.

$$\text{VC-helyzetben } F_{1cs} = 0,8VF_1 \text{ Hz} \quad (\text{ha } VF_1 \geq 500 \text{ Hz})$$

$$F_{1cs} = VF_2 \text{ Hz} \quad (\text{ha } VF_1 < 500 \text{ Hz})$$

$$F_{2cs} = \frac{2VF_2 + 1600}{3} \text{ Hz} \quad (\text{ha } VF_2 < 1600 \text{ Hz})$$

$$F_{2cs} = VF_2 \text{ Hz} \quad (\text{ha } 1600 \text{ Hz} \leq VF_2 \leq 2000 \text{ Hz})$$

$$F_{2cs} = \frac{2VF_2 + 2000}{3} \text{ Hz} \quad (\text{ha } VF_2 > 2000 \text{ Hz})$$

A V intenzitása fokozatosan nullára csökken a [t] előtt.

A [t] hang szintézise # CV-helyzetben

1. szelet: $T = 10 \text{ ms}$, $AC = 10 \text{ dB}$, K_1 értéke a V-től függ, $K_2 = 4500 \text{ Hz}$. Az $AO = 2 \text{ dB}$, $F_1 = 350 \text{ Hz}$, $F_2 = 1600 \text{ Hz}$, $F_3 = 2500 \text{ Hz}$, $B_1 = 90 \text{ Hz}$, $B_2 = 240 \text{ Hz}$, $B_3 = 500 \text{ Hz}$. A zörej színezetét az $AK = 17 \text{ dB}$ -lel biztosítjuk!

2. szelet: $T = 15 \text{ ms}$, $AC = 8 \rightarrow 2 \text{ dB}$, K_1 ugyanaz, mint az 1. szeletben, $K_2 = 4500 \text{ Hz}$, $AK = 17 \text{ dB}$.

A V csatoló szeletében az intenzitást hirtelen felvezéreljük, $AO = 0 \rightarrow 14 \text{ dB}$.

Palatális, zöngés zárhang

[J]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben 110/30 ms, VCV-helyzetben 60/20 ms, hangsor záró helyzetben pedig 150/50 ms.

b) Hangsor eleji helyzetben a zöngé nulla intenzitásról indul és kb. 40 ms múlva maximumot ér el, amikor is intenzitása kb. 15 dB-lel kisebb, mint a V-jé. Ettől kezdve a zöngé fokozatosan kb. 4 dB-nyit csökken a zárfelpattanásig. Magánhangzó utáni helyzetben a zöngé – a V-hez való csatolódás miatt – a V-nél kb. 10 dB-lel alacsonyabb intenzitásértékről indul, és fokozatosan csökken a zárfelpattanásig. A zárfelpattanás hangsor eleji és VCV-helyzetben egyforma. Intenzitása kb. 5 dB-nyit emelkedik ki a zöngéből, frekvenciaszerkezete: $F_1 = 300 \text{ Hz}$, $F_2 = 2100 \text{ Hz}$, $F_3 = 3000 \text{ Hz}$. A zár nyitódásával egyidőben a zöngés elemre enyhe – a [c]-hez hasonló – zörej is szuperponálódik. Hangsor záró helyzetben a zárfelpattanás szerkezete módosul. Intenzitása a V-nél kb. 10 dB-lel alacsonyabb értékig növekszik, majd a képzés utolsó 25 ms-ában fokozatosan nullára csökken. A frekvenciaszerkezetben formánsok nem találhatók, a zörejelem a hang végén megerősödik.

c) A V formánsai mind CV-, mind pedig VC-helyzetben szorosan illeszkednek a [J] zárfelpattanásának formánsaihoz.

A [J] hang szintézise # CV-helyzetben

1. szelet: $T = 45 \text{ ms}$, $AO = 0 \rightarrow 8 \text{ dB}$,

2. szelet: $T = 40 \text{ ms}$, $AO = 8 \rightarrow 4 \text{ dB}$,

3. szelet: $T = 30 \text{ ms}$, $AO = 0 \rightarrow 8 \text{ dB}$, $F_1 = 300 \text{ Hz}$, $F_2 = 2100 \text{ Hz}$, $F_3 = 3000 \text{ Hz}$.

A formánsokat a V formánsai felé kezdjük hajlítani! Az $AC = 3 \rightarrow 7 \text{ dB}$, K_1 ugyanaz, mint a [c]-nél, $K_2 = 4000 \text{ Hz}$.

A V formánsait és intenzitását a 3. szelethez csatlakoztatva hajlítjuk tovább a tiszta fázishoz.

Palatális, zöngétlen zárhang

[c]

Az akusztikai szerkezet jellemzése

a) A hang néma fázisát hangsor eleji helyzetben nem érzékeljük, VCV- és VC #

helyzetben a néma fázis kb. 80 ms, illetve 110 ms. A zárfelpattanás időtartama # CV- és VCV-helyzetben kb. 40 ms, VC #-helyzetben megnyúlik, kb. 80 ms-ra.

b) Hangsor eleji és VCV-helyzetben a zárfelpattanás kezdeti intenzitása 15–20 dB-lel, hangsor záró helyzetben 20–22 dB-lel gyengébb, mint a V-é. Ez az intenzitás, szinte a képzés végéig megmarad, csak a V előtti 15–20 ms-os részben – mint a zárfelpattanás lecsengő szakasza – csökken 4–6 dB-nyit. A zörej frekvenciaspektruma 2000 Hz-től 4000 Hz-ig tartalmaz komponenseket. A C és a VF_2 között az alábbi kölcsönhatás mutatható ki:

– ha $VF_2 \leq 1500$ Hz, akkor a [c] zörejsávja 2000–3000 Hz-ig terjed,

– ha $VF_2 > 1500$ Hz, akkor a [c] zörejsávja 3000–4000 Hz-ig terjed.

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban:

CV- és VC-helyzetben $F_{1cs} = 0,8 VF_1$ Hz (ha $VF_1 \geq 500$ Hz)

$F_{1cs} = VF_1$ Hz (ha $VF_1 < 500$ Hz)

$F_{2cs} = 1900$ Hz (ha $VF_2 \leq 1500$ Hz)

$F_{2cs} = 2300$ Hz (ha $VF_2 > 1500$ Hz)

A [c] hang szintézise # CV-helyzetben

1. szeptet: $T = 20$ ms, $AC = 6 \rightarrow 10$ dB, K_1 -et a V által meghatározott sáv felső pontjáról indítjuk és kb. 500 Hz-et lefelé mozgatjuk, $K_2 = 4500$ Hz.

2. szeptet: $T = 20$ ms, $AC = 6 \rightarrow 4$ dB, a K_1 -et tovább vezéreljük a zörejsáv alsó frekvenciájáig, $K_2 = 4500$ Hz. A V csatoló szeptetében $T = 70$ ms, az $AO = 2 \rightarrow 12$ dB.

Veláris, zöngés zárhang

[g]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben 130/30 ms, VCV- és hangsor záró helyzetben pedig 80/30 ms, illetve 140/40 ms.

b) Hangsor elején a zöngé nulla intenzitásról indul, és a képzési idő negyedénél éri el a maximumát, amely kb. 13 dB-lel alacsonyabb, mint a V intenzitása. Ettől kezdve a zöngé fokozatosan 4–7 dB-nyit csökken a zárfeloldódásig. Magánhangzó utáni helyzetben a zöngé intenzitása a kapcsolódási ponton mintegy 18 dB-lel alacsonyabb, mint a V-é, és erről az értékről mintegy 4–6 dB-nyit csökken a zárfeloldódásig. A zárfelpattanás hangsor eleji és VCV-helyzetben megegyezik. Intenzitása fokozatosan növekszik mintegy 5 dB-nyit, a formánsok és a zörej folyamatosan alakulnak ki a levegőáramlás megindulásával. A zárfeloldódáskor az $F_1 = 300$ Hz, F_2 értéke kb. 10–15 %-kal a mindenkori VF_2 felett van, az $F_3 = 2600$ Hz, a zörej gyenge intenzitású, frekvenciaszerkezete ugyanaz, mint a [k] hangé. Hangsor záró helyzetben a zárfelpattanás hosszabb, mivel egy lecsengő, befejező szakasza is van.

c) A V formánsai VCV-helyzetben egymásra hatást mutatnak. A V_1F_1 a 300 Hz-es ponthoz hajlik, a V_1F_1 -jének a mozgását a V_2F_2 -je határozza meg a következők szerint:

– a $V_1F_{2cs} = V_1F_2$ Hz (ha $V_1F_2 < 1500$ Hz)

– a $V_1F_{2cs} \rightarrow V_2F_2$ Hz (ha 1300 Hz $V_1F_2 < 1900$ Hz és $V_2F_2 < 1300$ Hz)

– a $V_1 F_{2cs} = 1,1 V_1 F_2$ Hz (ha $V_1 F_2 > 1900$ Hz). Ekkor $V_1 F_3$ viszont lefelé hajlik szintén a $V_1 F_{2cs}$ felé, így a két formáns összegeként egy nagy sávzélességű energiamaximum keletkezik. A V_1 fokozatosan csökkenti az intenzitását és beleolvad a [g] zöngéjébe.

A [g] hang szintézise # CV-helyzetben

1. szelet: $T = 30$ ms, $AO = 0 \rightarrow 4$ dB, $F_1 = 250 \rightarrow 206$ Hz.

2. szelet: $T = 60$ ms, $AO = 4 \rightarrow 2$ dB,

3. szelet: $T = 30$ ms, $AO = 2 \rightarrow 6$ dB, $AC = 6$ dB, K_1 és K_2 ugyanaz, mint a [k]

hangnál, $AK = 30$ dB. Formánsokat csak a V csatoló szeletébe adunk be.

A V intenzitását nulláról indítjuk, $AO = 0 \rightarrow 12$ dB.

Veláris, zöngétlen zárhang [k]

Az akusztikai szerkezet jellemzése

a) A hang néma fázisát hangsor eleji helyzetben nem érzékeljük. VCV- és VC # helyzetben a néma fázis időtartama kb. 100 ms. A zárfelpattanás kb. 40 ms-os, és két egyforma részből tevődik össze, az intenzív felpattanásból és a lecsengő zörejből. Hangsor záró helyzetben a [k] hang zárfelpattanása erősen megnyúlik, kb. 80 ms-ra.

b) Hangsor eleji és VCV-helyzetben a zárfelpattanás 10–15 dB-lel, hangsor végi helyzetben 15–20 dB-lel gyengébb, mint a V. A lecsengő zörejelem erről az értékről csökken közel nullára a képzés végére. Az intenzív zárfelpattanás frekvencia spektrumában egy erősebb és egy gyengébb zörejjóc figyelhető meg. Az előbbi frekvenciaértéke erősen függ a V-től. Az utóbbi kb. 4500 Hz-en helyezkedik el. A C és a VF_2 -je között az alábbi kölcsönhatás mutatható ki:

– a $CK_1 = 1,1 VF_2$ Hz

Hangsor záró helyzetben a lecsengő zörejelem megnyúlik, a zárfelpattanás alsó zörejjóca pedig enyhén változik a VF_2 -jének a függvényében a következők szerint:

– ha $VF_2 < 1300$ Hz $\rightarrow CK_1 = \frac{1400 + VF_2}{2}$ Hz

– ha $VF_2 \approx 1400$ Hz $\rightarrow CK_1 = 1400$ Hz

– ha $VF_2 > 1500$ Hz $\rightarrow CK_1 = 0,8 VF_2$ Hz

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisban a következők:

CV-helyzetben $F_{1cs} = 0,8 VF_1$ Hz

$F_{2cs} = VF_2$ Hz

VCV-helyzetben a $V_1 F_2$ -jének mozgását a $V_2 F_2$ befolyásolja:

– $V_1 F_{2cs} = 1,1 V_1 F_2$ Hz (ha $V_1 F_2 > 1800$ Hz)

– $V_1 F_{2cs} = 0,8 V_1 F_2$ Hz (ha $1500 \text{ Hz} \leq V_1 F_2 \leq 1800 \text{ Hz}$)

– $V_1 F_{2cs} = V_1 F_2$ Hz (ha $V_1 F_2 < 1500$ Hz)

VC-helyzetben a

$$- VF_{1cs} = 0,8 VF_1 \text{ Hz}$$

$$- VF_{2cs} = VF_2 \text{ Hz}$$

A [k] hang szintézise # CV-helyzetben

1. szelet: $T = 20 \text{ ms}$, $AC = 14 \text{ dB}$, $K_1 = 1,1 VF_2 \text{ Hz}$, $K_2 = 4500 \text{ Hz}$, $AK = 30 \text{ dB}$.

2. szelet: $T = 20 \text{ ms}$, $AC = 12 \rightarrow 2 \text{ dB}$, a többi adat nem változik.

Bilabiális, nazális, zöngés zárhang

[m]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben 100–150 ms, VCV- és VC #helyzetben kb. 80 ms, illetve 150–200 ms.

b) Hangsor elején a hang intenzitása a képzési idő negyedéig növekszik, majd közel egyenletes szintű. A V előtti szakasz intenzitásintjét a V típusa befolyásolja. Általában kimondható, hogy a [m] intenzitásintje az [i], [ɛ:], [y], [ø], [ɐ] magánhangzók előtt kb. 4 dB-lel, az [a:], [ɔ], [o], [u] előtt 6–10 dB-lel alacsonyabb, mint a V intenzitása. Lazább ejtésnél a hang intenzitásában nincs állandósult szint, hanem az intenzitás 0-ról indulva fokozatosan növekszik egészen a V-ig. VCV-helyzetben a hang intenzitása állandónak mondható, 5–10 dB-lel alacsonyabb, mint a V szintje. CV #helyzetben a hang intenzitásintje közvetlenül a V utáni szakaszban a V típusától függ. Az [i] után a legkisebb a különbség a V és a [m] intenzitása között – mintegy 5 dB –, az [o] és [u] utáni helyzetben mintegy 12 dB. Az intenzitás további időbeni alakulása az ejtéstől függ. Ha az artikuláció során a zárat felpattanjuk, akkor a [m] hang intenzitása szinte egyenletes, esetleg a hang utolsó negyedében – a zárfelpattanás miatt – kissé megemelkedik, majd hirtelen csökken nullára. Ha a zárat nem pattantjuk fel, akkor a hangsor végi [m] intenzitása egyenletesen csökken a hang végéig. A formánsszerkezet képzésében a nazális üreg mint nyílt, az orális mint zárt üreg vesz részt. A [m] formánsai tompultabak, nagyobb sáv szélességgel rendelkeznek, mint ahogy azt a magánhangzóknál megszoktuk. Hangsor eleji helyzetben az $F_1 = 250 \text{ Hz}$, az F_2 a 800–1300, az F_3 az 1900–2300, az F_4 a 2900–3200 Hz-es sávban mozog a V típusától függően, hiszen a [m] ejtések már a V-re artikulálunk. Minél magasabb a VF_2 , annál magasabb értékek felé tolnak el a [m] formánsai a megadott sávokon belül. VCV-helyzetben a V_1 -nek hasonló hatása van, így ha $V_1 F_2$ és $V_2 F_2$ erősen különbözik egymástól, akkor a [m] F_2, F_3, F_4 formánsai erős hajlást mutatnak a két magánhangzó között (pl. a *mami* szóban). Hangsor záró helyzetben a VF_2 és CF_2 között a következő kölcsönhatás mutatható ki:

$$- \text{ha } VF_2 \leq 1200 \text{ Hz} \rightarrow CF_2 = VF_2 \rightarrow 1200 \text{ Hz}$$

$$- \text{ha } VF_2 > 1200 \text{ Hz} \rightarrow CF_2 = 1200 \text{ Hz}$$

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban:

$$\text{CV-helyzetben } VF_{1cs} = 350 \text{ Hz} \quad \text{ha } VF_1 \geq 350 \text{ Hz}$$

$$VF_{1cs} = VF_1 \text{ Hz} \quad \text{ha } VF_1 < 350 \text{ Hz}$$

$$VF_{2cs} = CF_2 \text{ Hz}$$

A V formánsainak hajlása szinte ugrásszerűen zajlik le, mintegy 20 ms alatt! VCV- és VC-helyzetben a V_1 formánsai szorosan csatolódnak a [m] formánsaihoz.

A [m] hang szintézise #CV-helyzetben

1. szelet: $T = 40$ ms, $AO = 0 \rightarrow 6$ dB, $F_1 = 250$ Hz, $F_2 = a$ V típusától függ, $F_3 = 1800$ Hz, $B_1 = 180$ Hz, $B_2 = 100$ Hz, $B_3 = 500$ Hz, $AN = 0 \rightarrow 6$ dB, $N_1 = 250$ Hz.

2. szelet: $T = 40$ ms, $AO = 6 \rightarrow 10$ dB, $AN = 6 \rightarrow 8$ dB, a többi adat ugyanaz.

3. szelet: $T = 40$ ms, $AO = 10$ dB, $AN = 8$ dB, a többi adat ugyanaz.

A V csatoló szeletében, amely kb. 20 ms-os kell, hogy legyen, az intenzitást ugrásszerűen meg kell emelni, $AO = 14$ dB, a nazális komponenseket megszüntetjük, $AN = 8 \rightarrow 0$ dB, $N_1 = 250 \rightarrow 200$ Hz.

Labiodontális, zöngés zárhang

[m]

Az akusztikai szerkezet jellemzése

a) A hang csak VCV-helyzetben fordul elő, időtartama kb. 200 ms.

b) A hang [m] összetevőjének az intenzitása a V_1 -hez csatlakozó ponton kb. 8 dB-lel alacsonyabb, mint a V-é. A képzés első 100–130 ms-ában az intenzitáscsökkenés kb. 10 dB/100 ms, majd 6 dB/40 ms. Az ezután következő [v] összetevő intenzitása fokozatosan növekszik, így csatlakozik a V-hez. A [m] elem formánsszerkezete a hang első részében: $F_1 = 300$ Hz, az F_2 a 800–1200 Hz-es sávban mozog a V_1 -től függően (lásd a [m] hangot), $F_3 = 2100$ Hz, a második szakaszban $F_1 = 300 \rightarrow 350$ Hz, $F_2 = 1000$ Hz, $F_3 = 2100 \rightarrow 2500$ Hz. A sáv szélesség az F_1 , F_2 -nél nagy, kb. 100–200 Hz. A [v] elem frekvenciaszerkezete hasonló a [v] hangéhoz.

c) A V formánsai ugyanúgy csatolódnak a C-hez, mint a [m] esetében. A V_1 -ben az F_1 sáv szélessége fokozatosan növekszik, a magánhangzó nazalizálódik a [m] előtt.

A [m] hang szintézise VCV-helyzetben

1. szelet: $T = 90$ ms, $AO = 6$ dB, $AN = 14$ dB, $F_1 = 350$ Hz, F_2 a V_1 -től függ, $F_3 = 2100$ Hz, $B_1 = 100$ Hz, $B_2 = 100$ Hz.

2. szelet: $T = 50$ ms, $AO = 6$ dB, $AN = 14 \rightarrow 4$ dB, a többi adat nem változik.

3. szelet: $T = 40$ ms, $AO = 2 \rightarrow 10$ dB, $F_1 = 350 \rightarrow 400$ Hz, F_2 értékét a V_2 határozza meg, $F_3 = 2500$ Hz.

A V_1 csatoló szeletében $AN = 0 \rightarrow 12$ dB, $B_1 = 12 \rightarrow 100$ Hz, $B_2 = 35 \rightarrow 100$ Hz, $AO = 12 \rightarrow 4$ dB.

A V_2 csatoló szelete ugyanolyan, mint a V után.

Dentális, dentalveoláris, zöngés zárhang

[n]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben kb. 110 ms, VCV- és VC #helyzetben kb. 80 ms, illetve kb. 140 ms.

b) Hangsor elején és VCV-helyzetben az intenzitás hasonlóan alakul, mint a [m] hangnál, hangsor záró helyzetben a [n] intenzitása kb. 6 dB-lel alacsonyabb, mint a V-é, és nullához tart. Pontos ejtésnél az utolsó 30 ms-ban a zárfelpattanás okozhat egy kb. 2 dB-nyi intenzitásemelkedést a végső lecsengés előtt. A formánsszerkezet kialakításában

a nazális üreg mint nyílt, a dentálveoláris részhez tapadt nyelvvel lezárt orális üreg mint zárt rendszer vesz részt. Hangsor eleji és V utáni helyzetben a formánsok: $F_1 = 250$ Hz, F_2 értéke a V-től függ, $F_3 = \text{kb. } 2600$ Hz. A VF_2 -je és a CF_2 -je között a következő kölcsönhatás mutatható ki:

- ha $VF_2 > 1600$ Hz $\rightarrow CF_2 = 1600$ Hz
- ha $1250 \leq VF_2 \leq 1600$ Hz $\rightarrow CF_2 = VF_2$ Hz
- ha $VF_2 < 1250$ Hz $\rightarrow CF_2 = 1250$ Hz

A [n] hang formánsai állandó értékűek a hangon belül, nem hajlanak a V-hez. Ez annak tudható be, hogy a nazális üreg méretei állandók. VCV-helyzetben, ha a C_1F_2 és V_2F_2 erősen különbözik egymástól, akkor a CF_2 hajlást mutat a V_1 -től a V_2 felé a korábban megadott sávon belül.

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban:

- CV-helyzetben $F_{1cs} = 400$ Hz (ha $VF_1 \geq 400$ Hz)
 $F_{1cs} = VF_1$ Hz (ha $VF_1 < 400$ Hz)
 $F_{2cs} = CF_2$ Hz

A formánsok hajlása a V-ben hosszú idő alatt (kb. 100 ms) megy végbe, így a V tiszta fázisának időtartama lerövidül.

- VC-helyzetben $F_{1cs} = 400$ Hz (ha $VF_1 \geq 400$ Hz)
 $F_{1cs} = VF_1$ Hz (ha $VF_1 < 400$ Hz)
 $F_{2cs} = CF_2$ Hz

A [n] hang szintézise # CV-helyzetben

1. szelet: $T = 50$ ms, $AN = 0 \rightarrow 16$ dB, $N_1 = 250$ Hz.
2. szelet: $T = 50$ ms, $AN = 16$ dB, $N_1 = 250$ Hz.

A V csatoló szeletében az $AN = 14 \rightarrow 0$ dB, $AO = 0 \rightarrow 14$ dB, $N_1 = 250 \rightarrow 200$ Hz.

Palatális, zöngés, nazális zárhang

[ɲ]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben kb. 90 ms, VCV- és VC # helyzetben kb. 50 ms, illetve 130 ms.

b) A hang formánsszerkezete $F_1 = 350$ Hz, $F_2 = 1100$ Hz, $F_3 = 2400$ Hz, $F_4 = 3400$ Hz, $B_1 = 150$ Hz. Az F_1 intenzitása lényegesen nagyobb, mint a többi formánsé. A formánsszerkezetre a V nincs hatással.

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban:

- CV- és VC-helyzetben $F_{1cs} = CF_1$ Hz
 $F_{2cs} = CF_3$ Hz
 $F_{3cs} = CF_4$ Hz

A V átmeneti fázisa hosszú, mintegy 100 ms, ha nagy formánsközöket kell a képzés során áthidalni (pl. a *nyúl* szóban).

A [ɲ] hang szintézise # CV-helyzetben

1. szelet: $T = 50$ ms, $AO = 0 \rightarrow 4$ dB, $F_1 = 250 \rightarrow 300$ Hz, $F_2 = 1300$ Hz, $F_3 = 2600$ Hz, $B_1 = 100$ Hz, $B_2 = 240$ Hz, $B_3 = 500$ Hz, $AN = 0 \rightarrow 10$ dB.

2. szelet: $T = 50$ ms, $AO = 4 \rightarrow 6$ dB, F_1 -et tovább növelhetjük, ha a VF_1 megkívánja, a többi adat nem változik.

A V csatoló szeletében az $AN = 10 \rightarrow 0$ dB, $AO = 6 \rightarrow 14$ dB, $F_2 = 2000 \rightarrow VF_2$ Hz, $F_3 = 2900 \rightarrow VF_3$ Hz, az F_1 csatolása folyamatos.

Veláris, zöngés, nazális zárhang

[ɲ]

Az akusztikai szerkezet jellemzése

a) A hang hangsor eleji helyzetben nem fordul elő, VCV- és VC #helyzetben az időtartama 180/40 ms, illetve 280/50 ms.

b) A hang nazális zöngés elemből és zöngés zárfelpattanásból áll. VCV-helyzetben a magánhangzóhoz csatlakozó ponton a nazális elem intenzitása kb. 4 dB-lel kisebb, mint a V_1 -é, majd 4 dB/100 ms-os csökkenés következik. A zárfelpattanás előtt intenzitásminimum található, a csökkenés kb. -10 dB. A nazális elem formánszerkezete: $F_1 = 350$ Hz, $F_2 = a V_1 F_2$ -től függ, úgy, mint a [n] hangnál, $F_3 = 2600$ Hz. Az F_1 és F_2 sáv szélessége kb. 150 Hz. A zárfelpattanás szerkezete ugyanolyan, mint a [g] hangnál.

c) A V-ben az F_1 sáv szélessége fokozatosan növekszik, így a V a [ɲ] előtt nazalizálódik.

A [ŋ] hang szintézise VCV-helyzetben

1. szelet: $T = 60$ ms, $AO = 10 \rightarrow 8$ dB, $F_1 = 350$ Hz, F_2 a V-től függ, $F_3 = 2600$ Hz, $B_1 = 150$ Hz, $B_2 = 100 \rightarrow 150$ Hz, $AN = 8$ dB.

2. szelet: $T = 40$ ms, $AO = 8 \rightarrow 2$ dB, $AN = 8 \rightarrow 4$ dB, a többi adat nem változik.

3. szelet: $T = 40$ ms, $AO = 2$ dB, $AN = 4 \rightarrow 0$ dB, $F_1 = 300 \rightarrow 250$ Hz, a többi adat nem változik.

4. szelet: $T = 40$ ms, $AO = 0 \rightarrow 4$ dB, $AC = 0 \rightarrow 4$ dB, K_1 és K_2 ugyanolyan, mint a [k] hangnál, $AK = 30$ dB.

A V_1 -ben az F_1 sáv szélességét fokozatosan növeljük, ez adja meg a [ŋ] hang sajátos hangzásának az előkészítését.

Labiodentális, zöngés középréshang

[v]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben kb. 100 ms, VCV- és VC #helyzetben kb. 60 ms, illetve kb. 180 ms.

b) Hangsor elején a hang fokozatosan növeli intenzitását és kb. 30 ms múlva maximumot ér el, akkor kb. 15 dB-lel gyengébb, mint a V. Ezután 4 dB/50 ms-os csökkenés, majd 4 dB/20 ms-os intenzitásnövekedés következik. VCV-helyzetben a hang intenzitása először csökken, majd növekszik. A minimumnál kb. 16 dB-lel gyengébb, mint a V_1 . Hangsor záró helyzetben az intenzitás a hang első felében ugyanúgy csökken, mint a

VCV-helyzetben, ezután pedig az ejtéstől függ. Erős ejtésnél egy intenzív zöngés „sva”-szerű elem következik, gyenge ejtésnél a hangerősség fokozatosan nullára csökken.

A [v] formánsszerkezete: $F_1 = 300$ Hz, $F_2 = a$ V-től függő, $F_3 = 2500$ Hz, $B_1 = 60$ Hz, $B_2 = 250$ Hz, $B_3 = 700$ Hz. Erős ejtésnél a zöngés elemre zörej is szuperponálódik, amelynek frekvenciatartománya kb. 1000-tól 10000 Hz-ig terjed. A CF_2 és VF_2 között az alábbi kölcsönhatás mutatható ki:

- ha $VF_2 \leq 1500$ Hz $\rightarrow CF_2 = 800$ – 1500 Hz-es sávban mozog; minél alacsonyabb a VF_2 , annál alacsonyabb a CF_2 a megadott sávban
- ha $VF_2 > 1500$ Hz $\rightarrow CF_2 = 1500$ Hz

VCV-helyzetben a CF_2 mozog a közrefogó magánhangzók F_2 -jének összekötő vonala mentén. Hangsor záró helyzetben a hang formánsszerkezete: $F_1 = 350$ Hz, $F_2 = 1200$ Hz, $F_3 = 2400$ Hz, a zörej frekvenciája az 1000–10000 Hz-es sávban egyenletesen oszlik meg, és a képzés végére dominánssá válik.

c) A V formánsai a C-hez való csatlakozásnál mindig a C formánsaihoz simulnak.

A [v] hang szintézise # CV-helyzetben

1. szeptet: $T = 40$ ms, $AO = 0 \rightarrow 3$ dB, $F_1 = 350 \rightarrow 250$ Hz, az F_2 a V-től függ, $F_3 = 2540$ Hz, $B_1 = 60$ Hz, $B_2 = 250$ Hz, $B_3 = 750$ Hz. A hangot zörejesíthetjük, $AC = 0 \rightarrow 8$ dB, $K_1 = 1600$ Hz, $K_2 = 4500$ Hz. A zörej hangszínezését az $AK = 17$ dB-lel biztosítjuk.

2. szeptet: $T = 40$ ms, $AO = 3 \rightarrow 2$ dB, a formánsokat a V formánsaihoz kezdjük hajlítani, a zörej adatai nem változnak.

3. szeptet: $T = 20$ ms, $AO = 2 \rightarrow 4$ dB, $AC = 6 \rightarrow 0$ dB, a formánsokat tovább hajlítjuk a V-hez, a sávzélességeket lecsökkentjük.

Labiodentális, zöngétlen középréshang

[f]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben kb. 100 ms, VCV- és VC #-helyzetben kb. 140 ms.

b) Hangsor elején a hang intenzitása fokozatosan növekszik, a képzési idő háromnegyedénél éri el a maximumot, amely 20–25 dB-lel marad el a V intenzitásától. Ezután 4–5 dB-es intenzitáscsökkenés következik. VCV- és VC #-helyzetben a zörej kb. az előbbi maximumérték intenzitásával indul, majd 4–5 dB-t csökken, illetve nullára csökken. A [f] hang nem rendelkezik frekvenciagócokkal, frekvenciaspektrumában a komponensek az 1000–10000 Hz-es tartományban egyenletesen oszlanak meg. A zörejsáv alsó határa mozog a VF_2 függvényében, hiszen a [f] képzésekor már a V-re artikulálunk. Ha VF_2 magasabb, mint 1000 Hz, akkor a zörejsáv alsó határa VF_2 -je felé tolódik.

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban:

$$F_{1cs} = 400 \text{ Hz} \quad (\text{ha } VF_1 \geq 400 \text{ Hz})$$

$$F_{1cs} = VF_1 \text{ Hz} \quad (\text{ha } VF_1 < 400 \text{ Hz})$$

$$F_{2cs} = 0,9 VF_2 \text{ Hz} \quad (\text{ha } VF_2 \geq 1000 \text{ Hz})$$

$$F_{2cs} = VF_2 \text{ Hz} \quad (\text{ha } VF_2 < 1000 \text{ Hz})$$

A [f] hang szintézise # CV-helyzetben

1. szeptet: $T = 50$ ms, $AC = 0 \rightarrow 20$ dB, K_1 értéke a V-től függ, $K_2 = 3400$ Hz, $AK = 18$ dB.

2. szeptet: $T = 50$ ms, $AC = 18 \rightarrow 12$ dB, K_1 -et – ha kell – tovább hajlítjuk, a többi adat nem változik.

A V intenzitását fokozatosan növeljük a csatoló szeptetben, $AO = 0 \rightarrow 14$ dB.

Dentális, dentalveoláris, zöngés középreshang

[z]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben kb. 100 ms, VCV- és VC #-helyzetben kb. 70 ms, illetve 150 ms.

b) A hang zöngés elemből és az erre szuperponálódott zörejből áll. Hangsor eleji helyzetben intenzitása a képzés harmadáig növekszik, amikor kb. 10–15 dB-lel kisebb, mint a V erőssége. A zörejképződés ezután kezdődik el, mialatt a zöngés elem intenzitása viszont csökken, így a hang összintenzitása a 70.ms-ig kb. 4 dB-nyit csökken. Ettől kezdve mind a zöngés, mind a zörejelem növeli amplitúdóját, és a képzés végére a hang intenzitása kb. 6 dB-nyit emelkedik. VCV-helyzetben a zörejelem a hang teljes időtartama alatt jelen van. A zöngés elem a V_1 -ből csatlakozva folyamatosan csökkenti intenzitását a hang közepéig, ezzel egyidőben a zörejelem fokozatosan növekszik. A kettő együttes hatására a hang felénél az intenzitás kb. 15 dB-lel kisebb, mint a V_1 -é. Ezután mind a zöngés, mind pedig a zörejelem intenzitása növekszik, aminek együttes hatására a hang intenzitása kb. 6 dB-nyit emelkedik a képzés végére. Hangsor záró helyzetben a [z] hang intenzitása a képzés végére nullára csökken. A zöngés elem formánsszerkezete: $F_1 = 300$ Hz, F_2 az 1300–1800 Hz-es sávban mozog a V-től függően, $F_3 = 2800$ Hz. Az F_2 , F_3 intenzitása kicsi az F_1 -éhez képest. A zörejelem frekvenciastruktúrája megegyezik a [s] hangéval, intenzitása azonban kb. 5 dB-lel gyengébb. A VF_2 és CF_2 között a következő kölcsönhatás mutatható ki:

– ha $VF_2 < 1300$ Hz $\rightarrow CF_2 = 1300$ Hz

– ha $1300 \text{ Hz} \leq VF_2 \leq 1800 \text{ Hz} \rightarrow CF_2 = VF_2$ Hz

– ha $VF_2 > 1800$ Hz $\rightarrow CF_2 = 1800$ Hz

VCV-helyzetben a CF_2 mindig a V_1F_2 , V_2F_2 összekötő vonala mentén mozog a megadott sávban belül.

c) Mivel két orális zöngés hang kapcsolódásáról van szó, a folyamatos artikulációból következik, hogy a formánsoknak folyamatos átmenettel kell csatlakozniuk. A V formánsai a C formánsai által elért pontról folytatják a csatlakozást.

A [z] hang szintézise # CV-helyzetben

1. szeptet: $T = 30$ ms, $AO = 0 \rightarrow 4$ dB, $F_1 = 260 \rightarrow 230$ Hz, $F_2 = 1600$ Hz, $F_3 = 2800$ Hz, $B_2 = 200$ Hz, $B_3 = 500$ Hz.

2. szeptet: $T = 40$ ms, $AO = 4 \rightarrow 2$ dB, F_1 -et és F_2 -t a V felé hajlítjuk, $AC = 2 \rightarrow 4$ dB, $K_1 = 4000$ Hz, $K_2 = 9000$ Hz.

3. szelet: $T = 30$ ms, $AO = 2 \rightarrow 4$ dB, F_1 és F_2 -t tovább hajlítjuk, $AC = 4 \rightarrow 6$ dB, $K_1 = 4000$ Hz, $K_2 = 9000$ Hz.

A V csatoló szeletében a zörejt megszüntetjük, $AC = 4 \rightarrow 0$ dB, a többi összetevő mozgását tovább folytatjuk.

Dentális, dentalveoláris, zöngétlen középréshang

[s]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben 130–170 ms között van, az ejtéstől függően, VCV- és VC #helyzetben kb. 120 ms, illetve kb. 220 ms.

b) Hangsor elején a zörejt nulla intenzitásról indul és kb. 30 ms-nál éri el a maximumot, amely kb. 8 dB-lel alacsonyabb, mint a V erőssége. A következő szakaszban az intenzitás állandó, majd a képzés utolsó 30 ms-ában közel nullára csökken. VCV- és VC #helyzetben a zörejt intenzitása a V-hez csatlakozó ponton mintegy 15 dB-lel kisebb, mint a V-é. A továbbiakban pedig hasonlóan alakul, mint a # CV-helyzetben. A zörejt 3000 Hz és 12000 Hz között tartalmaz frekvenciakomponenseket, intenzív zörejt talátható 4000 Hz környékén és 8000 Hz-en. A zörejsáv alsó góca a V típusától függően mozog; hiszen a [s] hang képzésekor már ajkainkkal a V-re artikulálunk. Az [u] előtti helyzetben a zörejsáv alsó góca kb. 3200 Hz-en, [i] előttiben kb. 4500 Hz-en van.

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban:

CV-helyzetben $F_{1cs} = 400$ Hz (ha $VF_1 \geq 400$ Hz)

$F_{1cs} = VF_1$ Hz (ha $VF_1 < 400$ Hz)

$$F_{2cs} = \frac{VF_2 + 1700}{2} \text{ Hz}$$

VC-helyzetben $F_{1cs} =$ ugyanaz, mint a CV-helyzetben

$$F_{2cs} = \frac{2VF_2 - 1700}{3} \text{ Hz}$$

A [s] hang szintézise # CV-helyzetben

1. szelet: $T = 40$ ms, $AC = 0 \rightarrow 16$ dB, $K_1 = 4000$ Hz, $K_2 = 8000$ Hz.

2. szelet: $T = 30$ ms, $AC = 16$ dB, $K_1 = 4000$ Hz, $K_2 = 8000$ Hz.

3. szelet: $T = 40$ ms, $AC = 14 \rightarrow 4$ dB, $K_1 = 4000$ Hz-ről a V által meghatározott értékre hajlítjuk, $K_2 = 8000$ Hz.

A V csatoló szeletében az intenzitást fokozatosan növeljük, $AO = 0 \rightarrow 12$ dB.

Alveoláris, posztalveoláris, zöngés középréshang

[3]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben kb. 100 ms, VCV- és VC #helyzetben kb. 80 ms, illetve 150 ms.

b) A hang zöngés elemből és az erre szuperponálódott zorejtből áll. Hangsor eleji

helyzetben az intenzitása a képzési idő egyharmadáig növekszik, amikor is kb. 12 dB-lal gyengébb, mint a V-é. A következő harmadban az intenzitás 2–5 dB-nyit csökken, majd 4–6 dB-nyit nő. A zörej a hang kezdetétől fokozatosan növeli az intenzitását, és kb. a 40 ms után dominánssá válik a zöngés részhez képest. VCV-helyzetben a zöngés elem a V_1 -hez csatlakozik, intenzitása kb. 5–11 dB-lal alacsonyabb, mint a V_1 -é és ezen a szinten marad a képzés végéig. Hangsor záró helyzetben a [3] hang intenzitása fokozatosan nullára csökken. A zöngés elem formánszerkezete: $F_1 = 250$ Hz, az F_2 a V típusától függően veszi fel az értékét az 1600–2200 Hz-es sávban, $F_3 = 2500$ Hz. Minél alacsonyabb a VF_2 , annál közelebb van a CF_2 az 1600 Hz-hez. A zörej frekvenciaszerkezete megegyezik a [j] hangéval, csak kb. 5 dB-lal kisebb intenzitású.

c) Mivel két orális zöngés hangról (V és C) van szó, a V formánsai folyamatosan simulnak a [3] formánsaihoz.

A [3] hang szintézise # CV helyzetben

1. szelet: $T = 30$ ms, $AO = 0 \rightarrow 4$ dB, $F_1 = 350 \rightarrow 250$ Hz, F_2 a V-től függ, $F_3 = 2400$ Hz. A zörej adatai: $AC = 0 \rightarrow 2$ dB, $K_1 = 1800$ Hz, $K_2 = 4000$ Hz, $AK = 16$ dB.

2. szelet: $T = 30$ ms, $AO = 4 \rightarrow 2$ dB, $AC = 4 \rightarrow 10$ dB, a többi adat nem változik.

3. szelet: $T = 40$ ms, $AO = 2 \rightarrow 8$ dB, $F_1 = 250 \rightarrow 290$ Hz, $AC = 10 \rightarrow 4$ dB, a többi adat nem változik.

