

A magyar beszéd előállítása számítógéppel

KISS GÁBOR—OLASZY GÁBOR

A dolgozat az MTA Nyelvtudományi Intézetében 1979–82-ben kifejlesztett Interaktív Beszédszintetizáló Rendszer (INBERE) felépítését és működését ismerteti. A rendszer PDP 11/34-es számítógépből és OVE III beszédszintetizátorból áll. A felhasználó a szintézist a FOPROX nevű, saját fejlesztésű beszédszintetizáló programunkkal végezheti. A programmal párbeszédes módon kommunikálhatunk a VEZER nyelv segítségével.

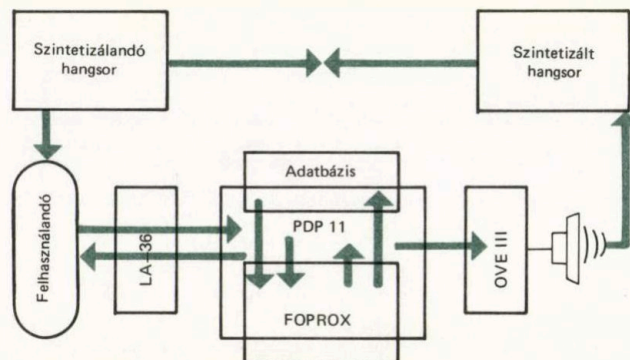
A rendszerrel tetszőleges hangsorokat lehet előállítani, vizsgálni, lejátszani, nyújtani, zsugorítani stb. A rendszer főleg olyan fejlesztő, kutató helyeken alkalmazható, ahol beszédakusztikai vizsgálatokkal foglalkoznak tudományos, ipari vagy egyéb felhasználás céljából.

Többek között az INBERE felhasználásával készítették el a szerzők a SZAMOK és KEPLET programokat, amelyekkel első ízben lehetett magyar számokat és matematikai képleteket kimondatni a számítógéppel. Az előbbiek tapasztalatai alapján fejlesztették ki 1982-ben az UNIVOICE elnevezésű, valós időben üzemelő, magyar nyelvű „text to speech” szintetizáló rendszer alapváltozatát, amelyet a 8. Budapesti Akusztikai Konferencián mutattak be.

(Érkezett: 1982. március 31.)

A mesterséges beszéd alkalmazási területei

Napjainkban — különösen a fejlett ipari országokban — egyre jobban előtérbe kerül a mesterséges beszéd alkalmazása a tudományos kutatás, az ipar és az élet számos egyéb területén.



1. ábra. Beszédszintetizáló rendszer

Magyarországon a magyar beszéd számítógéppel történő előállításának nincs nagy múltja; egyrészt azért, mert az ehhez szükséges nyelvészeti-fonetikai kutatások csak az utóbbi időben terjedtek ki erre a területre, másrészt, mert a szintézis technikai feltételei csak az utóbbi években teljesültek.

Dolgozatunkban bemutatjuk az MTA Nyelvtudományi Intézetének fonetikai osztályán, az 1979–1982. években kidolgozott INteraktív BEszédszintetizáló Rendszert (INBERE), amelynek segítségével (nemcsak magyar nyelvű) szavakat, mondatokat lehet előállítani nyelvészeti-fonetikai, orvosi, pszichológiai stb. kutatások kiszolgálására, ipari alkalmazásokhoz és egyéb társadalmi felhasználásra.

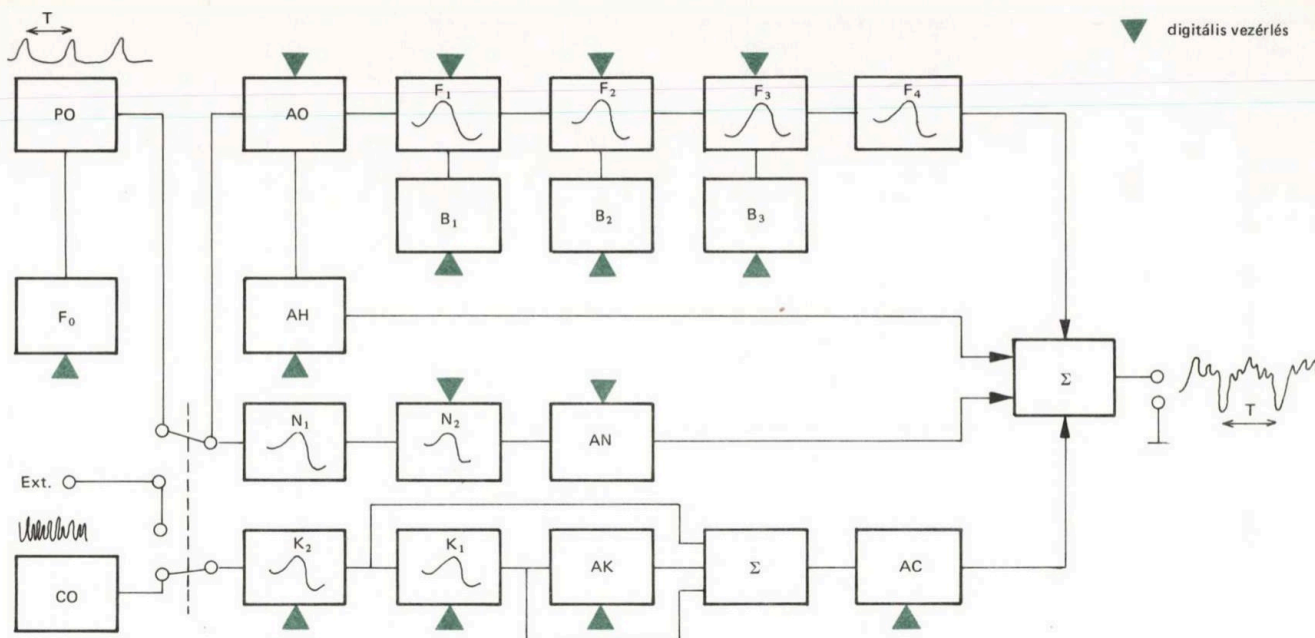
Előbb azonban szóljunk néhány szót a mesterséges beszéd felhasználásáról. A számítógéppel generált mesterséges hangsorok felhasználása új módszereket kínál a kutatómunkában, új távlatokat nyújt az ipar, a kereskedelem, a hírközlés stb. számára.

