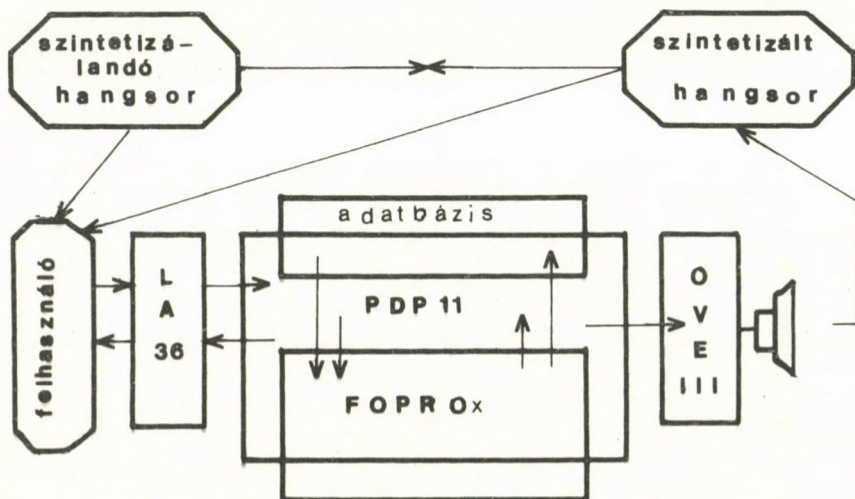


INTERAKTÍV BESZÉDSZINTETIZÁLÓ RENDSZER SZÁMÍTÓGÉPPEL ÉS OVE III SZINTETIZÁTORRAL

Kiss Gábor—Olaszy Gábor

Dolgozatunkban bemutatjuk az MTA Nyelvtudományi Intézete fonetikai osztályának laboratóriumában kidolgozott INTERaktív BESzéd-szintetizáló Rendszert (INBERE). A rendszerrel jó minőségű (nemcsak magyar) hangsorokat lehet előállítani, elsősorban fonetikai alapkutatások 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) beszéd-szintetizátorból áll. A számítógép konfigurációjából felhasználjuk a beszéd-szintetizálás során a 64 Kbyte-os központi memóriát, a két floppy disk egységet és az LA-36-os konzol írógépet. A beszéd-szintetizátor 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 beszéd-szintetizátort mutatjuk be.

Az OVE III beszéd-szintetizátor ismertetése

Az OVE III manuálisan és digitálisan vezérelhető analóg beszéd-szintetizátor. 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 beszédzintetizátor 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

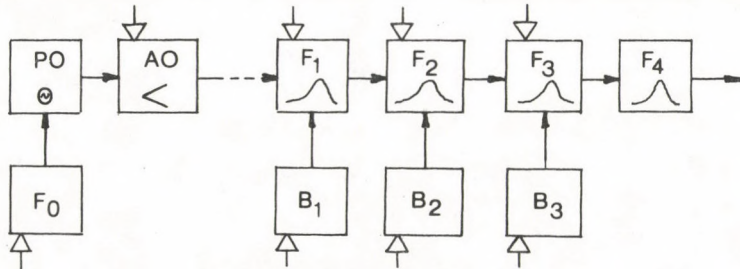
san külső kapcsolók, dinamikusan számítógép segítségével vezérelhető. (A készüléket Gunnar Fant és munkatársai tervezték, és a svéd FONEMA cég gyártja.)

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 következő 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

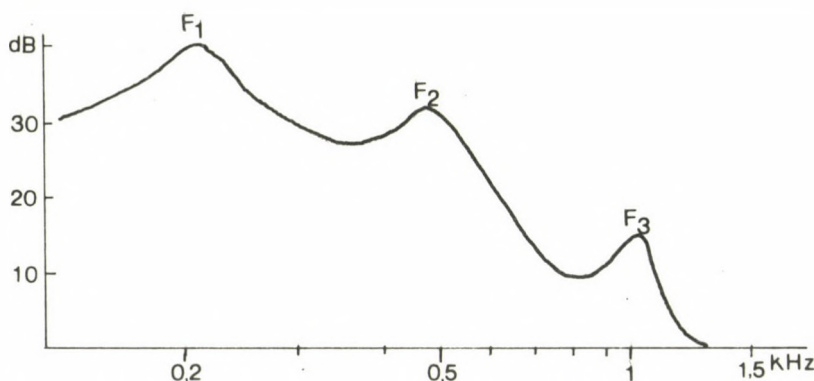
$B_1 = 12 - 188$ Hz-ig 12 Hz-es lépésemelkedéssel 16 lépésben;

$B_2 = 31\text{--}470\text{ Hz}$ -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).