Alveoláris, posztalveoláris, zöngétlen középreshang

[ʃ]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji és VCV-helyzetben kb. 130 ms, VC #helyzetben kb. 200 ms.

b) Hangsor elején a zörej nulla intenzitásról indul, és kb. 30 ms múlva maximumot ér el, amely kb. 8 dB-lal kisebb, mint a V intenzitása. Ezután a hang intenzitása állandó, majd a képzés utolsó 30 ms-ában kb. 15 dB-t csökken. VCV- és VC #helyzetben a hang intenzitása kb. 12 dB-lal kisebb, mint a V erőssége, a továbbiakban pedig hasonlóan alakul, mint a CV-helyzetben. A zörej 1800 Hz és 6000 Hz között tartalmaz frekvencia-komponenseket, az intenzív zörejjóc a 2500–3500 Hz-es sávban található. A [ʃ] képzésekor ajkainkkal már a V-re artikulálunk, így a V hatással van a zörejsáv alsó határára a következők szerint:

$$CK_{1cs} = \frac{VF_2 + 1800}{2} \text{ Hz} \quad (\text{a zörej mindig az 1800 Hz-es értékről kezd hajlani a } VF_2 \text{ felé})$$

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban:

$$\text{CV-helyzetben} \quad F_{1cs} = 400 \text{ Hz} \quad (\text{ha } VF_1 \geq 400 \text{ Hz})$$

$$F_{1cs} = VF_1 \text{ Hz} \quad (\text{ha } VF_1 < 400 \text{ Hz})$$

$$F_{2cs} = \frac{VF_2 + 1800}{2} \text{ Hz}$$

$$\text{VC-helyzetben} \quad F_{1cs} = \text{ugyanúgy alakul, mint a CV-helyzetben}$$

$$F_{2cs} = \frac{2VF_2 + 1800}{3} \text{ Hz}$$

A [f] hang szintézise # CV-helyzetben

1. szelet: T = 30 ms, AC = 0 → 20 dB, K₁ = 1800 Hz, K₂ = 4000 Hz, AK = 16 dB.

2. szelet: T = 70 ms, AC = 20 dB, K₁ = 1800 Hz, K₃ = 4000 Hz, AK = 16 dB.

3. szelet: T = 30 ms, AC = 20 → 10 dB, K₁-et a V felé hajlítjuk, K₂ = 4000 Hz, AK = 16 dB.

Palatális, zöngés középréshang

[j]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben kb. 100 ms, VCV-helyzetben nehezen meghatározható, kb. 40 ms, hangsor záró helyzetben kb. 80 ms.

b) Hangsor elején a hang intenzitása a képzés 70 %-áig állandó, kb. 10 dB-lal kisebb, mint a V-jé, a V előtti átmeneti szakaszban az intenzitás 4–6 dB-nyit emelkedik. VCV- és VC #helyzetben a [j] intenzitása beleolvad a V-be, a két hang nehezen elkülöníthető az intenzitásgörbe alapján. Formánsszerkezete: F₁ = 250 Hz, F₂ = 2400 Hz, F₃ = 3000 Hz. A formánsok sávszélessége nagy.

c) A [j] hang formánsai folyamatosan hajlanak a V formánsai felé, a két hang határát nehéz megállapítani.

A [j] hang szintézise # CV-helyzetben

1. szelet: T = 40 ms, AO = 0 → 4 dB, F₁ = 220 Hz, F₂ = 2400 Hz, F₃ = 2900 Hz, B₁ = 100 Hz, B₂ = 155 Hz, B₃ = 500 Hz.

2. szelet: T = 30 ms, a többi adat nem változik.

3. szelet: T = 30 ms, AO = 4 → 8 dB, a formánsokat a V-hez kezdjük hajlítani, a sávszélességeket lecsökkentjük.

Palatális, zöngétlen réshang

[ç]

Az akusztikai szerkezet jellemzése

a) A hang csak hangsor végi helyzetben fordul elő, időtartama kb. 230 ms.

b) A hang zöngétlen zárhanggal párosulva keletkezik. A zörej intenzitása a zárhang utáni ponton kb. 10 dB-lal alacsonyabb, mint a V-jé, majd egy 12 dB/100 ms-os csökkenő szakasza van. Ezután az intenzitás nem változik, majd a képzés utolsó 30 ms-ában nullára csökken. A zörej frekvenciasávja 2000 Hz-től 6000 Hz-ig terjed, és erősen változik a képzés folyamán. A hang kezdetén az alsó zörejgóc 2000 Hz-en van, majd a következő 100 ms-ban 3100 Hz-re, a felső 3000 → 4000 Hz-re csúszik fel. A következő 100 ms-ban a zörejgócok nem változnak, a hang végére azonban az alsó kb. 300 Hz-cel, a felső kb. 200 Hz-cel lejjebb csúszik.

c) A V formánsai úgy alakulnak, mint a [ç] hang előtti zárhanghoz.

A [ç] hang szintézise

1. szelet: T = 100 ms, AC = 14 → 10 dB, K₁ = 2000 → 3000 Hz, K₂ = 3000 → 4000 Hz.

2. szelet: $T = 100$ ms, $AC = 10$ dB, $K_1 = 3000$ Hz, $K_2 = 4000$ Hz.

3. szelet: $T = 30$ ms, $AC = 8 \rightarrow 0$ dB, $K_1 = 2900 \rightarrow 2700$ Hz, $K_2 = 4000 \rightarrow 2800$ Hz.

Pharyngális, zöngés középréshang

[ŋ]

Az akusztikai szerkezet jellemzése

a) A hang csak VCV-helyzetben fordul elő, időtartama kb. 90 ms.

b) A hang zöngés és zörejes elem kombinációjából áll, a V_1 utáni szakaszban gyenge intenzitású, mintegy 15 dB-lel kisebb, mint a V, majd erőssége fokozatosan nő a V_2 felé. A zöreje a zöngés elemmel párhuzamosan növekszik. A zöngés elem formánsai: $F_1 = 250$ Hz, F_2 és F_3 értékei a közrefogó magánhangzók ugyanezen formánsait összekötő egyenes mentén mozognak, azonban intenzitásuk nagyon kicsi. A zöreje frekvenciaszerkezete hasonló, mint a [h] hangnál.

c) A V-k formánsai közül csak az F_1 mutat némi hajlást a 300 Hz-es pont felé.

A [ŋ] hang szintézise

$T = 90$ ms, $AO = 3$ dB, $AC = 10 \rightarrow 20$ dB, a zöreje ugyanaz, mint a [h] hangnál.

Pharyngális, zöngétlen középréshang

[h]

Az akusztikai szerkezet jellemzése

a) A hangot csak # CV-helyzetben ejtjük, időtartama kb. 110 ms.

b) A hang intenzitása nulláról indul és a képzés végéig fokozatosan növekszik. A V előtti szakaszban intenzitása kb. 30 dB-lel alacsonyabb, mint a magánhangzóé. A zöreje frekvenciaszerkezetére az jellemző, hogy egyetlen, kb. 1000 Hz-es sávzélességű zorejsávból áll, amelynek középső frekvenciáját a VF_2 -je határozza meg, a következők szerint:

$$CK_{1cs} = VF_2 \text{ Hz}$$

c) A V formánsai nem hajlanak, mivel a [h] hang ejtésekör már a következő V-re artikulálunk.

A [h] hang szintézise

1. szelet: $T = 60$ ms, $AC = 0 \rightarrow 6$ dB, $K_1 = 1100$ Hz, $K_2 = 2540$ Hz. Ha VF_2 1500 Hz, akkor a K_2 értékét arányosan növeljük, a K_1 marad. A zöreje színezetét az $AK = 25$ dB-es értékkel adjuk meg.

2. szelet: $T = 50$ ms, $AC = 6 \rightarrow 14$ dB, a többi adat nem változik.

A V amplitúdóját fokozatosan, de rövid idő alatt kell felvezérelni, $AO = 0 \rightarrow 14$ dB.

Veláris, zöngétlen réshang

[x]

Az akusztikai szerkezet jellemzése

a) A hang # CV-helyzetben nem fordul elő, VCV-helyzetben az időtartama kb. 160 ms, hangsor záró helyzetben pedig kb. 260 ms.

b) A hang intenzitása a képzés alatt állandó, 10–15 dB-lel alacsonyabb, mint a V-é. Hangsor végi helyzetben a hang a V-nél kb. 20 dB-lel alacsonyabb intenzitásértékről indul, majd egy 8 dB/50 ms-os emelkedés következik. Az intenzitás a képzés végére fokozatosan nullára csökken. A zöreje frekvenciaszerkezete a V típusától függ, hasonlóan,

mint a [k] hangnál. Ha a V_1 és V_2 F_2 -je értékben eltérő egymástól, akkor a [x] hang zöreje is értelemszerűen változik a képzés folyamán. Ha a zörejt frekvenciája magas (kb. 1500 Hz feletti), akkor a hangot „ich-laut”-nak, ha alacsony akkor „ach-laut”-nak szokták nevezni, de mindkét hang ugyanazon zöngétlen veláris réshang két változata.

c) A V formánsai úgy csatlakoznak a [x] hanghoz, mint a [k]-hoz.

A [x] hang szintézise VCV-helyzetben

1. szelet: T = 60 ms, AC = 5 → 10 dB, a zörejt úgy alakítjuk ki, mint a [k] hangnál.

2. szelet: T = 60 ms, AC = 10 dB, a többi adat nem változik. A zörejt – ha kell –, hajlítjuk.

3. szelet: T = 40 ms, AC = 10 dB, a zörejt tovább hajlítjuk.

Bilabiális, labiodentális, zöngés zár-rés hang

[b̥v]

Az akusztikai szerkezet jellemzése

a) A hang csak VCV-helyzetben fordul elő, időtartama kb. 180 ms.

b) A hang a [b] és [v] hangok igazodása révén keletkezik úgy, hogy a [b] zárfelpattanási zöreje elmarad és helyébe a [v]-szerű hangelem kerül. A képzés első 140 ms-ában az intenzitás kb. 15 dB-lal kisebb, mint a V_1 -é, majd a [v] elem 4–6 dB-nyit növekvő intenzitású szakasza következik. A hang első részében csak a fojtott zöngé szerepel, majd belépnek a [v]-re jellemző formánsok és a zörejelem is.

c) A V_1 formánsai úgy hajlanak a [b̥v] hanghoz, mint a [b]-hez, a V_2 -é pedig, mint a [v]-hez.

A [b̥v] hang szintézise

1. szelet: T = 100 ms, AO = 3 dB.

2. szelet: T = 20 ms, AO = 3 dB, $F_1 = 200 \rightarrow 346$ Hz, $F_2 = 500 \rightarrow 100$ Hz, $F_3 = 1050 \rightarrow 2540$ Hz, $B_1 = 12 \rightarrow 96$ Hz, $B_2 = 100$ Hz, $B_3 = 500 \rightarrow 750$ Hz.

3. szelet: T = 20 ms, $F_2 = 1000 \rightarrow 1400$ Hz, $B_2 = 100 \rightarrow 300$ Hz, a többi adat nem változik.

4. szelet: T = 40 ms, AO = 4 → 6 dB, a formánsokat tovább csatoljuk a V_2 -höz, az AC = 10 → 2 dB, AK = 20 dB, a zörejt ugyanaz, mint az [f] hangnál.

A V_1 és V_2 csatolását lásd a [b] és [v] hangoknál!

Bilabiális, labiodentális, zöngétlen zár-rés hang

[p̥f]

Az akusztikai szerkezet jellemzése

a) Hangsor eleji helyzetben a néma fázist nem érzékeljük, a hang időtartama 110 ms, VCV- és VC #helyzetben a hang kb. 250 ms, illetve kb. 300 ms, amiből a néma fázis kb. 100 ms.

b) A hang a [p] és [f] kombinációjából keletkezik, intenzitása hirtelen növekszik, kb. 30 ms alatt éri el a maximumot, melynek intenzitásszintje kb. 12 dB-lal alacsonyabb, mint a V-é. A zörejt kb. 300 Hz-től 4000 Hz-ig tartalmaz frekvenciakomponenseket.

c) A V formánsai úgy csatlakoznak a C-hez, mint a [p], illetve a [f] hangnál.

A [p̥f] hang szintézise # CV-helyzetben

1. szelet: T = 5 ms, AC = 14 dB, $K_1 = 1500$ Hz, $K_2 = 2800$ Hz.

2. szelet: $T = 100$ ms, $AC = 14 \rightarrow 20$ dB, $K_1 = 1800$ Hz, $K_2 = 2600$ Hz, $AK = 18$ dB.

Dentális, dentalveoláris, zöngés zár-rés hang

[d̥z]

Az akusztikai szerkezet jellemzése

a) A hang # CV-helyzetben csak nevekben fordul elő, VCV- és VC #-helyzetben az időtartama kb. 200 ms, illetve kb. 240 ms. A hang két szakaszból áll, a fojtott zöngéből, amely kb. 100 ms-os és a [z] elemből.

b) A zöngé úgy kapcsolódik a V-hez, mint a [d] esetében, intenzitása a [z] elem előtti szakaszban kb. 18 dB-lal alacsonyabb, mint a V-jé. A [z] elem frekvenciaszerkezete olyan, mint a [z] hangé.

c) A V formánsai úgy csatlakoznak a [d̥z] hanghoz, mint a [d] és [z]-hez.

A [d̥z] hang szintézise VCV-helyzetben

1. szelet: $T = 100$ ms, $AO = 4$ dB, $F_1 = 250 \rightarrow 206$ Hz.

2. szelet: $T = 40$ ms, $AO = 4$ dB, $AC = 0 \rightarrow 14$ dB, $K_1 = 4000$ Hz, $K_2 = 8000$ Hz.

3. szelet: $T = 60$ ms, $AO = 0 \rightarrow 6$ dB a formánsokat a [z] hang szerint adjuk meg és hajlítjuk a V felé, $AC = 14 \rightarrow 10$ dB, $K_1 = 4000$ Hz, $K_2 = 8000$ Hz.

Dentális, dentalveoláris, zöngétlen zár-rés hang

[t̥s]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji helyzetben kb. 80 ms, mivel a néma fázist nem érzékeljük, VCV- és VC #-helyzetben az időtartama 130/60 ms, illetve 250/120 ms.

b) A zörej intenzíven indul, kb. 14 dB-lal kisebb intenzitással, mint a V erőssége, a V előtti közvetlen szakaszban elveszti erejét, erőssége szinte nulla lesz. Hangsor záró helyzetben a hang megnyúlik, és intenzitása fokozatosan csökken nullára. A zörej 5000 Hz-től 9000 Hz-ig tartalmaz frekvenciakomponenseket, amelyek eloszlása viszonylag egyenletes. A zörej frekvenciaszerkezetére a V típusának nincs módosító hatása.

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban:

CV-helyzetben $VF_{1cs} = 400$ Hz (ha $VF_1 \geq 400$ Hz)

$VF_{1cs} = VF_1$ Hz (ha $VF_1 < 400$ Hz)

$VF_{2cs} = \frac{VF_2 + 2000}{2}$ Hz

VC-helyzetben $VF_{1cs} = 450$ Hz (ha $VF_1 \geq 450$ Hz)

$VF_{1cs} = VF_1$ Hz (ha $VF_1 < 450$ Hz)

$VF_{2cs} = \frac{2VF_2 + 2000}{3}$ Hz

A [t̥s] hang szintézise # CV-helyzetben

1. szelet: $T = 20$ ms, $AC = 10 \rightarrow 18$ dB, $K_1 = 3900$ Hz, $K_2 = 6700$ Hz.

2. szelet: $T = 30$ ms, $AC = 18$ dB, a többi adat nem változik.

3. szelet: $T = 30$ ms, $AC = 18 \rightarrow 2$ dB, a többi adat nem változik.
A V csatoló szeletében az intenzitást fokozatosan növeljük, $AO = 0 \rightarrow 14$ dB.

Alveoláris, posztalveoláris, zöngés zár-rés hang
[d̥ʒ]

Az akusztikai szerkezet jellemzése

- a) A hang időtartama általában 120–140 ms. Ebből a fojtott zöngé kb. 80 ms.
b) A hang a [d] hangra emlékeztető fojtott zöngéből és a [ʒ], majd [t̥ʃ] hang elemeiből tevődik össze. A zörej intenzitása a képzés végére olyan nagy, hogy a zöngés elem szinte teljesen elnyomja.
c) A V formánsai úgy hajlanak a [d̥ʒ] hanghoz, mint a [d], illetve a [t̥ʃ] hangokhoz.

A [d̥ʒ] hang szintézise # CV-helyzetben

1. szelet: $T = 30$ ms, $AO = 4$ dB, $F_1 = 245 \rightarrow 206$ Hz.
2. szelet: $T = 40$ ms, $AO = 4 \rightarrow 2$ dB.
3. szelet: $T = 30$ ms, $AO = 0 \rightarrow 4$ dB, $K_1 = 2400$ Hz, $K_2 = 3500$ Hz.
4. szelet: $T = 40$ ms, $AO = 4$ dB, $AC = 16 \rightarrow 4$ dB, az F_1 -et 200 Hz-ről csatoljuk a V-hez, F_2 és F_3 ugyanaz, mint a [ʒ] hang hasonló szeletében.

Alveoláris, posztalveoláris, zöngétlen zár-rés hang
[t̥ʃ]

Az akusztikai szerkezet jellemzése

- a) A hang időtartama hangsor eleji helyzetben kb. 80 ms, mivel a néma fázist nem érzékeljük, VCV- és VC #helyzetben 140/70 ms, illetve 200/100 ms.
b) A zörej intenzíven indul, erőssége kb. 10 dB-lel kisebb, mint a V-é, majd kb. 10 dB-nyit csökken, és a V előtti közvetlen szakaszban intenzitása közel nullára esik. Ez egy néma fázis jellegű, kb. 20 ms-os elemként jelentkezik a [t̥ʃ] hangban. A zörej 1800 Hz-től kb. 7000 Hz-ig tartalmaz frekvenciakomponenseket, intenzívebb góc található az 1800–3000 Hz-es sávban. A góc alsó frekvenciája a követő V típusától függően csúszkál a következők szerint:

- ha $VF_2 \leq 1500$ Hz, akkor a zörejgóc alsó pontja 1700–1800 Hz-en van,
– ha $VF_2 > 1500$ Hz, 2300–2400 Hz-en van.

c) A V formánsainak csatlakozási értéke a C-hez az átmeneti fázisokban:

CV-helyzetben $VF_{1cs} = 350$ Hz (ha $VF_1 \geq 350$ Hz)
 $VF_{1cs} = VF_1$ Hz (ha $VF_1 < 350$ Hz)

$$VF_{2cs} = \frac{VF_2 + 1800}{2} \text{ Hz (ha } VF_2 \leq 1500 \text{ Hz)}$$

$$VF_{2cs} = \frac{VF_2 + 2300}{2} \text{ Hz (ha } VF_2 > 1500 \text{ Hz)}$$

VC-helyzetben $VF_{1cs} = 450$ Hz (ha $VF_1 \geq 450$ Hz)

$$VF_{1cs} = VF_1 \text{ Hz} \quad (\text{ha } VF_1 < 450 \text{ Hz})$$

$$VF_{2cs} = \frac{2VF_2 + 1800}{3} \text{ Hz} \quad (\text{ha } VF_1 \leq 1500 \text{ Hz})$$

$$VF_{2cs} = \frac{2VF_2 + 2300}{3} \text{ Hz} \quad (\text{ha } VF_1 > 1500 \text{ Hz})$$

A [tʃ] hang szintézise # CV-helyzetben

1. szelet: T = 20 ms, AC = 10 → 18 dB, K₁ = 1800 Hz, K₂ = 2300 Hz, AK = 16 dB.

2. szelet: T = 40 ms, AC = 18 → 8 dB, a többi adat nem változik.

3. szelet: T = 20 ms, AC = 6 → 0 dB, a többi adat nem változik.

A V csatoló szeletében az intenzitást fokozatosan növeljük, AO = 0 → 14 dB.

Dentális, dentalveoláris, zöngés oldalsóhang

[l]

Az akusztikai szerkezet jellemzése

a) A hang időtartama hangsor eleji és hangsor végi helyzetben 80–120 ms között van az ejtéstől függően, VCV-helyzetben kb. 80 ms.

b) A hang intenzitása hangsor eleji helyzetben a képzés feléig emelkedik, amikor kb. 8 dB-lal gyengébb, mint a V. Ezután az intenzitás alakulását a V típusa befolyásolja a következők szerint:

– ha $VF_2 > 1000$ Hz, akkor a [l] intenzitása a V előtt erősen csökken,

– ha $VF_2 < 1000$ Hz, akkor a [l] intenzitása a V előtt nem változik.

VCV-helyzetben a [l] intenzitása kb. 8 dB-lal kisebb, mint a V₁-é, hangsor záró helyzetben az intenzitás a hang feléig növekszik, mintegy 3–5 dB-nyit, majd ezután fokozatosan nullára csökken. A [l] hang formánszerkezete: F₁ = 350 Hz, az F₂ a V típusától függ, F₃ = 2700 Hz. A C és VF₂ között a következő kölcsönhatás mutatható ki:

– ha $VF_2 < 1000$ Hz → CF₂ = 1300 Hz

– ha $1000 \text{ Hz} \leq VF_2 \leq 1600 \text{ Hz}$ → CF₂ = VF₂ – 100 Hz

– ha $VF_2 > 1600$ Hz → CF₂ = 1600 Hz

VCV-helyzetben a [l] F₂-je, ha a közrefogó magánhangzók F₂-je között nagy különbség van, akkor az előbbi sávban mozogva szinte összeköti ezeket.

c) A V formánsainak csatlakozási értékei a C-hez az átmeneti fázisokban:

CV- és VC-helyzetben $F_{1cs} = 500 \text{ Hz}$ (ha $VF_1 \geq 500 \text{ Hz}$)

$F_{1cs} = VF_1 \text{ Hz}$ (ha $VF_1 < 500 \text{ Hz}$)

$F_{2cs} = \frac{VF_2 + 1600}{2} \text{ Hz}$ (ha $VF_2 > 1600 \text{ Hz}$)

$F_{2cs} = VF_2 \text{ Hz}$ (ha $1000 \leq VF_2 \leq 1600 \text{ Hz}$)

$$F_{2cs} = \frac{VF_2 + 1000}{2} \text{ Hz} \quad (\text{ha } VF_2 < 1000 \text{ Hz})$$

CV-helyzetben a VF_1 hajlása gyors (kb. 20 ms), az VF_2 -jé lassú (kb. 80 ms). A csatlakozó V hirtelen intenzitásugrással indul a [l] után, az ugrás mértéke 4–10 dB között van a V-től függően. Ez az intenzitásugrás lényeges jellemzője a [l] → V átmenetnek.

A [l] hang szintézise # CV-helyzetben

1. szelet: T = 40 ms, AO = 0 → 10 dB, $F_1 = 350$ Hz, $F_2 = 1300$ Hz, $F_3 = 2700$ Hz.

2. szelet: T = 40 ms, AO = 10 → 12 dB, az F_2 -t hajlítjuk a V által meghatározott érték felé, a többi adat nem változik.

3. szelet: T = 40 ms, AO-t csökkentjük, vagy állandó értéken tartjuk a V típusától függően, az F_2 -t a VF_2 felé simítjuk, a többi adat nem változik.

A V csatoló szeletében intenzitásugrást kell létrehozni, AO = 16 dB.

Dentális, dentialveoláris, zöngés tremuláns

[r]

Az akusztikai szerkezet jellemzése

a) A hang hangsor eleji helyzetben kb. 110 ms, ebből 90 ms a zár előtti zöngés elem és kb. 20 ms a zár ideje, VCV-helyzetben és VC #helyzetben a hang időtartama kb. 40–80 ms a perdületek számától függően, illetve kb. 100 ms.

b) Hangsor eleji helyzetben a [r] zöngés kezdő elemből és zárból áll. Hangsúlyozott ejtésnél több perdület is előfordulhat. A hang kezdeti, zöngés elemének intenzitása hirtelen növekszik, és a maximumnál csak 2–3 dB-lel gyengébb, mint a V. A továbbiakban a zár kialakulásáig az intenzitás nullára csökken, a zár ideje alatt hangképzés nincs. A zár feloldódása után a hang ismét növeli intenzitását, közel a V szintjéig. Ha több pergetés van, akkor a folyamat ismétlődik előlről. Minél több a pergetések száma, annál rövidebb a zöngés elem időtartama, amelynek formánsszerkezete: $F_1 = 500$ Hz, $F_2 = 1400$ Hz, $F_3 = 2800$ Hz. Az F_1 és F_2 sávzélessége kb. 100 Hz. A [r] hang formánsszerkezetére a V-nek nincs lényeges hatása, ha a VF_2 magas (pl. az [i]-nél), akkor CF_2 feljebb csúszik. A V előtti zöngés zárfeloldódás egybeolvad a V kezdeti fázisával, és formánsai a V formánsai felé tolódnak. VCV-helyzetben a [r] hang a V_1 -hez kapcsolódó zöngés elemmel kezdődik, majd szerkezete ugyanolyan, mint a korábbiakban volt.

c) A V formánsai folyamatosan csatolódnak a [r] hang formánsaihoz, a formánsok hajlása nagyrészt a V-ben történik, az átmeneti fázis a tiszta fázisig viszonylag hosszú, mintegy 100 ms.

A [r] hang szintézise # CV-helyzetben

1. szelet: T = 45 ms, AO = 0 → 12 dB, $F_1 = 500$ Hz, $F_2 = 1400$ Hz, $F_3 = 2800$ Hz, $B_1 = 90$ Hz, $B_2 = 90$ Hz, $B_3 = 500$ Hz.

2. szelet: T = 50 ms, AO = 12 → 2 dB, a többi adat nem változik.

3. szelet: T = 60 ms, AO = 0 → 12 dB, a formánsokat az előbbi értékekről enyhén hajlítjuk a V formánsai felé.

A V csatoló szeletében a formánsokat tovább hajlítjuk a tiszta fázishoz, AO = 12 → 14 dB.

Az analízis és a szintézis tapasztalatainak összegezése

A magyar mássalhangzók és a # CV, VCV, VC # hangkapcsolódási formák akusztikai szerkezetének, analízis szintézissel történt meghatározásán túl, a munka során sok olyan fontos adatot és törvényszerűséget sikerült feltárni, amely lényegesen hozzájárul a magyar mássalhangzók és a mássalhangzó–magánhangzó–kapcsolódások – mint hangsorépítő elemek – akusztikai szerkezeti képének teljessé tételéhez, valamint a szintézis gyakorlati munkájához.

Magánhangzók

Hangsorépítés során a magánhangzóknek három előfordulása van, hangsor eleji, hangsor belseji és hangsor záró helyzet. Ezen helyzetekben a magánhangzók kialakulási és lecsengési – befejező – intenzitásmenete különbözőképpen alakul. Hang kezdésénél az intenzitás meredeken fut fel (kb. 30 dB/20 ms), a magánhangzó befejezésekor a néma fázissal rendelkező és a gyenge intenzitású hangok előtt (pl. a [h], [f], [t]) az intenzitáscsökkenés lassúbb (kb. 30 dB/60 ms), hangsor záró helyzetekben a magánhangzó még ennél is lassabban hal el, intenzitásésése mintegy 30 dB/100 ms.

A VCV típusú elemekben, ha a két magánhangzó formánsai nagyon különböznek egymástól, számos esetben a V_2 visszahat a V_1 formánsainak mozgására (pl. a [g] esetében). Ezért az azonos idejű szintetizáló rendszerek kidolgozásánál nem elégséges csak a hangsor egymás utáni közvetlen elemeit vizsgálni, hanem a hangsorépítés legalább három egymás mellett álló elemét kell összehasonlítani, hogy a megfelelő akusztikai paramétereket is tudjuk a számítógéppel számítani. Ez a jelenség komoly segítséget adhat a beszédfelismerés kutatásához is, mivel a háromelemű kapcsolatok első elemének vizsgálatakor nagy valószínűséggel esetleg már előre meg tudjuk határozni a hangsor következő vagy az azután következő elemének típusát.

A hangsorbeli magánhangzók időtartamának helyes megválasztása – a hangsor többi eleméhez képest – döntően befolyásolja a szintetizált szó vagy mondat hangzását.

Automatikus hangsorépítésnél annál természetesebben kell az időszerkezetre vonatkozó szabályrendszert kidolgozni, minél nagyobb szótagszámú hangsort akarunk előállítani. Ezért az automatikus szintézishez olyan kompromisszumot kell kötnünk, hogy egy-egy magánhangzóból több hosszúságút kell az adatbázisban tárolnunk, és ezekből kell a számítógépnek kiválasztani azt az elemet, amelyik a hangsor adott részébe a szabályok szerint beleilleszthető. Tapasztalataim szerint a szintézissel előállított magyar hangsorok időtartamviszonyainak behatóbb vizsgálatából származó eredmények segíthetik a jövőben a beszédfelismerésre irányuló kutatásokat olyan formában, hogy az egyes hangok tartamának méréséből és egymás utáni összeállításából olyan időábrákat, építőelemkombinációkat állíthatunk össze, amelyek jellemzőek lehetnek a magyarra. Így számítógéppel esetleg megállapítható egy beszédfolyamból, hogy magyar vagy más nyelvű szövegről van-e szó.

Mássalhangzók

A zöngés–zöngétlen zárhangpárok mind időszerkezetükben, mind pedig frekvenciaszerkezetükben lényegesen különböznek egymástól, annak ellenére, hogy képzési he-

lyük megegyezik. Ezért, ha ezen zárhangok akusztikai szerkezetét pontosan, a fizikai valósághoz hűen akarjuk megadni, akkor a zöngés és a zöngétlen zárhangokat külön-külön kell tárgyalnunk.

A hangsor eleji és VCV-kapcsolatban szereplő [p], [t], [k] hangok intenzitás-, idő- és frekvenciaszerkezetét az 1. táblázatban hasonlítom össze.

1. táblázat

a hang # CV- és VCV- helyzet- ben	a zár-felpattanás				a követő V hatása a zár-felpattanásra
	V-hez viszonyított intenzi- tása (-dB)	időtartam (ms)	távolsága a követke- ző V-től, VOT (ms)	frekvencia- szerkezte	
[p]	10–15	4–5	6–8	zörej jellegű, az 500–5000 Hz-es sávot egyenletesen tölti ki, góccok nincsenek	a V nem befolyásolja a zár-felpattanás frekvenciaszerkezetét
[t]	10–15	10–12	12–15	a zörejsáv szűkül, 1600–4000 Hz, két zörejgóc van; 1600–2000 Hz és 4000 Hz	a V gyengén visszahat a [t] képzésére, az alsó zörejgóc az 1600–2000 Hz-es sávban mozog a $V F_2$ -jétől függően
[k]	15–18	20–30	30–40	a felpattanás két határozott zörejgócból áll: az alsó 800 és 2400 Hz között mozog, a felső 4500 Hz-en van	a V erősen befolyásolja a [k] képzését. Az alsó zörejgóc helye mindig a $V F_2$ -je felett van kb. 10 %-kal

A [p], [t], [k] hangoknál a szintézisből eredő tapasztalatok azt mutatták, hogy a néma fázisok időtartama egységes lehet mindhárom hangnál. Ez a hangok hangzását nem befolyásolja. Ezért a táblázatban csak a zár-felpattanások akusztikai szerkezetét hasonlítottam össze. A vizsgálatok során kitűnt továbbá, hogy a hangsor végi [p], [t], [k] hangok akusztikai szerkezte mind intenzitás-, mind idő-, mind pedig frekvenciaszerkezetében különbözik a hangsor eleji vagy hangsor belseji helyzetűekétől. A hangsor végi [p],

[t], [k] hangok zárfelpattanása kisebb intenzitású, a felpattanás és a lecsengő zörej időtartama lényegesen megnyúlik, és ez fontos tényező a hang hangzása szempontjából. Ezen hangok zárfelpattanásainak akusztikai szerkezetét a 2. táblázatban foglalom össze.

2. táblázat

a hang VC #-helyzetben	a zárfelpattanás				a megelőző V hatása a zárfelpattanásra
	a megelőző V-hez viszonyított intenzitása (-dB)	időtartama (ms)	lecsengő részének időtartama (ms)	frekvenciaszerkezete	
[p]	15–20	10–15	40–60	zörej jellegű ugyanolyan, mint a #CV-, VCV-helyzetben	a V-nek nincs hatása a zörej frekvenciájára
[t]	25	20–25	70–80	ugyanaz, mint a #CV-, VCV-helyzetben, csak az alsó zörejgóc mozgási sávja szűkebb	a V gyengén hat a zörej alsó gócéra, amely 1700–1980 Hz-es sávban mozog a $V F_2$ -jétől függően
[k]	20–25	a zárfelpattanás nem különíthető el a lecsengő szakasz zörejétől	100–120	ugyanaz, mint a #CV-, VCV-helyzetben, csak az alsó zörejgóc helye szűkebb sávban mozog	a V erősen hat a zörej alsó gócéra helyére, ami 900–1800 Hz-es sávban mozog a $V F_2$ -jétől függően

A [b], [d], [g] hangok szintézisének tapasztalatai arra mutatnak, hogy a zöngés zárhang zárfelpattanás előtti úgynevezett zöngé része lényeges szerepet játszik a hang minősége szempontjából. Ez a zöngés rész más és más intenzitás-, időszerkezettel rendelkezik a [b], a [d] és a [g] hang esetében. A 3. táblázatban a hangsor eleji CV elemekben szereplő [b], [d], [g] mássalhangzók akusztikai szerkezetét hasonlítom össze. A hangsor végi [b], [d], [g] hangokból a zárfelpattanás után a levegőkiáramlás és a pillanatnyi kép-

a hang # CV- hely- zetben	a z ö n g e			a z á r f e l p a t t a n á s			a követő V hatása a zárfelpattanásra
	V-hez vi- szonyított intenzitása (-dB)	intenzitás- szerkezete	időtár- tama (ms)	időtár- tama (ms)	intenzitás- szerkezete	frekvencia- szerkezete	
[b]	15–20	nulláról indul és fokozatosan nő, erős ejtésnél a képzés harmadánál maximuma van	60–90	10–15	a zöngéből folytatódva rohamosan nő, szinte eléri a V normál intenzitását	periodikus jellegű, három intenzív góccal jellemezhető: 350 Hz, 600–1600 Hz-es sáv és 2400 Hz	a középső góc a 600–1600 Hz-es sávban mozog a $V F_2$ -jétől függően
[d]	15–20	intenzíven indul és intenzitása nem változik a zárfelpattanásig	60–90	10–15	a zöngéből alig 3–4 dB-nyit emelkedik ki	periodikus jellegű, három góccal jellemezhető: 400 Hz 1300–1800 Hz-es sáv, 2600 Hz	a középső góc az 1300–1800 Hz-es sávban mozog a $V F_2$ -jétől függően
[g]	15–20	nulláról indul, 25 ms-nál maximuma van, majd az intenzitás csökken a zárfelpattanásig	100–110	20–30	a zöngéből kiindulva fokozatosan növekszik és folyamatosan megy át a V-be	periodikus, három frekvenciagóccal jellemezhető: 300 Hz, 600–2400 Hz-es sáv, 2600 Hz; jellegzetes színezetét a [k]-hoz hasonló zörej adja meg	a V erősen hat a [g]-re, a középső góc frekvenciája mindig a $V F_2$ -je felett van kb. 10 %-kal

zöszervi konfiguráció eredményeként minden esetben erősebb vagy gyengébb „sva”-szerű elem jellemzi a hangot. Ez határozza meg a hang hangsor végének megszokott hangzását.

A [p], [t], [k] és [b], [d], [g] hangok szintézise kimutatta, hogy a zárfelepattanások létrehozása során nem szabad a szintetizátor zöngés és zörej áramköreit mereven használni, ragaszkodva a megszokott elnevezések – pl. zárfelepattanási zörej – szó szerinti értelmezéséhez. Figyelembe kell venni ugyanis, hogy a természetes beszéd hangjait egy biológiai rendszerrel (idegimpulzusok, izmok, tüdő stb.) hozzuk létre, a szintézis során viszont egy technikai rendszerrel (áramkörök, oszcillátor, zajgenerátor stb.) kell a beszédhangok akusztikai szerkezetét felépítenünk. Ebből következik, hogy pl. helytelen, ha a [b] vagy [d] hang zárfelepattanási zörejét – mivel az zöngés jellegű – a szintetizátor zöreje (AC, AK, K_1 , K_2) áramkörével akarjuk létrehozni. Ugyanakkor szintén nem vezet megfelelő eredményre, ha pl. a magánhangzó előtt álló [t] hang zárfelepattanását csak a zörej körrel akarjuk létrehozni. Az egyes zárhangok zárfelepattanási zörejeinek szintetizálásánál törekedjünk arra, hogy a valóságban lezajló fizikai folyamatot, vagyis a zárfelepattanás akusztikai tartalmát a szintetizátor zöngés és zörej áramkörével a lehető legjobban utánozzuk le. Például egy [b] hang periodikus jellegű zárfelepattanásának előállításánál csak a zöngés kört (AO, FO, F_1 , F_2 , F_3 , B_1 , B_2 , B_3) alkalmazhatjuk. Hangsor eleji vagy belseji [t] hangnál viszont a zöngés és zörej kört együttesen kell használnunk, ugyanis a [t] zárfelepattanásának frekvenciagócait csak a zöngés körrel tudjuk megfelelő intenzitásban előállítani. Ha a [t] zárfelepattanását reprezentáló szelet időtartamát megfelelően rövidre választjuk, akkor a zöngés elem is egy hirtelen intenzitású zörejnek felelhet meg. Ugyanakkor, ha hangsor végi [t] hangot akarunk előállítani, akkor csak a zörejkört szabad használnunk, mivel a hang akusztikai szerkezetére csak a hosszúra nyúlt tiszta zörej a jellemző. Hasonló megfontolások alapján kell a többi zárhang szintézisét is elvégezni.

A [m] és [n] hangok vizsgálatai azt mutatják, hogy a két hang akusztikai szerkezeete közel azonos. Intenzitásuk a hangsoron belül közel egyforma, frekvenciaszerkezetükben sincs olyan lényeges különbség, mint például két különböző magánhangzóé között. Ezért nehéz a [m] és [n] hangot szintetizálni. Tapasztalataim szerint hangsoron belül a két hangot megkülönböztető jegyet az őket követő magánhangzó hordozza. A [m] utáni V-ben ugyanis a formánsmozgások – a tiszta fázisig – igen rövid idő alatt (kb. 20 ms) zajlanak le, míg a [n] utáni V-ben ugyanez a jelenség az előbbi érték 4–5-szöröse (80–100 ms). Az átmeneti fázisoknak ez a nagymértékű különbözősége határozza meg döntően, hogy a hangsorban [m] vagy [n] hangot hallunk.

A [m] és [n] hangok nazális hangszínezetét a nazális komponensen kívül az első formáns viszonylag nagy sáv szélessége (100–150 Hz) adja. Ezeknek a hangoknak a nazális színezete már a megelőző magánhangzóban kezd kialakulni, vagyis a V a képzési idejének már a második felében nazalizálódik. Ez az akusztikai képen úgy jelentkezik, hogy – feltehetően az uvula nyitódásának pillanatában – a magánhangzó első formánsának a sáv szélessége megnő. Ha a szintézis során a V-ben az F_1 sáv szélességét hirtelen megnöveljük, a hang nazalizálódni fog. Az előbbi feltevés tehát igaz lehet, noha ezt még röntgenvizsgálatokkal ellenőrizni kell.