A kutatómunkában a mesterséges beszéd hatékonyan felhasználható, mivel a számítógéppel szabályok alapján előállított hangsorok akusztikai szerkezetének bármelyik összetevőjét a kutatás céljának megfelelően, tudatosan megváltoztathatjuk, és így olyan akusztikai tartalmat adhatunk beszédhangoknak, szavaknak, mondatoknak, amelyet a természetes beszéddel nem tudunk létrehozni. Néhány példa ezekből az alkalmazási lehetőségekből: a magyar (vagy más) beszéd akusztikai szerkezetének feltárása, percepciók kísérletek, kontrasztív nyelvészeti kutatások, pszichológiai vizsgálatok.

Az orvosi kutatások számos ágában felhasználható a szintetizált beszéd. A beszédaudiometria új ága alakulhat ki azzal, hogy a hallástereszték anyagát szintetizált szavakból állítjuk össze. A szintetizált szavak egyes akusztikai paramétereinek kiemelésével, időtartamviszonyainak megváltoztatásával vagy egyes frekvencia-komponenseinek áthelyezésével például olyan „betegség típus-orientált” teszteket lehet előállítani, amelyekkel behatárolható a pontos diagnózis.

Az orvosi és rehabilitációs gyakorlat számos területén használható a szintetizált beszéd. A hallássérültek gyakoroltatására olyan „beszélő” készülékeket szerkeszthetünk, amelyek segítségével a beteg maga gyakorolhat, javíthatja hallását.

Az ilyen öngyakoroltató készülékkel olyan eszközt adunk a rehabilitálandó beteg kezébe, amit mással (pl. magnetofon) nem lehet helyettesíteni.



2/a. ábra. AZ OVE III beszédszintetizátor

A némák vagy beszédsérültek is „beszélhetnek”, ha olyan hordozható, mikroprocesszoros készülék áll rendelkezésükre, amelynek billentyűzetével szavakat, mondatokat lehet összeállítani, azokat például telefonálásnál meg lehet szólaltatni.

A vakok életének és munkájának könnyítésében is jelentős szerepet kaphat a gépi beszéd alkalmazása. Gondoljunk csak az „olvasó” számítógépekre, amelyekkel könyvek, újságok nyomtatott szövegeit hallhatóvá tehetjük. A beszélő számítógép-terminál az eddig csak képernyőn vagy nyomtatásban megjelenő információt hangos beszéddel adja meg. Így a gépekkel például vak dolgozók is tudnak kommunikálni. Számos olyan esetben, ahol a vak ember számára az információszerezés problémát jelent, csak a kis méretű „beszélő” készülékkel segíthetünk. Például a látássérült személy saját hőmérsékletét, vérnyomását, testsúlyát stb. csak olyan műszerrel tudja megmérni, amely hangosan közli a mért eredményt.

Az *iparban* is széles skálája található a mesterséges beszéd alkalmazhatóságának. A postai gyakorlatban számítógéppel előállított beszéddel működtethetők a pontos idő, műsortájékoztató, tudakozó, számváltozást bejelentő stb. szolgáltatások. Az ipari robotok gyors elterjedésével a termelés szerkezete megváltozik, központi irányítótermékből kísérik figyelemmel a szakemberek a munka folyamatát. Sokkal hatásosabb a vizuális jelzési formánál, ha a robotok „beszéddel” közlik a szükséges információkat kezelőikkel. A szárazföldi, vízi és légi közlekedésben számos műszaki probléma csak a szintetizált beszéd alkalmazásával oldható meg.

A *társadalom egyéb területein*, a kereskedelemben, a háztartásokban, játékok készítésénél (pl. speak and spell), a nyelvtanulásban, a helyesírás, a helyes kiejtés oktatásában, az anyanyelvi nevelés egészében szintén helyet kaphat a mesterséges beszéd alkalmazása.

A beszédszintetizáló rendszerünk hardver és szoftver részeit a következőkben ismertetjük (1. ábra).

Az INBERE hardver részei

Rendszerünk a PDP 11/34-es számítógépből és a svéd gyártmányú OVE III beszédszintetizátorból áll. A számítógép konfigurációjából a beszédszintetizálás során felhasználjuk a 64 kb-ot központi tárat, a két hajlékonylemez-egységet és az LA-36-os konzolirógépet. A beszédszintetizátor kapcsolatát a számítógéppel 16 bites paralel interfész (DR-11K) biztosítja. A rendszer működésének megértéséhez és a mesterséges beszéd eredményes előállításához szükséges az OVE III beszédszintetizátor alapos ismerete, ezért először ezt mutatjuk be.

