

**KANDIDÁTUSI ÉRTEKEZÉS  
MŰSZAKI LEÍRÁS**

**A BRAILAB BESZÉLŐ SZÁMÍTÓGÉPCSALÁD**

**Írta:  
ARATÓ ANDRÁS**

**Budapest  
1992**

## TARTALOM

### BEVEZETÉS

#### 1. fejezet

#### ELŐZMÉNYEK

1.1. TAKTILIS KIJEZŐK VAKOKNAK

1.2. SZINTETIKUS BESZÉDDEL KIJEZŐ SEGÉDESZKÖZÖK

#### 2. fejezet.

#### KÖVETELMÉNYEK, CÉLKITŰZÉSEK

2.1. LÁTÁSSÉRÜLTEK KOGNITÍV FUNKCIÓI TAKTILIS ÉS BESZÉDKIJEZŐKKEL

2.2. A BESZÉDÉRTHETŐSÉG ÉS TERMÉSZETESSÉG DILEMMÁJA SEGÉDESZKÖZÖKBEN

2.3. ERGONÓMIAI ÉS BESZÉDTECHNOLÓGIAI KÖVETELMÉNYEK BESZÉDKIJEZŐS ESZKÖZÖKBEN

#### 3. fejezet.

#### A KUTATÓ-FEJLESZTŐ MUNKA ÉS MÓDSZEREI

#### 4. fejezet.

#### ÚJ MŰSZAKI TUDOMÁNYOS EREDMÉNYEK

#### 5. fejezet.

#### A MEGVALÓSÍTOTT TEXT-TO-SPEECH RENDSZEREK

5.1. ÁTLAPOLÁSOS HANG ADATBÁZISOK

5.1.1. A MEA-8000 HANG ADATBÁZISA

5.1.2. A PCF-8200 HANG ADATBÁZISA

5.2. A HANG ADATBÁZISOK FEJLESZTŐ RENDSZEREI

5.3. ASCII -> METABRAILLE -> DIÁD KONVERTÁLÓ PROGRAM

5.4. MIKROINTONÁCIÓ, INTONÁCIÓ, ÉNEK

#### 6. fejezet.

#### A BRAILAB SEGÉDESZKÖZÖK

6.1. A BRAILAB BASIC

6.1.1. A BRAILAB BASIC SEGÉDESZKÖZ FELÉPÍTÉSE

6.1.2. A TEXT-TO-SPEECH RENDSZER BEÉPÍTÉSE A BRAILAB BASIC-BE

6.1.3. AZ ELSŐ MAGYAR OLVASÓGÉP BELSŐ FELÉPÍTÉSE

6.2. A BRAILAB PLUS

6.2.1. A BRAILAB PLUS SEGÉDESZKÖZ FELÉPÍTÉSE

6.2.2. A TEXT-TO-SPEECH RENDSZER BEÉPÍTÉSE A BRAILAB PLUS-BA

6.2.3. A BRAILAB PLUS BRAILLE PUBLISHING RENDSZERE

6.3. A BRAILAB PC

6.3.1. A BRAILAB PC SEGÉDESZKÖZ FELÉPÍTÉSE

6.3.2. AZ ÚJ TEXT-TO-SPEECH RENDSZER BEÉPÍTÉSE AZ MSDOS OPERÁCIÓS RENDSZERBE

#### 7. fejezet.

#### A BRAILLE RÖVIDÍRÁS ÉS A TEXT-TO-SPEECH KIVÉTELSZÓTÁRA

#### 8. fejezet.

#### A BRAILAB GÉPEK ALKALMAZÁSAI

#### VÁLASZ

Gordos Géza, a műszaki tudomány kandidátusa bírálatára

## BEVEZETÉS

Ez a műszaki leírás nem gyártási dokumentációja a BraiLab beszélő számítógépeknek, hanem azzal a céllal készült, hogy a szerző (Arató András) lényeges közreműködésével, vakok számára elkészített informatikai segédeszközök - mint műszaki alkotás - alapján a műszaki tudományok kandidátusa tudományos fokozat elnyerésére pályázhasson. A leírásban a megfelelő helyen jelezni fogom, hogy mi volt a szerző saját munkája, és miben volt irányítója az eszközök létrehozásának, ill. milyen kész berendezést használt fel.

A berendezéseket és a működésükhöz szükséges szoftver modulokat blokkvázlat szintjén ismertetem. A programok teljes forrásnyelvi listáját nem mellékelem, melyek összességében kb. negyvenezer Z80, I8086 assembler és C programnyelvi sort tennének ki. A leírásban a hangsúlyt a megoldások újszerűségére és hasznosságának megvilágítására helyezem.

A mellékletekben azok az oktatási segédletek, kézikönyvek találhatók, melyeket elsősorban vakoknak és tanáraiknak írtam, s melyek tanácsot adnak ahhoz, hogyan kell az általam kidolgozott új módszereket a BraiLab és BraiLab Plus beszélő segédeszközökben használni. (A. B.) A mellékletek tartalmazzák még a kivétel szótár leírasi szabályrendszerét, valamint egy minta kivételstótárt a BraiLab PC külső kivételei és rövidítései számára. (C.)

A Pedagógiai Technológia című folyóiratba írt cikkem áttekintést nyújt azokról a számítástechnikai és mikroelektronikai segédeszközökről, melyek a vak emberek tanulását és munkáját könnyíthetik meg. (D.) A rehabilitációs tanácsadás céljából készült könyvben a vakoknak szóló számítástechnikai eszközöket használó szakmákat ismertetem fontos metodikai tanácsokkal szolgálva a pályaválasztó sérült fiataloknak, nevelőiknek és szüleiknek. (E.)