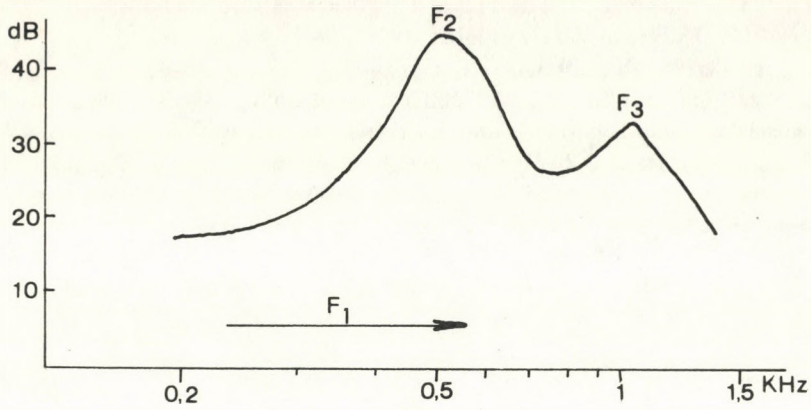
5. ábra

Az AO = 30 dB betáplálásával létrehozott szűrőkarakterisztika

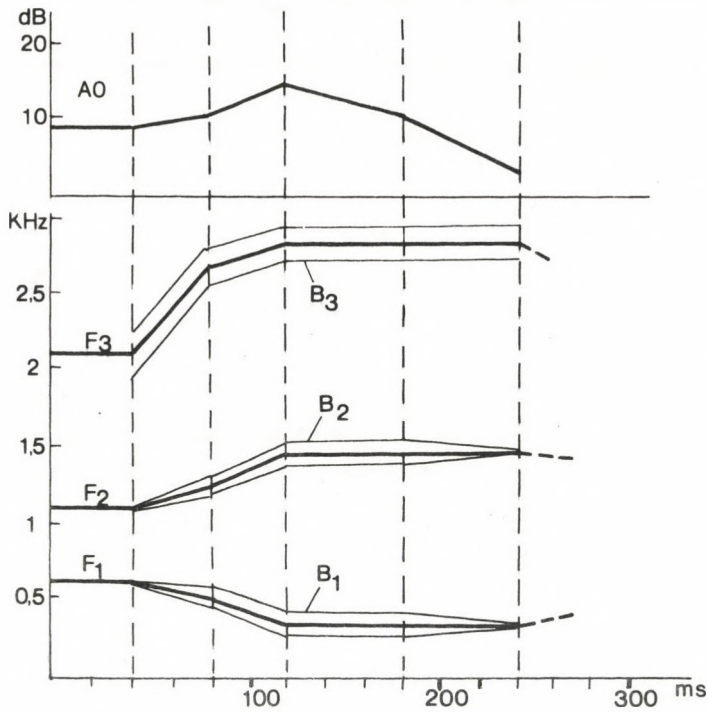
A görbén jelentkező csúcsok jelentik a formánsok legalacsonyabb frekvenciaértékeit. Az alapállásnál ezek a következők: $F_1 = 200\text{ Hz}$, $F_2 = 504\text{ Hz}$, $F_3 = 1008\text{ Hz}$. Ha az AO betáplálása után valamely zöngés hangra megállapított formánsértékeket is beadunk a szintetizátorba, akkor 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 hangszetek között (vö. Olasz 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

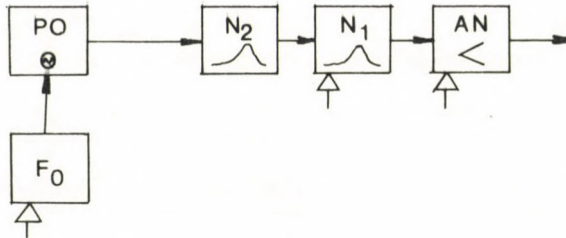


7. ábra
Az AO intenzitás, az F_1 , F_2 , F_3 és a B_1 , B_2 , B_3 mozgásának megtervezése a szintézishez

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).