A szintézissel újszerű adatokhoz jutottunk a magyar zörejh hangok akusztikai szerkezetét illetően is. Zörejh hangok előállításához az OVE III-ban két zörejgóc (K_1 , K_2) és egy antiformáns (AK) áll rendelkezésre. A természetes ejtésű zörejh hangok akusztikai szerkezetének vizsgálatából több adat állt rendelkezésünkre, mint amennyit a szintetizá-

4. táblázat

Természetes				Szintetizált					
a hang #CV- helyzet- ben	a V-hez vi- szonyított intenzitás (-dB)	időtartam (ms)	frekvenciasá- vok, góccok (Hz)	a V-hez vi- szonyított intenzitás (-dB)	időtartam (ms)	góccok		AK (dB)	Megjegyzés
						K ₁ (Hz)	K ₂ (Hz)		
[f]	20–25	100	1000–10000	20–25		1000–2000	3400	18	
[s]	5–10	130–170	3000–12000 góc: 6000– –9000	15–18	130–170	3800	8500	0	
[$\hat{t}s$]	10–15	50–90	5000–9000	10–20	50–90	3900	6800	0	
[ʃ]	5–10	150	1800–6500 góc: 2500– –3500	20–25	150	1800	4000	16	
[$\hat{t}ʃ$]	8–12	70–80	1800–7000 góc: 3000	20–25	70–80	2000–1700	3800	16	
[ç]	10	240	2000–4000	15	240	2000–3000	3000–4000	20–2	szóvégi helyzet
[h]	20–30	120	1000 Hz sáv- szélességű zöreje	35–40	120	1800	2600	25	
[x]	10–15	160	a V F ₂ -jétől függ	20–25	160	2400–2000	4500–4000	28	

torba lehet táplálni. Például a [f] hang zörejsávja a természetes beszédben az 1800–4000 Hz-es és a 4600–6000 Hz-es sávba esik. Arra, hogy ezekből a széles zörejsávokból melyek vagy mely részek fontosak a hang hangzása szempontjából, a szintézis adta meg a feleletet. A két zörejtű és az antifórmáns segítségével sikerült a zörejhangokat megfelelő hangzásban előállítani. Így megállapíthattuk azokat a zörejsávokat, góciókat, amelyek tipikusan jellemzőek az említett hangokra, s ezáltal világossá vált, hogy melyek azok a frekvenciakomponensek a hang spektrumában, amelyek nem játszanak lényeges szerepet a hangzás szempontjából.

A 4. táblázatban, összehasonlítással, megadom a zörejhangok természetes ejtésű változataiból elemzett intenzitás-csúcsokat, időtartamokat, frekvenciasávokat, góciókat és feltüntettem az ugyanezen hangok szintetizált változataiban mért intenzitásértékeket, valamint az előállításukhoz használt K_1 -, K_2 - és AK-értékeket. A táblázat csak a zörejek adatait tartalmazza, az esetleges néma fázisokat nem tüntettem fel!

Mint látható, a szintetizált zörejhangok intenzitása a hangsorban sok esetben lényegesen kisebb értékű, mint a természeteseké, mivel spektrumukban csak azok a legjellemzőbb frekvenciakomponensek szerepelnek, amelyeket a K_1 , K_2 és AK segítségével létrehoztunk. A szintetizált szavakban ezek a zörejhangok – kisebb intenzitásuk ellenére is – jól hangznak, ami bizonyítja, hogy az a feltevésünk, hogy a zörejhangok spektrumában is léteznek fontos, meghatározó erejű összetevők (mint pl. a magánhangzóknál az F_1 és F_2) és léteznek olyanok, amelyek a hang hangzását nem nagyon befolyásolják, esetleg csak az egyéni ejtéstől függenek (mint pl. a magánhangzóknál az F_3 , F_4 , F_5).

Érdekes jelenségre mutatott rá a [l] mássalhangzó szintézise is. Kimutatható, hogy a magánhangzóhoz való kapcsolódás előtt a [l] hangban – a nyelv felpattanása miatt – intenzitásminimum keletkezik. Ennek az intenzitásminimumnak lényeges szerepe van a hang hangzása szempontjából. Az intenzitásminimum értéke a V típusától függ, a legnagyobb az [i] előtt, a legkisebb az [a:], [ɔ], [u] előtt. A [l] hang szintézisénel ezt figyelembe kell venni.

*

Összefoglalásképpen elmondhatjuk, hogy a fenti munka elvégzése előrelépés a magyar mássalhangzók és hangsorbeli elhelyezkedésük szegmentális akusztikai szerkezetének feltárásához, és a magyar beszéd folyamatos szintézisének megalapozásához.

Irodalom

- A magyar nyelv szóvégműtató szótára. Szerk. PAPP Ferenc. Budapest 1969.
 BOLLA Kálmán: A magyar magánhangzók analízise és szintézise. MFF 1. 1978, 53–68.
 BOLLA Kálmán: A fonetikus írás. MFF 2. 1978, 7–24.
 BOLLA Kálmán: A magyar hosszú mássalhangzók képzése. MFF 7. 1981, 7–56.
 FANT, G.: Acoustic theory of speech production. 's-Gravenhage 1960.
 FANT, G.: Speech sounds and features. Massachusetts 1973.
 FÓNAGY Iván–SZENDE Tamás: Zárhangok, réshangok, affrikáták hangszínképe. NyK LXXXI, 1969, 281–344.
 FLANAGAN, J.L.–RABINER, L.R.: Speech synthesis. Stroudsburg 1973.
 GORDOS Géza: Újabb eredmények a gépi beszédfeldolgozásban. Híradástechnika XXX, 1979, 375–8.
 GÓSY Mária: A [b], [d], [g] mássalhangzók percepció vizsgálat. MFF 10. 1982, 84–108.

- KASSAI Ilona: A hangkapcsolatokról. MFF 2. 1978, 66–73.
 KASSAI Ilona: Magánhangzó–mássalhangzó találkozások. MFF 4. 1979, 80–119.
 KISS Gábor–OLASZY Gábor: Interaktív beszéd-szintetizáló rendszer számítógéppel és OVE III. szintetizátorral. MFF 10. 1982, 21–46.
 MAGDICS Klára: A magyar beszédhangok akusztikai szerkezete. Budapest 1965.
 SZENDE Tamás: A beszéd-folyamat alaptényezői. Budapest 1976.
 TARNÓCZY, T.: Resonance data concerning nasals, laterals and trills. Word 4. 1948.
 TARNÓCZY, T.: Die akustische Struktur stimmlosen Engelaute. ALinguH IV, 1954.
 TARNÓCZY, T.–VICSI, K.: Some remarks on the perception of voiceless stop-consonants. Acustica 43. 1979, 167–73.

THE ANALYSIS AND SYNTHESIS OF THE ACOUSTIC STRUCTURE OF HUNGARIAN CONSONANTS AND CONSONANT–VOWEL COMBINATION TYPES

Gábor Olaszzy

1. In Hungarian the most frequent 2-element segmental combinations are the consonant-vowel combinations CV and VC. Accordingly, the combinations of the type #CV and VC # are the most important basic units to build up synthesized Hungarian sound sequences. The unit VCV is a combination of the two types mentioned above.

2. This paper is the first summary of the complete temporal, intensity and frequency structure of the Hungarian consonants and the #CV, VCV, VC # combinations (regarded as the basic units of sound sequences) and one example of their synthesis. The results of the synthesis and the data gained at the experiments give us a better understanding of the segmental acoustic structure of the Hungarian speech and help us lay the foundations of Hungarian speech synthesis. The author presents this study as a summary, which, on the one hand, gives guidance how to develop a relatively rapid and acoustically acceptable synthesis of Hungarian utterances and to work out an automatic synthesising system for Hungarian texts. On the other hand, experts of different fields (phoneticians, physicists, engineers, computation technicians, programmers, linguist-engineers, speech therapists, teachers, physicians, etc.) can find all the acoustic data necessary for the description of the Hungarian consonants and the combinations mentioned above.

3. The analysis of the acoustic structure of the Hungarian consonants and consonant-vowel combinations was carried out for the consonants [b], [p], [d], [t], [ʃ], [c], [g], [k], [m], [ŋ], [n], [ɲ], [ɳ], [v], [f], [z], [s], [ʒ], [ʃ], [j], [ç], [h], [ɦ], [x], [çv], [pʃ], [dz], [ts], [dʒ], [tʃ], [l], [r], and also for all their combinations with the vowels [a:], [ɔ], [o], [u], [y], [ø], [i], [ɛ:], [e]. The total number of combinations examined was about 864.

We could not carry out the analysis of the consonants and the above mentioned combinations without synthesis because of their intricate, and in some cases unexplored, acoustic structure. We have also found that in many cases seemingly unimportant data are essential factors in the acoustic structure of the sounds and sound combinations examined (e.g. the sound [l]). Therefore we had to use the method „analysis by synthesis” throughout the whole work (cf. Figure 1). We used our own FOPROx program in the speech synthesis.

4. The global results of our work are the sound descriptions. There is a separate description for each consonant which includes the following:

The description of the acoustic structure

- a) Data characterizing the temporal structure of the consonant.
- b) Data describing the intensity and frequency structure of the consonant.
- c) Data describing the intensity and frequency structure of the vowel the consonant in question combines with during the transition phase of the CV and VC combinations.

The synthesis of the sound with an OVE III synthesizer

With the data obtained in the analysis-synthesis each consonant has been synthesized as part of two different words. I have made intensity curves and spectrograms of the natural and synthesized variants of the sample words, and in a diagram I have also present the final input data the computer used in the synthesis (cf. Figure 2). The sample words recorded on tape, the intensity curves and the spectrograms mentioned above can be found in the phonetics laboratory of the Institute of Linguistics of the Hungarian Academy of Sciences.

5 The result obtained so far are satisfactory. Using the sound descriptions we synthesized Hungarian words and sentences of good quality for the first time in 1980. Our further goal is to develop the automatic synthesis of Hungarian.

A [b, d, g] MÁSSALHANGZÓK PERCEPCIÓS VIZSGÁLATA

Gósy Mária

Bevezetés

Számos tanulmány, amely a felpattanó zárhangok azonosítási problémájával foglalkozik, elsősorban azt a kérdést igyekeznek megválaszolni, hogy ezeknek a mássalhangzóknak melyik összetevője az, amelyik alapján a hallgató a képzési helyet felismeri. Ennek a felismerésnek a birtokában történik az adott fonémaosztályba sorolás (pl. Delattre és mtsai JASA 27. 1955; Fant 1969; Stevens—Blumstein JASA 64. 1978; Dorman—Raphael JASA 67. 1980; Blumstein—Stevens JASA 67, 1980; stb.). Az eredmények szerint két tényező a meghatározó: a zárhang zörejsomójának (tehát a zárfelpattanásnak) és az átmenetnek a frekvenciaértéke.

Kísérletsorozatot terveztem a magyar felpattanó zárhangok percepciós vizsgálatához. Céлом annak a meghatározása, hogy a három mássalhangzó F_2 -értékének változtatása miként befolyásolja a felismerésüket. Választ kerestem arra is, hogy a zárhangot követő magánhangzó minősége hogyan hat a mássalhangzó azonosítására, és hogy a fonémahatár közelében élesebb-e a megkülönböztetés.

A kísérlet nyelvi anyagát a [b, d, g]-nek az összes magánhangzóval alkotott szintetizált CV-hangkapcsolatai adták (a hosszú—rövid magánhangzópárok közül csak a hosszúak szerepeltek). A zárhangoknak egyetlen magánhangzóval alkotott kapcsolata könnyen azonosítható (szemben például az izolált ejtésű mássalhangzókkal); ugyanakkor a CV-szekvencia szótagértékű, e tekintetben tehát nyelvi státusa van. Másrészt teljesíti azt az igényt, hogy más — a beszédmegértésben fontos — tényezőket, mint amilyen például az egyes szavak hírértéke, gyakorisága, az alanyok nyelvi szintje stb. kiküszöböl.

A szintézist OVE III beszéd szintetizátorral végeztük, amelyet PDP 11/34-es típusú számítógép vezérelt. A szintetizálásban a kiindulást a „legjobb” hangzású hangkapcsolatok létrehozása jelentette (akusztikai szerkezetüket alapszerkezetnek nevezem). Ezek egyértelműen a magyar *bá, gú, dé* stb. hangsorok benyomását keltették a hallgatókban. A mássalhangzókat három szeletből állítottam elő: az első kettő a zárhang zöngé részének, a harmadik a zárfelpattanásnak felel meg. Ezt követi a magánhangzóhoz való átmenetet reprezentáló hangszelet, majd pedig a magánhangzó tiszta fázisát képviselő szeletek (a [b] mássalhangzó kapcsolatainak alapszerkezetét az 1. táblázat, a [d]-ét a 2., a [g]-ét a 3. táblázat tartalmazza). A magánhangzók szintézisében figyelembe vettem a magyar magánhangzókra megadott legújabb formánsértékeket (Bolla MFF 1. 1978); felhasználtam a mássalhangzók és a mássalhangzó—magánhangzó-átmenetek szintetizálási szabályait (Olaszy MFF 10. 1982). (A magánhangzók tiszta fázisának hangszeleteit a 4. táblázat mutatja.) A hiányzó értékeket természetes férfi ejtésből elemeztem. A hangkapcsolatok alaphangmagasság-értékét nem változtattam, 116 Hz-en tartottam.

A kísérlet tesztanyagát úgy állítottam össze, hogy az F_2 értékét változtattam a hangkapcsolatoknak a zárfelpattanást és az átmenetet képviselő hangszeletében. A kísérleti személyeknek az elhangzott hangkapcsolatokat kellett leírniuk. Valamennyi vizsgálatot ugyanazzal a 18 alannyal végeztem. A kísérlet résztvevői nők, férfiak (fele-fele arányban), életkoruk, iskolázottsági szintjük és foglalkozásuk különböző.

1. táblázat

A hang- szület azono- sítója	Tényle- ges idő- adat (ms)	Viszony- lagos (dB) erősség	F ₀ (Hz)	F ₁ (Hz)	F ₂ (Hz)	F ₃ (Hz)
B1	90	1-3	116			
B2	90	3-2	116			
A:3	20	10	116	356-449	1200	2329-2614
A3	16	10-14	116	400-519	1008	2614
0:3	12	6-12	116	346	800	2397
U:3	12	8	116	300	692	2397
B 0'3	12	10	116	356-400	1425	2397
Y:3	12	6-12	116	231	1600	2468
E3	12	4-12	116	400	1600	2075
E'3	15	10	116	346	1550	2770
I:3	16	8-10	116	252	1695	2468-2691
A:	39	10-12	116	449-734	1200-1307	2614
A	40	14	116	519-599	1008	2614
0:	39	12-14	116	346-400	800	2397
U:	39	9-14	116	300	692	2397
B 0'	39	10-14	116	400	1425-1600	2397-2329
Y:	39	12-14	116	231	1600-1695	2468
E	39	12-14	116	436-519	1647-1745	2075-2329
E'	39	10-12	116	346-400	1550-2016	2770
I:	52	10-14	116	252	1695-2198	2691-3020

	A hang- szelet azono- sítója	Tényle- ges idő- adat (ms)	Viszony- lagos (dB) erősség	F ₀ (Hz)	F ₁ (Hz)	F ₂ (Hz)	F ₃ (Hz)
D1	A; U; O'	92	2	116			
	A, O; Y; E', I:	92	3	116			
	E	46	4	116	231-218		
D2	A; O'	24	2-1	116			
	A, O; Y; E', I:	24	3-2	116			
	U:	24	2-0	116			
	E	78	4-2	116	206		
D	A:3	12	1-4	116	400-504	1345	2770
	A3	6	2	116	400	1307	2770
	O:3	21	0-6	116	400	1307	2397
	U:3	7	6	116	252	1307	2397
	O'3	12	1-3	116	356-400	1695	2691
	Y:3	6	2	116	231	1958	2329
	E3	15	0-4	116	378-436	1796	2329
	E'3	15	0-4	116	400	1796	2770
	I:3	6	2	116	252	2016	3020
	D	A:	78	14-12	116	504-755	1345-1307
A		39	12-14	116	462-582	1307-1008	2770-2614
O:		39	8-12	116	400	1307-800	2397
U:		35	14	116	252-300	1307-692	2397
O'		35	14	116	400	1695-1600	2691-2468
Y:		39	6-12	116	231	1958-1695	2468
E		55	14	116	449-550	1796	2329
E'		39	14-12	116	400	1796-2016	2770
I:		12	6-12	116	262	2016-2198	3020

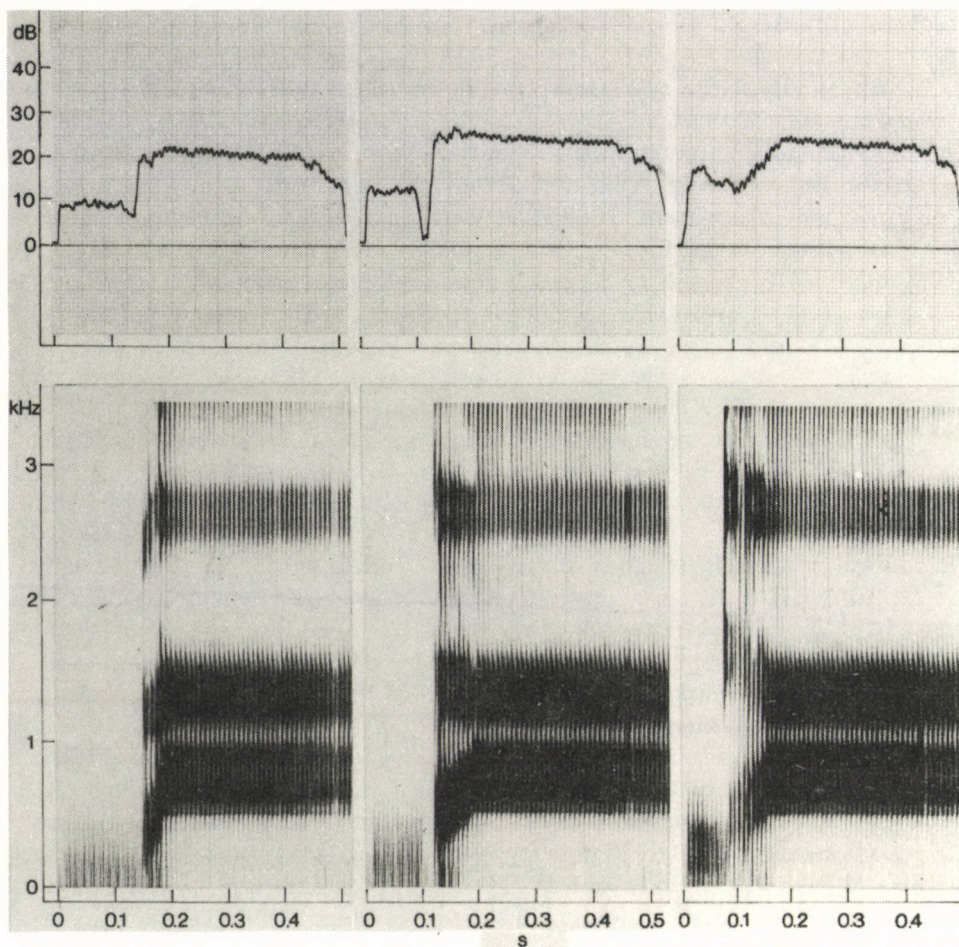
3. táblázat

	A hang- szelet azono- sítója	Tényle- ges idő- adat (ms)	Viszony- lagos erősség (dB)	F ₀ (Hz)/ AK (dB)	F ₁ /K ₁ (Hz)	F ₂ /K ₂ (Hz)	F ₃ (Hz)
G1	A:	60	0-4	116	252-231		
	A, O; U; O', Y; E, E', I:	52	0-4	116	252-231		
G2	A:	44	4-2	116	231-206		
	A, O; O', Y; E, E', I:	52	4-2	116	231-206		
	U:	60	4-2	116	231-206		
G	A:3	52	2-4/2-6	116/30	212-449/ 1385	1600/4525	2691
	A3	60	2-4/4	116/30	212-436/ 1099	1131/4525	2614
	O:3	52	2-6/4	116/30	252-300/ 1099	1068/4525	2397
	U:3	52	2-6/0-6	116/30	206-252/ 1037	1099/4525	2397
	O'3	52	2-4/6-2	116/30	259-378/ 1037	1903/4525	2468
	Y:3	32	2-8/4	116/30	231/1099	1848/4525	2468
	E3	39	2-6/6-0	116/30	206-346/ 2075	1958/4525	2614
	E'3	12	2-6/6-0	116/30	252-300/ 1099	2198/4525	2770
	I:3	16	2-8/0-2	116/30	252/2934	2468/4525	6041
	A:	78	6-12	116	449-755	1600-1307	2691-2614
A	75	6-12	116	462-582	1131-1008	2614	
O:	60	6-12	116	300-400	1068-800	2397	
U:	90	6-14/6-0	116/30	300/1037	1099-692/ 4525	2397	
G	O'	90	4-10	116	400	1903-1600	2468-2329
	Y:	55	4-14	116	231	1849-1695	2468
	E	65	7-14	116	367-519	1958-1796	2614-2329
	E'	52	6-12	116	300-400	2198-2016	2770
	I:	55	0-14	116	252	2468-2198	6041

A hang- szelet azono- sítója	Tényle- ges idő- adat (ms)	Viszony- lagos erősség (dB)	F ₀ (Hz)	F ₁ (Hz)	F ₂ (Hz)	F ₃ (Hz)
A:1	208	12-10	116	755	1307	2614
A:2	156	10-0	116	755	1307	2614
A1	184	14	116	599	1008	2614
A2	144	14-0	116	599	1008	2614
O:1	224	14	116	400	800	2397
O:2	184	14-0	116	400	800	2397
U:1	104	14	116	300	692	2397
U:2	104	14	116	300	692	2397
U:3	104	14-0	116	300	692	2397
O'1	104	10-14	116	400	1600	2329
O'2	104	14	116	400	1600	2329
O'3	104	14-0	116	400	1600	2329
Y:1	224	14	116	231	1695	2468
Y:2	224	14-0	116	231	1695	2468
E1	88	14-12	116	550	1796	2329
E2	88	12-10	116	550	1796	2329
E3	88	10-0	116	550	1796	2329
E'1	224	12-14	116	400	2016	2770
E'2	184	14-0	116	400	2016	2770
I:1	224	14	116	252	2198	6041
I:2	224	14-0	116	252	2198	6041

Percepciósi eredmények

Az [a:]-magánhangzós kapcsolatok F₂-értékeit és a percepciósi válaszokat a *bá* esetében az 5.a, a *dá*-nál az 5.b, a *gá*-nál az 5.c táblázat tartalmazza. Az alapszerkezetek (vö. 1. ábra) meghangosítása 100 %-os felismerést adott a *bá* és a *dá*, 80 %-ost a *gá* esetében. A *bá* alapszerkezet második formánsának változtatása 1037 Hz-től 1270 Hz-ig 100 %-os [b]-azonosítást hozott. Amíg az F₂ értéke nem érte el a magánhangzó második formánsának értékét, a [b] felismerése jó volt. Amikor azonban megegyezett a magánhangzóéval vagy meghaladta azt, akkor a [b] felismerése csökkent: 1345 Hz-nél 74 %-ra. Ebben az akusztikai szerkezetben a [b] azonosítását már nem a formánsfrekvencia, hanem az intenzitásviszonyoknak a [b]-re jellemző menete biztosítja (az összes töb-



1. ábra

A *bá*, *dá*, *gá* szintetizált hangkapcsolatok intenzitásgörbéje és hangszínképe

bi komponenssel együtt); hiszen az F_2 -érték a [d] mássalhangzóra jellemző. Ennek alátámasztására azt a változtatást végeztem, hogy a mássalhangzó teljes időtartamának harmadik harmadában a viszonylagos erősséget 10 dB-ről 2–4 dB-es emelkedésre választottam, ami a *dá* hangkapcsolat intenzitáskomponenseit közelíti meg. A percepció ítéletek szerint ez a módosított mássalhangzó már csak 58 %-ban hallatszott [b]-nek, 21–21 %-ban [d]-t és [g]-t értettek a kísérleti személyek. (A *gá*-azonosítás is az intenzitásviszonyok módosulásának tudható be.) Az F_2 -érték további növelésével csökken a *bá*-azonosítás (67 %), növekszik a [g]-érték (28 %). Ez azzal magyarázható, hogy a zárfelepattanás és a magánhangzó második formánsának különbsége kezdi elérni a [g]-re jellemző értéket. A *dá*-alapszerkezetben eltoltt F_2 -k eredményei azt mutatják, hogy sokkal szűkebb az a frekvenciasáv, amely a [d]-azonosítást biztosítja. 1270 Hz-es F_2 esetén már 58 %-ra

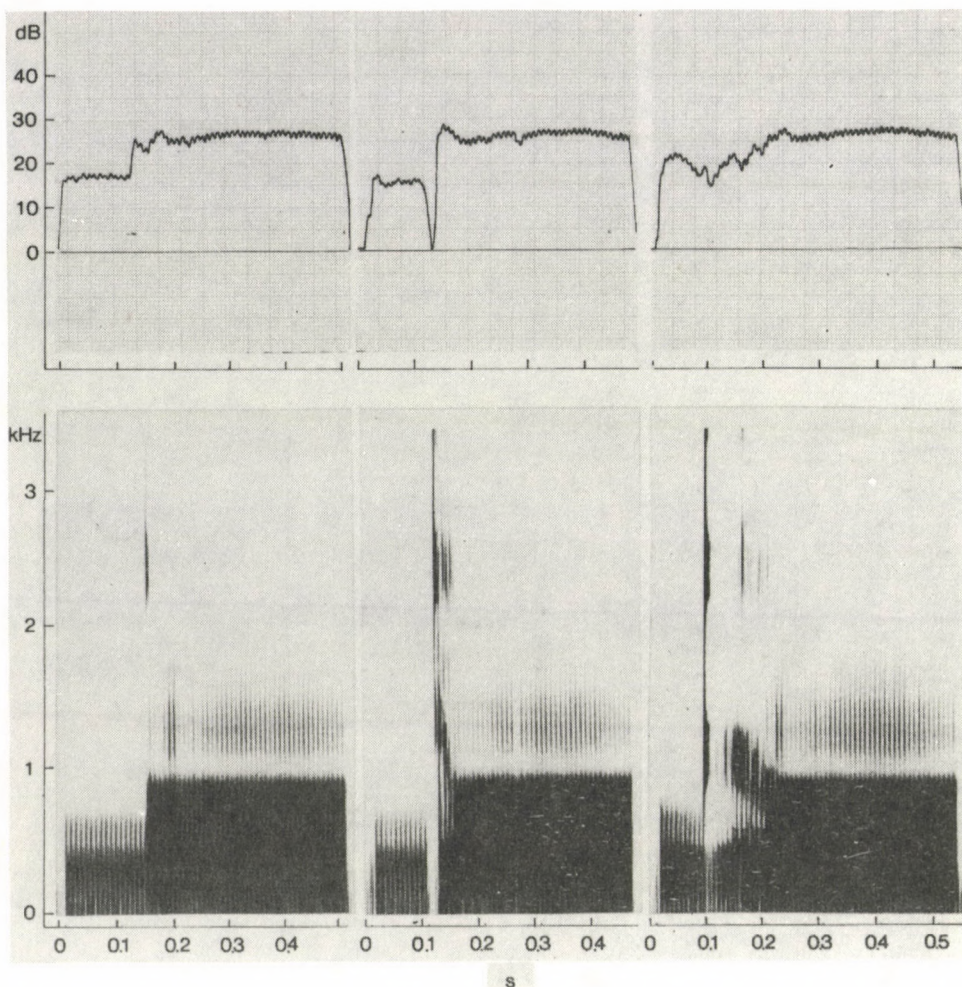
csökken a [d]-értés, 42 %-ban [b]-t hallanak. A *gá*-alapszerkezetben végzett F_2 -változtatások minden esetben magas százalékban (70–80 %) [g]-felismerést eredményeztek.

Az [ɔ]-magánhangzós CV-kapcsolatok F_2 -értékeit és az ezekre kapott percepciósi ítéleteket a *ba* esetében a 6.a táblázat, a *da*-nál a 6.b táblázat, a *ga*-nál pedig a 6.c táblázat mutatja. Az alapszerkezetek meghangosítása a *ba* és a *ga*-nál 100 %-os, a *da*-nál 83 %-os azonosítást hozott. A [b] felismerése akkor jó, ha a zárfelpattanás második formánsának értéke nem haladja meg az 1068 Hz-et, illetve ha a frekvenciaérték alulról csatlakozik a magánhangzó 1008 Hz-es F_2 -jéhez. Az adatok szerint ugyan 924 és 979 Hz között kissé csökken a [b] azonosítása (de csak a *da* alapszerkezetben belül, itt tehát más paraméterek percepcióra gyakorolt hatását kell feltételeznünk). A *da* felismerése az 1200–1300 Hz-es F_2 -tartományban jó. 900 Hz-nél alacsonyabb második formáns esetén jelentősen növekszik a [b]-értés. A *ga* második formánsát a [b]-re és a [d]-re jellemző értékekre toltam el; az eredmények minden esetben 100 %-os [g]-felismerést hoztak.

Az [o:]-magánhangzós hangkapcsolatok F_2 -értékeit és a válaszokat a *bó*-nál a 7.a a *dó*-nál a 7.b, a *gó*-nál pedig a 7.c táblázat tartalmazza. Valamennyi alapszerkezet meghangosítása 100 %-os azonosítást hozott. Az [o:] magánhangzót megelőző zárhangot a kísérleti személyek akkor azonosították [b]-nek, ha az F_2 600–900 Hz-es tartományban realizálódott; de még 980 Hz esetén is 70 %-os volt a [b] értéke. [d]-t ismertek fel, ha az F_2 az 1200–1300 Hz-es tartományban mozgott; ettől alacsonyabb frekvenciasávban csökken a [d]-ítéletek aránya, növekszik a [b]-értés. Az $F_2 = 1068$ Hz-nél 75 %-ban [g]-t azonosítottak. Ez feltételezi, hogy ez az érték a legjellemzőbb a [g] mássalhangzóra (a *gó* alapszerkezet F_2 -jét ennek alapján választottam 1068 Hz-re). Az eltolt F_2 -k minden esetben 100 %-os [g]-felismerést eredményeztek.

Az [u:]-magánhangzós CV-kapcsolatok második formánsának frekvenciaértékeit és a percepciósi válaszokat a *bú*-nál a 8.a táblázat, a *dú*-nál a 8.b táblázat, a *gú*-nál a 8.c táblázat tartalmazza. Az alapszerkezetek meghangosítása (vö. 2. ábra) a *bú* és a *dú* esetében 100 %-os, a *gú* esetében 85 %-os felismerést hozott. A [b] azonosítása az alapszerkezeti 692 Hz-es F_2 -nél alacsonyabb és magasabb frekvenciákon is jó; csak az $F_2 = 800$ Hz-nél bizonytalanodnak el az ítéletek. 872 Hz-nél már csak 53 %-ban értenek [b]-t, 21 %-ban [d]-t és 26 %-ban [g]-t azonosítanak. A frekvencia emelésével tovább csökken a [b] és növekszik a [d]-felismerések aránya. 951 és 979 Hz-nél csaknem fele-fele arányban [g]-t és [d]-t hallottak a kísérleti alanyok: 45–50 %, illetve 40–45 %-ban. Az F_2 értékének további emelésével csökken a [g]- és növekszik a [d]-ítéletek száma. A kapott arányok pontosan mutatják – a *bú* alapszerkezetén belül is – a másik két zárhangra jellemző F_2 frekvenciahelyét. A *dú*-értés viszonylag nagy F_2 -tartományban jelentkezik: 1000–1300 Hz között. Lényeges eltérést 734 Hz-nél tapasztaltam: itt 42 %-ban hallottak [d]-t, 53 %-ban már [b]-t azonosítottak. A [b]-re jellemző F_2 elérésével (692 Hz) felszökik a [b]-ítéletek aránya 85 %-ra. A *gú* alapszerkezetben eltolt F_2 -k 78–79 %-ban [g]-felismerést hoztak.

Az [ɸ:]-magánhangzós hangkapcsolatok F_2 -értékeit és az ezekre kapott ítéleteket a *bő* esetében a 9.a táblázat, a *gő*-nél pedig a 9.b táblázat tartalmazza. A *dő* alapszerkezetében az F_2 -t 1800 Hz-re tolván 100 %-os volt a [d] felismerése. Mindhárom alaphangkapcsolat meghangosítását (vö. 3. ábra) 100 %-osan azonosították a kísérleti személyek. Ha a *bő* alapszerkezetben növelem az F_2 értékét, egyre csökken a [b] és növekszik a [d]-értés (33–67 %). A [d]-re jellemző 1647 Hz-es értékénél a kísérleti alanyok 100 %-ban

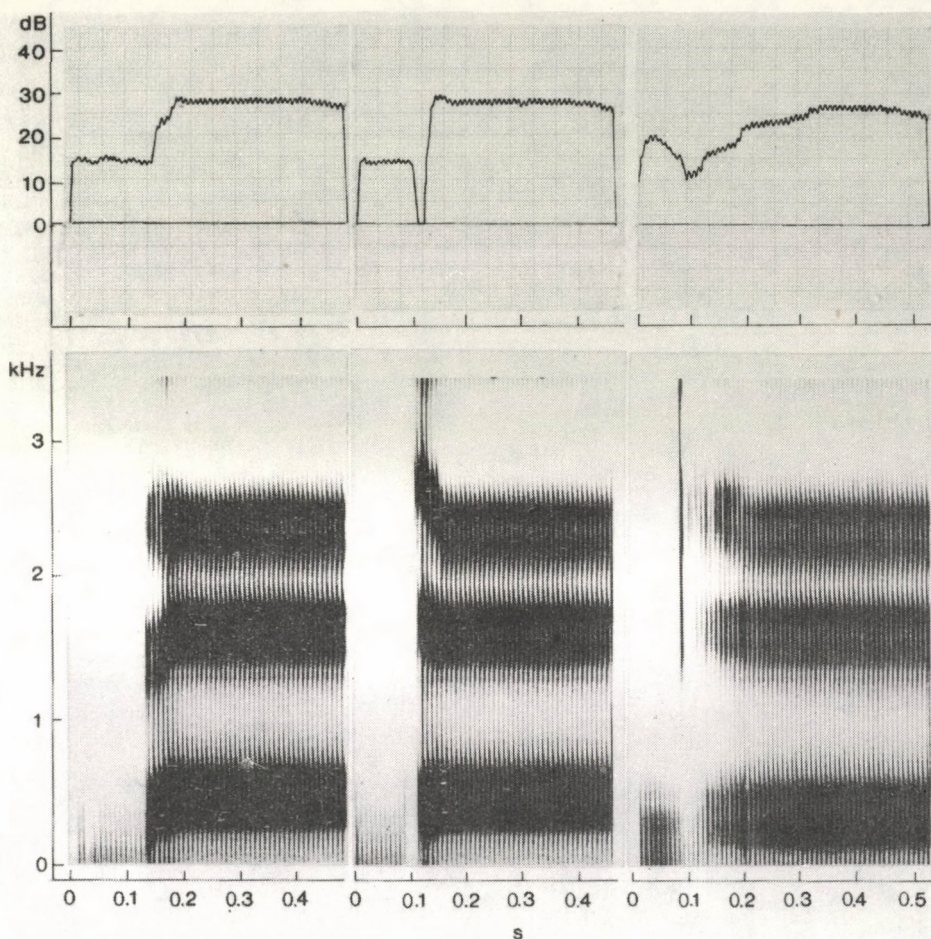


2. ábra

A *bű, dú, gű* szintetizált hangkapcsolatok intenzitásgörbéje és hangszínképe

a dentális zárhangot azonosítják. Az F_2 további emelésével a kísérleti személyek ítélete megoszlik a [d] és a [g] között; 1903 Hz-nél 42 %-ban [d]-t, 58 %-ban [g]-t értenek (ez már a [g]-re jellemző F_2). A *gű* alapszerkezetben a második formáns értékét csökkentve bizonytalanodik a [g]-értés, a kísérleti alanyok egy része [d]-t hall. 1695 Hz-nél már csak 39 % a [g]- és 61 % a [d]-ítéletek aránya.

Az [y:]-magánhangzós CV-kapcsolatok változtatott második formánsait és a perceptiós válaszokat a *bű* esetében a 10.a táblázat, a *dű*-nél a 10.b táblázat, a *gű*-nél pedig a 10.c táblázat tartalmazza. A *bű* és a *gű* alapszerkezetének meghangosítása 100 %-os felismerést hozott, a *dű*-é csak 65 %-ost. A *bű* alapszerkezetben az $F_2 = 1600$ Hz-től alacsonyabb értékeken egyértelműen jó a [b] azonosítása; ettől az értéktől felfelé azonban

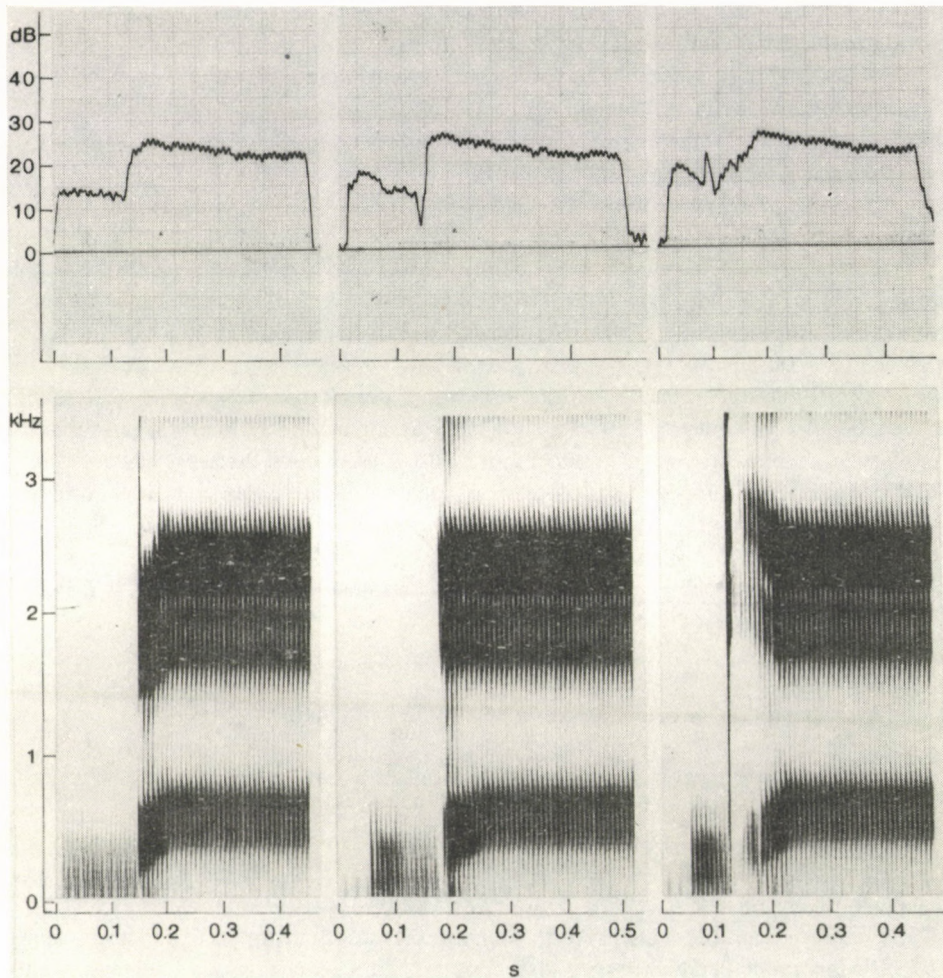


3. ábra

A *bõ*, *dõ*, *gõ* szintetizált hangkapcsolatok intenzitásgörbéje és hangszínképe

ugrásszerűen csökken a [b]-ítéletek aránya. 1695 Hz-nél a kísérleti alanyok harmada [b]-t, harmada [d]-t, harmada pedig [g]-t hallott. Mintegy 50 Hz-cel magasabb F_2 -nél 78 %-ra emelkedik a [d] azonosítása. A *dõ* alapszerkezetben csak 1900 Hz fölött növekszik meg a [d]-felismerés, s csökken az alacsonyabb frekvenciákon viszonylag magas százalékarányú [g]-értés. Eredményeim szerint az [y:] magánhangzó előtt akkor legbiztosabb a dentális zárhang azonosítása, ha az F_2 értéke 1958 Hz körül van. A *gõ* alapszerkezetben változtatott F_2 -k kismértékben befolyásolják a [g] felismerését (az eredmények 60–95 %-os [g]-értést mutatnak).