Az OVE III beszédszintetizátor

Az OVE III manuálisan és digitálisan vezérelhető, analóg beszédszintetizátor. Segítségével beszédhangok és hangsorok állíthatók elő. A szintetizátor statikusan külső kapcsolókkal, dinamikusan számítógéppel 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 zörejszerű (3) áramkörrel, valamint a hehezetesező (4) zörejszerű forrással rendelkezik. A zöngés és nazális kört impulzusgenerátor (PO), a zörejszerű zajgenerátor (CO) gerjeszti. A szintetizátor felépítését – a digitális vezérlések helyének feltüntetésével – a 2/a ábra mutatja. A generátorokat, valamint az említett áramköröket és azok belső értékeit 16 paraméter segítségével – beleértve az időtartamot is – adhatjuk meg (2/b ábra).

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

Sorsz.	Paraméter	Címbytek				Paraméterbytek								Tartomány	Lépés	Funkció
1.	AO	0	0	0	0									$-\infty, 2-31,5$ dB	0,5 dB	zöngés hang
2.	AC	0	0	0	1									$-\infty, 2-30$ dB	2 dB	zöreje hang
3.	AH	0	0	1	0									$-\infty, 2-30$ dB	2 dB	hehezetes hang
4.	AN	0	0	1	1									$-\infty, 2-30$ dB	2 dB	nazális hang
5.	F0	0	1	0	0									50–315 Hz	0,8%	hangmagasság (dallam)
6.	F1	0	1	0	1									200–1260 Hz	0,8%	első formáns
7.	F2	0	1	1	0									500–3200 Hz	0,8%	második formáns
8.	F3	0	1	1	1									1000–6300 Hz	0,8%	harmadik formáns
9.	N1	1	0	0	0									200–1230 Hz	3%	nazális formáns
10.	AK	1	0	0	1									0–31,5 dB	0,5 dB	antiformáns
11.	K1	1	0	1	0									1000–6200 Hz	3%	alsó zöreigóc
12.	K2	1	0	1	1									2500–15700 Hz	3%	felső zöreigóc
13.	B1	1	1	0	0									0–188 Hz	12 Hz	1. formáns sávzélessége
14.	B2	1	1	0	1									0–470 Hz	31 Hz	2. formáns sávzélessége
15.	B3	1	1	1	0									0–750 Hz	250 Hz	3. formáns sávzélessége

2/b ábra. Az OVE III beszéd szintetizátor paramétertáblázata

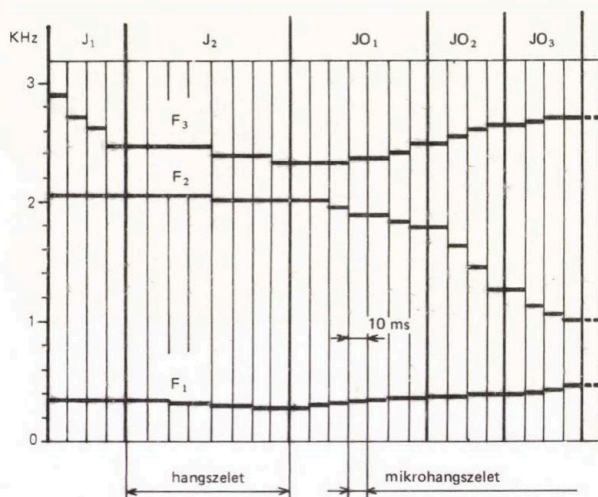
Az OVE III digitális vezérlése

Mint már említettük, a szintetizátort 16 bites paralel „interfész”-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 címbyte, 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 központi tárból kell programutalítással beírni. Miután a program beírta a biteket az OBR-be, az interfész kiadja a „kész” jelet, és az adatok a szintetizátor tároló regiszterébe (SZB) íródnak. A szintetizátor válaszjellel 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. bitre a logikai „igen” szintet (+5V), akkor képes új adatokat elfogadni, az előbbi folyamat kezdődhet elől. A bitek kiosztása az OBR-ben a következő:

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

A hangokat és hangsorokat meghatározott időtartamú szeletekből (3. ábra) építjük fel, amelyek mikrohangszeletekből állnak [4]. A 3. ábrán a JÓ... hangsor mikrohangszeleteit, azok belső frekvenciaértékeit és a mikrohangszeletekből összeálló hangszeleteket láthatjuk.



3. ábra. A JÓ... hangsor mikrohangszeletei

Egy ilyen mikrohangszelet akusztikai szerkezetének felépítéséhez több paraméter megadására (pl. AO, F0, F1, F2, F3 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 építőkövei a következők: AO, F0, F1, F2 és F3. Ezeket egymást követő assembler szintű programlépésekkel tápláljuk be a szintetizátorba. Egy ilyen lépés végrehajtása kb. 2 μ s ideig tart. Így egy mikrohangszeletet az építőkövek számától függően kb. 50–60 μ s alatt lehet

generálni. Ez az idő növekedhet, ha magasabb szintű programnyelven írott programmal működtetjük a szintetizátort. A következőkben nézzük meg, hogy a bemenő adatokból (1. táblázat) a számítógép hogyan hozza létre az előbb említett, kitarított hangzású [a:] hangot.

1. táblázat. Bemenő adatok

Frekvencia	Digitális kód		
$F_0 = 79 \text{ Hz}$	1 1 1 1	1 0 1 1	1 0 1 1 1 1 1 1
$F_1 = 650 \text{ Hz}$	1 1 1 1	1 0 1 0	0 1 0 1 1 0 1 1
$F_2 = 1250 \text{ Hz}$	1 1 1 1	1 0 0 1	1 0 1 1 1 1 1 1
$F_3 = 1950 \text{ Hz}$	1 1 1 1	1 0 0 0	1 0 1 0 0 0 1 1

Azokat a programlépéseket, amelyekkel a bemenő adatokat sorba betápláljuk a szintetizátorba, a 2. táblázaton tüntettük fel.