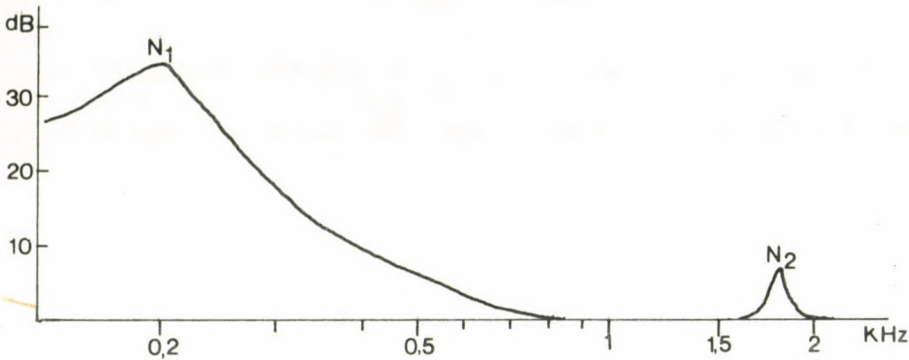


8. ábra

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).

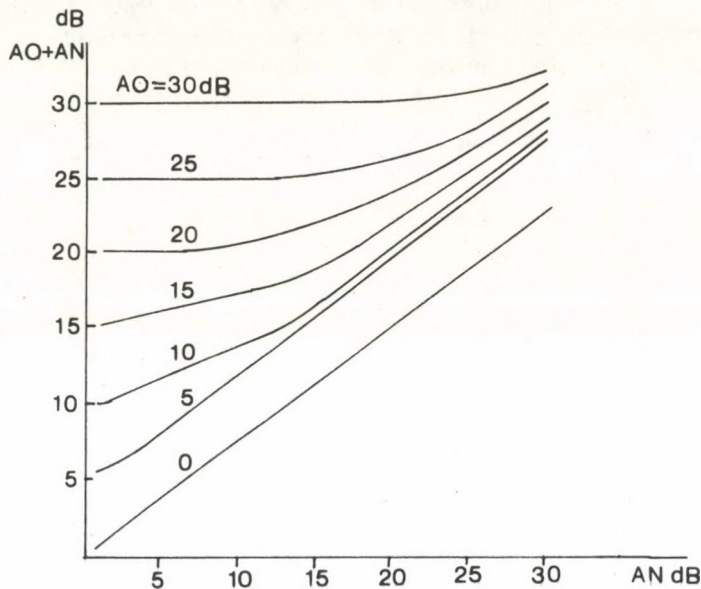


9. ábra

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.

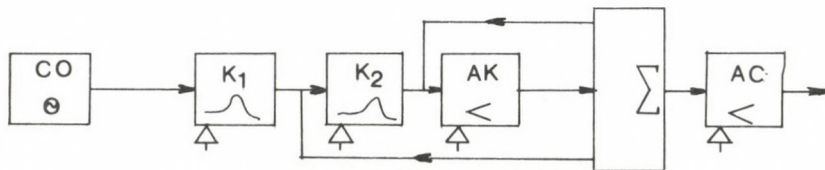


10. ábra

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 hangszegletekben 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

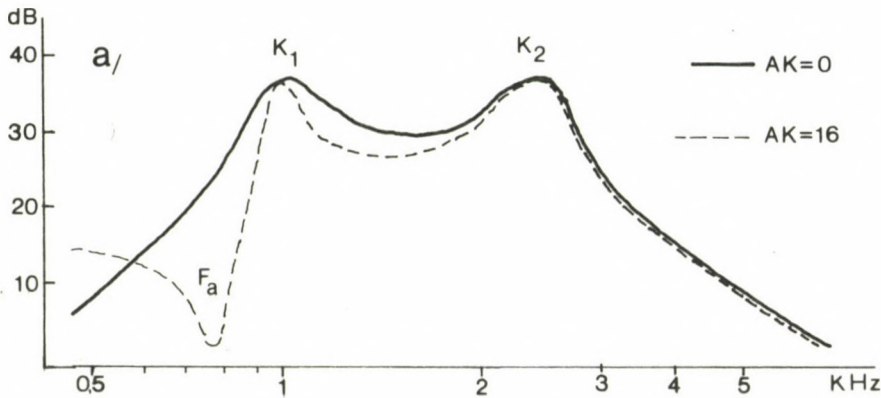
A szűrő erősítésének mértékét az AC változóval adhatjuk meg 30 dB-es átfogási sávban 2 dB-es lépésekben. A szűrő két pólusának helye változtatható, értéküket a következő határok között választhatjuk meg:

$K_1 = 1008-6218$ Hz-ig 3 %-os lépésemelkedéssel 64 lépésben;

$K_2 = 2540-15668$ 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 zörejhangoknak nincs komponensük, másrésztől segítségével lehet a zörejhangok frekvenciaszerkezetére jellemző szűrőkarakterisztikát létrehozni. Az AK erősítést illetően három helyzetet különböztetünk meg.