Az [e]-magánhangzós kapcsolatok F_2 -értékeit és a percepciós ítéleteket a *be* esetében a 11.a táblázat, a *de* esetében pedig a 11.b táblázat tartalmazza. Az alapszerkeze-



4. ábra

A *be*, *de*, *ge* szintetizált hangkapcsolatok intenzitásgörbéje és hangszínképe

tek meghangosítása (vö. 4. ábra), a *be*-nél 100 %-os, a *de*-nél 83 %-os, a *ge*-nél pedig 95 %-os felismerést hozott. A *be* alapszerkezetben jól azonosítható a [b], amíg az F_2 értéke nem éri el a [d]-re jellemző 1796 Hz-es értéket. Ekkor 42 %-ban [b]-t, 37 %-ban [d]-t és 21 %-ban [g]-t hallanak a kísérleti alanyok. Az F_2 növelésével tovább csökken a [b] és növekszik a [d] és a [g] felismerése. 1849 Hz-nél gyakorlatilag nem azonosítanak [b]-t, 61 %-ban [d]-t és 33 %-ban [g]-t hallanak. A *de* alapszerkezet jellemző F_2 -je 1796 Hz, alacsonyabb értékekre tolásával csökken a [d] felismerése. Minél nagyobb a magánhangzó tiszta fázisú F_2 -jének és a zárfelattanás második formánsának a különbsége, annál nagyobb mértékben hallanak a kísérleti alanyok [b]-t. Az F_2 növelése mintegy

100 Hz-en belül még 74–79 %-ban [d]-értést eredményez, az 1958 Hz elérésével azonban ugrásszerűen megnövekszik a [g]-ítéletek aránya.

A zárhangok [ɛ:] magánhangzóval alkotott hangkapcsolatainak F_2 -értékeit és a percepció válaszokat a *bé* esetében a 12.a táblázat, a *dé* esetében a 12.b táblázat, a *gé*-nél pedig a 12.c táblázat tartalmazza. Az alapszerkezetek meghangosítása a *dé* és a *gé*-nél 100 %-os, a *bé*-nél 85 %-os felismerést hozott. Viszonylag szűk az az F_2 -sáv, amelyben egyértelmű a [b] azonosítása. Az alapszerkezeti 1600 Hz-es értéket alig 50 Hz-cel meghaladva már 100 %-ban [d]-t értettek. 1800 és 1900 Hz között ugyancsak 100 %-os a [d] felismerése (a *dé* alapszerkezeten belül). 1950 Hz-nél 78 %-os a [g]-értés a *bé* alapszerkezetében, a *dé*-ben ugyanilyen arányban csak 2136 Hz-nél hallanak [g]-t. A *gé* hangkapcsolat percepciójára jellemző, hogy 1900 és 2200 Hz között jó a zárhang felismerése: 70–100 %.

Az [i:] magánhangzós CV-hangkapcsolatok F_2 -értékeit és a percepció ítéleteket a *bi* esetében a 13.a táblázat, a *di*-nél a 13.b táblázat, a *gi*-nél a 13.c táblázat tartalmazza. Az alapszerkezetek meghangosítására a *gi*-nél 100 %-os, a *di*-nél 90 %-os, a *bi*-nél csak 72 %-os felismerést kaptam. Az F_2 értékét növelve a *bi* alapszerkezetben egyre csökken a [b]- és növekszik a [d]-értés. 2200 Hz-nél ugrásszerűen átcsap a felismerés [g]-be. A *di* alapszerkezetben 1958 Hz-nél alacsonyabb F_2 esetén csökken a [d]-ítéletek aránya, ettől magasabb frekvenciákon jó a dentális zárhang azonosítási aránya (75 %). A *gi*-ben elolt F_2 80–100 %-os [g]-felismerést hozott.

Az akusztikum, az artikuláció és a percepció kapcsolata

1. A három magyar zárhanggal végzett kísérletsorozat eredményei az alábbi tendenciákról tanúskodnak:

a) A [b] mássalhangzó felismerése a CV-hangkapcsolatban egyértelműen a második formáns frekvenciaértékétől függ, ami azt jelenti, hogy a bilabiális képzési helynek akusztikai szinten az F_2 800–1800 Hz-es tartománya felel meg. Mégpedig úgy, hogy a hátul képzett magánhangzók második formánsához a mássalhangzó F_2 -je az átmenetben alulról vagy kissé magasabb frekvenciaértékről felülről csatlakozik. Az elől képzett magánhangzókánál csak alulról csatlakozhat, nem lehet azonos értékű a magánhangzó második formánsával.

b) A [d] mássalhangzó felismerése palatális magánhangzók előtt elsősorban a második formáns értékétől, veláris magánhangzók előtt mind a második formánstól, mind az intenzitásértékektől függ. Az F_2 frekvenciatartománya: 900–2200 Hz; jellemzője, hogy a második formáns felülről vagy alulról csatlakozik a magánhangzó F_2 -jéhez úgy, hogy labiális magánhangzók esetében nem lehet azonos értéken a magánhangzóéval, míg az illabiálisoknál meg is egyezhet azzal.

c) A [g] mássalhangzóra jellemző valamennyi hangkapcsolatban, hogy nem kizárólag az F_2 értéke a döntő tényező a percepcióban. Az elől képzett magánhangzók előtt mind a második formáns, mind az intenzitásviszonyok, mind az F_1 mozgása és az időviszonyok meghatározók. A hátul képzett ma-

5.a táblázat

F ₂ (Hz) /BA:3/BA:/	[b] (%)	[d] (%)	[g] (%)
/1037/1037-1307/	100	—	—
/1069/1069-1307/	95	5	—
/1099/1099-1307/	100	—	—
/1199/1199-1307/	100	—	—
/1270/1270-1307/	89	11	—
/1345/1345-1307/	74	21	—
/1385/1385-1307/	67	5	28
/1425/1425-1307/	22	44	33

5.b táblázat

F ₂ (Hz) /DA:3/DA:/	[b] (%)	[d] (%)	[g] (%)
/1345/1345-1307/	—	100	—
/1270/1270-1307/	42	58	—
/1190/1199-1307/	90	10	—
/1165/1165-1307/	78	22	—

5.c táblázat

F ₂ (Hz) /GA:3/GA:/	[b] (%)	[d] (%)	[g] (%)
/1600/1600-1307/	—	17	83
/1425/1425-1307/	—	17	83
/1345/1345-1307/	—	22	78
/1270/1270-1307/	—	21	79
/1165/1165-1307/	21	10	69

6.a táblázat

F ₂ (Hz) /BA3/BA/	[b] (%)	[d] (%)	[g] (%)
/1000/1000/	100	—	—
/1037/1037-1000/	100	—	—
/1068/1068-1000/	100	—	—
/1099/1099-1000/	63	37	—
/1131/1131-1000/	16	68	16
/979/979-1000/	100	—	—

6.b táblázat

F ₂ (Hz) /DA3/DA/	[b] (%)	[d] (%)	[g] (%)
/1307/1307-1000/	—	100	—
/1131/1131-1000/	58	42	—
/1099/1099-1000/	74	21	5
/1068/1068-1000/	61	39	—
/1000/1000/	83	17	—
/979/979-1000/	47	53	—
/951/951-1000/	58	42	—
/924/924-1000/	64	31	5
/898/898-1000/	78	22	—
/872/872-1000/	78	17	5
/848/848-1000/	84	16	—
/824/824-1000/	80	20	—
/800/800-1000/	80	20	—
/777/777-1000/	90	10	—
/755/755-1000/	90	10	—
/734/734-1000/	95	5	—

6.c táblázat

F ₂ (Hz) /GA3/GA/	[b] (%)	[d] (%)	[g] (%)
/1131/1131-1000/	—	—	100
/1307/1307-1000/	—	—	100
/1000/1000/	—	—	100

7.a táblázat

F ₂ (Hz) /BO:3/BO:/	[b] (%)	[d] (%)	[g] (%)
/800/800/	100	—	—
/777/777-800/	75	5	20
/755/755-800/	95	5	—
/734/734-800/	90	10	—
/713/713-800/	78	17	5
/600/600-800/	80	10	10
/872/872-800/	70	15	15
/898/898-800/	70	20	10
/924/924-800/	53	31	16
/951/951-800/	67	28	5

7.b táblázat

F_2 (Hz) /DO:3/DO:/	[b] (%)	[d] (%)	[g] (%)
1307/1307-800	—	100	—
1099/1099-800	31	63	5
1068/1068-800	—	25	75
1037/1037-800	20	48	32
1000/1000-800	37	53	10
979/979-800	70	10	20
951/951-800	72	18	10
924/924-800	74	10	16
900/900-800	80	10	10

7.c táblázat

F_2 (Hz) /GO:3/GO:/	[b] (%)	[d] (%)	[g] (%)
898/898-800	—	—	100
848/848-800	—	—	100
800/800	—	—	100
700/700-800	—	—	100
951/951-800	—	—	100
1068/1068-800	—	—	100

8.a táblázat

F_2 (Hz) /BU:3/BU:/	[b] (%)	[d] (%)	[g] (%)
599/599-692	100	—	—
692/692	100	—	—
800/800-692	100	—	—
872/872-692	53	21	26
898/898-692	30	55	15
924/924-692	35	35	30
951/951-692	5	45	50
979/979-692	15	40	45
1008/1008-692	15	50	35
1037/1037-692	5	79	16
1068/1068-692	—	78	22
1100/1100-692	17	67	17
1131/1131-692	—	78	22
1300/1300-692	—	83	17

8.b táblázat

F ₂ (Hz) /DU:3/DU:/	[b] (%)	[d] (%)	[g] (%)
/1307/1307-692/	—	100	—
/1000/1000-692/	5	90	5
/734/734-692/	53	42	5
/713/713-692/	50	50	—
/692/692/	85	15	—

8.c táblázat

F ₂ (Hz) /GU:3/GU:/	[b] (%)	[d] (%)	[g] (%)
/1099/1099-692/	—	15	85
/1037/1037-692/	5	17	78
/692/692/	5	16	79

9.a táblázat

F ₂ (Hz) /BO'3/BO'/	[b] (%)	[d] (%)	[g] (%)
/1425/1425-1600/	100	—	—
/1467/1467-1600/	100	—	—
/1510/1510-1600/	83	17	—
/1554/1554-1600/	74	26	—
/1600/1600/	33	67	—
/1695/1695-1600/	—	100	—
/1745/1745-1600/	—	100	—

9.b táblázat

F ₂ (Hz) /GO'3/GO'/	[b] (%)	[d] (%)	[g] (%)
/1903/1903-1600/	—	5	100
/1849/1849-1600/	—	28	72
/1800/1800-1600/	—	11	89
/1745/1745-1600/	5	22	73
/1695/1695-1600/	—	61	39
/1647/1647-1600/	5	67	28

10.a táblázat

F ₂ (Hz) /BY3/BY/	[b] (%)	[d] (%)	[g] (%)
/1600/1600–1700/	100	–	–
/1647/1647–1700/	60	28	12
/1695/1695–1700/	34	33	33
/1745/1745–1700/	5	78	17
/1796/1796–1700/	11	72	17
/1550/1550–1700/	78	5	17
/1510/1510–1700/	95	5	–

10.b táblázat

F ₂ (Hz) /DY3/DY/	[b] (%)	[d] (%)	[g] (%)
/1695/1695/	22	50	28
/1647/1647–1700/	37	32	31
/1600/1600–1700/	47	27	26
/1550/1550–1700/	67	28	5
/1510/1510–1700/	67	5	28
/1467/1467–1700/	68	16	16
/1745/1745–1700/	10	56	34
/1796/1796–1700/	–	50	50
/1849/1849–1700/	–	50	50
/1900/1900–1700/	–	60	40
/1958/1958–1700/	–	65	35

10.c táblázat

F ₂ (Hz) /GY3/GY/	[b] (%)	[d] (%)	[g] (%)
/1849/1849–1700/	–	–	100
/1800/1800–1700/	10	18	72
/1700/1700/	11	11	78
/1600/1600–1700/	34	5	61
/1900/1900–1700/	–	5	95

11.a táblázat

F ₂ (Hz) /BE3/BE/	[b] (%)	[d] (%)	[g] (%)
/1600/1600–1796/	100	–	–
/1695/1695–1796/	95	5	–
/1796/1796/	42	37	21
/1849/1849–1796/	5	61	33
/1903/1903–1796/	5	45	50
/1958/1958–1796/	5	42	53
/2016/2016–1796/	–	45	55
/2329/2329–1796/	–	16	79

11.b táblázat

F ₂ (Hz) /DE3/DE/	[b] (%)	[d] (%)	[g] (%)
/1695/1695–1796/	56	44	–
/1745/1745–1796/	58	42	–
/1796/1796/	17	83	–
/1849/1849–1796/	5	79	16
/1903/1903–1796/	–	74	26
/1950/1950–1796/	–	53	47
/2016/2016–1796/	–	17	83
/2540/2540–1796/	–	5	95
/2770/2770–1796/	–	5	95

12.a táblázat

F ₂ (Hz) /BE'3/BE'/	[b] (%)	[d] (%)	[g] (%)
/1600/1600–2016/	78	22	–
/1647/1647–2016/	–	100	–
/1695/1695–2016/	25	70	5
/1745/1745–2016/	–	100	–
/1796/1796–2016/	5	80	15
/1950/1950–2016/	–	22	78
/2016/2016/	–	22	78
/2075/2075–2016/	–	10	90
/2136/2136–2016/	–	10	90



12.b táblázat

F ₂ (Hz) /DE'3/DE'/	[b] (%)	[d] (%)	[g] (%)
1796/1796–2016	–	100	–
1745/1745–2016	–	100	–
1695/1695–2016	28	72	–
1647/1647–2016	–	100	–
1600/1600–2016	56	44	–
1550/1550–2016	80	20	–
1850/1850–2016	–	100	–
2016/2016	–	70	30
2136/2136–2016	–	20	80
2200/2200–2016	–	10	90

12.c táblázat

F ₂ (Hz) /GE'3/GE'/	[b] (%)	[d] (%)	[g] (%)
2190/2198–2016	–	–	100
1950/1950–2016	–	28	72
1900/1900–2016	–	30	70
1600/1600–2016	16	47	37

13.a táblázat

F ₂ (Hz) /BI:3/BI:/	[b] (%)	[d] (%)	[g] (%)
1695/1695–2200	72	28	–
1958/1958–2200	50	45	5
2000/2000–2200	63	37	–
2200/2200	5	12	83
2263/2263–2200	–	–	100
2329/2329–2200	–	–	100

13.b táblázat

F_2 (Hz) /DI:3/DI:/	[b] (%)	[d] (%)	[g] (%)
/2016/2016–2200/	–	75	25
/1958/1958–2200/	–	90	10
/1900/1900–2200/	17	55	28
/1849/1849–2200/	11	67	22
/1796/1796–2200/	22	61	17
/1745/1745–2200/	40	50	10
/1695/1695–2200/	50	50	–
/1647/1647–2200/	39	44	17
/1600/1600–2200/	58	31	11

13.c táblázat

F_2 (Hz) /GI:3/GI:/	[b] (%)	[d] (%)	[g] (%)
/2468/2468–2200/	–	–	100
/1600/1600–2200/	–	16	84

gánhangzók előtt az intenzitásmenet, az F_1 menete, az időviszonyok alapvetően fontosak, s csak másodsorban az F_2 értéke. A [g] F_2 -jének frekvenciatartománya elől képzett magánhangzók előtt nagyobb vagy egyenlő, mint 1740 Hz. (A felső értékhatárról nincsenek pontos adataim.)

A [b, d, g] mássalhangzók második formánsának akusztikai tartományát ábrázoltam a különböző magánhangzók mellett (5. ábra); a grafikonon jelöltem a jó felismerést mutató frekvenciasávot és az ún. átmeneti mássalhangzók értékeit is. Átmenetinek nevezem azokat a mássalhangzókat, amelyeket a kísérleti alanyok egyik része az egyik, másik része egy másik fonémaosztályba sorol. A függőleges tengelyen a frekvencia, a vízszintes tengelyen a három zárhang szerepel. Téglalapok szemléltetik azt a tartományt, amelyben az adott mássalhangzó azonosítása a megfelelő fonémaosztályba 70 % fölötti volt. A téglalapokból kijövő vékony vonal a 70 %-nál kisebb arányú azonosítást jelzi. A kereszttel az alapszerkezetek F_2 -értékeit jelöltem. (Ott, ahol a vékony vonal hiányzik, nincsenek a felismerésre vonatkozó kísérleti adataim.)

2. Az F_2 tartománya képletben kifejezve (vö. még 6. ábra):

ha $F_2 \leq 1345$ Hz \longrightarrow *bá*

ha $F_2 = 1300$ – 1345 Hz \longrightarrow *dá*

ha $F_2 \geq 1200$ Hz \longrightarrow *gá*

ha $F_2 \leq 1068$ Hz \longrightarrow *ba*

ha $F_2 = 1200-1300$ Hz	→	<i>da</i>
ha $F_2 = 980-1300$ Hz	→	<i>ga</i>
ha $F_2 \leq 800$ Hz	→	<i>bó</i>
ha $F_2 = 1200-1300$ Hz	→	<i>dó</i>
ha $F_2 = 700-900$ Hz	→	<i>gó</i>
ha $F_2 \leq 800$ Hz	→	<i>bú</i>
ha $F_2 \geq 900$ Hz	→	<i>dú</i>
ha $F_2 = 850-1150$ Hz	→	<i>gú</i>
ha $F_2 \leq 1550$ Hz	→	<i>bő</i>
ha $F_2 = 1650-1800$ Hz	→	<i>dő</i>
ha $F_2 \geq 1740$ Hz	→	<i>gő</i>
ha $F_2 \leq 1600$ Hz	→	<i>bü</i>
ha $F_2 = 1900-1958$ Hz	→	<i>dü</i>
ha $F_2 = 1700-1900$ Hz	→	<i>gü</i>
ha $F_2 < 1796$ Hz	→	<i>be</i>
ha $1950 \text{ Hz} > F_2 \geq 1796 \text{ Hz}$	→	<i>de</i>
ha $F_2 \geq 1958$ Hz	→	<i>ge</i>
ha $F_2 \leq 1600$ Hz	→	<i>bé</i>
ha $F_2 = 1647-1850$ Hz	→	<i>dé</i>
ha $F_2 = 1900-2200$ Hz	→	<i>gé</i>
ha $F_2 \leq 1700$ Hz	→	<i>bí</i>
ha $F_2 = 1958-2200$ Hz	→	<i>dí</i>
ha $F_2 = 2450$ Hz körül van	→	<i>gí</i>

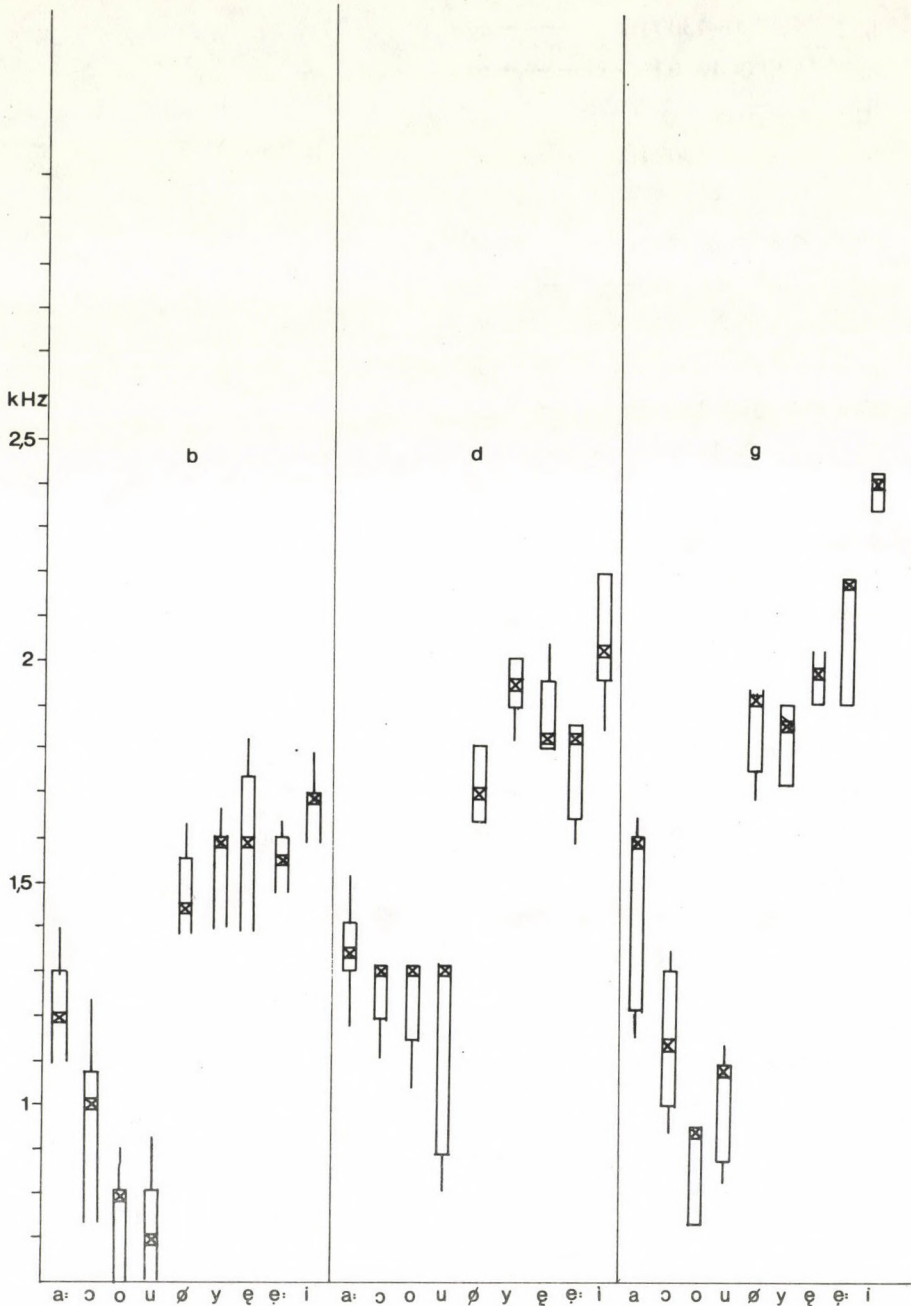
Az, hogy a [g] mássalhangzó felismerése szempontjából az F_2 a hátul képzett magánhangzóknál másodlagos, azt jelenti, hogy a többi komponens nagyobb szerepet játszik a képzési hely azonosításában. Ugyanakkor a jó hangzás szempontjából természetesen fontos az F_2 helyes megválasztása is.

3. A percepcióis ítéletek feldolgozásakor kiderült, hogy az egyes zárhangok felismerése a következő magánhangzó minőségétől is függ (vö. Ainsworth L&S LI, 1968). A kísérleti személyek is különböző nehézségűnek ítélték a feladatot a hangkapcsolat magánhangzójától függően.

A percepcióis válaszok alapján a következő tendenciák rajzolódtak ki a magánhangzó minőségének a zárhang azonosítására gyakorolt hatásáról:

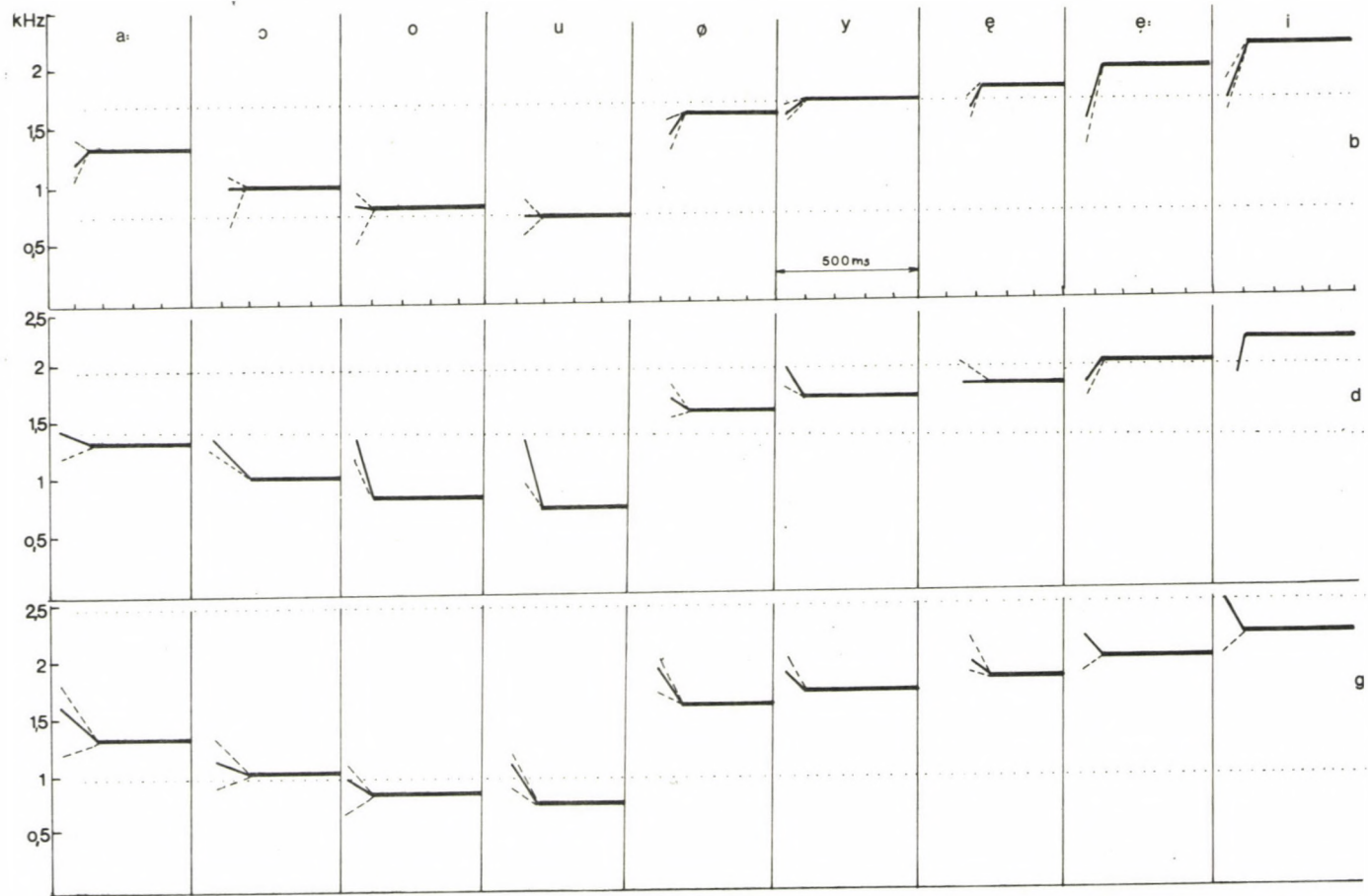
a) a labiális magánhangzók előtt könnyebb a zárhangok felismerése, mint az illabiálisok előtt;

b) a hátul képzettek előtt könnyebb, mint az elől képzettek előtt;



5. ábra

A [b, d, g] mássalhangzók magánhangzós kapcsolataiban az átmenet második formánsának frekvenciatartománya (a 70 %-nál jobb felismerést eredményező értékeket a téglalapok, a rosszabbakat a vonalak jelzik)



A vizsgált zárhang+magánhangzós kapcsolatok átmenetének F_2 -értéke, illetve a még jó felismerést eredményező frekvenciatartomány

c) a középső nyelvválásúak előtt könnyebb, mint az alsóbb vagy felsőbb nyelvválásúak előtt;

d) az alsóbb nyelvválásúak előtt könnyebb, mint a felső nyelvválásúak előtt.

Mindebből az következik, hogy a [b, d, g]-t legkönnyebben a középső nyelvválású, hátul képzett, ajakkerekítéses magánhangzók előtt és legnehezebben az elől képzett, felső nyelvválású, ajakrészeselek előtt azonosították. Legkönnyebb tehát az [o:]s és legnehezebb az [i:]s hangkapcsolatok felismerése volt. A vizsgált zárhangok azonosításának „nehézségi” sorrendje a követő magánhangzó típusától függően: [o:, φ:, ɔ, a:, ɛ, ɛ:, u:, y:, i:].

Megvizsgáltam továbbá, hogy melyik zárhang milyen magánhangzó(k) szomszéd-ságában azonosítható legkönnyebben és legnehezebben és mi(k) lehet(nek) ennek az oka(i). Feltételezésem az volt, hogy minél nagyobb területen történik a CV-szekvencia képzése, annál könnyebb a zármasalhangzó felismerése. Nehezebb volt ui. elkülöníteni a bilabiális és a dentális zárhangot pl. az [ɛ] magánhangzó előtt, mint ugyanezeket az [ɔ] előtt. A zárhangok képzési helye és a magánhangzó minősége között az alábbi összefüggéseket láttam:

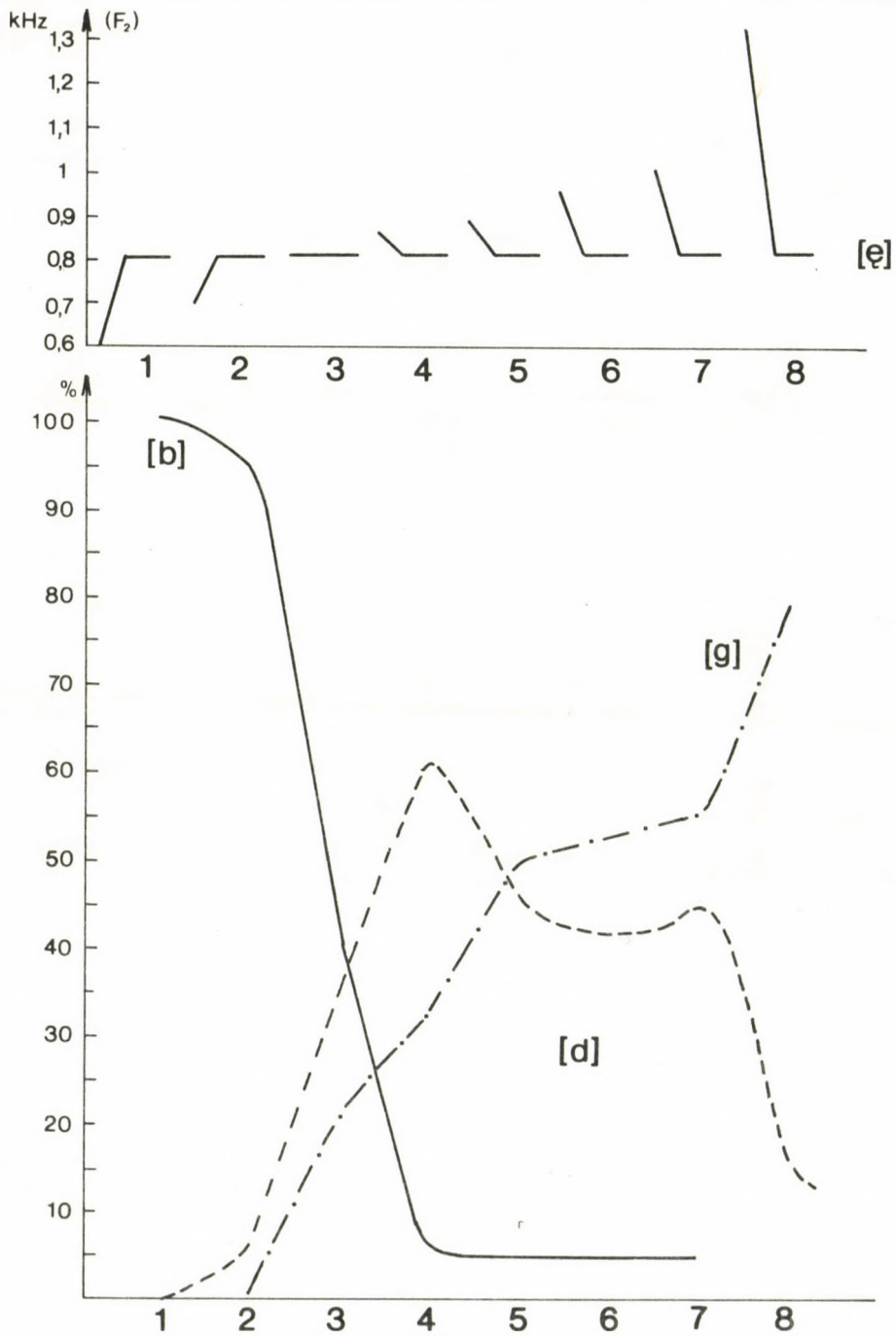
a) a bilabiális és a dentális képzési terület elkülönítése könnyebb, ha a magánhangzó képzése ettől relatíve „távol” esik: veláris magánhangzók és az elől képzettek közül a középső nyelvválású ajakkerekítések;

b) a dentális és a veláris képzési terület elkülönítése könnyebb volt az ettől relatíve „távol” eső összes középső és alsó nyelvválású magánhangzó előtt. (Az alapszerkezetek meghangosításaira nem adódott 100 %-os felismerés a *bí, bé, a dí, dü, de* és a *gá, gú* hangkapcsolatok esetében.)

Azoknak a hangsoroknak a dekódolási folyamata, amelyek egy-két fonéma azonosítását teszik szükségessé, általában „csak” a fonémaszintig jut el; ami a köznapi beszédben rendszerint nem történik meg, hiszen a természetes kommunikációban nem különálló hangkapcsolatokat, szótagokat azonosítunk, hanem nagyobb egységeket. A CV-szekvenciák dekódolásában a megértési folyamat sajátos visszajelző rendszere csak részben működhetik, mivel a fonémaszint „feletti” szintek működési eredményéről nincs visszajelzés, nincs mód a korrekcióra.

4. A beszédhangok percepciójával kapcsolatban az a felfogás uralkodik, hogy különbözőképpen azonosítjuk a magánhangzókat és a mássalhangzókat. A magánhangzók esetében ún. nem kategoriális, a mássalhangzók esetében pedig ún. kategoriális percepcióról beszélünk. A megkülönböztetés Liberman és munkatársai nevéhez fűződik (Liberman—MacNeilage—Harris—Cooper 1963), és a beszédmegértés motoros elméletét támasztja alá (vö. Gósy MFF 8. 1981). Az elmélet lényege, hogy a megértés szorosabb kapcsolatban van az artikulációval, mint az akusztikummal: ennek következtében percepciósan csak az létezik, ami artikulációsan is létezik — a megértés ui. a motoros komponens nem megnyilvánuló közbeiktatásával jön létre. A mássalhangzók egy részének, elsősorban a zárhangoknak a kategoriális percepciójára jellemző, hogy azonosításuk „hirtelen csap át” egyik fonémából a másikba. A fonémahatár közelében élesebb a megkülönböztetés, mint a fonémakategória belsejében.

A magyar zármássalhangzók percepció vizsgálatát magától értetődően felveti a kategoriális észlelés kérdését. A kísérletek alapján az látszik, hogy a beszédhangok percepciója kétségtelenül kategoriális abban az értelemben, hogy az elhangzott



7. ábra

A *be* hangkapcsolat átmenetének különböző F_2 -értékei (1–8-ig), illetve a meghangosításukra adott percepciós válaszok grafikonja

szekvencia meghatározott részét (pl. egy beszédhangot) mindig egy adott anyanyelvi fonémaosztályba soroljuk be. Ez magában foglalja azt, hogy a dekódolás döntési alapja az artikulációs, következésképpen akusztikus sajátosságok észlelésével az anyanyelvi fonéma-rendszer. A kategoriális percepció egy adott nyelv-közösség egy kisebb csoportját tekintve azonban csak akkor mutatható ki egyértelműen, ha a kísérleti nyelvi anyag artikulációs különbsége egyetlen akusztikai paraméterben kifejezhető. A magyar zármássalhangzók képzési helyében megvalósuló artikulációs különbség nem egyetlen akusztikai komponens függvénye (vö. 7. ábra). A kategoriális percepció tehát nem jelentkezik egyértelműen.

Irodalom

- AINSWORTH, W.A.: Perception of stop consonants in synthetic CV syllables. *L&S* 11, 1968, 139–56.
- BLUMSTEIN, S.E.—STEVENS, K.N.: Perceptual invariance and onset time spectra for stop consonants in different vowel environments. *JASA* 67. 1980, 648–61.
- BOLLA Kálmán: A magyar magánhangzók analízise és szintézise. *MFF* 1. 1978, 53–68.
- DELATTRE, P.—LIBERMAN, A.—COOPER, F.: Acoustic loci and transitional cues for consonants. *JASA* 27. 1955, 769–73.
- DORMAN, M.F.—RAPHAEL, L.J.: Distribution of acoustic cues for stop consonant place of articulation in CVC syllables. *JASA* 67. 1980, 1333–6.
- FANT, G.: Stops in CV syllables. *Speech Transmission Laboratory, QPRS. Stockholm* 4. 1969, 1–25.
- GÓSY Mária: A szegmentális hangszerkezet percepciójáról. *MFF* 8. 1981, 87–104.
- LIBERMAN, M.—COOPER, F.S.—HARRIS, K.S.—MACNEILAGE, P.F.: A motor theory of speech perception. *Proc. Comm. Sem. II. Stockholm* 1963.
- OLASZY Gábor: A magyar mássalhangzók és a mássalhangzó–magánhangzó-kapcsolódások akusztikai szerkezetének analízise és szintézise. *MFF* 10. 1982, 46–82.
- STEVENS, K.—BLUMSTEIN, S.: Invariant cues for place of articulation in stop consonants. *JASA* 64. 1978, 1358–68.

THE PERCEPTUAL ANALYSIS OF [b, d, g]

Mária Gósy

The present study examines the criteria for the perception of voiced stops in Hungarian. The author analyzed the sounds [b, d, g] in synthesized CV combinations. The author varied the value of the second formant (F_2) in the sound section representing the plosion and the transition to the following vowel. She tested the synthesized sound combinations on 18 persons of different age and educational background.

The primary aim of many studies on the identification of stops is to determine the component of the stops which enables the listener to identify their place of articulation. According to the results two factors seem to be essential: the frequency of the locus of the stop and that of the transition. The experiments with the Hungarian stops did not yield such a general rule. The analyses indicate the following tendency (cf. figures 5 and 6).

a) The perception of the consonant [b] primarily depends on the frequency of the second formant, which means that the frequency of F_2 corresponds to the bilabial place of articulation on the acoustic level. The frequency range of realization is indicated by the formulae beside each vowel: e.g. if $F_2 < 1796$ Hz $\rightarrow be$, if $F_2 \leq 1550$ Hz $\rightarrow b\check{o}$, if $F_2 \leq 1345$ Hz $\rightarrow bá$, if $F_2 \leq 800$ Hz $\rightarrow bú$.

b) Before palatal vowels the perception of [d] primarily depends on the value of the second formant; before velar vowels it depends both on F_2 and the intensity values. E.g. if $1950 \text{ Hz} > F_2 \geq 1796 \text{ Hz} \rightarrow de$; if $F_2 = 1650-1800 \text{ Hz} \rightarrow d\acute{o}$; if $F_2 = 1300-1345 \text{ Hz} \rightarrow d\acute{a}$; if $F_2 \geq 900 \text{ Hz} \rightarrow d\acute{u}$.

c) Not only the frequency of F_2 is essential in the perception of the consonant [g]. Before front vowels F_2 , the intensity relations, the movement of F_1 and the temporal relations are also important. Before back vowels the intensity curve, the movement of F_1 and the temporal relations are the most important factors, while the value of F_2 is only of secondary importance. E.g. if $F_2 \geq 1958 \text{ Hz} \rightarrow ge$, if $F_2 \geq 1740 \text{ Hz} \rightarrow g\acute{o}$; if $F_2 \geq 1200 \text{ Hz} \rightarrow g\acute{a}$; if $F_2 = 850-1150 \text{ Hz} \rightarrow g\acute{u}$.