Látható, hogy 28 programlépésre volt szükség az előbbi öt paraméter betáplálására. Ez egyetlen hangrészlet (mikrohangszelet) létrehozására volt példa. Mint már említettük, a hangokat és hangsorokat is ilyen hangrészletek összekapcsolt sorozatával kell létrehoznunk. 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 szintetizálásakor tehát az előbbi lépéssorozat zajlik le több százszor egymás után, mindig változó adattal.

2. táblázat. Az adatok szintetizátorba táplálásának programlépései

Lépés	Oktális kód	Utasítás	Az OBR tartalma	Eredmény
1. 2. 3.	012737 177400 76774	mov # 177400, @ # 76774	AO	AO megszólal
4. 5. 6.	012737 040000 167770	mov # 040000, @ # 167770	14. bit	int. enab. bit beírása
7. 8. 9.	012737 175677 767774	mov # 175677, @ # 767774	F_0	$AO + F_0$ szól
10. 11. 12.	012737 040000 167770	mov # 040000, @ # 167770	14. bit	int. enab. bit beírása
13. 14. 15.	012737 175133 767774	mov # 175133, @ # 767774	F_1	$AO + F_0 + F_1$ szól
16. 17. 18.	012737 040000 167770	mov # 040000, @ # 167770	14. bit	int. enab. bit beírása
19. 20. 21.	012737 174577 767774	mov # 174577, @ # 767774	F_2	$AO + F_0 + F_1 + F_2$ szól
22. 23. 24.	012737 040000 167770	mov # 040000, @ # 167770	14. bit	int. enab. bit beírása
25. 26. 27.	012737 174243 767774	mov # 174243, @ # 767774	F_3	$AO + F_0 + F_1 +$ $+ F_2 + F_3$ szól
28.	000000	halt		

INFORMÁCIÓ-ELEKTRONIKA 1982/6

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

#11, #AB, #XX, #E1

A paraméter. A szintetizátor működését meghatározó 16 paraméterek a szimbolikus neve, amelyek két karakterből állnak. A paramétert mindig kettőspont követi. A nagy tömegű adatbevitel megkönnyítése érdekében a számítógép elfogadja a paramétert követő karakterek 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, AO, FO, F1, F2, F3, B1, B2, B3, AN, N1, AC, K1, K2, AK, AC

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, 234-123

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

Az egyes utasításokat a következő csoportokban ismeretjük:

- I. Adatbeviteli utasítás: BEA
- II. Szeletszerkesztő utasítások: EVA, ETI
- III. Megszólaltató 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 szintaxisa,
- b) az utasítás szemantikája,
- c) példák az utasítás helyes használatára,
- d) az utasítás helytelen használata után megjelenő hiba-üzenetek.

I. Az adatbeviteli 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 használatakor ügyelni kell arra, hogy új szelet definíciója esetén az első paraméternek mindig a TM-nek kell lennie; a már definiált szelet TM-értékét nem lehet a BEA utasítással változtatni (lásd IDA utasítás). Az n-értékének az utasítással megengedett hossza szab határt.

c) ~BEA 1A1(TM:<50>,AO:<1-12>,FO:<112-132>,F1:<654>)

~BEA 1A1(F2:<1131-1234>,F3:<2137>)

~BEA PAUSA(TM:<500>)

~

d) ~BEA 1A1(TM:<140>)

HIBA! MAR DEF. SZELETNEK NEM LEHET UJ "TM:"-JE; OKE?

~BEA 1A2(AO:<12-15>,FO:<120>)

HIBA! UJ SZELET ESETEN ELSO A "TM:" !

~BEA 1A1(FQ:<132-143>)

HIBA! AZ FQ: KOMP. HIBAS !

~BEA 1A1(AO:<12-33>)

HIBA! AO > 31 !

~BEA 1A3**TM:<120>0

HIBA! A JEL MEGADASA NEM HELYES !

~

II. Szeletszerkesztő utasítások

1. Az EVA utasítás

- a) EVA (azonosító₁) azonosító₂
- b) Az azonosító₁-gyel jelölt szelet elnevezése azonosító₂-re. Az addig azonosító₁ nevű szeletre ezután azonosító₂ névvel hivatkozhatunk. Gyakran használatos utasítás, ugyanis sem az operatív, sem az adatbázisban nem lehet két egyforma azonosítójú szelet.

c) ~EVA (PAUSA)SZUN

~

d) ~EVA (S1) IS1

HIBA! A NEVVÁLTOZTATÁS ROSSZ !

~EVA S1(S1)

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 (1A,2A,Z1,Z2,SZ1,SZ2,SZ0,0L,L1) #AZ

~ETI (1,2,3,4,5,SZUN,1,2,3,4,5,SZUN,1,2,3,4,5) #15

d) ~ETI (C1,C1,C1,I,KK1,KK,KK2)CK

HIBA! A JELEK OSSZEVONASA HELYTELEN !

~

III. Megszólaltató utasítás

a) LEO (azonosító [,azonosító] n₁) [I]

LEO (#jel [, #jel] n₂) [I]

- b) A kijelölt szeletek egymás utáni megszólaltatása, lejátszása. Egy utasítással egymás után maximum 300 szelet szólaltatható meg, tehát $1 \leq n_1 \leq 300$; az n₂ értékét pedig #jelekkel összevont szeletek száma határozza meg; itt sem lehet a lejátszandó szeletszám 300-nál több. A lejátszandó szeletek számának másrésztől határt szab az a tény, hogy a parancssor nem lehet 130 karakternél hosszabb. Az I használata esetén a hangsorból csupán a szupraszegmentális szerkezet lejátszására kerül sor.

c) ~LEO (1A1,1A2,1A3,1S1*,1S2*,1S3*,T1,T2,T3)

~LEO (#15,#15,#15)

~LEO (#AA,#BB,#CC/I

~

d) ~LEO (Y1,Y2,Y3,1X,2X,3X,1W1,2W2,3W3)

HIBA! Y1,Y2 NEM DEF. SZELET !