a) Ha az AK értéke < 20 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 zörejhangot 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 $K_1 = 1008$ Hz, a $K_2 = 2540$ 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 (dB) = $\frac{K_1}{K_2} \times 10 + AC$ (dB) képlettel számíthatjuk,

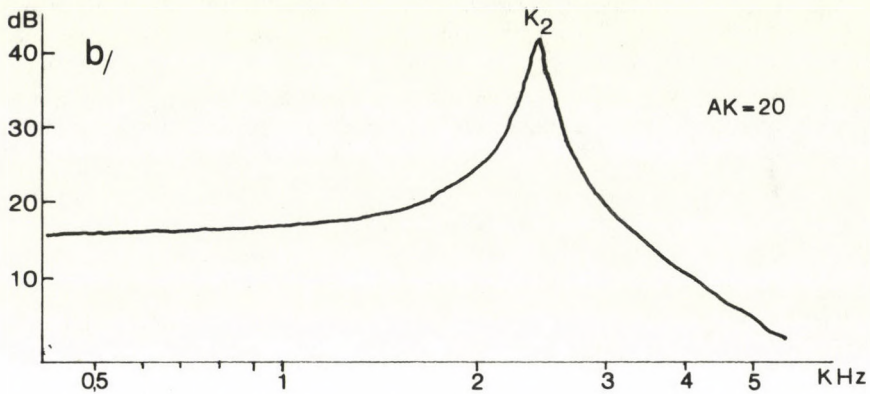
– ha a $K_1 = K_2$ -vel, akkor a kiemelés mértéke 10 dB,

– 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 dB értéknél eléri azt.

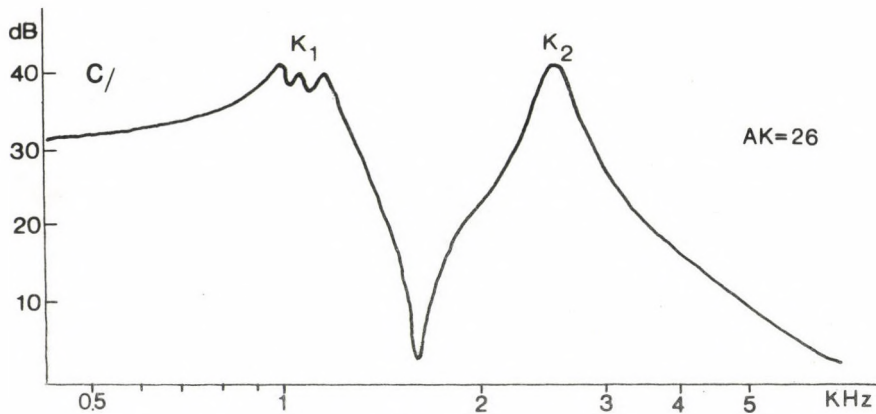
b) Ha az AK értéke egyenlő 20 dB-lel, akkor az F_a értéke megegyezik a K_1 értékével, így azt közömbösíti. Ekkor a szűrő karakterisztikájában csak a K_2 kiemelése játszik szerepet, az alsó sáv átvitele egyenletes (12.b ábra).



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

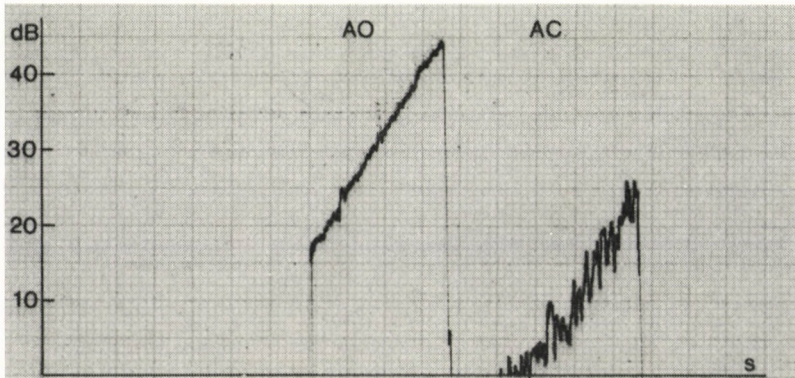
A zörejkör szűrőjének átviteli karakterisztikája különböző AK-értékekénél

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 lépésemelkessel 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	$AO+F_0+$
21.	767774			$+F_1+F_2$ 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	$AO+F_0+F_1+$
27.	767774			$+F_2+F_3$ 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 szintetizálásakor 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.