I have set up the following hypothesis based on the acoustic parameters of the sound sequences examined and the perceptual tests: a) the bigger „articulatory area” the given sound combination is realized at, the easier its perception is (thus, it is easier to perceive *ge* than *be* or *de*; the same applies to *bú* vs. *gú*, etc); b) the articulatory, and consequently the acoustic characteristics of the following vowel are essential in the perception of the preceding stops.

Our perceptual experiments with the Hungarian stops did not support the theory of categorical identification while they did not contradict the motor theory of speech perception.

These conclusions are true of all the Hungarian combinations of the form [b, d, g] + V.

A MAGYAR BESZÉD AUTOMATIKUS SZINTÉZISÉNEK ELSŐ LÉPCSŐJE

Kiss Gábor—Olaszy Gábor

A magyar beszéd mesterséges előállítására irányuló törekvések új keletűek. Magyarországon csak az utóbbi években alakultak ki azok a feltételek, amelyek lehetővé tették, hogy a magyar beszéd mesterséges előállítására irányuló munkák megkezdődjenek. Köztudott, hogy a szintetizált beszéd a tudományos kutatás, a kommunikációs szolgáltatások, az anyanyelvi oktatás, a számítógépipar, valamint a társadalom számos területén felhasználható. Az igények egyre jobban sürgetik, hogy a magyar beszédre vonatkozóan is megszülessenek azok a nyelvészeti-technikai kutatási eredmények, amelyek alapján számítógépeket, mikroprocesszoros berendezéseket meg lehet „tanítani” magyarul beszélni. Az ilyen kutatások mind nyelvészeti-fonetikai, mind pedig technikai szempontból igen bonyolultak, és speciális képzettségű szakembereket igényelnek. Magyarországon többek között a Posta Kísérleti Intézetben (PKI) és az MTA Nyelvtudományi Intézetének fonetikai osztályán folynak kísérletek a magyar beszéd mesterséges előállítására. A PKI munkatársa, Takács György már 1978-ban kifejlesztett egy olyan számítógépprogramot és hardware-t, amellyel úgynevezett félszintetizált beszédet lehet használni a megváltozott telefonszámok automatikus közlésére. Ez a rendszer számokat tud kimondani 0-tól 9-ig és néhány mondatot. A félszintetizált beszéd lényege abban áll, hogy a számokat, valamint az információ szolgáltatásához szükséges mondatokat természetes emberi beszédből állítják össze, számítógépes mintavételezéssel rögzített szótagok vagy szórészek összekapcsolásával.

A magyar beszéd szabályokon alapuló szintéziséhez az MTA Nyelvtudományi Intézete fonetikai osztályán folyó alapkutatások eredményei jelentik az első lépéseket. Munkánk során olyan számítógépprogramot (FOPRO) hoztunk létre (vö. Kiss—Olaszy MFF 10. 1982), amelynek segítségével el tudtuk végezni azokat a szintetizálási munkákat, amelyek a magyar beszédhangok akusztikai szerkezetének teljes feltárásához kellettek. A munka során létrehoztunk egy olyan adatbázist (vö. Olaszy MFF 8. 1981), amely tartalmazza a magyar beszéd legfőbb alkotóelemeit képező # CV, VCV és VC # hangkapcsolódások és néhány CC-kapcsolat akusztikai szerkezetének létrehozásához szükséges adatokat, így alap lehet a későbbi, automatikus beszédszintézis hardware és software rendszerének kidolgozásához. Az automatikus szintézishez azt kell elérni, hogy a számítógép (vagy mikroprocesszor) a megadott szabályok és adatok segítségével a mesterséges beszédet a természetes beszéd tempójában közvetlenül az utasítás megkapása után azonos idejű üzemmódban generálni tudja. Természetesen az ilyen „beszélő” rendszerek vagy gépek elkészítéséhez szükséges szellemi és műszaki ráfordítás mértéke arányban áll a kitűzött feladat szintjével. Általában még a fejlett ipari országokban is — ahol a mesterséges beszédet már számos területen alkalmazzák — leginkább a korlátozott, adott szószámot igénylő feladatoknál használják a számítógépbeszédet. A meghatározott szóképlet előállítása nagyságrendekkel kisebb feladat, mint a tetszőleges szöveget meghangosító, ún. „text to speech” szintézis. Az előbbinél ugyanis csak a megadott szavak előállításához szükséges adatokat és szabályokat kell a számítógép számára kidolgozni, ebből következik, hogy viszonylag kis adatmennyiséggel kell dolgozni.

Az utóbbi megoldásnál a szintetizálni kívánt nyelvre vonatkozó szinte összes szabályt és adatot – a kivételeket is – ki kell dolgozni, ami komoly és hosszú kutatómunkát igényel, nem is beszélve a hatalmas adattömegekről, amit a számítógépnek kezelnie és mozgatnia kell a szintéziskor.

Osztályunkon a magyar beszéd automatikus szintézisének első lépcsőjeként az MFF 8-ban ismertetett adatbázis felhasználásával „Számok” néven készítettünk automatikus szintetizáló programot.

A szintézishez PDP 11/34-es számítógépet és OVE III típusú beszéd szintetizátort használtunk. A program és a számítógép segítségével bármilyen számot ki tudunk mondani a szintetizátorral 1-től 1 billióig. A számítógép értelmezi az alapvető számtani műveleteket és az azokkal felépített képleteket is, kívánságra azokat ki is mondja, közben kiszámítja a képlet végeredményét, és azt hangosan közli. Például a $3617\text{-es szám leütése után a számítógép a következőt mondja: háromezer-hatszáz tizenhét. A } 2 \times (49 + 97) / 4 = ?$ karaktersorozat beütése vagy megadása után a következőt halljuk: *kettőször – negyvenkilenc – plusz – kilencvenhét – per – négy – egyenlő – hetvennégy* vagy a 864–523 jelsorozatra kimondja a telefonszámot. Az alábbiakban ismertetjük a „Számok” program előkészítésének fázisait, rendszerét és működését.

Az akusztikai építőkockák megtervezése

Ha számítógéppel minden számot meg akarunk szólaltatni 1-től 1 billióig, akkor első lépésben egy olyan mátrixot kell készítenünk, amelyik tartalmazza a számok felépítésének összes variációs lehetőségeit. A műveleti jelek megszólaltatásához a mátrixba be kell építeni az összeadás, kivonás, szorzás, osztás szavait, valamint az *egyenlő* kifejezést is. A mátrix első oszlopában helyezkednek el azok a számok és számrészek, amelyek alapelemként szolgálnak az egy- és többjegyű számok előállításához. Összesen 23 ilyen alap-építőkockát különböztetünk meg. Ezek a következők: *1, 2, 3, 4, 5, 6, 7, 8, 9, 10, -tizen-, 20, huszon-, 30, 40, 50, 60, 70, 80, 90, 100, 1000* és *millió*. Ezek helyes kapcsolódásaiból felépülhet minden szám, amely 1 és 999999999 közé esik. A helyes kapcsolódások szintaktikai szabályai:

<SZ 1>	= egy kettő három négy öt hat hét nyolc kilenc
<SZ 11>	= tíz húsz
<SZ 12>	= harminc negyven ötven hatvan hetven nyolcvan kilencven
<SZ 111>	= tizen huszon
<SZ AZ>	= száz
<EZER>	= ezer
<MILLIÓ>	= millió
<SZ 10>	<SZ 11> <SZ 12> <SZ 12> <SZ 12> <SZ 1>

<SZ 100>	= <SZ 1> <SZÁZ>
	<SZ 1> <SZÁZ> <SZ 10>
<SZ 100>	= <SZ 100> <EZER>
	<SZ 100> <EZER> <SZ 100>
<SZ 1000000>	= <SZ 100> <MILLIÓ>
	<SZ 100> <MILLIÓ> <SZ 1000>
SZÁM	= <SZ 10000.000>
	<SZ 1000>
	<SZ 100>
	<SZ 10>
	<SZ 1>

Az alapelemeket 1-től 23-ig számozott azonosítóval jelöljük. A mátrix első sorában ugyanezek az elemek és a műveleti jelek foglalnak helyet. Ezek képviselik a kettő vagy többbelemű számsor következő elemét (1. ábra).

A mátrix első oszlopában lévő alapelemek sorainak és az első sorban lévő „következő” számok, illetve műveleti jelek oszlopainak találkozásánál található az egyes számkombinációk létrehozásához szükséges csatoló elemek. Ezeket is azonosítóval láttuk el 24-től felfelé. Természetesen minden vízszintes sorbeli elem nem találkozhat minden függőleges oszlopbelivel. Például az *egy* után csak a *száz*, *ezer* vagy *millió* állhat. Az *1000* után viszont az *1, 2, 3, 4, 5, 6, 7, 8, 9, 10, tizen, huszon, 20, 30, 40, 50, 60, 70, 80, 90*-nel kezdődő elemek adnak értelmes számot. A mátrix természetesen tartalmazza a hangsor eleji és hangsor végi szünetet is mint építőelemet.

Tulajdonképpen a mátrix első oszlopában lévő 23 számelem felhasználásával már minden számot össze lehetne állítani. Ehhez csak a megszólaltatni kívánt számhoz tartozó elemeket kellene a programnak egymás után kapcsolnia. Például a 717-es számot a *hét – száz – tizen – hét* elemek összekapcsolásával kapnánk meg. Ezt azonban csak az írás szintjén tehetjük meg így. A beszéd szintjén, hogy a számok hangzását a leghűbben megvalósítsuk, figyelembe kell venni a hasonulásokat, a kapcsolódási szabályokat, valamint azokat az időszerkezeti szabályokat is, amelyek a számhangsorok képzésére érvényesek. Vizsgálataink során kimutattuk, hogy a számhangsorok kimondásakor ugyanazon elemek időtartama egy számon belül változó, attól függően, hogy milyen számelem környezetében van. Például a *huszonkettő* és a *huszonnyolc* számok esetében a második szótagokban lévő [o] hangok időtartama nem ugyanaz. Továbbá a mérések kimutatták, hogy egyes számelemek magánhangzóinak időtartama más a hangsor eleji helyzetben és más hangsor záró helyzetben. Például a *húszezer-kettőszázhusz* számban az első *húsz* időtartama rövidebb, mint a másodiké.

Szintetizálási kísérleteink igazolták, hogy az ilyen és ehhez hasonló jelenségek figyelembevétele lényegesen javítja a gépi beszéd minőségét.

Az akusztikai építőkockákba dallamot nem terveztünk, tehát a kimondott számok és műveletek egyenlő alaphangmagasságon szólalnak meg. Az előbbiekből kialakított mátrixból kiolvasható, hogy a számok és a műveletek meghangosításához maximum 340 építőelemre van szükség. Mivel az egyes csatlakozási formák és hasonulások nemcsak egy számkombináció előállításánál fordulnak elő, hanem többnél is; a tényleges elemszám a 340-nél lényegesen kevesebb. Például a [t] és [s] hang találkozásakor kialakuló

A k ö v e t k e z ő s z á m

	1	2	3	4	5	6	7	8	9	10	20	30	40	50	60	70	80	90	10-en	20-on	100	10 ³	10 ⁶	szó vége					
Számok	A z o n o s í t ó k																												
1	1																								25	26	27	28	
2	2																									30	31	32	33
3	3																									35	36	37	38
4	4																									40	41	42	43
5	5																									45	46	47	48
6	6																									45	46	47	48
7	7																									45	46	47	48
8	8																									45	51	52	53
9	9																									45	51	52	53
10	10																										54	55	56
20	11																										57	58	59
30	12	24	29	34	39	44	34	49	50	29																	51	52	53
40	13	60	61	63	64	65	63	63	66	61																	60	67	68
50	14	60	61	63	64	65	63	63	66	61																	60	67	68
60	15	69	70	71	72	73	71	71	74	70																	75	69	76
70	16	60	61	63	64	65	63	63	66	61																	67	60	68
80	17	69	70	71	72	73	71	71	74	70																	75	69	76
90	18	60	61	63	64	65	63	63	66	61																	67	60	68
10-en	19	60	61	63	64	65	63	63	66	61																			
20-on	20	77	78	79	80	81	79	79	66	78																			
100	21	83	84	85	86	87	85	88	89	84	84	85	85	86	87	85	88	89	84	84	85						90	83	91
10 ³	22	24	29	34	39	44	34	49	50	29	29	34	34	39	44	34	49	50	29	29	34								29
10 ⁶	23	24	29	34	39	44	34	49	50	29	29	34	34	39	44	34	49	50	29	29	34								29
szó eleje		24	29	34	39	44	34	49	50	29	29	34	34	39	44	34	49	50	29	29	34								

1. ábra
A lehetséges kapcsolódásokat realizáló mátrix

hasonulás ugyanaz az *ötszáz*, *hatszáz* és a *hétszáz* szavakban. Ebből következik, hogy a mátrix egyes belső csatoló elemei egyformák lesznek. Így kialakul az a végleges építőelemszám – 119 –, amellyel a feladatot meg tudjuk oldani (vö. 1. ábra).

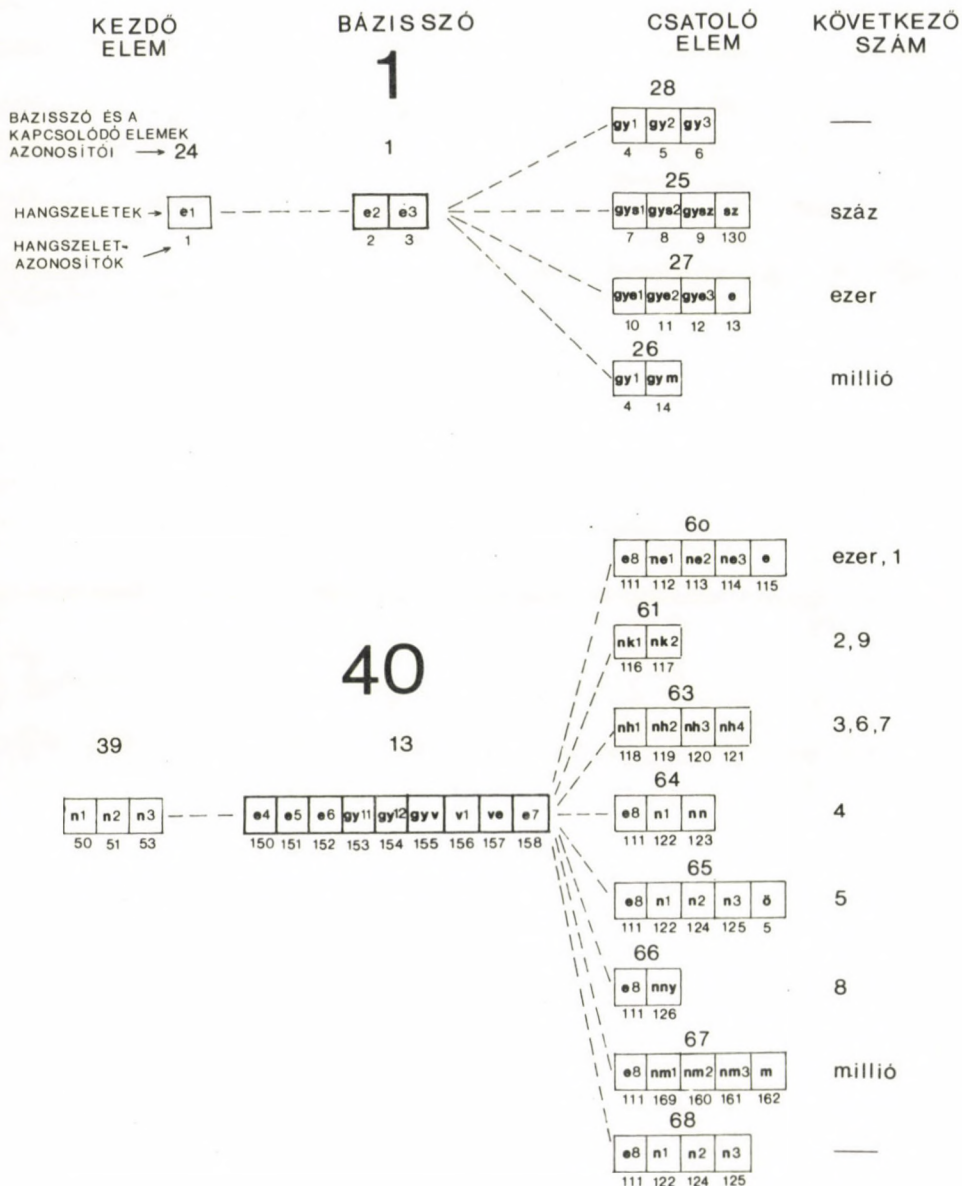
Az adatbázis elkészítése

Az automatikus szintézishez a számítógépprogramnak szüksége van az előbb említett mátrix elemeit realizáló hangszetelek adataira. A hangszeteleket az MFF 8-ban ismertetett adatbázis alapján állítottuk össze. A mátrixot képviselő adatbázist úgy alakítottuk ki, hogy a SZÁMOK programnak a lehető legkevesebb lépésre legyen szüksége a szintézishez. Ez lényeges szempont, hiszen az azonos idejű üzemmód megköveteli, hogy a műveletek számát a legminimálisabbra csökkentjük. Csak így lehet biztosítani, hogy a számítógép el tudja végezni a szám értelmezését, a helyiértékek megállapítását, a szükséges bázisszavak és a csatoló elemek kiválasztását, összeállítását, összekapcsolását a számhoz, valamint a műveleti jelek előhívását és az azokkal járó ragok hozzákapcsolását. Az optimális adatbázis kialakításához bevezettük az ún. bázisszó-építőkockát és a kapcsolódó csatoló elemet. A 23 alapelem mindegyikére meghatároztuk azt a hangsorrészt, amelyik minden előfordulási variációban ugyanaz, hangzása állandó. **E z a b á z i s s z ó.** A bázisszóhoz kapcsolódik a kezdő elem és a változó akusztikai részeket tartalmazó csatoló elem. Így minden számot a bázisszóval és az eléje és utána kapcsolt elemmel valósítunk meg. A bázisszó mindig állandó, a kapcsolódó elem attól függően változik, hogy a szám után mi következik. Ha a meghangosítandó szám több számjegyű, akkor azt a megfelelő bázisszavak és kapcsolódó elemek sorozatából állítjuk elő. Példaként az 1-es és a 40-es szám bázisszavát és a számhoz kapcsolódható elemeket a 2. ábrán mutatjuk be.

A 2. ábra szerinti felosztást elvégeztük mind a 23 alapelemre, így megkaptuk a mátrix tényleges elemeit. Az első oszlopba kerül tehát a 23 bázisszó és a szóeleji szünet. A többi oszlopot pedig a megfelelő kezdő és csatoló elemmel töltjük meg. Ezen elemek azonosítóiból leolvasható, hogy mely számkombinációknál használjuk ugyanazokat a kezdő vagy csatoló elemeket. Ezek azonosítója egyenlő egyforma.

A bázisszavakat és a kapcsolódó elemeket hangszetelekből építettük fel. A hangszeteleket szintén 1-től kezdődően azonosítóval láttuk el (vö. a 2. ábrán). Természetesen ezeket az azonosítókat a program más helyen tárolja, mint a bázisszavakét és a kapcsolódó elemekét. A hangszeteleknél is előfordult az, hogy ugyanazon hangszetet fel lehetett használni különböző bázisszavak vagy csatoló elemek felépítésénél. A hangszetek számának ilyen optimalizálásával a memóriagényt sikerült csökkenteni. A bázisszavak, a kezdő és csatoló elemek és a műveleti jelek felépítéséhez összesen 209 hangszetet használtunk fel.

A számítási műveleteket, vagyis képleteket a KEPLET nevű számítógépprogram hangosítja meg. A programot legegyszerűbben úgy mutathatjuk be, ha leírjuk működését, amely a következő lépcsőkben valósul meg egy adott képlet meghangosítása során.



2. ábra
Példa a szintetizált számok felépítésére

1. Kezdeti betöltés, a floppy disken tárolt adatbázisból a szükséges szeletek beolvasása a központi memóriába.
2. A „~” (tilde) karakter kiírása, evvel jelzi a számítógép, hogy készen áll a meghangsítandó képlet leolvasására.
3. A felhasználó által a képlet beolvasása.
4. A képletsorozat szintaktikai elemzése: a számok, műveleti jelek elkülönítése.
5. A képlet értékének kiszámítása, lengyel formára hozás segítségével, az eredménynek a szintaktikai sorozathoz való csatolása.
6. A szintaktikai egységek sorozata alapján a megszólaltatandó szeletek, hangrészek összeválogatása.
7. A kiválasztott hangrészek egymás utáni lejátszása, meghangsítása.
8. A feladat 2-es ponttól való folytatása.

*

A KEPLET program tapasztalatai alapján 1981–82-ben továbbléptünk, és elkészítettük az UNIVOICE szintetizáló rendszert, amelynek segítségével nemcsak számokat, de bármilyen magyar szöveget real time meghangsíthatunk. A program 365 hangszelet segítségével minden magyar V, VV, CV, VCV, VC és CC hangsorépítő elemet elő tud állítani, értelmezi a leggyakoribb hasonulásokat, valamint a megadott hangsorra kijelentő és felszólító dallamformát is rá tud ültetni. Ez a magyar nyelvű folyamatos text-to-speech szintézis első lépése. Az UNIVOICE rendszert a 8. Budapesti Akusztikai Kollokviumon mutattuk be. A rendszerről az MFF egyik későbbi számában számolunk be részletesen.

Irodalom

- KISS Gábor—OLASZY Gábor: Interaktív beszéd szintetizáló rendszer számítógéppel és OVE III szintetizátorral. MFF 10. 1982, 21–46.
 OLASZY Gábor: Hangsorok számítógépes formánsszintézisének előkészítése. MFF 8. 1981, 147–60.

THE FIRST STEP IN THE AUTOMATIC SYNTHESIS OF HUNGARIAN

Gábor Kiss—Gábor Olasz

This paper gives an account of the first Hungarian real time speech synthesizing program. The results of the research done at the Phonetics Department of the Institute of Linguistics of the Hungarian Academy of Sciences constitute the acoustic basis of the synthesis. A PDP 11/34 computer and an OVE III synthesizer were used in the synthesis. The research was focussed on the formant synthesis of the Hungarian speech by rule. Our main goal in making the program was to try out the data of the data basis gathered as a result of the research done so far. The program makes numbers audible, from 1 to 999 million. When we press e.g. the key 3617, on the keyboard of the typewriter and then press „return”, the computer immediately starts to feed the data into the synthesizer and the number becomes audible. There are 23 basic elements in the data basis from which any number can be made up. These are the Hungarian equivalents of (words or word-parts) 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90. ...-teen, twenty-..., 100, 1000 and million.

In addition to these, the data basis contains those elements of the sound sequence which enable us to consider the rules of assimilation and sound combination necessary for combining the basic elements.

Because of the assimilations, the initial number of the basic elements (23) has increased considerably (to 119). Each element is composed of sound sections and the sound sections are composed of micro-sound sections. 209 sound sections were stored in the memory of the computer. The computer combines the sound sections and orders them according to the number of digits as soon as it gets the command. Thus, any number can be made audible immediately. An improved part of the program (KEPLET) is able to interpret operations and mathematical formulas using these operations and can also make them audible. In case of formulas, while the program makes the formula in question audible, it also calculates and gives the result. Our program only works with integers.

On the basis of the program KEPLET we developed the UNIVOICE synthesizing system in 1982 (presented at the Acoustic symposium on 6-th May 1982). The new system is not only capable of making numbers audible but it can make any Hungarian text audible real time. With 365 sound sections the program is able to generate any Hungarian element of the sound sequence; it can interpret the most frequent assimilations and can build an interrogative or an assertive intonation pattern on the given sound sequence.

FOLYAMATOS BESZÉDSZINTETIZÁLÓ RENDSZER MAGYAR NYELVEN (VOXON)

Bolla Kálmán

A magyar beszéd akusztikai szerkezetének az elemzése és szintézise témában végzett kutatásaim újabb — a kísérleti fonetika hazai fejlődése szempontjából is lényeges — eredményéről számolhatok be. 1982 márciusában elkészült az első magyar nyelvű imperszonális beszédszintetizáló rendszer, amelyet Voxonnak (vox-sonus) neveztem el. A Voxont az MTA Nyelvtudományi Intézete igazgatóságának március 29-én tett bejelentés után április 2-án a fonetikai osztály munkatársainak is bemutattam. A Voxonnal bármilyen magyar nyelvű hangsor, tetszés szerinti beszédszöveg előállítható.

A Voxon felépítése és használata

A folyamatos beszédszintetizáló rendszer három fő részből áll: a) hangszelettárból, az elemi akusztikus szerkezeti egységek adataiból, azaz adatbázisból; b) az elemek hangokká, hangsorokká szervezéséhez szükséges fonetikai kódból és c) a műveleti programokból. Ez utóbbit Kiss Gábor programozó készítette.

A szintetizáló rendszer megtervezéséhez az alábbi körülményekkel kellett számolnunk:

a) A nyelvnek mint sajátos emberi jelzőrendszernek az egyik lényeges összetevőjét, pontosabban az akusztikus realizációs részrendszerét akarjuk mesterségesen előállítani. Ebből az következik, hogy a szintézis kezdő és befejező szakaszában a hangtestben fellelhető nyelvi—fonetikai struktúrákból kell kiindulnunk, illetőleg azokhoz kell eljutnunk.

b) A beszédhangsor fizikai-akusztikai szerkezete és a nyelv fonológiai—fonetikai organizációja nem fedik egymást. Az akusztikai és a nyelvi—fonetikai struktúra különbözősége folytán az akusztikus folyamatban más-más szerkezeti alapelemeket kapunk aszerint, hogy a fizikai azonosság—eltérés vagy a nyelvi—fonetikai azonosság—eltérés alapján határozzuk-e meg a folyamat diszkrét egységeit. A bonyolult szerkezeti felépítésű fonetikai struktúrák, amelyek egyben elvont típusok is, konkrét fizikai megvalósulásukban az akusztika törvényszerűségeinek megfelelően szerveződő szerkezetekben léteznek. A beszédhangsor mesterséges előállításában tehát figyelembe kell vennünk a nyelvi sajátosságok mellett a fizikai valóság tényeit is, mivel a beszéd folyamat mindig olyan konkrét fizikai folyamatváltozással valósítható meg, amelynek megvannak a saját belső mozgásformái és törvényszerűségei.

c) A beszédszintetizáló rendszer felépítését meghatározó tényezők harmadik csoportját a szintézisben felhasználható technikai rendszer képezi. A műszaki feltételek határozzák meg az alkalmazható technológiát, az elérhető minőséget, biztosítják a szintézis hatékonyságát, finomságát stb. Minthogy egyetlen technikai rendszer sem pontos analógja az emberi beszédképző tevékenységnek, továbbá mivel minden rendszer más-képpen működik, különbözőképpen állítja elő a beszéd akusztikus elemeit, a beszédszintetizáló rendszerek általában nem alkalmasak sem a beszédképzés mechanizmusainak megfelelő működtetésre, sem pedig a beszéd folyamat akusztikus elemzéséből nyert

valós fizikai paraméterek közvetlen betáplálásával való működtetésre. A fizikai adatokat alkalmazni kell a technikai rendszerre, s csak bizonyos ártérítéssel használhatók a szintézishez.

A fonetikai struktúra legkisebb szegmentális elemét beszédhangnak nevezzük. Erőről részletesebben A beszédhang fonetikai definíciója címmel az MFF 9-ben írtam (Bolla MFF 9. 1981, 9). Ezzel szemben az akusztikai struktúra legkisebb szerkezeti egységének a hangszeletet tekintjük. A hangszelet a beszédhangsor akusztikai szerkezetében az akusztikus összetevők folyamatváltozásainak a figyelembevételével kikülöníthető homogén felépítésű szelvény, szegmentum. A hangszelet több fizikai paraméterből épül fel. A szegmentum egyneműségét az összetevők adatainak az állandósága, illetőleg nem változó irányú, egyenes vonalú változása adja. A beszédhang, fizikai alkatát tekintve, bonyolultabb képződmény. Egy-egy hang terjedelmében az összetevők változása már több irányú, azaz több törésponttal jellemezhető. Így a beszédhang néhány hangszeletből építhető fel. Előfordul, hogy a hangszelet és a beszédhang egybeesik, például a [m], [n], [l] hangok bizonyos fonetikai helyzetekben egy szelettel is megvalósíthatók.

A hangszelet felépítésében az alábbi fizikai összetevők játszanak szerepet:

- az időtartam (T)
- a zöngés alapfrekvencia (F_0)
- a zöngés alaprezgés erőssége (I_0)
- a formánsok frekvenciája (F_1, F_2, F_3, F_4)
- a formánsok erőssége (I_1, I_2, I_3, I_4)
- a formánsok frekvenciájának sávszélessége (B_1, B_2, B_3, B_4)
- a zörejjócok frekvenciasávja (K_1, K_2, K_3, K_4)
- a zörejjócok erőssége (I_1, I_2, I_3, I_4)

A hangtest akusztikai szerkezetében a felsorolt összetevők kombinációiból négyféle elem-típus jön létre:

- a) szünet (P),
- b) zöngés struktúrájú elem (Z),
- c) zörejsztruktúrájú elem (K) és
- d) összetett vagy kevert struktúrájú elem (ZK).

Szünetnek (P) nevezzük a beszédhangsorban előforduló nulla intenzitású szakaszokat. Hangzásszünetek nemcsak szólamok, beszédszakaszok között vannak, hanem egy-egy beszédhang tartományán belül is fellelhetők, mint például a zöngétlen zárhangok néma fázisában vagy a zöngétlen affrikáták kezdő szakaszában.

A zöngés struktúra (Z) kialakításában az időtartamon kívül az alapfrekvenciának, az alaprezgés intenzitásának, a formánsfrekvenciának, a formánsintenzitásnak és a formáns-sávszélességnek van szerepe. Ezen összetevők adatainak a variációiból n számú különböző struktúra képezhető (ld. a magánhangzók és a szonánsok akusztikai szerkezetét).

A zörejsztruktúrájú elemekben (K) az időtartam mellett a zörejjócok frekvenciaadatainak, a zörejj erősségének és a zörejj jellegének van jelentősége. A zörejj jellegét a hangforrás milyensége határozza meg. Eszerint megkülönböztetünk impulzív és turbulencia zörejjeket (vö. [p, t, k] ↔ [f, s, ʃ, h]).

Az összetett struktúrájú elemekben (ZK) keverednek, különböző arányban és

formában ötvöződnek a zöngés és zörejes összetevők. (Vö. a zöngés zörejhangokat: [b, d, g, v, z, ʒ].)

Mindhárom ismertetett akusztikai elemtípusban számtalan eltérő minőségű és különböző tartamú szegmentum keletkezhet, míg a szünetelem csak mennyiségi variálást enged meg. Minden nyelv az emberi hangképző apparátussal létrehozható akusztikai alapelemekből „válogatja össze” elemtárát és ennek jellegzetes felhasználásával, azaz meghatározott szabályok szerinti kombinálásával „alkotja meg” hangsorait. Az idézőjel használatát az indokolja, hogy a valóságban mindez másképpen történik. A beszéd folyamata, azaz a hangalak létrehozása feltehetően nem ennyire analitikusan megy végbe, hanem globálisan, a jelentéssel bíró nyelvi elemek (morfémák, szavak) hangtestének a megalkotására irányuló tevékenységként. A hangszeletek és a beszédhangok is a hangfolyam tudományos elemzésének a termékei, objektíve létező hangszerkezeti tényezők. A beszéd folyamat akusztikai szintéziséhez tehát a természetes hangsorokból ki kell elemeznünk a kváziazonos formában ismétlődő szegmentumokat és alapszerkezeteket, valamint meg kell ismernünk az elemek kapcsolódásának, kombinációinak szabályosságait, törvényszerűségeit is.

A Voxon beszéd szintetizáló rendszerbe 550 alapelem, hangszelet adatait építettem be. Ebből 27 hangszelet szolgál a magánhangzók különböző típusainak az előállítására, 89-ből pedig a mássalhangzókat szintetizáljuk. A hangok időtartambeli variálását, a rövid–hosszú szembenállás fonetikai kifejezését, a megfelelő hangszelet duplázásával érjük el. A beszédhangok fonetikai jellege, akusztikai minősége a magyarban is változik a fonetikai helyzet függvényében. A pozíciót meghatározó tényezők közül kettőnek: a) a hangkörnyezetnek és b) a szóban elfoglalt helynek az akusztikai szerkezetet módosító hatását tartottam a legjelentősebbnek, s így a rendszerbe beépítendőnek. A hangkörnyezeti sajátosságok kifejezésére szolgálnak az ún. átmeneti szeletek. Ezek száma a legnagyobb: 414. A hangátmeneteknek nemcsak az akusztikai struktúra képzésében van jelentőségük, hanem a beszédészlelésben és beszédértésben is, ezért a szintetizáló rendszerünk kidolgozásában arra törekedtünk, hogy a Voxonnal biztosítani tudjuk e probléma behatóbb tanulmányozhatóságát a beszéd szintézis által is. A szóban elfoglalt hely szerint minden hang három helyzetnek: szóeleji, szóbelseji és szóvégi helyzetnek megfelelő minőségben állítható elő. Ezt a kezdő és a befejező szeletek változtatásával valósítja meg a rendszer. E pozicionális sajátságoknak bizonyos tagoló és határjelző funkciójuk van a magyarban. Végül az elemtárban minden magánhangzónak megtalálható egy nyíltabb és egy zártabb változata is.

A szintézis hardware és software részével külön nem foglalkozunk, mivel ez megtalálható Kiss Gábor és Olasz Gábor e füzetben közölt ismertetésében (Kiss–Olasz MFF 10. 1982, 21), csupán a szintetizálási folyamat egészének könnyebb áttekinthetősége miatt mutatunk rá arra, hogy a kielemezett fizikai paraméterek hogyan, milyen transzformációval használhatók a szintetizáló rendszer bemenetén.

A hangszelet időtartama TM időadat-bevitellel, IDA adatváltoztató utasítással és IMI hangsoridőtartam-nyújtó és -zsugorító művelettel adható meg, illetőleg alakítható, módosítható a szintézis folyamán. (Ld. IDA-táblázat.)

A zöngés szeletstruktúrák mesterséges előállításához az alapfrekvenciát FO-val, az alaprezgés erősségét AO-val, nazálisoknál AN-nel, az első három formáns frekvenciáját F1, F2, F3-mal, illetőleg a nazálisoknál N1-gyel visszük be, míg a formáns-sáv-

IDA \ TM	3	4	5	6	7	8	9	10
10	3	4	5	6	7	8	9	10
20	6	8	10	12	14	16	18	20
30	9	12	15	18	21	24	27	30
40	12	16	20	24	28	32	36	40
50	15	20	25	30	35	40	45	50
60	18	24	30	36	42	48	54	60
70	21	28	35	42	49	56	63	70
80	24	32	40	48	56	64	72	80
90	27	36	45	54	63	72	81	90
100	30	40	50	60	70	80	90	100
110	33	44	55	66	77	88	99	110
120	36	48	60	72	84	96	108	120
130	39	52	65	78	91	104	117	130
140	42	56	70	84	98	112	126	140
150	45	60	75	90	105	120	135	150
160	48	64	80	96	112	128	144	160
170	51	68	85	102	119	136	153	170
180	54	72	90	108	126	144	162	180
190	57	76	95	114	133	152	171	190
200	60	80	100	120	140	160	180	200
210	63	84	105	126	147	168	189	210
220	66	88	110	132	154	176	198	220
230	69	92	115	138	161	184	207	230
240	72	96	120	144	168	192	216	240
250	75	100	125	150	175	200	225	250

A hangzási idő meghatározása IDA- és TM-érték alapján (elméleti érték)
(A táblázatot Kiss Gábor készítette.)

11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
22	24	26	28	30	32	34	36	38	40	42	44	46	48	50
33	36	39	42	45	48	51	54	57	60	63	66	69	72	75
44	48	52	56	60	64	68	72	76	80	84	88	92	96	100
55	60	65	70	75	80	85	90	95	100	105	110	115	120	125
66	72	78	84	90	96	102	108	114	120	126	132	138	144	150
77	84	91	98	105	112	119	126	133	140	147	154	161	168	175
88	96	104	112	120	128	136	144	152	160	168	176	184	192	200
99	108	117	126	135	144	153	162	171	180	189	198	207	216	225
110	120	130	140	150	160	170	180	190	200	210	220	230	240	250
121	132	143	154	165	176	187	198	209	220	231	242	253	264	275
132	144	156	168	180	192	204	216	228	240	252	264	276	288	300
143	156	169	182	195	208	221	234	247	260	273	286	299	312	325
154	168	182	196	210	224	238	252	266	280	294	308	322	336	350
165	180	195	210	225	240	255	270	285	300	315	330	345	360	375
176	192	208	224	240	256	272	288	304	320	336	352	368	384	400
187	204	221	238	255	272	289	306	323	340	357	374	391	408	425
198	216	234	252	270	288	306	324	342	360	378	396	414	432	450
209	228	247	266	285	304	323	342	361	380	399	418	437	456	475
220	240	260	280	300	320	340	360	380	400	420	440	460	480	500
231	252	273	294	315	336	357	378	399	420	441	462	483	504	525
242	264	286	308	330	352	374	396	418	440	462	484	506	528	550
253	276	299	322	345	368	391	414	437	460	483	506	529	552	575
264	288	312	336	360	384	408	432	456	480	504	528	552	576	600
275	300	325	350	375	400	425	450	475	500	525	550	575	600	625

A hangzási idő meghatározása IDA- és TM-érték alapján (elméleti érték)

(A táblázatot Kiss Gábor készítette.)

szélesség és a formánsintenzitás együttesen kezelhető a B1, B2 és B3 kód szerinti értékek adataival. Az F4-et a szintetizátor automatikusan adja hozzá a zöngés szerkezethez.

A z ö r e jstruktúrájú szeletek szintézise AC, AH, AK, K1 és K2 bemeneti adatokkal történik. Ezen kódok bemeneti értékeiből kapjuk meg a zörej frekvenciasávját, erősségét, képezhetők a zörejgócok és közelíthető meg az impulzív és a turbulencia zörejnek megfelelő szerkezet. A természetes beszédben előforduló zörejstruktúrák sokféleségét nehezebb az OVE III-mal „hangzáshűen” előállítani, mint a zöngés szerkezeteket. Az elemzési adatok felhasználása is nagyobb transzponálást kíván.

A k e v e r t struktúrájú szeletek szintézisében mindhárom fentebb ismertített szintézisblokkot felhasználjuk.

A Voxon fonetikai kódját a hangszeletek azonosítói és a hangok „hívójelei” alkotják. Az azonosítók 1–3 karakterből állnak és jelzik, hogy: a) milyen beszédhanghoz tartozik a szelet, b) milyen helyet foglal el a beszédhang fonetikai szerkezetében és c) milyen helyzetű a szelet a hangsorban elfoglalt helye szerint. Például a JO azonosítóból megtudjuk, hogy az [o] hang első, a [j] mássalhangzó utáni átmeneti szeletét megvalósító adatokat tartalmazza, a 2T a [t] hang zárfelpattanásának a lecsengő szakaszát jelöli szóbelseji helyzetben, míg a T~ ugyanezt, csak szóvégi helyzetben.

A hangok „hívójeleit” az LA 36 típusú konzol írógép betűiből és írásjeleiből állítottam össze. Az angol ábécé felhasználásával kialakított grafikai szimbólumokat a magyar beszédhangok fonetikai jelölésére használok úgy, hogy ez nem felel meg teljesen sem a nemzetközi fonetikus írás, sem pedig a magyar helyesírás rendszerének, hanem mindkettőből összeválogattam azokat az elemeket, amelyek a konzol írógéppel kivitelezhetők.