~LEO (#MA,#MA,#MA,#MA,#MA)

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 levő és az utasításban felsorolt szeletek elhelyezése a hajlékony lemezen levő 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 adatbázisban csak a már ott meglevő azonosítóktól különböző azonosítóval jelölt szeletet lehet elhelyezni (l. EVA utasítás),
- adott adatbázisban 2048 szelet helyezkedhet el, valamint az # jelek száma nem haladhatja meg az 512-t.

c) ~MOR (S11,S22,ST,T0,T,TT00,D,01,P0,P,P*)

~MOR (#PA)

d) ~MOR (A1,A2,A3)

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

~

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 levő azonosítóktól különböző azonosítóval jelölt szeletet lehet felhozni az adatbázisból. Az n_1 , n_2 értéke megegyezik a LEO utasításnál megadott értékkel.

c) ~JON (0A',A',A'J\$,A'J\$,JJ)

~JON (#DT,#PA,#ES)

d) ~JON (SSS0,SSS,S1,ST1,ST\$,T0,T,T0,01,012,0P,P0,P,P\$)

HIBA!*** SAJNOS >>SSS0 << SZELET MÄR VAN A GEPBEN.

~JON (I@I,ID,D00,D,D0*,0*,0*26,0*#)

HIBA! SAJNOS A >>I@I << SZELETET NEM TALALOM

~

3. Az 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) ~AGI (ELSO,MASOD,HARM)

~AGI (12,23,65)

d) ~AGI (JAV1,JAV2,JAV3)

HIBA! SAJNOS >>JAV1 < 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 változtatása. Az egyes szeletek idejét a következő képlet határozza meg:

a szelet időtartama =

$$= \frac{\text{a szelet TM-értéke}}{10} \cdot \text{a szelet IDA-értéke.}$$

Ha az IDA utasításban az érték kisebb, mint 10, a szelet rövidebb ideig, ha az érték nagyobb, mint 10, a szelet hosszabb ideig hallható (mint a BEA utasításban meghatározott TM-érték.) Az n_1 , n_2 értékét lásd a LEO utasítás-

nál. Minden szelet definiálásakor a szelet IDA-értékét a gép automatikusan 10-ben határozza meg.

c) ~IDA (A'J\$,A'J\$)<5>

~IDA (#PA)<50>

~

d) ~IDA (#AJ)45

HIBA! 45 ERTEK ROSSZ !

~ICA (#AJ)<45>

HIBA! UT.-NEVE ROSSZ !

~

2. Az IMI utasítás

a) IMI <érték>

b) Az IMI utasítással lehet közvetett úton váltztatni, 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:

a szelet lejátszási időtartama =

$$= \text{IMI-érték} \cdot \frac{\text{a szelet TM-értéke}}{10} \cdot \text{a szelet IDA-értéke.}$$

A számítógépben mindaddig, amíg meg nem változtatják, az IMI-érték = 100.

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) ~IMI <25>

~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 (@@F,F,UZ,U,U,TT,T00,T,T@)

ZOL (#S0)@

d) ~ZOL (R,R123,RAA,AAA,AAN,AN,N,N*)

HIBA! R,R12 NEM DEF. SZELET !

~

2. A TOM utasítás

a) TOM <érték>

b) Az adatbázis tartalomjegyzékének listászerű kiírása. Az értékkel kell megadni, hogy az adatbázis hányadik elemétől hányadik eleméig történjék 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 <2-100>

d) ~TOM <2-100>

HIBA! UT.-NEVE ROSSZ !