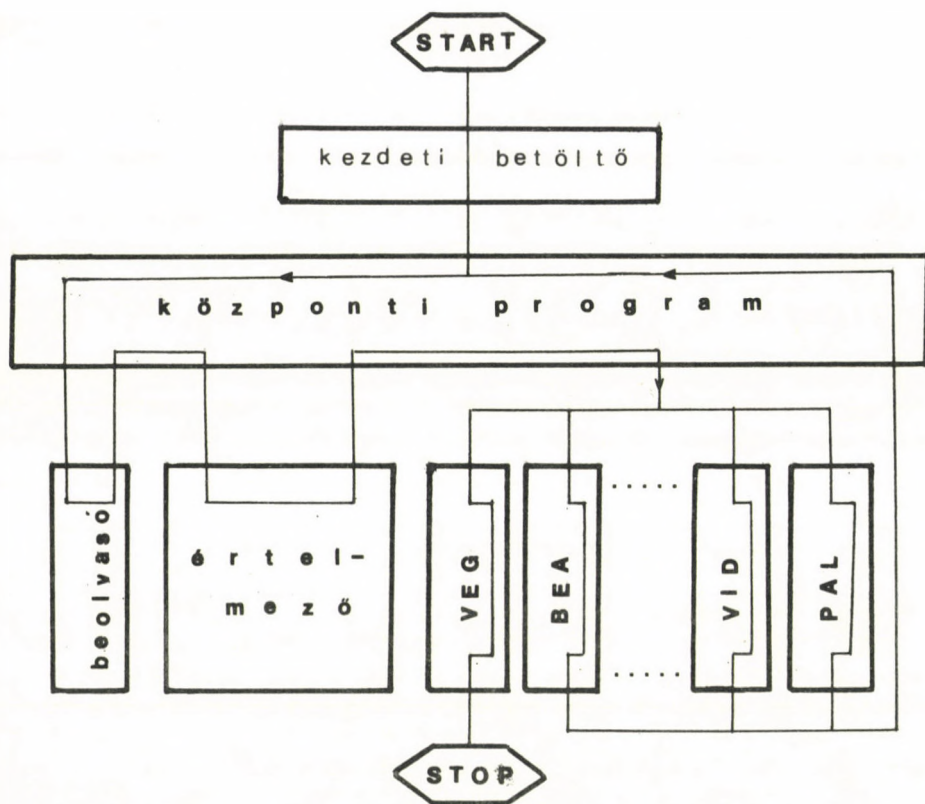
Software részek

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 előállításánál szükséges dokumentációs műveleteket végrehajtani. Az egyes kérések (utasítások) milyenségét a felhasználó a három karakteres utasításnévet 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 MFF 8. 1981) egy hang vagy hangrész, amelyben a paramétereknek nincs töréspontja, vagyis minden (a hangszeleteket alkotó aktív) paraméter értéke egy számmal vagy számpárral (tól-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 összeideje 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 hibaüzenetek.

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 másrésztől 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>,AD:<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' OISSZEVONA'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:TSZOROS IDO' ALATT VALO' LEJA'R\A'TSZA'SA

*IMI <500>

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

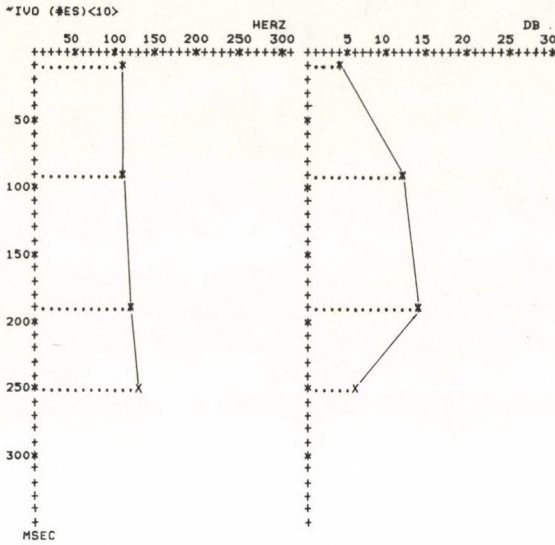
*IMI <100>

*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?' HANGSOROT 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- -12	112- -119		400		2016		2691						
* E'2	70	15		12- -14	119- -126		400		2016		2691						
* E'S	60	10		14- -6	126- -137		400- -300		2016		2691						
* S1	80	10											14- -20			1903- -1796	3020
* S2	90	10											20- -4			1796	3020

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

*NOR (#ES)

HIBAI 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:PDPI1

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.