A magyar hangok számítógépes szimbólumai:

A magánhangzók

[a:] — ~ A', A', A' ~
 [ɔ] — ~ A, A, A~
 [o] — ~ O, O, O~
 [u] — ~ U, U, U~
 [y] — ~ Y, Y, Y~

[i] — ~ I, I, I~
 [e:] — ~ E', E', E' ~
 [e] — ~ E, E, E~
 [ø] — ~ Q, Q, Q~

A mássalhangzók

[b] — ~ B, B, B~
 [p] — ~ P, P, P~
 [m] — ~ M, M, M~
 [d] — ~ D, D, D~
 [t] — ~ T, T, T~
 [n] — ~ N, N, N~
 [ʃ] — ~ D', D', D' ~
 [ç] — ~ T', T', T' ~
 [ʒ] — ~ N', N', N' ~
 [g] — ~ G, G, G~

[s] — ~ S, S, S~
 [ʒ] — ~ Z, Z, Z~
 [ʃ] — ~ S', S', S' ~
 [j] — ~ J, J, J~
 [ç] — ~ J', J' ~
 [çh] — ~ H', H' ~
 [h] — ~ H, H, H~
 [x] — ~ X, X~
 [dʒ] — ~ D'', D'', D'' ~
 [tʃ] — ~ C, C, C~

[k] - ~ K, K, K ~
 [v] - ~ V, V, V ~
 [f] - ~ F, F, F ~
 [z] - ~ Z, Z, Z ~

[dʒ] - ~ D', D', D' ~
 [tʃ] - ~ C', C', C' ~
 [r] - ~ R, R, R ~
 [l] - ~ L, L, L ~

A hosszúság hívőjele mind a magán-, mind pedig a mássalhangzóknál a pontosvessző (;), míg a pont (.) a szólamok és a beszédszakaszok közötti szünetek jelölésére szolgál.

A hívőjel leütésével a számítógép automatikusan választja ki és vonja össze az adott beszédhangot felépítő hangszeteleket a következő módon:

~ A' = ~ A', A'	A' = A'	A' ~ = A', A' ~
~ A = ~ A, A	A = A	A ~ = A, A ~
~ O = ~ O, O	O = O	O ~ = O, O ~
~ U = ~ U, U	U = U	U ~ = U, U ~
~ Y = ~ Y, Y	Y = Y	Y ~ = Y, Y ~
~ I = ~ I, I	I = I	I ~ = I, I ~
~ E' = ~ E', E'	E' = E'	E' ~ = E', E' ~
~ E = ~ E, E	E = E	E ~ = E, E ~
~ Q = ~ Q, Q	Q = Q	Q ~ = Q, Q ~
~ B = ~ B, B, 2B	B = B, 2B	B ~ = B, 2B, B ~
~ P = *P, P, 2P	P = *P, P, 2P	P ~ = *P, P, P ~
~ M = ~ M, M	M = M	M ~ = M, M ~
~ D = ~ D, D, 2D	D = D, 2D	D ~ = D, D ~
~ T = *T, T, 2T	T = *T, T, 2T	T ~ = *T, T, T ~
~ N = ~ N, N	N = 1N, N	N ~ = N, N ~
~ D' = ~ D', D'	D' = 1D', D'	D' ~ = 1D', D', D' ~
~ T' = *T', T', 2T'	T' = *T', T', 2T'	T' ~ = *T', T', T' ~
~ N' = ~ N', N'	N' = N'	N' ~ = N', N' ~
~ G = ~ G, G, 2G	G = G, 2G	G ~ = G, G ~
~ K = *K, K, 2K	K = K, K, 2K	K ~ = *K, K, K ~
~ V = ~ V, V	V = V	V ~ = V, V ~
~ F = ~ F, F, 2F	F = F, 2F	F ~ = F, F ~
~ Z = ~ Z, Z, 2Z	Z = 1Z, Z, 2Z	Z ~ = 1Z, Z, Z ~
~ S = ~ S, S, 2S	S = 1S, S, 2S	S ~ = 1S, S, S ~
~ Z^ = ~ Z^, Z^, 2Z^	Z^ = Z^, 2Z^	Z^ ~ = Z^, Z^ ~
~ S^ = ~ S^, S^, 2S^	S^ = 1S^, S^, 2S^	S^ ~ = 1S^, S^, S^ ~
~ J = ~ J, J	J = J	J ~ = J, J ~
	J' = J'	J' ~ = J', J' ~
	H' = H', 2H'	H' ~ = H', H' ~
~ H = ~ H, H, 2H	H = H, 2H	H ~ = H, H ~
	X = X, 2X	X ~ = X, X ~
~ D'' = ~ D'', D'', 2D''	D'' = 1D'', D'', 2D''	D'' ~ = 1D'', D'', D'' ~
~ C = *C, C, 2C	C = *C, C, 2C	C ~ = *C, C, C ~
~ D^ = ~ D^, D^, 2D^	D^ = 1D^, D^, 2D^	D^ ~ = 1D^, D^, D^ ~

$\sim C^{\wedge} = *C^{\wedge}, C^{\wedge}, 2C^{\wedge}$
 $\sim R = \sim R, R, 2R$
 $\sim L = \sim L, L$

$C^{\wedge} = *C^{\wedge}, C^{\wedge}, 2C^{\wedge}$
 $R = 1R, R, 2R$
 $L = L$

$C^{\wedge}\sim = *C^{\wedge}, C^{\wedge}, C^{\wedge}\sim$
 $R\sim = 1R, R, R\sim$
 $L\sim = L, L\sim$

A hosszúság jelének a leütésére a hangot alkotó szeletek közül egyet, a hosszúság fonetikai kifejezésében a leglényegesebbet, a gép megduplázva hív be. Ez a magánhangzóknál a tiszta fázis szelete, a zöngétlen zárhangoknál a néma fázis szelete, míg a többi mássalhangzónál az akusztikai szerkezet függvényében más és más szeletet kell kétszer szerepeltetni a hangban.

A magánhangzó és a mássalhangzó (VC), továbbá a mássalhangzó és a követő magánhangzó (CV) közötti átmeneti szakasz szeletazonosítóját (ezt az adott magánhangzó és a mássalhangzó szimbólumának összevonásával jelöljük; pl.: NA, AN, TE', E'T stb.) nem kell külön beütni, mivel ezek behívása egy szerkesztő program segítségével automatikusan történik.

A „Voxon beszél” hangsor előállítására például leütjük a következő fonetikai jeleket: <~A, ~V, O, K, S, O, N~, ~B, E, S, E'; L~>, mire a gép az alábbi azonosítók adatait hívja fel az adattárból: <<~V, V, VO, O, OK, *K, K, 2K, 1S, S, 2S, SO, O, ON, N, N~, ~B, B, 2B, BE, E, ES, 1S, S, 2S, SE', E', E', E'L, L, L~>>. A fonetikus jelekkel előhívott és megszerkesztett hangsor egyenletes hangmagasságon ($F_0 = 109$ Hz) szóval meg. A szintetizált hangsor intonációs megszerkesztése és variálása külön lépésben történik. Az intonáció ráépítése után mód van arra, hogy a hangsor szupraszegmentális szerkezetét külön vagy a szegmentális szerkezettel együtt szólaltassuk meg. Ennek az intonáció kutatásában vesszük nagy hasznát. A rendszer lehetővé teszi még a beszédsebesség (tempó) és a ritmus, valamint a hangfekvés változtatását is.

A műveleti programok biztosítják a beszédszintetizáló rendszer sokoldalú fonetikai felhasználását.

- BEA és IDA utasítással a már megszerkesztett s az operatív memóriában tárolt hangsor bármely hangszelvényében tudunk összetevőket (időtartam-, frekvencia-, intenzitás- stb. adatokat) változtatni, ami megkönnyíti az akusztikai struktúra fonetikai relevanciájának a vizsgálatát.
- IMI utasítással a hangzásidőt változtathatjuk, ami a beszédsebesség (tempó) variálásával végezhető fonetikai kísérletek lényeges eleme.
- ORS utasítással a már szintetizált hangsor szupraszegmentális szerkezetében a hanglejtést tudjuk változtatni. Anélkül, hogy hangszelvényként külön-külön kellene az F_0 -értékeket átírni, elég törésponttól töréspontig megadni az adatokat.
- ILI utasítással a lemezről újabb hangszelvények hívhatók fel az operatív memóriába, amelyekkel kiegészíthető és bővíthető a kísérletre előkészített hangsor.
- ETI-vel a már említett hangszelvény-összevonásokat végezzük el. Az összevonással a kísérlet céljának megfelelő hangszerkezetek képezhetők, amelyek különböző sorrendbe állítva még összetettebb szerkezeti formákat adnak. A hosszabb hangsoroknak háromkarakteres azonosítóval való összefogása leegyszerűsít és megkönnyít minden további műveletet (megszólaltatás, eltávolítás, kiírás stb.).
- LEO utasításra megszólaltathatók az operatív memóriában levő hangszelvények

- külön-külön vagy a belőlük összeállított sorok. LEO I utasításra csak a T; F_o - és I_o - adatokból képződő hangszerkezet szólal meg.
- g) ZOL utasítással a szintetizált hangok és hangsorok összes adatát kiírhatjuk táblázatba rendezett formában.
 - h) Az IVO utasítás a szólamoknak és a beszédszakaszoknak az időszerkezetét, a hanglejtését és a dinamikai felépítését ábrázoló diagram kirajzolására szolgál.
 - i) MOR utasítással a szintetizált beszédanyagot floppy diskre tehetjük el, s ezzel nemcsak az elért eredményt őrizzük meg, hanem a további kísérletekben könnyen felhasználható anyaghoz is jutunk.
 - j) VID utasítással az operatív memória, míg a
 - k) TOM utasítással a floppy disk tartalmáról tájékozódhatunk.
 - l) Végül AGI utasítással javíthatók a lemezen tárolt adatok,
 - m) EVA-val pedig a hangszeletek vagy az összevont tömbök azonosítóit változtathatjuk meg.

A fenti ismertetőből kitűnik, hogy a Voxon felépítése és működése a magyar nyelvészeti-fonetikai alap kutatás igényeinek és szükségleteinek a kielégítését szolgálja. Arra törekedtünk, hogy a szintézis szerkezetét a természetes emberi beszéd fonetikai strukturáltsága szerint építsük fel, és minden szintetizált tömbben lehetővé tegyük az összetevők tetszés szerinti változtatását, ami az „analízis szintézissel” módszer alkalmazhatóságát jelenti a felhasználásban.

A Voxonnal a szintetizálási munka a program indításával kezdődik. Rövid, öt szekundumos zenei szignál, amelyet Csapó Gyula zeneszerző komponált, jelzi, hogy a rendszer kész a felhasználásra.

Befejezésül közöljük a Voxon bemutatkozó szövegét, amely az ismertetett folyamatos beszéd szintetizáló rendszerrel előállított első beszédszöveg volt:

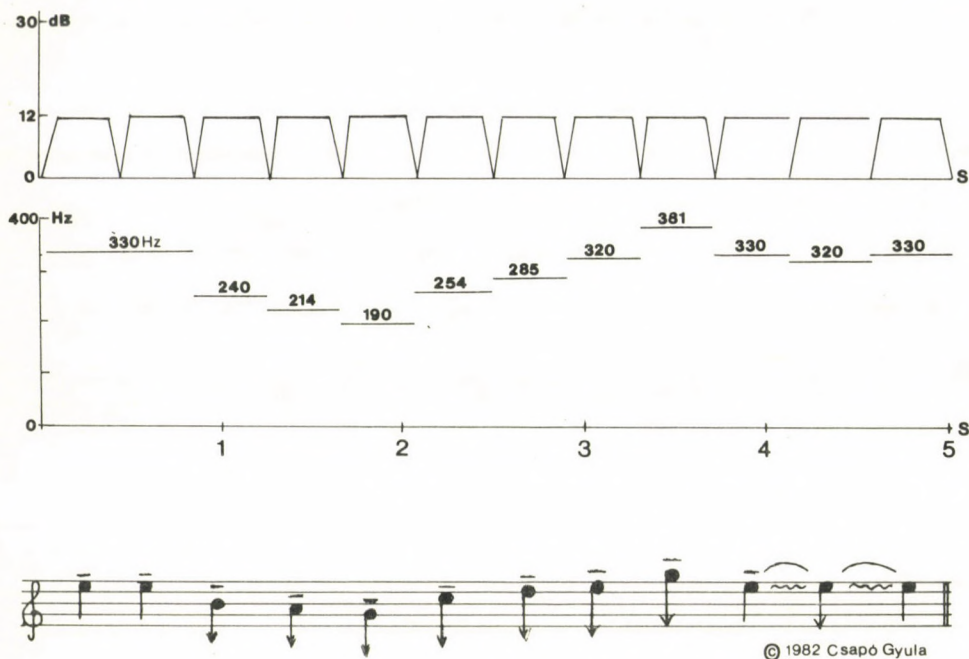
Figyelem! Figyelem! Jó napot kívánok.

A Voxon beszél. Szintetikus szöveget hallanak magyar nyelven.

A folyamatos beszéd szintetizáló rendszer Bolla Kálmán munkája. Programozó: Kiss Gábor. Készült a Magyar Tudományos Akadémia Nyelvtudományi Intézetének fonetikai osztályán 1982 márciusában. Köszönöm figyelmüket. Viszonthallásra.

Voxon

A munka befejezését jelző VEG utasításra ismét megszólal a szintetikus zenei szignál [a:] hanggal „énekelve”.



1. ábra
A Voxon zenei szignálja

Minivoxon

Mivel a rendelkezésre álló technikai rendszer – kis operatív memóriakapacitása miatt – maximálisan csak 5 s-os szekvenciák egyidejű megszólaltatására alkalmas, a Voxon real time üzemmódban nem használható. Kísérleteket végeztünk az adatbázis minimalizálására, aminek eredményeképpen eljutottunk a Minivoxonhoz. Ez 101 hangszereletről áll, 4,5 s memóriakapacitást igényel, használható real time üzemmódban, és megfelelő minőségű beszédszintézist biztosít.

Irodalom

- BOLLA Kálmán: A beszédhang fonetikai definíciója. in: Az amerikai angol beszédhangok atlasza. MFF 9. 1981, 9–11.
- DUDLEY, H.: Fundamentals of speech synthesis. in: Speech Synthesis. Szerk. FLANAGAN, J.L. és RABINER, L.R. Stroudsburg, Pennsylvania 1973, 43–58.
- FANT, G.: Analysis and synthesis of speech processes. in: Manual of Phonetics. Szerk. MALMBERG, B. Amsterdam 1968, 173–277.
- FANT, G.: The acoustics of speech. in: Speech Synthesis. Szerk. FLANAGAN, J.L. és RABINER, L.R. Stroudsburg, Pennsylvania 1973, 77–91.
- KASSAI Ilona: A magyar beszéd hangsorépítési szabályszerűségei. MFF 8. 1981, 63–86.
- KISS Gábor–OLASZY Gábor: Interaktív beszédszintetizáló rendszer számítógéppel és OVE III szintetizátorral. MFF 10. 1982, 21–46.

VOXON: A SYSTEM GENERATING IMPERSONAL HUNGARIAN SPEECH BY RULE

Kálmán Bolla

1. In the present study the author gives an account of a new achievement in the studies dealing with the analysis and synthesis of the acoustic structure of the Hungarian speech. This new result is important in the development of phonetics in Hungary. A speech synthesizing system generating impersonal speech by rule, called Voxon was completed in March 1982. After the announcement made to the board of directors of the Institute of Linguistics of the Hungarian Academy of Sciences, Voxon was presented to the colleagues on the 2nd April 1982.

2. Voxon consists of three parts: a) a collection of the structural elements of the sounds (the sound sections), i.e. the data basis; b) a phonetic code which combines the elements into sounds and sound sequences; c) the operational programs. The operational programs were made by Gábor Kiss programmer.

3. The Voxon system was created according to the needs and requirements of the basic phonetic-linguistic research in Hungary. Thus, our aim was to build up the structure of the synthesis which corresponds to the phonetic structure of natural human speech. All the common standard Hungarian sounds and sound combinations can be generated from the primary constituents of the segmental sound structure, the sound sections. Of the total 550 elements 27 are used to synthesize the vowel types and 89 are used to synthesize the consonant types. The relevant sound section is doubled in order to express temporal variation or quantity.

In Hungarian the phonetic value of a speech sound varies according to its environment (the neighbouring sounds) and its value is different when it is in word-initial, word-medial or word-final position. The influence the phonetic environment has on the acoustic structure of the sound in question manifests in the sound transitions. As the so called transitional phases of the sound sequence are essential in the build-up of an acoustic structure, in speech reception and perception, we decided to use separate elements corresponding to the possible transitions in order to be able to analyze the problem more closely. The number of these elements is 414. Finally, the data basis contains a more open and a more closed variant of each vowel.

Variation of sound quality and quantity depending on the position of the sound (word-initial, word-medial and word-final position) is realized in the system by changing the beginning and the end sections of the sound.

4. We choose the phonetic code, i.e. the identifiers of the sound sections (consisting of 1-3 characters), in a way which makes it easy to determine which sound the code belongs to, which phase it realizes and in which position.

The „call signals” of the sounds are the letters of the phonetic alphabet as they can be printed by an LA 36 console typewriter.

5. The sound sequence called from the mass storage device to the memory and structured by the phonetic signs is realized on an even tone. ($F_0 = 109$ Hz.) The structuring and the variation of the intonation pattern of the synthesized sound sequence is carried out as a separate step. After the intonation pattern is built up, it is possible to listen to the suprasegmental structure separately or together with the segmental structure. This is very useful in the study of intonation. The system allows us to vary the tempo, the rhythm and the register.

It also increases the importance and the usefulness of Voxon in phonetic research that any component of the synthesized sound or sound sequence can be changed according to the goals of the analysis, which extends the scope of the „analysis by synthesis” method.

6. As the capacity of the operative memory of the system available to us is small and thus the maximum duration of the sequences that can be sounded simultaneously is 5 s, Voxon cannot be used in real time mode. We have tried to minimize the data basis and have developed Mini-voxon. It consists of 101 sound sections, it requires a memory capacity of 4.5 s, it can be used in real time mode and the quality of speech synthesized with it is satisfactory.

A LENGYEL BESZÉDINTONÁCIÓ ELEMZÉSE SZINTÉZISSSEL

Földi Éva

Az intonáció mint a gondolatok és az érzelmek tükrözésének kifejezőeszköze fontos szerepet játszik a beszédben kifejezett közléstartalom megértésében. Szükség van tehát alkotóelemeinek feltárására, rendszerbe foglalására s a kutatási eredményeknek a gyakorlatban való alkalmazására. Itt elsősorban a nyelvoktatásra/-tanulásra kell gondolni.

A kísérleti fonetika rohamos fejlődése lehetővé tette a beszéd folyamat akusztikus jegyek alapján történő leírását s az intonációkutatás fellendülését. Míg korábban csak szubjektív, hallás utáni elemzést végeztek, most a korszerű elektrotechnikai eszközök alkalmazásával mód nyílt a beszédintonáció objektív, akusztikai-fonetikai vizsgálatára.

Az intonációval foglalkozó hazai és külföldi munkákból kitűnik, hogy az intonációról alkotott felfogások, nézetek igen eltérőek, s magát a terminust illetően sem egységes a kutatók álláspontja.

A sokféle intonációértelmezés két nagyobb csoportba sorolható: az egyik az intonáció ún. „szűkebb”, a másik az ún. „tágabb” értelmezését képviseli. Az intonáció szűkebb értelemben az alaphang frekvenciaváltozásaival előállított dallamot jelenti. E nézet szerint tehát a beszédintonációt egyetlen komponens hozza létre, s az alaphangfrekvencia változásait kísérő, pontosabban az azzal együtt ható, s egymással kölcsönhatásban levő akusztikai paramétereket (az intenzitás- és időtényezőkön alapuló elemeket) egymástól elválasztva, fonetikai részkategóriákként vizsgálják, melyeket többnyire mint szupraszegmentális vagy prozódiai tényezőket, esetleg mint hangtulajdonságokat említenek. (Ld. például: Lazicius 1944; Dłuska 1976; Szende 1976.) Ezt a felfogást képviseli a lengyel fonetikusok többsége is. (Pl.: Jassem 1962; Dłuska 1976; Wierzchowska 1980.)

A másik – az utóbbi időben egyre inkább előtérbe kerülő – nézet szerint a beszédintonáció több akusztikai összetevőből szerveződő fonetikai részrendszer, melyet az alaphangfrekvencia, az intenzitás és az időkomponensek (időtartam, tempó, ritmus), valamint a hangszín elemeinek változásai hoznak létre. Meg kell azonban jegyezni, hogy az intonációt alkotó elemek meghatározását illetően nem egységes a kutatók álláspontja, sokszor még egy nyelv intonációs rendszerére vonatkozóan sem (Zinder 1979, 272).

Dolgozatomban az intonáció, beszédintonáció kifejezéseket a Bolla Kálmán által alkalmazott szupraszegmentális hangszerkezet szinonimájaként használom (Bolla MFF 5. 1980, 41–2). Ebben az értelemben az intonáció, a beszédhangokhoz hasonlóan, több összetevőből álló komplex egység. A beszédhangok akusztikailag a hangmagassággal, az intenzitással, az időtartammal és a hangszínképpel jellemezhetők. Ezek a komponensek hozzák létre a hangszerkezetet is: „A beszéd szignálnak az intenzitáson, az alaphang frekvenciáján, az időtartamon és a hangszínkép változásain alapuló folyamatváltozásaiból különböző hangeffektusok alakíthatók ki, melyek mintegy ráépülnek a hangsor rövidebb-hosszabb metszeteire, szakaszaira. A nyelv az ebből adódó hangalaki különbségeket használja fel intellektuális, voluntatív, emotív stb. tartalmak kifejezésére.” (Bolla i.m. 41–2.)

A dolgozat kijelentő, kérdő és felszólító típusú lengyel mondatok szintézissel kiegészített elemzésének eredményeit tartalmazza. Az elemzés célja a beszédintonáció elemei közül négynek (tempó, ritmus, dallam, intenzitás) a vizsgálata volt. A szintézis egyrészt az analízis során mért adatok hitelességének ellenőrzésére szolgált, másrészt pedig a munka későbbi fázisában tervezett percepciós tesztelés anyagát képezi.

A hangfelvételek Poznańban, az LTA Akusztikai-Fonetikai Laboratóriumában (Pracownia Fonetyki Akustycznej IPPT PAN) készültek hat (három nő és három férfi) bemondóval. A vizsgálat alapja 150, a kutatási célnak megfelelően összeállított, egyszerű tö-, hiányos és bővített mondatból álló nyelvi anyag volt. A kísérlethez az akusztikus diagramok egy női adatközlő, Barbara Klusińska felvételéről készültek. A műszeres méréseket az MTA Nyelvtudományi Intézetének fonetikai laboratóriumában végeztük FFM 650 típusú alaphangmagasság-mérővel, IM 360-as intenzitásmérővel, 34 T típusú négycsatornás mingográf-felvétellel, PDP 11/34-es minicomputerrel, valamint OVE III típusú beszéd szintetizátorral.

A diagramokon kimért adatok alapján elkészítettem a mondatok intonációs szerkezetének fonetikus átírását a Bolla Kálmán által javasolt módszer szerint (Bolla MFF 2. 1978, 7–23). Az egyszerűség kedvéért nem a komplex, hanem a részleges lejegyzést választottam, vagyis nem a fonetikus, hanem a helyesírási jeleket használtam. Ezekről az átírt sémákról könnyen leolvashatók az intonációs szerkezet alkotóelemeinek adatai.

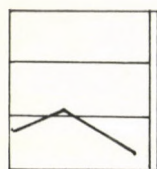
Az elemzést és a szintézist Bolla Kálmán koncepciója alapján végeztem. A szintézishez a programot Kiss Gábor programozó matematikus készítette.

A kijelentő mondatok. — A vizsgált lengyel kijelentő mondatok időtartama 0,36–1,75 s között változott, s átlagos $t_{em p}$ ójukra 9,3 h/s-ot mértem (a beszédsebességet az 1 s-ra jutó hangok számával fejeztem ki). A kérdő és a felszólító típusú mondatokhoz viszonyítva a kijelentő mondatok tempója volt a leglassúbb.

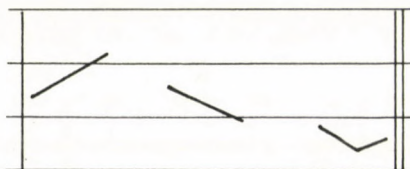
A mondatok $r i t m u s a$ egyenletesnek mondható, nagyobb változásokat a szótagok egymáshoz viszonyított időtartamadatok nem mutattak, s ez is oppozícióba állítja a kijelentést a kérdő és a felszólító mondatokkal. A szavak hangsúlyos szótagja (a lengyelben ez az utolsó előtti szótag) 60–110 ms-mal volt hosszabb a szomszédos szótagoknál.

A kijelentő mondatok $d a l l a m á t$ emelkedő–ereszkedő hangmenet jellemezte, amely háromféle variációban fordult elő:

1. Emelkedő–ereszkedő, s a mondat végén enyhén (50–60 Hz-et) emelkedő hangmenet, amelyet a 2–5 szótagból álló kijelentéseknél tapasztaltam, az egy szótagosoknál a mondatvégi emelkedés nem fordult elő:

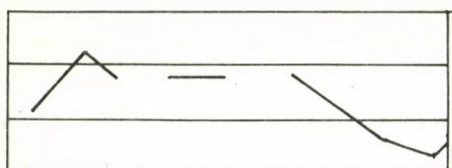


Nie. (Nem.)



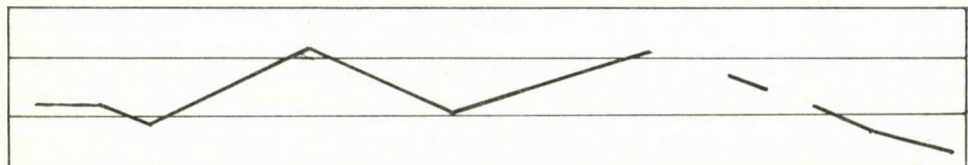
Na pocz- cie. (A postán.)

2. Emelkedő–ereszkedő–lebegő–ereszkedő–enyhén emelkedő dallamformájuk volt a 4–6 szótagos mondatoknak:



On pra- cu- je. (Ő dolgozik.)

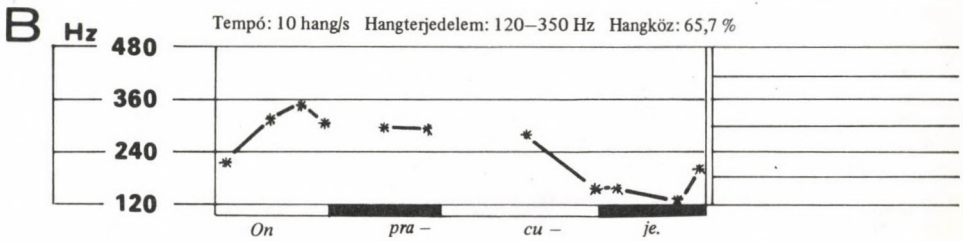
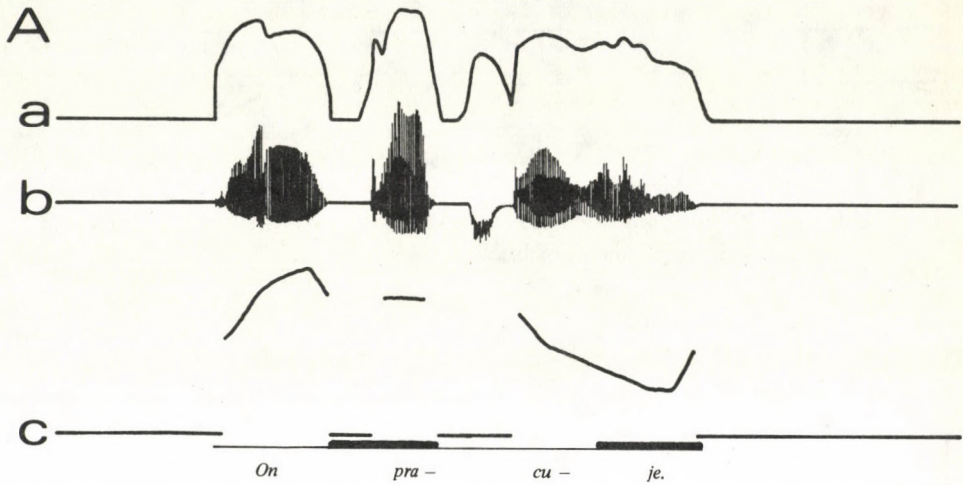
3. Lebegő/ereszkedő–emelkedő–ereszkedő tremolóféle hangmenet volt jellemző általában a 7-nél több szótagból álló mondatokra:



Na do- le bę- dzie- my cze- ka- li.
(Lent fogunk várni.)

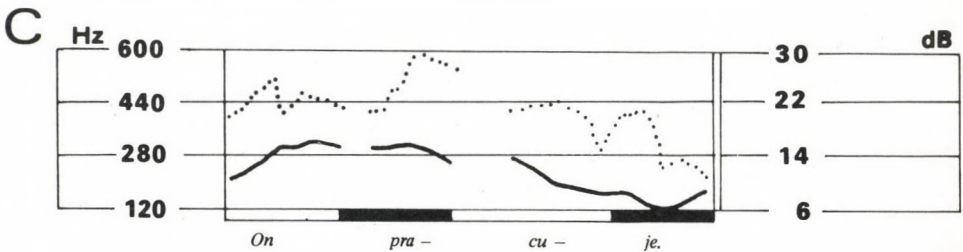
A kijelentő mondatok dallama a beszélő szubjektív hangterjedelmének (jelen esetben ez 120–480 Hz) mély vagy középső zónájából (220–300 Hz-ről) indult, s mindig a mély hangmagassági sávban (120–240 Hz) végződött. A frekvenciacsúcsok 230–375 Hz közötti értékekkel szerepeltek, vagyis a mély zóna felső határa és a felső zóna alsó határa között mozogtak, s rendszerint a mondat első vagy második szótagjára estek. A kérdéshez és a felszólításhoz viszonyítva a csúcsértékek kisebbek, s az alaphang meredek változása nem figyelhető meg. A kérdő és a felszólító mondatokhoz képest a kijelentő mondatok hangterjedelme bizonyult a legszűkebbnek: 120–375 Hz-es frekvenciasávban fordultak elő a kijelentések. Átlagos hangterjedelmük tehát 255 Hz, azaz a beszélő hangterjedelmének 65,5 %-a volt. A mondatok dallamára közepes–mély hangfekvés jellemző, a magas–mély nem mondható tipikusnak, csak szórványosan fordult elő.

A lengyel kijelentő mondatok *i n t e n z i t á* sváltozásának iránya erősödő–csökkenő és erősödő–egyenletes–csökkenő lehet, amelyek közül az előbbi volt a gyakoribb. Az intenzitásértékek 19–40 dB között ingadoztak, s a csúcsértékek (35–40 dB) az első



Hangmagasság:	225/325/350/300	300	280/155 ↓	155/120/200 Hz
Intenzitás:	33/37/33/31	41/39	32/29	31/29/22 dB
Időtartam:	210	200	290	200 ms

Az intonációs szerkezet tartama: 0,9 s; 9 hang



1. ábra

Az *On pracuje.* lengyel kijelentő mondat a) intonogramja, b) fonetikus átírása és c) szupraszegmentális szerkezetének szintézise

vagy a második, ritkábban pedig a harmadik szótagra estek. A minimumértékek (19–22 dB) mindig az utolsó szótagon fordultak elő. A frekvencia- és intenzitáscsúcsok általában az 1–2 szótagos mondatokban szerepeltek párhuzamosan. Az intenzitásváltozás foka szerint a mondatok dinamikája enyhén változó volt.

Az elemzés eredményeit szintézissel ellenőriztem. Az intonációs szerkezetet kialakító alapprofrekvencia-, intenzitás- és időtartamadatok számítógépbe vitele után hangfelvétel készült a szintetizált szerkezetről, s a géppel kirajzoltattam az intonációs sémákat. A szintézis megerősítette az elemzés során tapasztaltakat. (Ld. 1. ábra.)

A szintetizált szerkezetek megfigyelhetőbbé teszik a típusjegyeket, csak a szuprasegmentális emlékképet őrzik, s ezáltal nagyobb lehetőség nyílik az egy nyelven belüli, valamint az interlingvális egybevetésekre is.

A magyar és a lengyel kijelentő mondatok között eltérés csupán a hangterjedelemben van, a magyar anyanyelvűek hangterjedelme ugyanis kisebb.

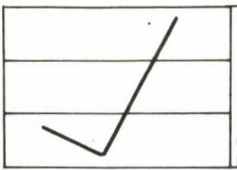
A kérdő mondatok. – A kérdő mondatok két típusát, az eldöntendő és a kiegészítendő kérdést vizsgáltam.

A) A lengyel eldöntendő kérdő mondatok időtartama 0,5 s (egy szótagos mondat) és 1,96 s (tizenegy szótagból álló kérdés) között váltakozott. A vizsgált mondatok átlagos t_{em} ójára 9,94 h/s-ot mértem, ami azt mutatja, hogy az eldöntendő kérdés beszédsebessége a kijelentéshez és a felszólításhoz viszonyítva gyorsabb, a kiegészítendő kérdéshez mérten pedig lassúbb.

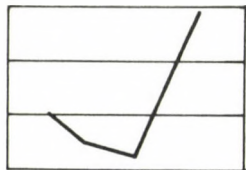
Az eldöntendő kérdő mondatok *r i t m u s a* változó, rendszerint gyorsuló–lassulónak mondható. A mondatvégi utolsó előtti vagy utolsó szótag (a dallam szökő szakasza) időtartama volt a leghosszabb, a szomszédos szótagok idejéhez képest 50–220 ms-mal nyúlt meg.

Az eldöntendő kérdő mondatokra az emelkedő/ereszkedő–szökő *d a l l a m*-forma a jellemző, amely az alábbi változatokban fordult elő:

1. Ereszkedő–szökő hangmenetük volt az egy szótagos kérdéseknek:

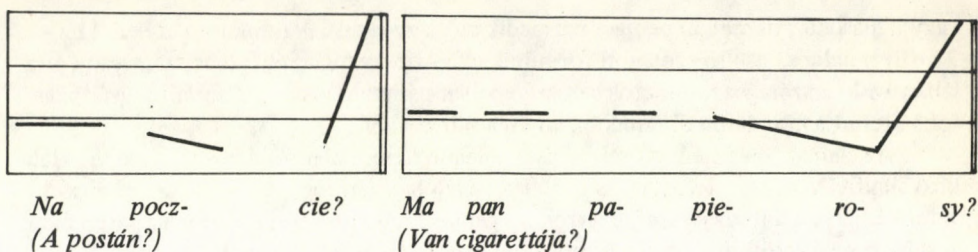


Tak? (Igen?)

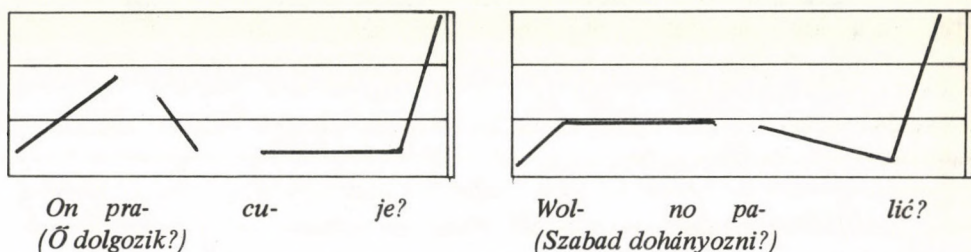


Piotr? (Péter?)

2. Lebegő–ereszkedő–szökő dallamforma jellemezte általában a 2–6 szótagból álló kérdéseket:



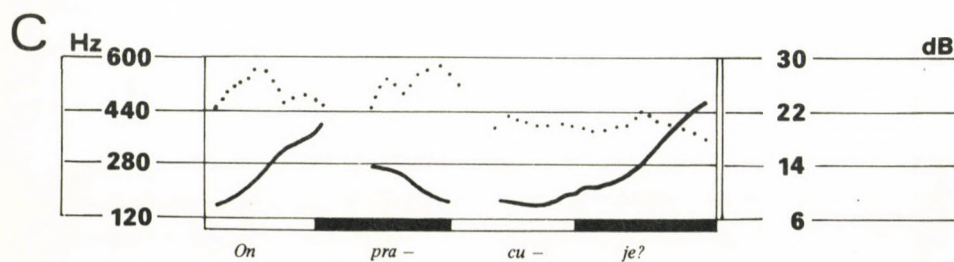
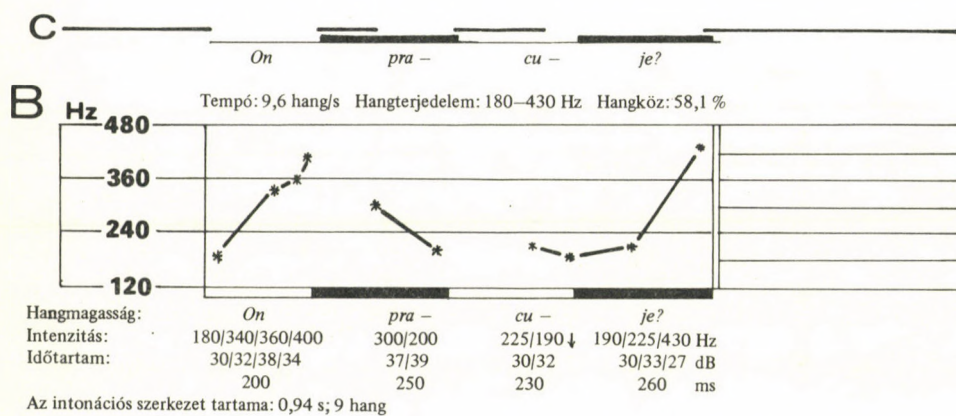
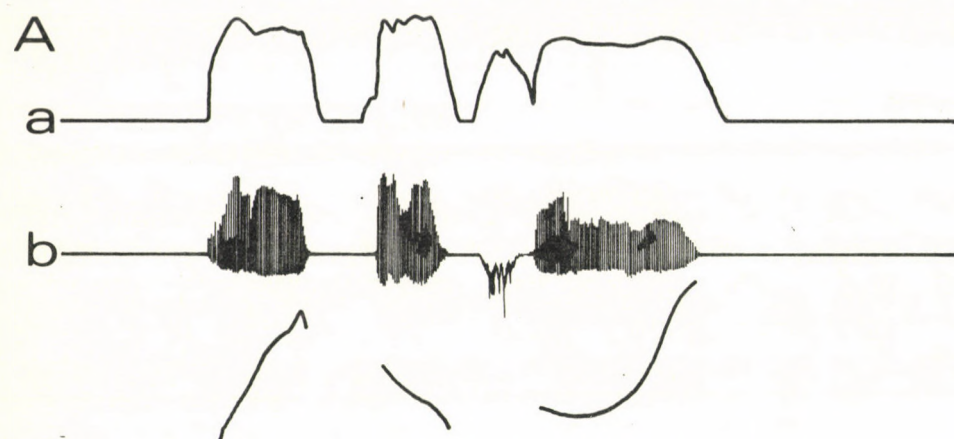
3. Emelkedő—ereszkedő/lebegő—szökő hangmenet figyelhető meg a négy vagy ennél több szótagból álló kérdések nagy részénél:



A lengyel eldöntendő kérdő mondatok hangmenete a dallam szökő szakasza előtt (a mondatvégi utolsó kettő vagy csak az utolsó szótag) igen változatos, sokszor mindhárom forma – ereszkedő, emelkedő, lebegő – váltakozásából alakul ki. Az eldöntendő kérdések dallama a mély vagy a középső hangmagassági sávból (200–320 Hz-ről) indult. Az alaphangcsúcsértéke mindig a mondat végén, vagyis az utolsó szótagon volt, s a felső zónában 420–460 Hz-es értéket vett fel. Az eldöntendő kérdés dallamának meredek változása tehát itt következik be. Az alaphang legmélyebb pontját, 150–200 Hz-et a mély hangmagassági sávban, a dallam szökő szakasza előtt mértem. Az eldöntendő kérdő mondatok hangterjedelmére a 150–460 Hz-es frekvenciasáv jellemző, átlagos hangközük 310 Hz, azaz 67,5 %-os. Ez a hangterjedelem a kijelentéshez viszonyítva tágabb, a kiegészítendő kérdő és a felszólító típusú mondatokhoz képest szűkebb. A lengyel eldöntendő kérdő mondatokra a mély–magas hangfekvés jellemző.