~

```

.RUN DX1:FOPR01
*PAL AZ *E'S* HANGSOR ELD'A'LLI'TA'SA,
*PAL AZ ADATBA'ZISBO'L AZ E'100, E'200 SZELETEK FELHOZATALA A KO:ZPONTI TA'RBA,
*JON (E'100,E'200)
*PAL A KO:ZPONTI TA'R TARTA;\LMMA'NAK KII'RATA'SA
*VID
      JELEK      POZICIO      IDOALAP
      E'100      1- 6      10
      E'200      7- 13      10

OSSZEVONT JELEK

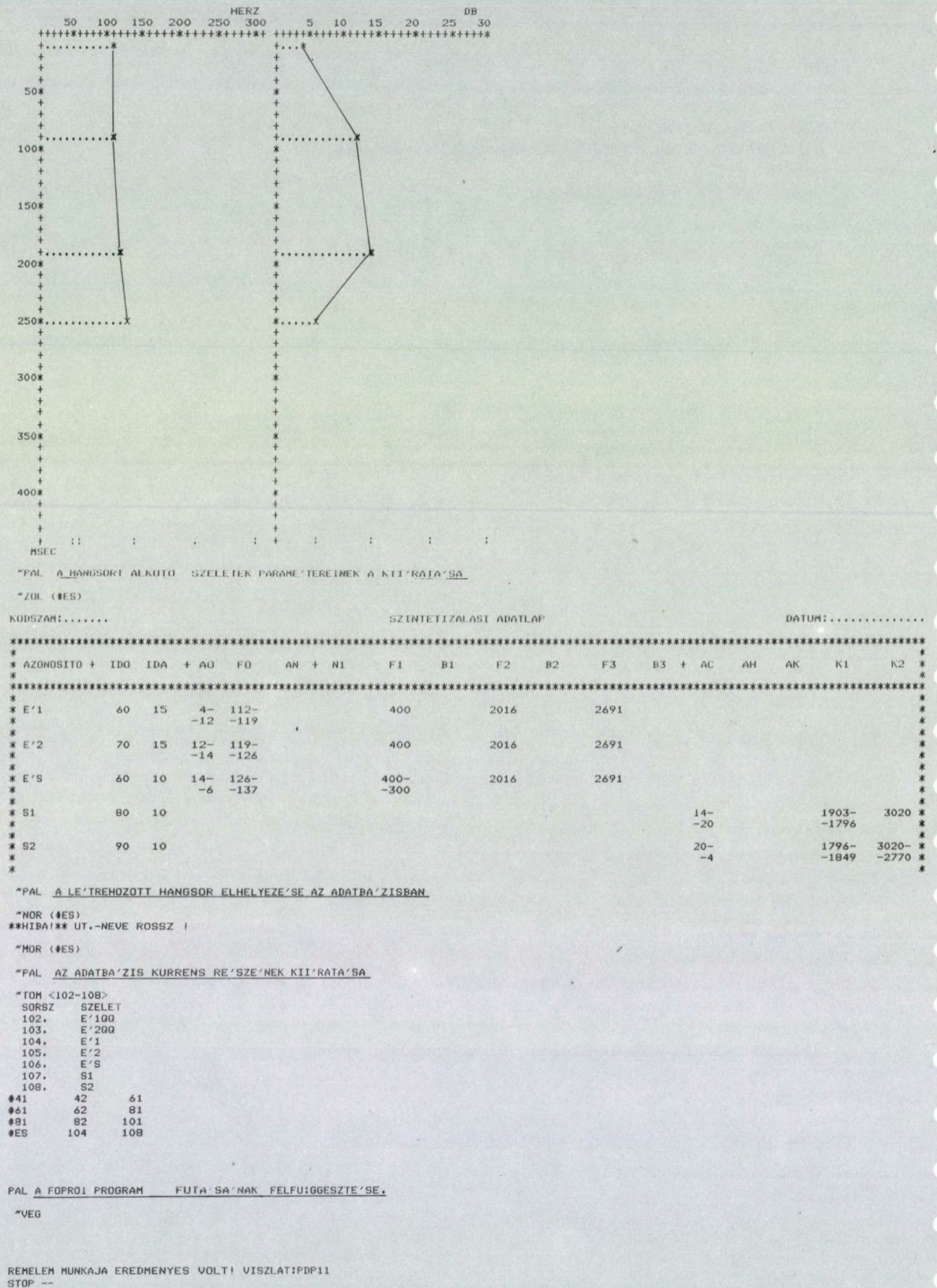
*PAL A FENTI KE'T SZELET A'TNEVEZE'SE
*EVA (E'100)E'1
*EVA (E'200)E'2
*PAL AZ *E'S* HANGSORT ALKOTO' TOVA'RBI SZELETEK LE'TREHÖZÄ'SA
*BEA E'S(TM:<60>,AD:<14-6>,F0:<100>,F1:<400-300>,F2:<2000>,F3:<2700>)
*BEA S1(TM:<80>,AC:<15-20>,K1:<1700-1800>,K2:<3000>)
*BEA S2(TM:<90>,AC:<20-5>,K1:<1805-1850>,K2:<3000-2800>)
*PAL AZ *E'S* HANGSORT ALKOTO' SZELETEK KO:ZU:S *ES* AZONOSI'TO' ALA VALO' O:SSZEVONA'SA
*ETI (E'1,E'2,E'S,S1,S2,ES
*PAL SZU:NETI,NAH,MA'E'HA SZELET LE'TREHÖZÄ'SA
*BEA PAUSA(TM:<500>)
*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

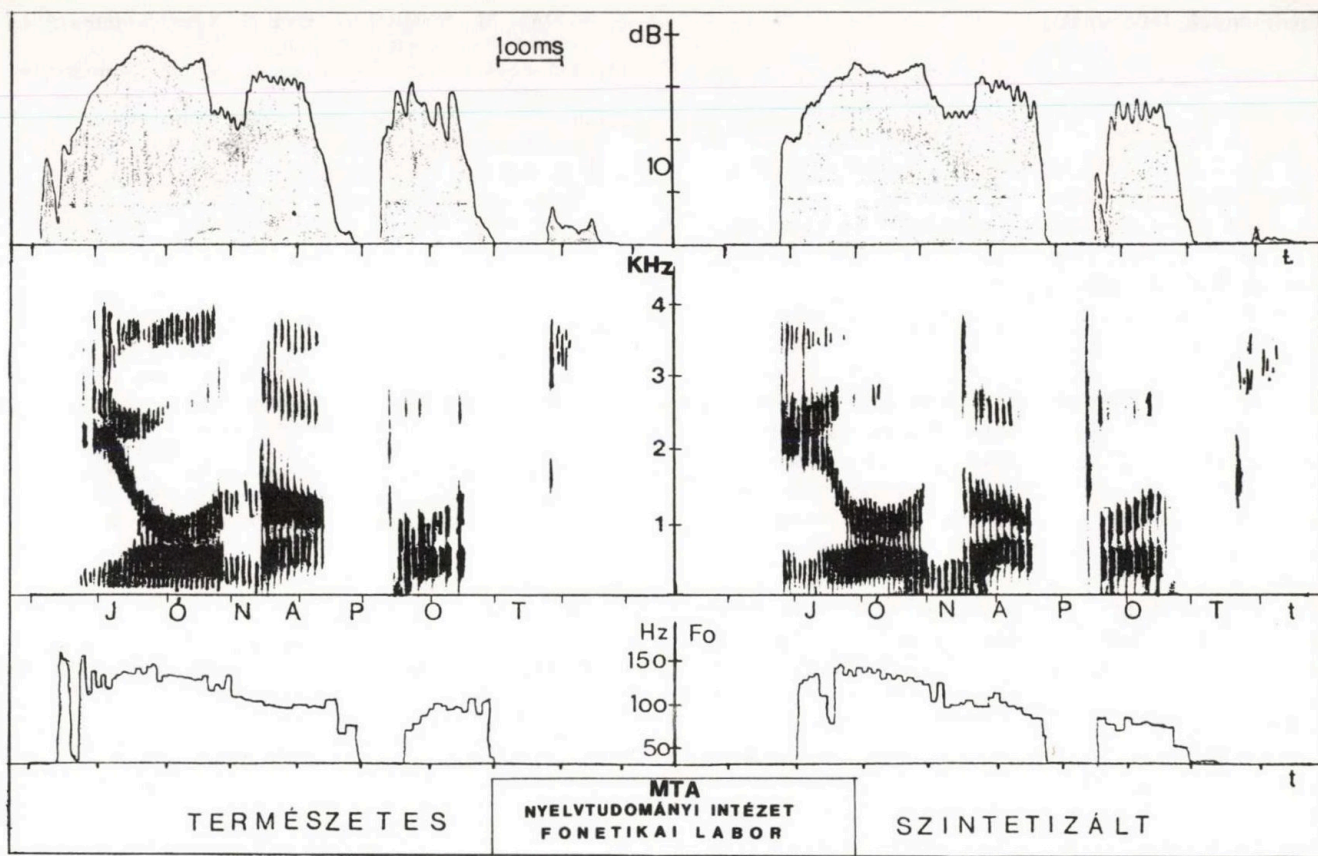
*PAL AZ E'S HANGSORRA KE'RDO' DALLAM FELVITELE
*BEA E'1(F0:<110-118>)
*BEA E'2(F0:<118-125>)
*BEA E'S(F0:<125-135>)
*PAL AZ ELKE'SZI'TETT HANGSOR HA'ROMSZORI LEJA'TSZA'SA,KO:ZBEN I SEC SZU:NETTEL
*LEO (#ES,#PP,#ES,#PP,#ES)
**HIBA!** #PP NEM DEF. ETI !
*LEO (#ES,#PA,#ES,#PA,#ES)
*PAL AZ E'1, E'2 SZELETEK 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' A;\LAIT VALO' LEJA'TSZA'SA
*IMI <500>
*LEO (#ES,#PA,#ES)
*IMI <100>
*PAL A HANGSOR ALAPHANG ( F0 ) E'S INTENZITA'S (AD) MENETE'NEK KIRAJZOLTATA'SA
*IVO (#ES)10
**HIBA!** 10 ERTEK ROSSZ !
*IVO (#ES)<10>