A mondatok az *i n t e n z i t á* sváltozás iránya szerint erősödő–csökkenőek, egyenletes–csökkenőek és csökkenő–erősödőek lehetnek. A leggyakrabban az erősödő–csökkenő forma fordult elő. A vizsgált kérdő mondatok intenzitásértékei 24–40 dB között mozogtak. A csúcspontok (37–40 dB) rendszerint az első és/vagy a második szótagon voltak, a minimumértékek (24–30 dB) az utolsó szótagra estek. A mondatok utolsó két vagy utolsó szótagja (ez az alaphangcsúcsértékének szakasza) erősebb, 33–38 dB-es pontokat is tartalmazott. Az intenzitásváltozás foka szerint az eldöntendő kérdő mondatok dinamikája enyhén változó.

A szintézis alátámasztotta az elemzés tapasztalatait. (Ld. 2. ábra)



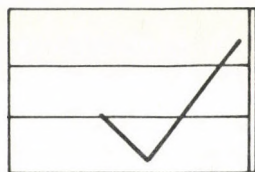
2. ábra

Az *On pracuje?* lengyel kérdő mondat a) intonogramja, b) fonetikus átírása és c) szupraszegmentális szerkezetének szintézise

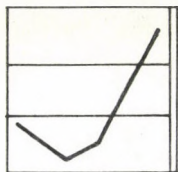
B) A kiegészítendő (vagy kérdőszavas) kérdő mondatok időtartama 0,46–1,75 s között volt, s 1–10 szótagból álltak. A vizsgált mondattípusok közül a kiegészítendő kérdés t e m p ója bizonyult a leggyorsabbnak, az átlagos beszédsebességre 10,8 h/s-ot mértem.

A mondatok r i t m u s a lassuló, ritkábban egyenletes/gyorsuló–lassuló volt. A mondatvégi utolsó szótag időtartama az eldöntendő kérdésekhez hasonlóan itt is nagyobb, de a megnyúlás mértéke kisebb.

A lengyel kiegészítendő kérdés d a l l a mára az emelkedő/szökő–eső/ereszkedő forma a jellemző, amely a kiegészítendő kérdő mondat hangmenetének ellentéte. A két kérdéstípus dallamformája csak az egy szótagos mondatoknál hasonló, amely ereszkedő–emelkedő/szökő volt:



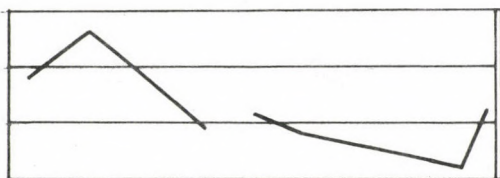
Kto? (Kicsoda?)



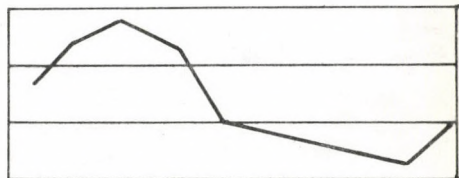
Ja? (Én?)

A kiegészítendő kérdés eső–ereszkedő hangmenete az alábbi változatokban szerepelt:

1. Emelkedő–eső–ereszkedő, s a mondat végén 60–140 Hz-et emelkedő hangmenet volt a leggyakoribb:

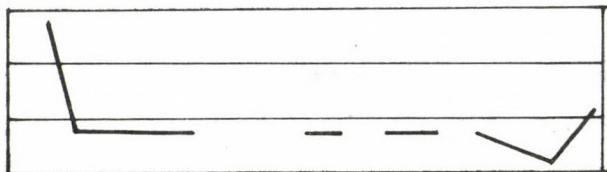


Kie- dy to by- to?
(Mikor volt ez?)

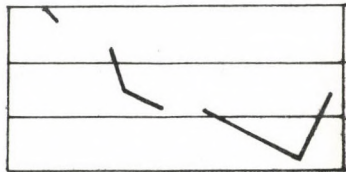


Na ja - k dłu- go?
(Mennyi időre?)

2. Eső–(lebegő)–ereszkedő–enyhén emelkedő dallamforma:

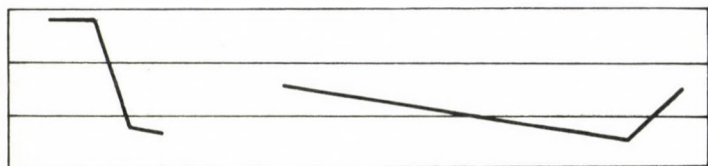


Cze- go jes- teś cie- ka wy?
(Mire vagy kíváncsi?)



Kto się po- my- lił?
(Ki tévedett?)

3. A lebegő–eső–ereszkedő hangmenet ritkán fordult elő:



Kto jes – t te- mu wi- nien?
(*Ki ebben a hibás?*)

A kiegészítendő kérdő mondatok dallama a beszélő szubjektív hangterjedelmének középső vagy felső, néhány esetben pedig a mély hangmagassági zónájából – 250–460 Hz-ről, illetve 225 Hz-ről – indult, s rendszerint a középső sávban fejeződött be. Az alapfrekvencia-csúcsokat, 420–480 Hz-et, mindig a mondatok első és/vagy második szótagján mértem. A minimumértékek – 125–200 Hz – az eldöntendő kérdő mondatokhoz hasonlóan, a kiegészítendő kérdéseknél is vagy a két utolsó szótag határán, vagy az utolsó szótagon fordultak elő. A hangmenet a kijelentésnél erősebben, az eldöntendő kérdésnél viszont gyengébben emelkedett, 260–330 Hz között volt a frekvenciaérték. A kiegészítendő kérdések hangterjedelme 125–480 Hz között mozgott, átlagos hangterjedelmük 355 Hz, hangközük tehát 73,3 %. Ez azt mutatja, hogy a kiegészítendő kérdő mondat hangterjedelme és hangköze nagyobb, mint a kijelentésé és az eldöntendő kérdésé, a felszólító mondatok azonban hasonló hangterjedelem- és hangközértékekkel szerepeltek. A kiegészítendő kérdő mondatok dallamára a mély–magas hangfekvés jellemző.

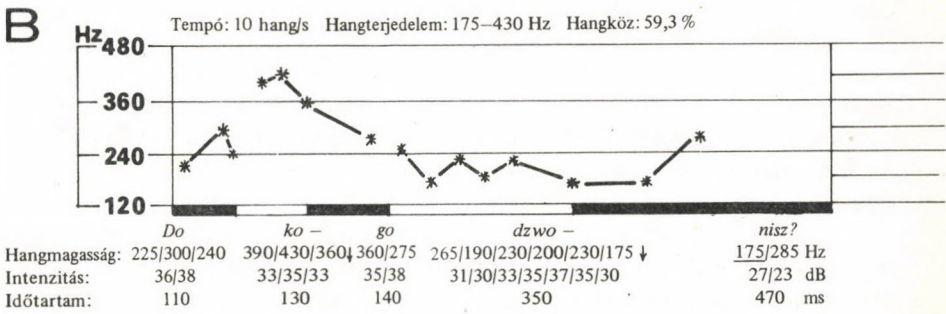
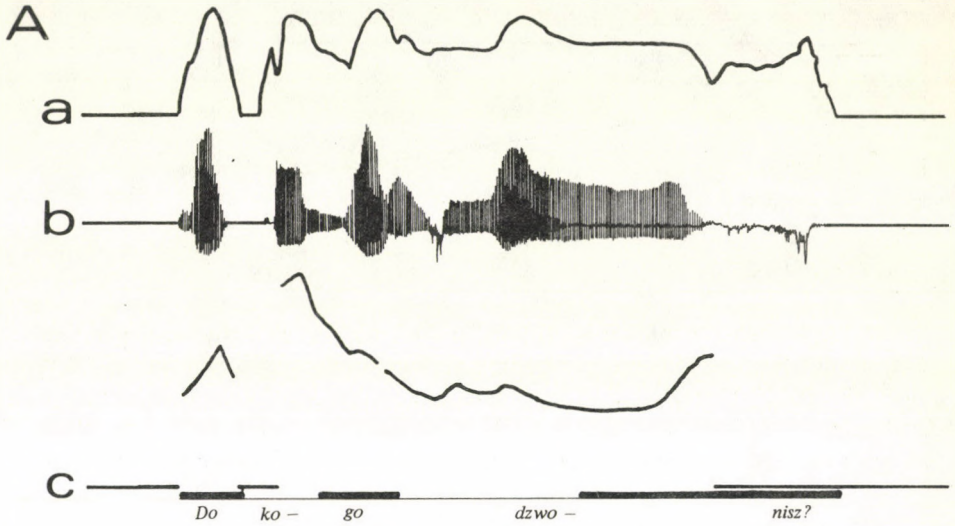
A vizsgált mondatok *i n t e n z i t á* sváltozásának iránya a legtöbb esetben csökkenő volt, de előfordult az erősödő–csökkenő és változó irány is. A mondatok 24–42 dB-es intenzitásértékeket vettek fel, s a csúcsok – 39–42 dB – mindig az első és/vagy a második szótagra estek. Az intenzitás- és az alapfrekvencia-csúcsértékek tehát egybeestek a hangmenet legmeredekebben változó szakaszában. Az intenzitás minimumértékei 20–33 dB között mozogtak, s mindig az utolsó szótagon voltak. A mondatvégi 60–140 Hz-es frekvenciaemelkedéssel párhuzamosan az utolsó szótag a minimumérték előtt erősebb, 30–36 dB-es pontot is tartalmaz.

A szintézis megerősítette az elemzés eredményeit. (Ld. a 3. ábra.)

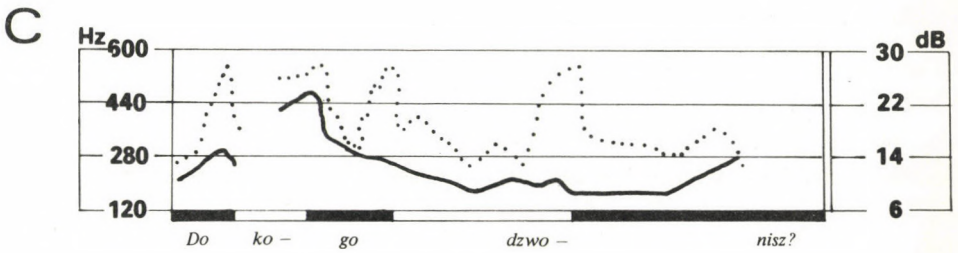
A lengyel és a magyar kérdő mondatok között a legnagyobb eltérés az eldöntendő kérdések hangmenetében figyelhető meg. A lengyelre az emelkedő, a magyarra pedig az ereszkedő dallamforma a jellemző, s csupán az egy, esetenként a két szótagból álló mondatok hangmenete hasonló a két nyelvben, vagyis emelkedő dallamú.

A felszólító mondatok. – A legrövidebb (egy szótagos) lengyel felszólító mondat időtartama 0,6 s, a leghosszabbé 1,56 s volt. Átlagos *t e m p* ójukra 9,5 h/s-ot mértem. Ez a kijelentésnél valamivel nagyobb, a két kérdéstípus beszédsebességénél pedig kisebb.

A felszólítások *r i t m u s a* erősen változó, rendszerint gyorsuló–lassuló volt. A mondat utolsó szótagjának időtartama általában hosszabb, mint a többié.



Az intonációs szerkezet tartama: 1, 2 s; 12 hang

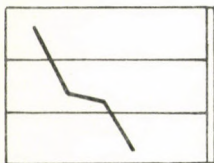


3. ábra

A *Do kogo dzwonisz?* lengyel kérdő mondat a) intonogramja, b) fonetikus átírása és c) szupraszegmentális szerkezetének szintézise

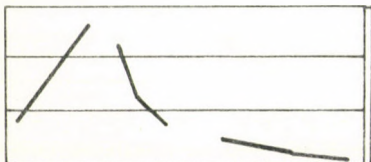
A lengyel felszólító mondatok d a l l a ma rendkívül változatos, s leginkább az emelkedő–ereszkedő hangmenettel jellemezhető, amely – a mondatokban kifejezett közléstartalomtól függően – sokféle variációban fordult elő. A vizsgált mondatok az alábbi formákkal írhatók le:

1. Eső–ereszkedő hangmenetük volt az egy szótagos üdvözléseknek:

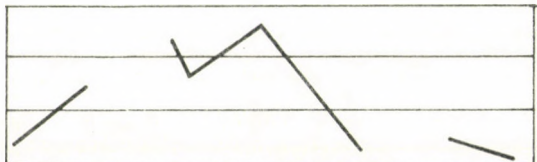


Cześć! (Szervusz!)

2. Emelkedő/szökő–eső–ereszkedő dallamforma jellemezte a 3–5 szótagból álló felszólító mondatok nagy részét:

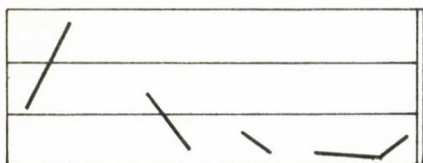


*Wi- tam pa- na!
(Üdvözlöm!)*

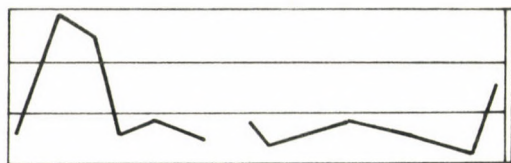


*Mam te- go- do- syć!
(Elegem van ebből!)*

3. Emelkedő/szökő–eső–ereszkedő, s a mondat végén 65–160 Hz-et emelkedő hangmenet figyelhető meg a négy vagy ennél több szótagos mondatoknál:

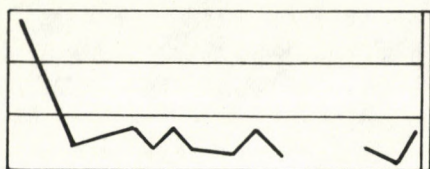


*Jesz- cze cze- go!
(Hát még mit nem!)*

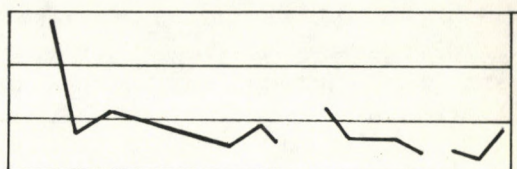


*Mo- wy o tym nie ma!
(Szó sincs róla!)*

4. Eső–emelkedő–ereszkedő–enyhén emelkedő dallam jellemezte a négy vagy az ennél több szótagból álló mondatok nagy részét:

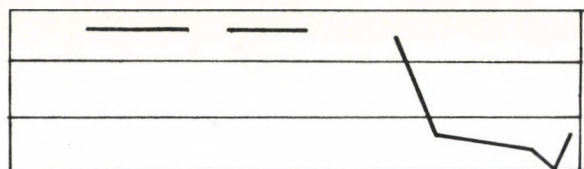


Nie bądź dzie- ckiem!
(Ne légy gyerek!)



Przejdź- my do te- ma- tu!
(Térjünk a tárgyra!)

5. A lebegő–eső–ereszkedő–enyhén emelkedő hangmenet a 4–6 szótagos mondatok egyharmadánál fordult elő:

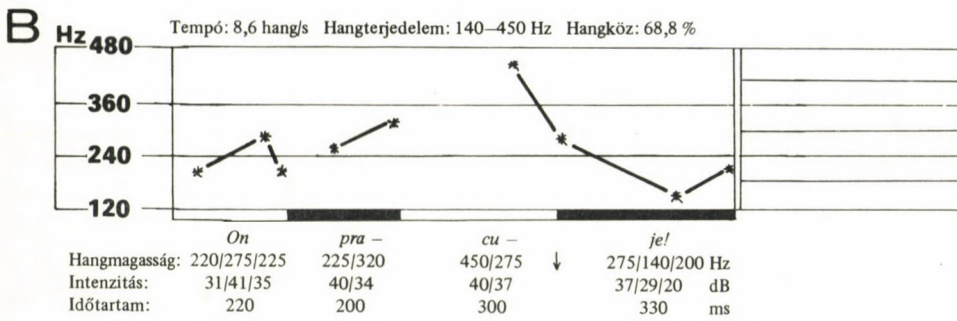
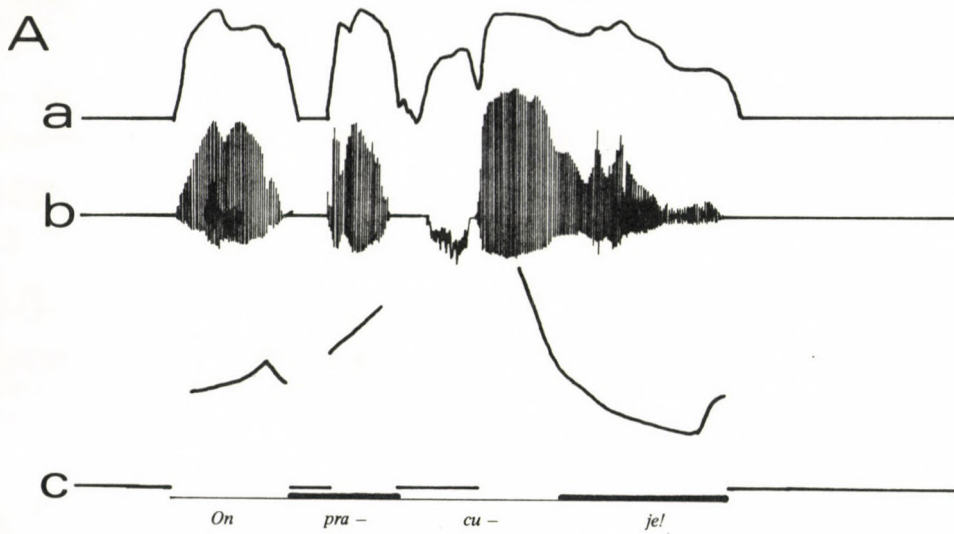


Co mnie to o- bcho- dzi!
(Mi közöm hozzá!)

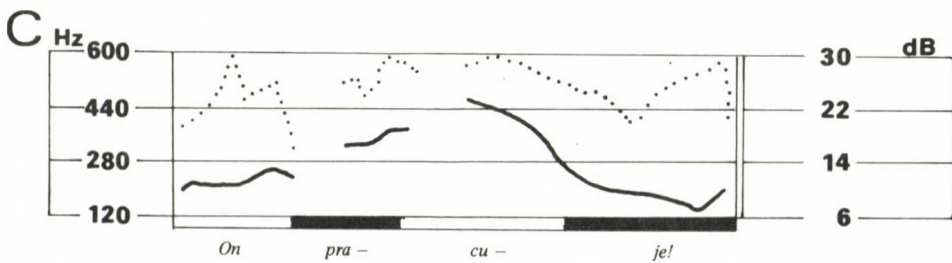
A lengyel felszólító mondatok dallama a szubjektív hangterjedelem mély vagy középső hangfekvési sávjából (200–325 Hz), az eső–ereszkedő mondatok dallama pedig a hangterjedelem felső zónájából (425–480 Hz-ről), a frekvenciacsúsról indult, s a mély vagy a középső zónában (125–325 Hz) ért véget. Az alaphang csúcsértéke 425–480 Hz között váltakozott, s általában az első vagy a második, esetenként pedig más szótagra esett. A minimumérték (125–160 Hz) leggyakrabban a mondat utolsó szótagján, ritkábban a két utolsó szótag határán volt. A dallam a mondat végén rendszerint 65–125 Hz-et emelkedett. A felszólító mondatok hangterjedelme 125–480 Hz, azaz 335 Hz, így a hangköz 73,3 %, ebben tehát megegyezett a kiegészítendő kérdéssel. A kijelentő és az eldöntendő kérdő mondatokhoz viszonyítva a felszólítás hangterjedelme tágabbnak bizonyult.

A felszólítás dallamformában is nagyon hasonlít a kiegészítendő kérdéshez, egyformának mondható hangmenetek is előfordultak. Az alapvető különbség a két mondat-típus között az, hogy míg a kiegészítendő kérdő mondat dallamcsúcsa csakis a mondat elején (első vagy/és második szótag) lehet, a felszólítás esetében az alapfrekvencia-csúcs helye változó, s a mondatvégi hangmenet-emelkedés a felszólításnál gyengébb, mint a kiegészítendő kérdésnél. A felszólító mondatokra a közepes–magas hangfekvés volt jellemző.

A vizsgált lengyel felszólító mondatok az i n t e n z i t á s változásának az iránya szerint csökkenő, erősödő–csökkenő, egyenletes–csökkenő vagy erősödő–egyenletes–csökkenő erősségűek lehetnek. Az intenzitásértékek 20–43 dB között voltak. A csúcsérték (37–43 dB) helye általában változó, gyakran a mondat első vagy/és második szótagjára esett, de sokszor előfordult más szótagon is. A minimumértékek (20–32 dB) mindig a mondat végén, az utolsó szótagon voltak. Az intenzitás-csúcs vagy egybe-



Az intonációs szerkezet tartama: 1,05 s; 9 hang



4. ábra

Az *On pracuje!* lengyel felkiáltó mondat a) intonogramja, b) fonetikus átírása és c) szupraszegmentális szerkezetének szintézise

esett a frekvenciacsúccsal (legtöbbször a három szótagból álló mondatoknál), vagy az alaphang csúcsát megelőző szótagra esett.

A szintézissel ellenőriztem az elemzés következtetéseit. (Ld. 4. ábra.)

A lengyel és a magyar felszólító mondatok intonációs szerkezete között elég nagy a hasonlóság. A leglényegesebbnek mondható különbség az, hogy a magyar felszólítás hangmenetére nem jellemző a mondatvégi emelkedés.

Összefoglalás. — Az intonációs szerkezetet létrehozó alkotóelemek közül négyet, a tempót, a ritmust, a dallamot és az intenzitást vizsgáltam. A szintézissel kibővített elemzésből a következő tapasztalatok szűrhetők le.

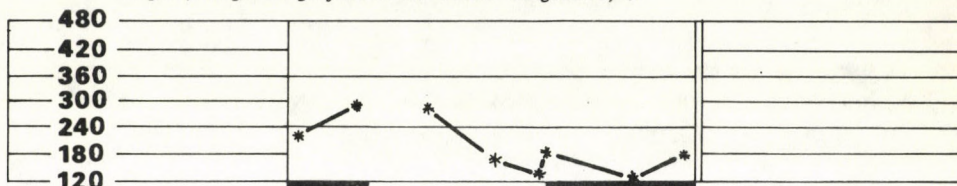
Az egyes mondattípusokra jellemző intonációs szerkezeteket az alkotóelemek értékváltozásainak nagyfokú variálódása alakítja ki. Az akusztikus összetevők változásai tehát különböző szerkezeteket hoznak létre, amelyek a nyelvben jelentés-megkülönböztető funkcióval rendelkeznek.

Az intonációs szerkezet minősége — a vizsgált alkotóelemek alapján — függ az alaphangfrekvencia- és az intenzitáscsúcs értékétől, s a mondatban elfoglalt helyétől. Az előbbi a dallamformát, az utóbbi pedig a dinamikát meghatározó tényező. Ezek szerint tehát viszonylag alacsony alaphangmagasság-csúcsérték jellemzi a lengyel kijelentő mondatokat; magas, a mondat végén levő alaphangfrekvencia-csúcs az eldöntendő kérdésnél figyelhető meg; magas, a mondat elejére eső frekvenciacsúcs tapasztalható a kiegészítendő kérdések esetében; ugyancsak magas, de „változó” helyű csúcs jellemző a felkiáltó mondatokra. Az intenzitásviszonyokat vizsgálva megállapítható, hogy a kijelentések csúcsértékei alacsonyabbak, mint a kérdő vagy a felkiáltó típusú mondatoknál. A legerősebb intenzitásértékek a felkiáltó mondatoknál szerepeltek. A tempó és a ritmus szintén befolyásolja az intonáció minőségét. Megállapítható tehát, hogy a vizsgált négy akusztikai paraméter alapján a mondatok intonációs szerkezete eltérő, s így egymással oppozícióban állnak, következésképpen különböző jelentéstartalmat fejeznek ki. Ez azt jelenti tehát, hogy a hangtest szuprasegmentális szerkezetét kifejező beszédintonáció — alkotóelemeinek természetéből következően — fizikai valóságában megragadható, leírható, nyelvi relevanciával rendelkező önálló fonetikai részrendszer.

Minták a lengyel beszédintonáció elemzésére és fonetikus lejegyzésére

1. *Na prawo. – Jobbra.*
2. *Na dole będziemy czekali. – Lent fogunk várni.*
3. *Jest już pociąg pospieszny. – Itt van már a gyorsvonat.*
4. *Czujesz różnicę? – Érzed a különbséget?*
5. *Mam czytać dalej? – Olvassak tovább?*
6. *Pojedziemy do Warszawy? – Elutazunk Varsóba?*
7. *Twoje te dżinsy? – Tied ez a farmernadrág?*
8. *Ma pan papierosy? – Van cigarettája?*
9. *Mamy jeszcze coś załatwić? – El kell még valamit intéznünk?*
10. *Czego jesteś ciekawy? – Mire vagy kíváncsi?*
11. *Co to ma znaczyć? – Mit jelentsen ez?*
12. *Kto jest temu winien? – Ki ebben a hibás?*
13. *O co chodzi? – Miről van szó?*
14. *Na jak długo? – Mennyi időre?*
15. *Czyja jest ta książka? – Kié ez a könyv?*
16. *Proszę mówić! – Tessék beszélni!*
17. *Mowy o tym nie ma! – Szó sincs róla!*
18. *W tym nic nie ma! – Ebben semmi sincs!*
19. *Co za szczęście! – Micsoda szerencse!*
20. *Co mnie to obchodzi! – Mi közöm hozzá!*
21. *Wszystkiego najlepszego! – Minden jót!*

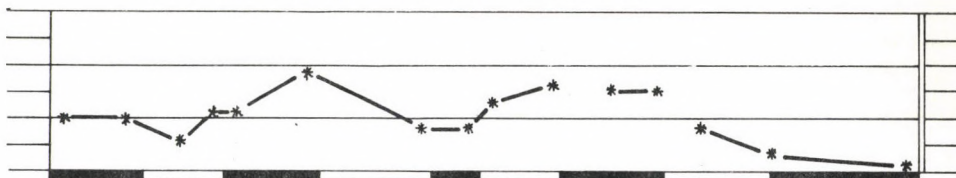
Tempó: 9,2 hang/s Hangterjedelem: 125–280 Hz Hangköz: 55,4 %



1.	<i>Na</i>	<i>pra -</i>	<i>wo.</i>
Hangmagasság:	235/280	275/220/130 ↓	175/125/180 Hz
Intenzitás:	30/37	35/40/35	31/23 dB
Időtartam:	140	340	280 ms

Az intonációs szerkezet tartama: 0,76 s; 7 hang

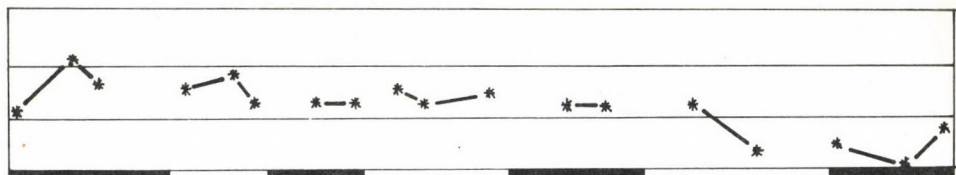
Tempó: 11,3 hang/s Hangterjedelem: 125–350 Hz Hangköz: 64,3 %



2.	<i>Na</i>	<i>do -</i>	<i>le</i>	<i>bę -</i>	<i>dzie - my</i>	<i>cze -</i>	<i>ka -</i>	<i>li.</i>
240 ↓	240/200/250	250/350 ↓	350/230 ↓	230 ↓	275/325	320	230/150 ↓	150/125 Hz
33/40	30/39/36	36	32/25/34	37	28/32	30/35	36/34	32/25/23 dB
160	160	170	220	80	150	180	220	260 ms

Az intonációs szerkezet tartama: 1,6 s; 18 hang

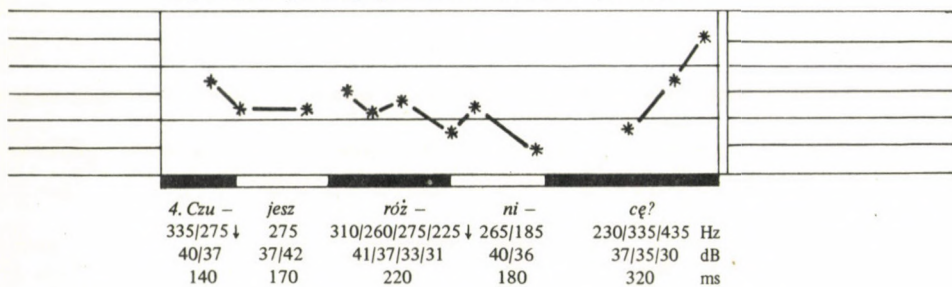
Tempó: 11,4 hang/s Hangterjedelem: 125–375 Hz Hangköz: 66,6 %



3.	<i>Jest</i>	<i>już</i>	<i>po -</i>	<i>ciąg</i>	<i>pos -</i>	<i>piesz -</i>	<i>ny.</i>
250/375/330	310/350/280	280	310/280/300	275	270/155	180/125/210 Hz	
25/37/27	34/35/30	40	26/35/29	35/26	37/30	28/23/20 dB	
290	190	170	270	250	350	230 ms	

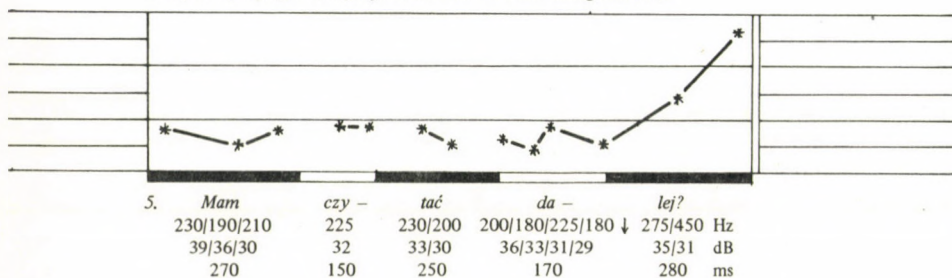
Az intonációs szerkezet tartama: 1,75 s; 20 hang

Tempó: 11,7 hang/s Hangterjedelem: 230–435 Hz Hangköz: 47,1 %



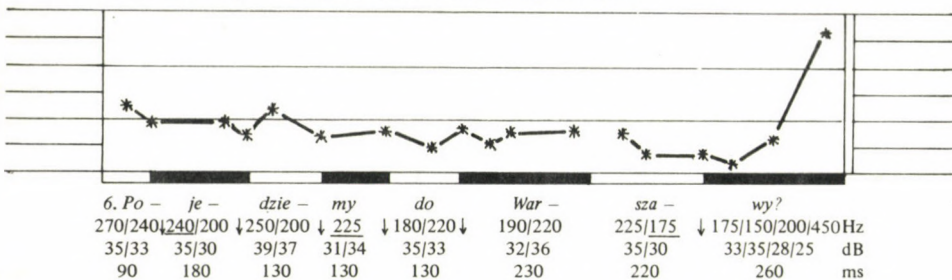
Az intonációs szerkezet tartama: 1,03 s; 12 hang

Tempó: 11,6 hang/s Hangterjedelem: 180–450 Hz Hangköz: 60 %



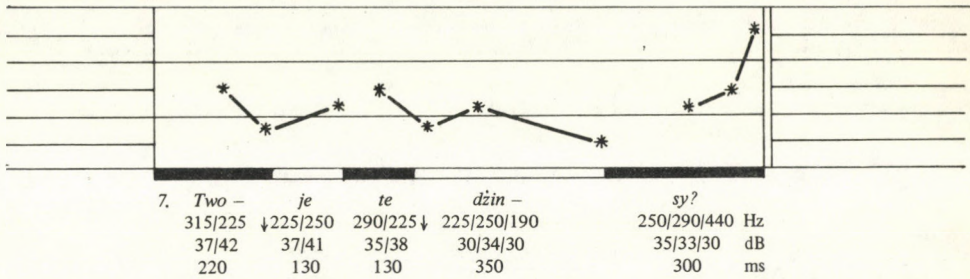
Az intonációs szerkezet tartama: 11,2 s; 13 hang

Tempó: 12,4 hang/s Hangerjedelem: 150–450 Hz Hangköz: 66,6 %



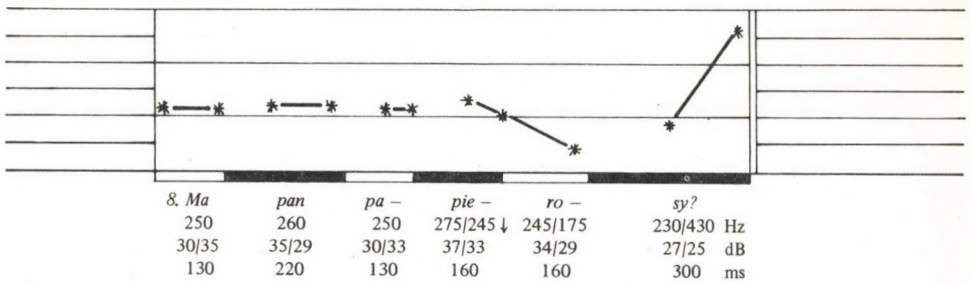
Az intonációs szerkezet tartama: 1,37 s; 17 hang

Tempó: 10,6 hang/s Hangterjedelem: 190–440 Hz Hangköz: 56,8 %



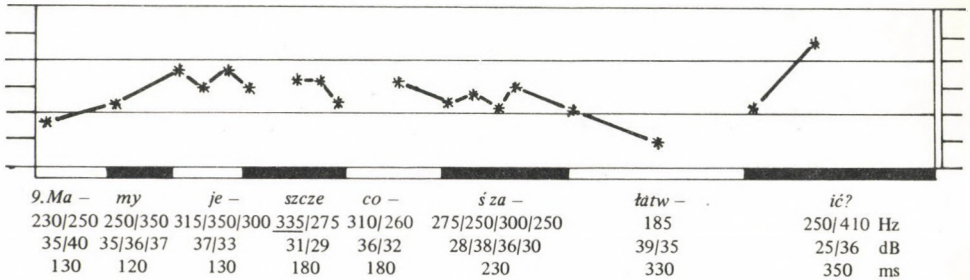
Az intonációs szerkezet tartama: 1,13 s; 12 hang

Tempó: 11,8 hang/s Hangterjedelem: 175–430 Hz Hangköz: 59,3 %



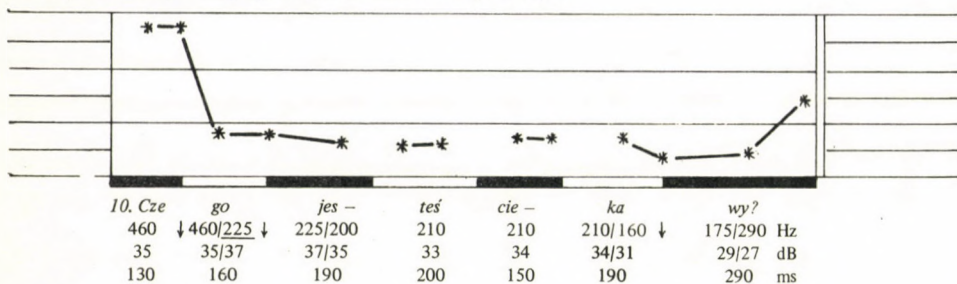
Az intonációs szerkezet tartama: 1,1 s; 13 hang

Tempó: 12,12 hang/s Hangterjedelem: 185–410 Hz Hangköz: 55 %



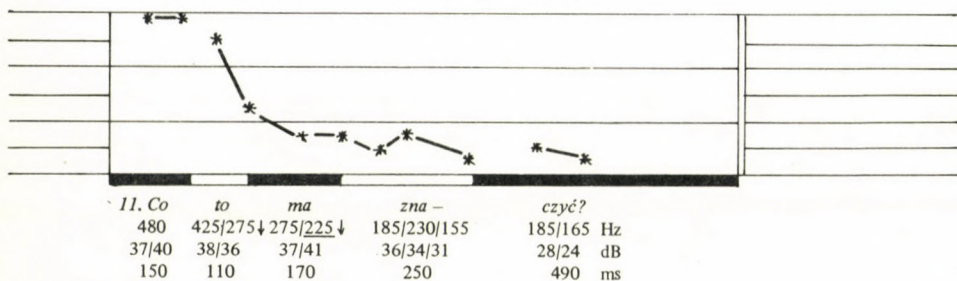
Az intonációs szerkezet tartama: 1,65 s; 20 hang

Tempó: 12,2 hang/s Hangterjedelem: 160–460 Hz Hangköz: 65,2 %



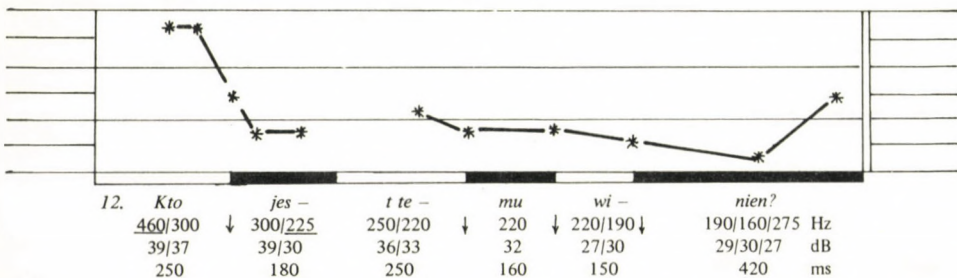
Az intonációs szerkezet tartama: 1,31 s; 16 hang

Tempó: 10,27 hang/s Hangterjedelem: 165–480 Hz Hangköz: 65,6 %



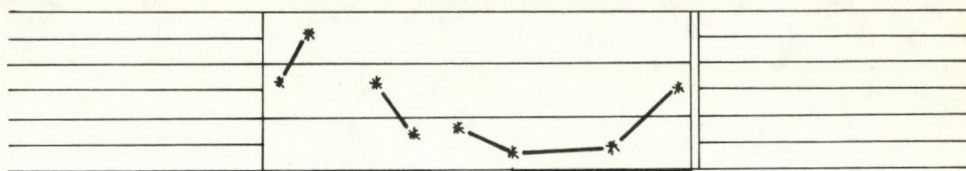
Az intonációs szerkezet tartama: 1,17 s; 12 hang

Tempó: 11,3 hang/s Hangterjedelem: 160–460 Hz Hangköz: 65,2 %



Az intonációs szerkezet tartama: 1,41 s; 16 hang

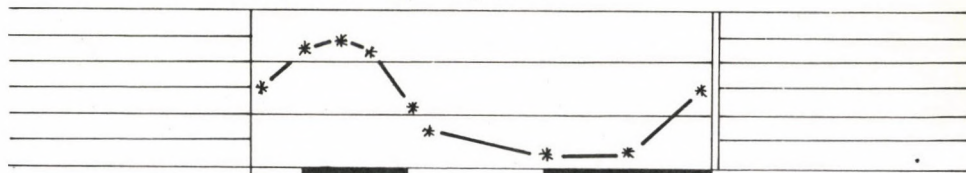
Tempó: 8,97 hang/s Hangterjedelem: 160–425 Hz Hangköz: 62,4 %