```

5. ábra. Az ÉS hangsor előállítása (folytatás a 352. oldalon)



5. ábra. Az ES hangsor előállítása (folytatás a 351. oldalról)



6. ábra. A természetes és szintetizált beszéd összehasonlítása

3. A PAL utasítás

a) PAL TETSZŐLEGES KARAKTERSOROZAT

b) A PAL utasítással megjegyzés, kommentár iktatható be a szintetizálásunk menetét dokumentáló leporellóra. A negyedik pozíción az utasítássorban itt is szünetnek kell állnia!

c) ~PAL AZ ADATOK AZ FD-116-OS LEMEZRE KERÜLTEK.

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 konzolrögépen. Az érték itt az időtengely léptékét adja meg, míg az N használata során az FO frekvenciátengelyén 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) ~IVO (1H,HH,HO,D,DO,OL,L1,L,L&)<10>

IVO (#IT,#OT)<10>N

d) ~IVO (11,12,13,14)N
HIBA!* N ERTEK.ROSSZ !

VII. Megállító utasítás

a) VEG

b) A VEG utasítás hatására a program futása befejeződik

c) ~VEG

REMELEM MUNKAJA EREDMENYES VOLT! VISZLAT:PDF11
STOP --

d) —

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

A rendszer használatának és működésének jobb megértése érdekében az 5. ábrán bemutatjuk az ES? hangsor előállításának menetét, az adatok betáplálását, mozgatóját, dokumentálását és a hangsor lejátszását.

Eredmények, felhasználás

A dolgozatunkban ismertetett INBERE az első olyan számítógépes szintetizáló rendszer Magyarországon, amellyel tetszőleges, jó minőségű magyar (illetőleg adatok birtokában bármilyen nyelvű) szavakat, mondatokat lehet előállítani (6. ábra).

Az MTA Nyelvtudományi Intézetének fonetikai osztályán ezzel a rendszerrel végeztük el többek között a magyar beszéd hangsorépítő elemeinek analizisét és szintézisét, amely alapját képezi egy automatikus, magyar nyelvű szintetizáló eszköz kidolgozásának. Eredményeink felhasználásával megtettük az első lépéseket a magyar beszéd automatikus és valós idejű szintéziséhez is. Ezzel a gép–ember kommunikációs kapcsolat legmodernebb formájának alapelveit próbáljuk kidolgozni a magyar nyelvre. Első kísérletként olyan számítógépprogramot állítottunk össze (KEPLET), amellyel a számítógép 250 hangszeletből és az ezek kombinációjából felépített 96 szórészből épít fel tetszőleges számokat 1-től 1 billióig. A program valós idejű üzemmódban dolgozik, ami azt jelenti, hogy a megadott utasítás vagy paran'cs után a számítógép azonnal kimondja a kért számot vagy számokat. Programunk kibővített változata a számokon kívül képes az alapvető számtani műveletek (összeadás, kivonás, szorzás, osztás) felhasználásával felépített képletek kimondására is. Ilyenkor a képlet kiszámítása után a számítógép hangosan közli a számítás végeredményét is. Programunkat tudományos konferencián és szakmai körökben is bemutattuk 1981-ben.

A KEPLET program tapasztalatait felhasználva, 1981-ben elkezdtük kidolgozni azt a rendszert, amely már nemcsak számokat, hanem tetszőleges – a magyar helyesírás szabályai szerinti betűkből összeállított – magyar szöveget is meg tud szólaltatni. Munkánk eredményeként 1982 májusában a 8. Budapesti Akusztikai Kollokviumon bemutattuk az UNIVOICE elnevezésű, magyar nyelvű, valós idejű „text to speech” szintetizáló rendszert.

Eredményeinket sikeresen alkalmazzák orvosi, logopédiai területen is. A beszédaudiometriai vizsgálat új formáját dolgoztuk ki a Heim Pál gyermekkórház gégeészeti osztályának munkatársaival. A halláskárosodás vizsgálatára az INBERE rendszerrel tesztanyagot állítottunk össze szintetizált szavakból, és ezt a gyakorlatban ki is próbáltuk. Az első eredmények nemcsak azt igazolták, hogy a szintetizált beszédet fel lehet használni a kisgyermek és felnőttek halláskárosodásának vizsgálatában, hanem felmerült az az újszerű gondolat is, hogy a mesterséges beszéd alkalmazásával új típusú szűrőteszt-rendszert lehetne kidolgozni a különböző halláskárosodások jobb behatárolására [6].

Az előbbi példákban látható, hogy a magyar beszéd szintézise terén végzett nyelvészeti-fonetikai, műszaki és számítástechnikai kutatásaink eredményeit néhány helyen már a gyakorlatban is alkalmazzák. Reméljük, hogy az alkalmazások területe a jövőben szélesedni fog.

Irodalom

- [1] AB FONEMA: OVE III a Speech Synthesizer Manual. Type 21001.
- [2] FANT, G.: Speech Sounds and Features. Massachusetts, 1973.

- [3] NYMAN, M.—NYMAN, G.: OVE III B-Puhsyntetisaattorin ohjausohjelman kayttoopas. Helsinki, 1978.
- [4] KISS G.—OLASZY G.: Mesterséges beszéd számítógéppel. Számítástechnika, 1980. X. XI. XII.
- [5] OLASZY G.: Hangsorok számítógépes formánsszintézisének előkészítése. Magyar Fonetikai Füzetek 8. 1981. 147–60.
- [6] FARKAS ZS.—HIRSCHBERG J.—GÖSY M.—OLASZY G.: On the Use of Synthesized Words in Audiology. Proc. of XVI-th International Congress of Audiology. 1982. Helsinki.

Резюме

В настоящей статье описывается структура и работа диалоговой системы синтеза речи ИНБЕРЕ. Система может применяться в основном теми организациями, которые занимаются анализом и исследованием акустики речи с целью развития технических средств для науки, промышленности и др. С помощью ИНБЕРЕ авторами были составлены те программы, которые впервые позволили ЭВМ произнести цифры и математические формулы на венгерском языке.

Summary

The structure and operation of the INBERE interactive speech-synthesizer is introduced in this article. This system is applicable at research organisations, making acoustic studies of speech in order to use them in industrial, scientific or other fields. Using INBERE author wrote those programs, which first enabled a computer to pronounce names of figures and mathematical formulas in Hungarian.

VLSI áramkörök uránszenyezésének kimutatása

1978 óta ismeretes „soft error” néven az a típusú hiba, amelyet a VLSI áramkörökben is különösen a nagy adatsűrűségű tárolóknál a nyomnyi mennyiségben jelen levő, természetes előfordulását radioaktív szennyezések okoznak. Az angol Atomenergia Kutató Intézet ezek kimutatására rendkívül nagy érzékenységu módszert fejlesztett ki. Az FTA (hasadásos autoradiográfia) elnevezésu technikával már 2 mg/tonna koncentrációjú urán jelenléte is észlelhető, így a mikroelektronikai ipar nyersanyagainak vagy féltermékeinek ezt a szennyezését az ellenőrzés időben jelezheti. A félvezetőgyártók előírásai szerint az alfa részecskék emissziója a tárolókhoz használt anyagoknál óránként maximálisan 0,001 részecske/cm² lehet. Az urán 235 izotópot felhasználó autoradiográfias eljárással ennek ötödrésze még kimutatható.

Electronic Design
1982. május 13., 52. oldal

Vertikális mágneslemez-írási technika

A japán kutatások nyomán kb. 1983-ra várható az új elvű, függőleges mágnesezéssel dolgozó mágneslemezek nagyban megvalósulása. Azáltal, hogy a hagyományos vízszintes-hosszanti rögzítéssel ellentétben az írás folyamán a mágnesezett területek sűrűségének növelésével a demagnetizált mező egyre kisebbedik, igen nagy adatsűrűség-növelés érhető el mind a hajlékony, mind a merev mágneslemezekben. Laboratóriumi termékeken már 80–100 ezer bit per inch sűrűségértékeket is sikerült elérni; a sávsűrűség inchenként 1000–2000-re növelhető. Állítólag 445 ezer bit/inch sűrűséget is elértek a Sony kutatói egy digitális-audio alkalmazásnál.

A nagyságrendnyivel megnövelt kapacitású lemezek piacáért már megkezdődött a verseny Európában és az USA-ban is, tekintve, hogy a mágneslemezek világpiaci forgalma a becslések szerint 1982-ben 12 milliárd, 1984-re pedig 18–19 milliárd dollárt ér el.

Electronics
1982. április 21., 89. oldal