13. O	co	cho –	dzi?
325/425	325/210	225/160 ↓	175/300 Hz
27/35	33/37	33/36	29/25 dB
90	200	160	330 ms

Az intonációs szerkezet tartama: 0,78 s; 7 hang

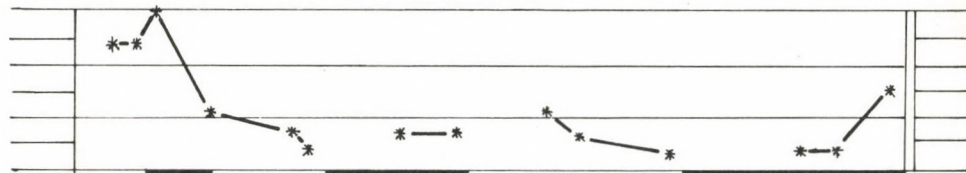
Tempó: 11,9 hang/s Hangerjedelem: 160–410 Hz Hangköz: 61 %



14. Na	ja –	k dlu –	go?
310/375 ↓	410/375/250	200/160	↓ 160/300 Hz
27/37	38/31/29	25/32	36/30 dB
90	190	260	300 ms

Az intonációs szerkezet tartama: 0,84 s; 10 hang

Tempó: 10,5 hang/s Hangterjedelem: 160–480 Hz Hangköz: 66,6 %



15. Czy – ja	jes –	t ta	ksiq –	žka?
410/480 ↓	480/250 ↓	250/225/175	225	250/200/175
38/36	40/37	35/30/35	30/37	36/33/27
140	120	210	260	400
				160/300 Hz
				33/27 dB
				400 ms

Az intonációs szerkezet tartama: 1,53 s; 16 hang

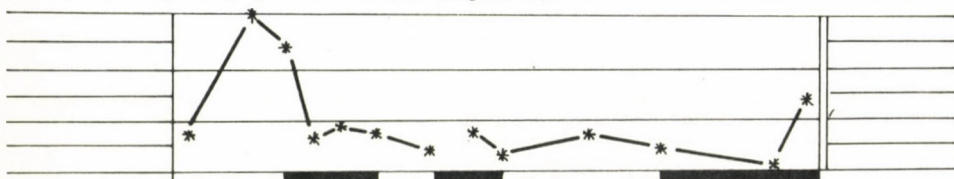
Tempó: 9,9 hang/s Hangterjedelem: 140–460 Hz Hangköz: 70 %



16. <i>Pro</i> –	<i>szę</i>	<i>mów</i> –	<i>ić!</i>
225	300/380 ↓	380/460/150/170 ↓	170/140 Hz
37/35	37/35	35/37/34/32	25/23 dB
170	200	280	360 ms

Az intonációs szerkezet tartama: 1,01 s; 10 hang

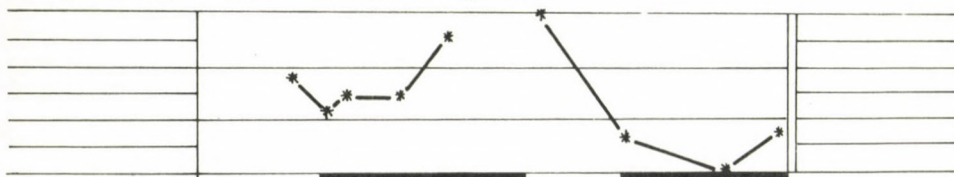
Tempó: 9,16 hang/s Hangterjedelem: 130–480 Hz Hangköz: 73 %



17. <i>Mo</i> –	<i>wy</i>	<i>o</i>	<i>tym</i>	<i>nie</i>	<i>ma!</i>
250/480/400 ↓	400/220/235/215 ↓	215/175	230/175 ↓	175/210/180 ↓	180/130/290 Hz
32/38/42	37/35/41/39	37/40	38/36	33/35/33	30/28/24 dB
200	170	110	120	290	310 ms

Az intonációs szerkezet tartama: 1,2 s; 11 hang

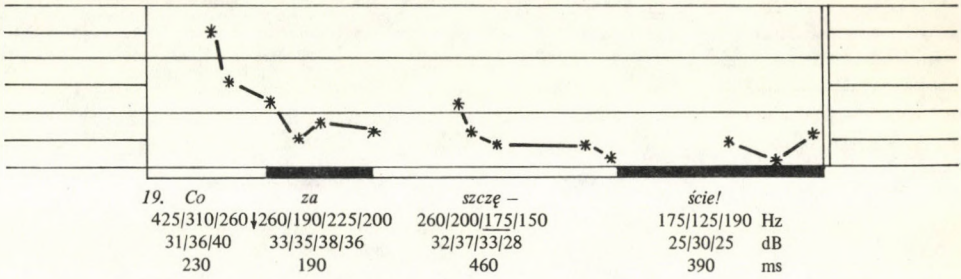
Tempó: 10,1 hang/s Hangterjedelem: 130–480 Hz Hangköz: 73 %



18. <i>W tym</i> –	<i>nie</i>	<i>nie</i>	<i>ma!</i>
340/250 ↓	250/310/435	480/200 ↓	200/130/210 Hz
38/36	36/38/30	38/36	35/34/27 dB
230	380	170	310 ms

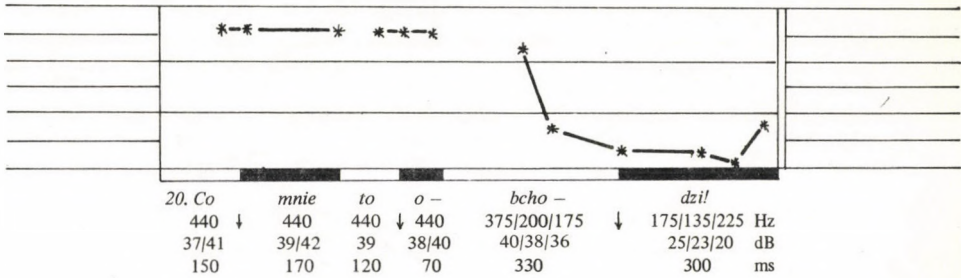
Az intonációs szerkezet tartama: 1,09 s; 11 hang

Tempó: 7,87 hang/s Hangterjedelem: 125–425 Hz Hangköz: 70,5 %



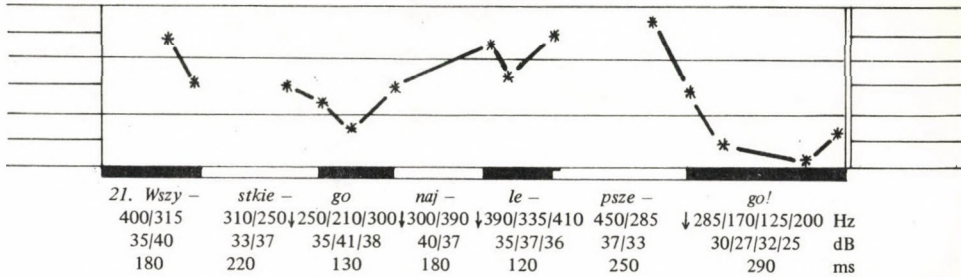
Az intonációs szerkezet tartama: 1,27 s; 10 hang

Tempó: 11,4 hang/s Hangterjedelem: 135–440 Hz Hangköz: 69,3 %



Az intonációs szerkezet tartama: 1,14 s; 13 hang

Tempó: 13,4 hang/s Hangterjedelem: 125–450 Hz Hangköz: 72,2 %



Az intonációs szerkezet tartama: 1,36 s; 19 hang

Irodalom

- BOLLA Kálmán: A fonetikus írás. MFF 2. 1978, 7–24.
- BOLLA Kálmán: A beszéd-folyamat intonációs elemzése és az intonáció fonetikus lejegyzése. MFF 3. 1979, 19–31.
- BOLLA Kálmán: A fonetikai szerkezetek interlingvális egybevetéséről (Problémavázlat). MFF 5. 1980, 40–69.
- BOLLA Kálmán: A fonetikus írás problémái. in: Fejezetek a magyar leíró hangtanból. Szerk. BOLLA Kálmán. Budapest 1982, 25–52.
- DĘUSKA, M.: Prozodia języka polskiego. Warszawa 1976.²
- DUKIEWICZ, L.: Intonacja wypowiedzi polskich. Wrocław 1978.
- FÖLDI Éva: A kérdés kifejezésének intonációs eszközei a lengyelben és a magyarban. MFF 5. 1980, 109–16.
- JASSEM, W.: Akcent języka polskiego. Wrocław 1962.
- LAZICZIUS Gyula: Fonétika. Budapest 1944.
- METTAS, O.: Les techniques de la phonétique instrumentale et l'intonation. Bruxelles 1971.
- SZENDE Tamás: A beszéd-folyamat alaptényezői. Budapest 1976.
- SZOBER, S.: Gramatyka języka polskiego. Warszawa 1953.
- WIERZCHOWSKA, B.: Fonetika i fonologia języka polskiego. Warszawa 1980.
- ZINDER, L.R.: Obščaja fonetika. Moskva 1979.²

THE ANALYSIS OF POLISH INTONATION BY SYNTHESIS

Éva Földi

This paper presents the results of the analysis and the synthesis of Polish assertive, interrogative and negative sentences. The goal of the analysis was to study four elements of the intonation of speech (tempo, rhythm, tune, intensity). The synthesis was partly the verification of the data measured during the analysis, and partly it provides material for a further stage of the study, the perceptual tests.

The recordings were made in the Acoustic Phonetics Laboratory of the Polish Academy of Sciences in Poznań. The corpus consisted of 150 sentences – simple, complex and incomplete – which were compiled as to fit the goals of the analysis. The recordings made with Barbara Klusińska were used in the experiment. The measurements were carried out with an FFM 650 Fundamental Frequency Meter, an IM 360 Intensity Meter, a 34 T four-channel mingograph, a PDP 11/34 computer and an OVE III speech synthesizer.

The analysis of the intonation pattern of the sentences yielded the following results:

1. I have found that the Polish assertive sentences have a characteristic relatively steady rhythm, a relatively narrow register, a steady slightly falling and a slightly rising–falling tone, i.e. there is no abrupt change in the fundamental frequency.
2. The question word questions and the yes-no questions examined are similar in the sense that compared to the assertive sentences their register is wider and their rhythm is usually accelerating–decelerating. The intonation pattern of the two types of interrogative sentences is different: the question word type is characterized by a rapidly falling–slightly falling–slightly rising tone or a rapidly rising–rapidly falling–slightly falling tone; yes-no questions, on the other hand, are realized on a slightly falling–rapidly rising tone or on a slightly rising–slightly falling–rapidly rising tone. Thus, the abrupt change in the fundamental frequency is at the beginning of the sentence in the former case, while it is at the end of the sentence in the latter one.
3. Imperative sentences are characterized by a wide register, a changing rhythm and a rapidly rising–slightly falling, rapidly falling–slightly falling or a level–rapidly rising–rapidly falling–slightly falling tone.

Comparing the intonation pattern of Polish and Hungarian sentences, I have found that the greatest difference is in the yes-no questions. The Hungarian intonation pattern in this case is usually slightly falling while it is rapidly rising in Polish. There is a greater similarity between the Hungarian and the Polish assertive and imperative sentences. The register of the Polish native speakers is wider, but there are no considerable differences in the tempo of speech.

**A MAGYAR FONETIKAI FÜZETEK ELSŐ TÍZ KÖTETÉNEK
BIBLIOGRÁFIÁJA
THE BIBLIOGRAPHY OF THE FIRST TEN ISSUES OF
THE HUNGARIAN PAPERS IN PHONETICS**

1. VIZSGÁLATOK A HANGTAN KÖRÉBŐL (Hangadás, hangfejlődés, hangrendszer, intonáció, beszédjavítás) / STUDIES IN PHONETICS (Production of sound, sound system, intonation, speech correction)
MFF 1. 1978, 128 lap, 10,2 (A/5) ív. (A sorozat száma: HU ISSN 0134-1545.)
2. LEÍRÓ HANGTANI TANULMÁNYOK / PAPERS IN DESCRIPTIVE PHONETICS
MFF 2. 1978, 114 lap, 9,1 (A/5) ív.
3. A BESZÉDINTONÁCIÓ NÉHÁNY ELMÉLETI, MÓDSZERTANI ÉS GYAKORLATI PROBLÉMÁJA / SOME THEORETICAL, METHODOLOGICAL AND PRACTICAL PROBLEMS OF INTONATION
MFF 3. 1979, 137 lap, 10,8 (A/5) ív.
4. HANGTANI TANULMÁNYOK / PAPERS IN PHONETICS
MFF 4. 1979, 200 lap, 15,8 (A/5) ív.
5. INTERLINGVÁLIS HANGTANI EGYBEVETÉSEK / INTERLINGUAL COMPARISON IN PHONETICS
MFF 5. 1980, 158 lap, 13,3 (A/5) ív.
6. MAGYAR HANGALBUM. A magyar beszédhangok artikulációs és akusztikus sajátosságai / A PHONETIC CONSPECTUS OF HUNGARIAN. The articulatory and acoustic features of Hungarian speech sounds
MFF 6. 1980, 179 lap, 14,1 (A/5) ív.
7. A BESZÉDHANG FONETIKAI MINŐSÉGE / PHONETICAL QUALITY OF SPEECH SOUNDS
MFF 7. 1981, 184 lap, 14,4 (A/5) ív. ISBN 963 8461 09 8
8. A SZEGMENTÁLIS HANGSZERKEZET VIZSGÁLATA / EXAMINATION OF THE SEGMENTAL LEVEL OF SPEECH
MFF 8. 1981, 168 lap, 13,75 (A/5) ív. ISBN 963 8461 13 6
9. AZ AMERIKAI ANGOL BESZÉDHANGOK ATLASZA (A beszédhangok artikulációs és akusztikus sajátosságai) / A PHONETIC CONSPECTUS OF AMERICAN ENGLISH (The articulatory and acoustic features of American English speech sounds)
MFF 9. 1982, 216 lap, 16,9 (A/5) ív. ISBN 963 8461 14 4
10. A BESZÉD AKUSZTIKAI ANALÍZISE ÉS SZINTÉZISE / THE ACOUSTIC ANALYSIS AND SYNTHESIS OF SPEECH-PRODUCTION
MFF 10. 1982, 163 lap, 12,8 (A/5) ív. ISBN 963 8461 16 0

*

BABAI László: Az emlékezeti hangoztatások (beszéltetések) jelentősége, szerkezeti és kivitelezési formái a dadogás terápiájában / Verbale Gedächtnisproben und Gespräche in der Therapie des Stotterns. MFF 1. 1978, 121-5.

- BALÁZS János: A nyelvközi (interlingvális) kutatási módszerek történeti áttekintése / A historical survey of interlingual research methods. MFF 5. 1980, 7–17.
- BALOGH Lajos: Kontrasztív hangstatisztikai vizsgálatok a regionális köznyelviség köréből / A statistical examination of sounds in regional standards. MFF 5. 1980, 77–90.
- BAŃCZEROWSKI Janusz: A lengyel és a magyar vokális rendszer egybevetésének néhány kérdése / Some aspects of the comparison of the Polish and Hungarian vowel systems. MFF 5. 1980, 117–30.
- BARTÓK János: A négyféle kérdő hanglejtés / Die viererlei fragende Tonhöhenbewegung. MFF 1. 1978, 97–104.
- Röviden a beszéd ritmusáról / A short survey of speech rhythm. MFF 3. 1979, 31–42.
- BOLLA Kálmán: Elnöki megnyitó (A Fonetika '77 konferencián) / Opening speech (At the scientific session: Fonetika '77). MFF 1. 1978, 7–9.
- A magyar magánhangzók akusztikai analízise és szintézise / On the analysis and synthesis of Hungarian vowels. MFF 1. 1978, 53–68.
 - Előszó / Preface. MFF 2. 1978, 5–7.
 - A fonetikus írás / Phonetic transcription. MFF 2. 1978, 7–24.
 - A magyar beszédhangok ajakartikulációjának kísérleti-fonetikai vizsgálata / Experimental-phonetic analysis of the labial articulation of Hungarian speech sounds. MFF 2. 1978, 31–51.
 - A magyar beszédhangok képzési konfigurációinak meghatározása palato- és lingvografikus kísérletekkel / Defining the articulatory configurations of Hungarian speech sounds with palato- and linguographic experiments. MFF 2. 1978, 51–66.
 - Elnöki megnyitó (A Fonetika '78 konferencián) / Opening speech (At the scientific session: Fonetika '78). MFF 3. 1979, 5–7.
 - A beszéd folyamat intonációs elemzése és az intonáció fonetikus lejegyzése / The intonational analysis of speech flow and the phonetic representation of intonation. MFF 3. 1979, 19–31.
 - Előszó / Preface. MFF 4. 1979, 5–7.
 - Az orosz magánhangzók analízise és szintézise / On the analysis and synthesis of Russian vowels. MFF 4. 1979, 33–80.
 - Elnöki megnyitó (A Fonetika '79 konferencián) / Opening speech (At the scientific session: Fonetika '79). MFF 5. 1980, 5–7.
 - A fonetikai szerkezetek interlingvális egybevetéséről (Problémavázlat) / On the interlingual comparison of phonetic constructions (An outline of the problems involved). MFF 5. 1980, 40–70.
 - Előszó / Preface. MFF 6. 1980, 5–7.
 - Magyar hangalbum. A magyar beszédhangok artikulációs és akusztikus sajátosságai / A phonetic conspectus of Hungarian. The articulatory and acoustic features of Hungarian speech sounds. MFF 6. 1980, 7–168.
 - Előszó / Preface. MFF 7. 1981, 5–7.
 - A magyar hosszú mássalhangzók képzése (Kinoröntgenografikus vizsgálat számítógéppel) / The articulation of Hungarian long consonants (Cineradiographic analysis with computer). MFF 7. 1981, 7–56.

- Előszó / Preface. MFF 8. 1981, 3–5.
- A magyar magánhangzók és rövid mássalhangzók képzési sajátosságainak dinamikus kinöröntgenográfiai elemzése / The articulation of Hungarian vowels and short consonants. A cineradiographic analysis. MFF 8. 1981, 5–63.
- Előszó / Preface. MFF 9. 1982, 7–9.
- Az amerikai angol beszédhangok atlasza (A beszédhangok artikulációs és akusztikus sajátosságai) / A phonetic conspectus of American English (The articulatory and acoustic features of American English speech sounds). MFF 9. 1982, 9–216.
- Előszó / Preface. MFF 10. 1982, 5.
- A magyar beszéd akusztikai szerkezetének analízise és szintézise. Kutatástörténeti áttekintés / The analysis and synthesis of the acoustic structure of Hungarian speech. A historical survey of the research. MFF 10. 1982, 7–21.
- Folyamatos beszéd szintetizáló rendszer magyar nyelven (VOXON) / VOXON: A system generating impersonal Hungarian speech by rule. MFF 10. 1982, 118–29.
- A lengyel beszédhangok képzési és akusztikus sajátosságairól / Articulatory and acoustic features of Polish speech sounds. Társszerző: Földi Éva. MFF 7. 1981, 91–140.
- A lengyel beszédhangok palato- és lingvografikus vizsgálata / Palato- and linguographic analysis of Polish speech sounds. Társszerző: Földi Éva. MFF 7. 1981, 140–56.
- A lengyel beszédhangok ajakartikulációja / The labial articulation of Polish speech sounds. Társszerző: Földi Éva. MFF 8. 1981, 104–47.
- Rendeződő cerebellaris dysarthria fonetikai vizsgálata / The phonetic examination of a successfully treated case of cerebellaris dysarthria. Társszerző: Pintér Nándor. MFF 3. 1979, 117–35.

DEME László: Grammatikai képlet és akusztikai képlet kapcsolatához / On the connection between grammatical and acoustic patterns. MFF 3. 1979, 7–14.

ERDÉLYI Alisa–KAPUSI Gyula: Az agy szerepe a beszéd folyamatok szervezésében / The role of the human brain in the organization of the speech mechanism. MFF 3. 1979, 68–73.

B. FODOR Katalin: Nyelvjárásaink intonációs vizsgálatáról / On the examination of Hungarian dialect intonation patterns. MFF 3. 1979, 73–80.

FÖLDI Éva: A kérdés kifejezésének intonációs eszközei a lengyelben és a magyarban / Intonational means of expressing questionhood in Hungarian and Polish. MFF 5. 1980, 109–17.

- A lengyel beszédintonáció elemzése szintézissel / The analysis of Polish intonation by synthesis. MFF 10. 1982, 129–53.
- A lengyel beszédhangok képzési és akusztikus sajátosságairól / Articulatory and acoustic features of Polish speech sounds. Társszerző: Bolla Kálmán. MFF 7. 1981, 91–140.
- A lengyel beszédhangok palato- és lingvografikus vizsgálata / Palato- and linguographic analysis of Polish speech sounds. Társszerző: Bolla Kálmán. MFF 7. 1981, 140–56.
- A lengyel beszédhangok ajakartikulációja / The labial articulation of Polish speech sounds. Társszerző: Bolla Kálmán. MFF 8. 1981, 104–47.

- FRINT Tibor: A hangképzés szervi eredetű zavara / Die Stimmstörung organischen Ursprungs. MFF 1. 1978, 18–25.
- GEREBEN Ferencné: Súlyos beszédhibások vizsgálata és terápiája / Analysis and therapy of children with serious speech defects. MFF 1. 1978, 113–7.
- GÓSY Mária: A szóhangsor kialakulása a gyermeknyelvben / From sound-sequences to words in early childhood (till 15 months). MFF 1. 1978, 25–37.
- Szavak és toldalékok hangalaki jellemzői a gyermeknyelvben / Characteristics of phonetic forms of words and affixes in the child's language. MFF 2. 1978, 90–100.
 - Egy percepciós hanglejtésvizsgálat / A perceptual experiment in intonation. MFF 3. 1979, 58–68.
 - Akusztikai paraméterek és nyelvi funkció a beszéddallam és a nyomaték percepciójában / Acoustic parameters and linguistic function in the perception of speech melody and stress. MFF 4. 1979, 119–37.
 - A [r] hang kialakulása a gyermeknyelvben / The evolution of the sound [r] in the child language. MFF 4. 1979, 172–82.
 - Az intonáció percepciója összehasonlító vizsgálatban / The perception of intonation from a confrontative point of view. MFF 5. 1980, 100–9.
 - A beszédhang kialakulása a gyermeknyelvben / Development of speech sounds in child language. MFF 7. 1981, 67–91.
 - A szegmentális hangszerkezet percepciójáról / On the perception of segmental level of speech. MFF 8. 1981, 87–104.
 - A [b, d, g] mássalhangzók percepciós vizsgálata / The perceptual analysis of [b, d, g]. MFF 10. 1982, 84–110.
- HAJDÚ Péter: Nyitó szavak a Magyar Fonetikai Füzetek első számához / To the first number of Hungarian Papers of Phonetics. MFF 1. 1978.
- Rokon nyelvi fonológiai rendszerek egybevetése / The comparison of the phonological systems of related languages. MFF 5. 1980, 17–22.
- HIRSCHBERG Jenő: A gyermek beszédfejlődésének orvosi vonatkozásai / Die ärztlichen Aspekte der Sprachentwicklung. MFF 1. 1978, 37–44.
- HONTI László: Az északi obi-ugor nyelvjárások fonémarendszerének összevetése / A comparison of the phoneme systems of northern obi-ugrian dialects. MFF 5. 1980, 90–100.
- IMRE Samu: A magyar nyelvjárások néhány fonológiai kérdése / Some phonological questions of Hungarian dialects. MFF 5. 1980, 22–9.
- KASSAI Ilona: Sva-jelenségek a magyar beszédben / Schwa phenomena in Hungarian speech flow. MFF 1. 1978, 92–7.
- A hangkapcsolatokról / On sound clusters. MFF 2. 1978, 66–73.
 - A gyermeknyelvi hasonulások egy fajtája: a távhasonulás / On one type of assimilation in child language. MFF 2. 1978, 100–8.
 - Az intonáció szerepe a gyermeknyelvben / The role of intonation in child language. MFF 3. 1979, 80–4.
 - Magánhangzó–mássalhangzó találkozások / Vowel–consonant clusters. MFF 4. 1979, 80–119.
 - Gyermeknyelvi dallamminták / Melodic patterns in child language. MFF 4. 1979, 147–72.

- Sva-jelenségek a francia és a magyar beszédben / Schwa phenomena in Hungarian and French speech. MFF 5. 1980, 139–52.
 - A hangsúly kialakulása a gyermeknyelvben / The formation of stress in child language. MFF 7. 1981, 156–75.
 - A magyar beszéd hangsorépítési szabályszerűségei / The phonotactic rules in Hungarian. MFF 8. 1981, 63–87.
 - A hadaró beszéd dallama / Melody in cluttering. Társszerző: Vassné Kovács Emőke. MFF 3. 1979, 93–101.
- KISS Gábor: A dinamikus röntgenográfiai vizsgálat számítógépes programja / The computer program for dynamic radiographic examinations. MFF 7. 1981, 56–9.
- Interaktív beszéd szintetizáló rendszer számítógéppel és OVE III szintetizátorral / An interactive speech synthesizing system with computer and OVE III synthesizer. Társszerző: Olasz Gábor. MFF 10. 1982, 21–46.
 - A magyar beszéd automatikus szintézisének első lépcsője / The first step in the automatic synthesis of Hungarian. Társszerző: Olasz Gábor. MFF 10. 1982, 110–8.
- KOTLEJEV, V. I.: A csuvas és az orosz magánhangzók egybevető akusztikai elemzése / A comparison of the acoustic features of Chuwash and Russian vowels. MFF 5. 1980, 130–9.
- A. MOLNÁR Ildikó: A hanghelyettesítések típusai a gyermeknyelvben 18–21 hónapos kor között / The types of sound-substitutions in child language between the age of 18 and 21 months. MFF 1. 1978, 44–53.
- Hezitációs jelenségek az élőbeszédben / Hesitational phenomena in speech. MFF 3. 1979, 49–58.
- MOHR János: A dadogók kezelésének tervezhetősége „terápiás programban” / „Therapeutisches Programm” für Stotterer – ist es planbar? MFF 1. 1978, 117–21.
- OLASZY Gábor: Szintetizált magyar magánhangzók formáns-intenzitás és formáns-sáv-szélesség értékei / Formant intensity and formant bandwidth values in the synthesized vowels of standard Hungarian. MFF 1. 1978, 68–77.
- Új műszer a beszéd intonációjának percepció vizsgálatához / An instrument for perceptual examination of intonation. MFF 2. 1978, 108–14.
 - A magyar beszéd intonációs vizsgálatának eszközfontetikai háttere / The technical background of the examination of intonation in Hungarian. MFF 3. 1979, 101–10.
 - A zöng szerepe az egyéni hangszínezet kialakításában / The participation of voice in the formation of individual timbre. MFF 4. 1979, 137–47.
 - Hangsorok számítógépes formáns szintézisének előkészítése / Preparation of the computer formant synthesis of sound sequences. MFF 8. 1981, 147–60.
 - A magyar mássalhangzók és a mássalhangzó–magánhangzó–kapcsolódások akusztikai szerkezetének analízise és szintézise / The analysis and synthesis of the acoustic structure of Hungarian consonants and consonant–vowel combination types. MFF 10. 1982, 46–84.
 - Interaktív beszéd szintetizáló rendszer számítógéppel és OVE III szintetizátorral / An interactive speech synthesizing system with computer and OVE III synthesizer. Társszerző: Kiss Gábor. MFF 10. 1982, 21–47.

- A magyar beszéd automatikus szintézisének első lépcsője / The first step in the automatic synthesis of Hungarian. Társszerző: Kiss Gábor. MFF 10. 1982, 00–00.
- PETŐ Zsigmond: Intonációs típushibák a főiskolai hallgatók orosz beszédében / Typical intonational mistakes in the Russian speech of Hungarian college students. MFF 5. 1980, 164–8.
- PINTÉR Nándor: Rendeződő cerebellaris dysarthria fonetikai vizsgálata / The phonetic examination of a successfully treated case of cerebellaris dysarthria. Társszerző: Bolla Kálmán. MFF 3. 1979, 117–35.
- SUBOSITS István: Energiaeloszlás a magyar [s] és [ʃ] hangok hangszínképében / Energieverteilung im Spektrum des ungarischen [s] und [ʃ]. MFF 1. 1978, 77–92.
- Monotónia a dadogók beszédében / Monotony in the speech of stammerers. Társszerző: Vinczéné Bíró Etelka. MFF 3. 1979, 84–93.
- SURÁNYI Ibolya: A refrén a versmondásban / The role of refrain in the recitation of poetry. MFF 5. 1979, 110–7.
- SZENDE Tamás: A „felsikló” hanglejtésforma létrejöttének magyarázata a hanglejtésképzés módozatai alapján / Interdependencies in stress production, intonation patterns and structuring of complex sentences in Hungarian. MFF 1. 1978, 104–7.
- Az ’elvont’ fonéma definíciói / Defining the phoneme: phenomenological aspects. MFF 2. 1978, 24–31.
- A hangtulajdonságok felhasználásának rétegei / Prosodic devices in communicative systems level of utilization in language. MFF 3. 1979, 14–9.
- A szünet és a junktúra / Pause and juncture. MFF 4. 1979, 7–33.
- ’Egyetemes fonetikai–fonológiai egységek’ kutatása és értéke az interlingvális fonológiai összevetésben / Phonetic-phonological units as universals in language and how to use them in confrontative investigations. MFF 5. 1980, 29–40.
- TARNÓCZY Tamás: A hangrésmozgás néhány tulajdonságáról / Some properties of the motion of the glottis. MFF 1. 1978, 9–18.
- P. TÁLOS Endre: Fonológiai szabálykölcsonzés a cigányban / Phonological rule-borrowing in Romani. MFF 5. 1980, 152–9.
- VALACZKAI László: Magyar és német beszédhangok néhány asszociatív sajátosságáról / Some associative peculiarities of Hungarian and German speech sounds. MFF 5. 1980, 159–64.
- VASSNÉ KOVÁCS Emőke: A hadaró beszéd dallama / Melody in cluttering. Társszerző: Kassai Ilona. MFF 3. 1979, 93–101.
- VÉRTES O. András: A szótag / The syllable. MFF 2. 1978, 73–90.
- A hang némely tulajdonságának történeti változásáról / On historical changes in the voice characteristics of Hungarian. MFF 3. 1979, 42–9.
- A történeti hangstiliztika egy alapkérdése / A fundamental question of historical phonostylistics. MFF 4. 1979, 182–5.
- Egy reformkori értekezés a dadogás gyógyításáról / An early Hungarian attempt at the surgical treatment of stammering. MFF 4. 1979, 185–90.
- A fonetikus írás egyetemessége és a kontrasztív fonetika / The universality of phonetic transcription and confrontative phonetics. MFF 5. 1980, 70–7.
- A szöveg hangalakjáról / Gábor Egressy (1808–1866) on the phonic form of the text. MFF 8. 1981, 160–4.

- VICSI Klára: Az időtartam szerepe néhány mássalhangzótípus hallás alapján történő megkülönböztetésében / The role of duration in perceptual identification of some types of consonants. MFF 7. 1981, 59–67.
- VINCZÉNÉ BÍRÓ Etelka: A fonetikai hibák javításának dinamizmusa / On the mechanism of correcting phonetic faults. MFF 1. 1978, 107–13.
- Monotónia a dadogók beszédében / Monotony in the speech of stammerers. Társ-szerző: Subosits István. MFF 3. 1979, 84–93.

ÚJ KIADVÁNYOK NEW PUBLICATIONS

BALÁZS János: Magyar deákság. Anyanyelvünk és az európai nyelvi modell.
Elvek és Utak. Budapest 1980. 656 lap.

Balázs János nagy munkája három szakaszra tagolódik: I. Az európai nyelvi modell a magyar deákság nyelvbölcseletében; II. A bizánci diákonosztól a magyar deákgig; III. Az európai nyelvi modell szintjei. Összesen 32 fejezetben, illetőleg többé-kevésbé önálló dolgozatban tárgyalja a „magyar deákságot”. Mit ért ezen a kifejezésen? Révai nyomán: a „magyar nemzetnek írástudását, és külömb-külobmb tanúságát, kiváltképpen pedig annak hazai nyelvén” (7).

A magyar fonetikust különösen érdeklik a következő fejezetek: Kempelen nyelvfilozófiai nézetei (40–5), Kis János pályaműve (94–108), A nyelvek jóhangzása (136–44), A nyelvek hathatósága (172–87), Hangok és betűk (363–407), A szövegszinttől a prozódiai szintig – A Halotti Beszéd vallomása (451–86), Az Ómagyar Mária-siralom prozódiai rejtelmek (486–514), A középkori latin versek és a magyar vers ritmusa (514–55), A magyar időmértékes vers kialakulása (555–77).

A szerzőnek – a hangtant is érintő – számos gondolata közül e rövid ismertetésben csak néhányat emelek ki.

1. Szenci Molnár Albert, Komáromi Csipkés György és más régi grammatikusaink – mint tudjuk – összevetették több hangunkat élő idegen nyelvekéivel: Balázs szerint az „ilyen nekigyürkőzéseket úgy tekinthetjük, mint első kísérleteket fonémáink izoglosszáinak tétova megrajzolására” (405); joggal tartja kívánatosnak ezzel kapcsolatban valamennyi beszédhang európai vagy akár eurázsiai izoglosszáinak megvizsgálását és térképi ábrázolását.

2. Szerzőnk felhívja a figyelmet Daniel Jenischnek 1796-ban megjelent pályamunkájára s e mű hazai hatására: ezt Pápay és Kiss János munkásságában találja meg s Kölcseynek Jenisch-kivonatában. Mint könyvének számos helyén, Balázs itt újat ad, ez esetben Simainak századeleji becses megjegyzéseihez képest is jelentősen többet (vö. MNy VIII, 1912, 316 kk.; XII, 1916, 395 kk.); helyénvaló, hogy megállapításairól megemlékeznek a magyar hangtan történetének monográfusa, mert a német tudós magyarországi hangtani nézetekre is hatott.

3. Prozódiai kérdések és viták szövevényében is kalauzunk a könyv. Arany János, Négyesy László, Gábor Ignác, Horváth János, Németh László, László Zsigmond, Vargyas Lajos és más kutatók nézeteinek vizsgálata nyomán lényegében a Gábor–Németh–Vargyas–Gáldi „nevével fémjelezhető hazai hagyományhoz” csatlakozik. E szerint az ősi ritmust követő magyar vers tagoló volt. Balázs rámutat arra, hogy ez a nézet volta-képpen Arany Jánostól származik, s bizonyítja nagy költőnk *tag* és *íz* műszavának a klasszikus ókorból való származását.

Talán elérkezik az idő, amikor a magyar verstan kutatója (az eddigi kezdeményeinknél sokkal szélesebb és mélyebb) eszközfonetikai vizsgálatokat is figyelembe vehet: monodáknak, közmondásoknak, verseknek, de általában is népnyelvi és művészi előadású

szövegeknek sokoldalú elemzését, amely kiterjedne a nyomatékra, az időtartamra, a tempóra, a hangmagasságra, a hanglejtésre és egyéb tényezőkre. Kétségtelen azonban, hogy a verstanunk mindezek nélkül is tiszteletre méltó eredményeket ért el —, ez kiderül Balázs munkájából is.

A könyvben megnyilvánuló problémaérzék, az érdekes gondolatmenetek, a hatalmas irodalmi, szakirodalmi tájékozottság, a filológiai akribia s a szerző szép magyar nyelve: mind, mind Balázs Jánosnak ismert erénye.

Vértes O. András

Fejezetek a leíró magyar hangtanból. Szerk. BOLLA Kálmán. Budapest 1982. 300 lap + 47 tábló.

A kötet tanulmányai a szegmentális hangszerkezet kísérleti-fonetikai kutatásainak eredményeit dolgozzák fel. A Bevezetésen (9–13) kívül 6 szerző 12 tanulmányát tartalmazza a könyv.

*A leíró hangtan vázlat*a (Bolla Kálmán) alapvető elméleti kérdésekkel foglalkozik; elemzi a nyelvészeti fonetika szakrendszerét, ezen belül az öt ágazatot: a kortikális, az artikulációs, az akusztikai, az auditorikus és a percepciós fonetikát (13–25). Mind elméleti, mind gyakorlati kérdéseket vet fel és válaszol is meg a beszédhangok és intonációs struktúrák artikulációs és akusztikus sajátosságainak ismeretében *A fonetikus írás problémái* (Bolla Kálmán). Az APhI-átírást bemutató táblázatok és a beszédintonáció fonetikus lejegyzésére javasolt minták szemléletesen segítik elő a hangjelölési módok egységesítésének igényét, amelytől a nyelvi valóság pontosabb, jobb megismerése várható (25–53). Az *Akusztikai alapfogalmak* című fejezet (Kassai Ilona) a hangjelenségeket általában, majd a beszédhanggal összefüggő akusztikai alapfogalmakat tekinti át két kisebb egységben: a fizikai és a szubjektív akusztika elemzésén belül (53–67). A beszéd — akusztikai szempontból — sajátos rezgéskép. Ennek a rezgésképnek a kialakulásához szükséges levegő nyomásának a szerepét vizsgálja Szende Tamásnak *A levegőnyomás szerepe a hangképzésben* című dolgozata (67–71). Vértes O. András tanulmánya (*A magyar beszédhangok akusztikai elemzésének kérdései*) egyrészt szakirodalmi áttekintést ad e témáról, másrészt a szerző saját kísérleti eredményeivel foglal állást a vitás kérdésekben (71–115). Ugyancsak saját kísérleti adatokon alapszik Kassai Ilona *A magyar beszédhangok időtartamviszonyai* című tanulmánya. A szerző a teljes magyar hangrendszert elemzi az időtartam szempontjából, a korábbi kutatások eredményeit egymással és saját adataival vetve össze (115–55). Az artikuláció és az akusztikum kapcsolatával foglalkozik röviden Vértes O. András *Az artikuláció akusztikus vetülete* című dolgozata (155–65). A beszédhangok korszerű artikulációs és akusztikus vizsgálatának módozatait ismerjük meg Bolla Kálmán munkájának (Magyar hangalbum) szöveges részéből (165–72). A hangtablók diagramjairól (valamennyi magyar magán- és mássalhangzó foto- és kinolabiogramja, palato- és lingvogramja, röntgenogramja és röntgensémája, illetve oscillogramja, hangszínképe és amplitúdómetSZete) a magyar beszédhangok artikulációs és akusztikus sajátosságai olvashatók le. *A beszéd megértése* (Pauka Károly) áttekinti e részterület csaknem valamennyi fontosabb problémakörét; képet ad a hallási rendszeréről, a hallásjelenségekről; kitekint a beszédmegértés elemi egységeire és a gyermek be-

szédmegértési képességének fejlődésére (175–33). A beszédhangok, a magyar fonémaállomány és fonémarendszer elemzését adja Szende Tamás *A mai magyar nyelv fonémái* című tanulmánya (233–67). A normától eltérő ejtésbeli sajátosságokkal foglalkozik Gósy Mária dolgozata (*Az élőbeszéd hibáiról*) az egyes hangok helytelen ejtésétől a közlésegség hangos megjelenítésében elkövetett hibákig (267–83). A tanulmánykötetet *A magyar leíró hangtan története az újgrammatikusoktól 1945-ig* címmel Vértes O. András munkája zárja, amely szerves folytatása a szerzőnek a közelmúltban megjelent „A magyar leíró hangtan története a kezdetektől az újgrammatikusokig” című művének. Az olvasó rövid összefoglalást kap az előzményekről, majd részletes leírását az új módszereknek, elméleti törekvéseknek.

Az egyes tanulmányok után álló részletes bibliográfia a résztémák fontosabb hazai és külföldi szakirodalmát mutatja be.

A tanulmánykötet hasznos információkat nyújt fonetikusoknak, nyelvészeknek, a határtudomány művelői és a beszéd hangjelenségei iránt érdeklődők számára.

Gósy Mária



Címünk:

A Magyar Tudományos Akadémia
Nyelvtudományi Intézete
Fonetikai Osztály
Budapest, I., Szentháromság u. 2. Pf. 19.
1250

Address for communications:

Department of Phonetics,
Institute of Linguistics,
Hungarian Academy of Sciences
Budapest, I. Szentháromság u. 2. Pf. 19.
H-1250 Hungary