A Magyar Fonetikai Füzetek 15. számában a MEA-8000 szintetizátorra készített egyik fejlesztőrendszeremet ismertetem (F.), míg a 17. számban a BraiLab-ba beépített full text-to-speech rendszeremet. (G.) Ez utóbbi cikkem megjelent nagyrészt változatlan formában a XI. ICPhS fonetika tudományi konferencia anyagai között is. (H.)

A Magyar Elektronika folyóiratban megjelent ismertetőm a KFKI-ban végzett tudományos kutatásaimról és annak eredményeiről számolt be már 1987-ben. (I.) Az 1988-ban megrendezett „Beyond Number Crunching” című közös osztrák magyar tudományos konferencián előadást és bemutatót is tartottam a BraiLab Plus-ba épített magyar és német nyelvű rendszereimről. (J.) Az 1990-ben megrendezett „Computers for Handicapped Persons” nemzetközi konferencián a braille nyomtatórendszeremmel kapcsolatos tudományos és gyakorlati eredményeimet ismertettem. (K.) Ugyan ennek a nemzetközi konferenciának 1992-es anyagai közt fog megjelenni az a cikkem, mely áttekintést ad a speciálisan vakoknak tervezett beszéd-szintetizátoros rendszereimről, valamint a legújabb kutatásaim gyakorlati megvalósulásáról a BraiLab PC gépemben. (L.)

Az (M.) mellékletben található az a szabadalmi leírás, mely az „Eljárás és berendezés hang szintetizálására előnyösen beszéd és/vagy ének szótár nélküli, a vezérléssel teljes egészében azonos idejű előállítására” címet viseli.

Ezúton is szeretném megköszönni azok közreműködését, akik segítségemre voltak a BraiLab gépek létrehozásában: először is feleségemnek, munkatársamnak, Vaspöri Teréznek, aki a programozási feladatokban segített a legtöbbet, és egyben az eszközök első számú felhasználója is, mert teljesen vak.

Továbbá Kiss Gábornak és Olasz Gábornak, akik a szintetizátor hanganyagának kialakításához adtak értékes tanácsokat. Lukács Józsefnek és Endrének, Sulyán Jánosnak, valamint mindazoknak, akik aktívan segítettek a BraiLab-ok létrehozásában.

# 1. fejezet

## ELŐZMÉNYEK

### 1.1. TAKTILIS KIJELEZŐK VAKOKNAK

A lehetséges számítógépes kijelzők nagy csoportját alkotják a tapintható kijelzők. Ezek legnagyobb része a braille írást követi, de van olyan, mely a látó írás tapintására szolgál. Ez utóbbi talán egyetlen számítógépes segédeszközként is használható típusa az Optacon (OPTical to TActile CONverter), melyet az MIT egyetemen fejlesztettek ki az 1970-es évek elején.

Ennek az eszköznek 6-szor 24 rezgő tűskéje jeleníti meg mozgatható kamera segítségével a síkírás egy-egy betűjét, vagy annak egy nem teljes részét. Számítógépes terminálhoz külön optikai előtétet kell hozzá használni, bár újabb típusai, melyek a 80-as évek közepén jelentek meg, már tartalmazznak karaktergenerátort, hogy ascii kód szintjén csatlakozhassanak a DTE (Data Terminal Equipment) és DCE (Data Communication Equipment) eszközök közé. Ez utóbbi fejlesztés gyakorlatilag az ernyő előtétek használhatatlanságának elismerése.

Az Optacon igen nagy függetlenséget biztosít a vak felhasználónak papíron írt anyagok olvasásához, bár az elérhető olvasási sebesség nem haladja meg egy első-második osztályos látó kisiskolás olvasási tempóját. Ezt a sebességet is csak kitartó tanulás után éri el a felhasználó. Magyarországi megjelenésekor 1980-ban az ára kétszázezer deviza Ft volt, ma ötszázezer Ft.

A „soft” braille kijelzők, melyek ideiglenesen, mozgó tűskékkel jelzik ki a számítógép szolgáltatott információkat, két fő technológiai vonulatba tartoznak. Korábban az elektromágneses rendszerek jelentek meg kis modulokból összeállítható 20, 40, vagy két sorban összesen 80 braille cellával. Már az elektromágneses soft braille kijelzők is készültek 8 pontos változatban, áthidalva ezzel a 6 pontos szokványos braille írás és a 7 és 8 bites kódtáblák közötti mély szakadékot. Megjegyezzük, hogy a 8 pontos braille nemzetközi szabványosítási törekvései, melyben magam is részt vállaltam, még a mai napig nem jutottak nyugvópontra.

Az újabb soft braille kijelzők már piezo elektromos elvűek, de ezzel az áruk mégsem csökkent jelentősen, csupán a megbízhatóságuk lett jó. (1 db. egyponthoz piezo modul ára kb. 4-5 USD). A magas ár miatt gyakran csak 6 pontot használnak egy modulban. A nagybetűk jelzésére a tűskék 1-2 Hz-es rezgését használják. A 8 pontos változatban az egyik kiegészítő tűske a kis-nagy betű jelzésére szolgál, míg a másik a programozáshoz szükséges speciális ascii jelek jelölésére. Mindkét kiegészítő tűske együttes jelenléte a kontroll karaktereket jelenítheti meg.

A braille hardcopy készítésére a sornyomtató ill. mátrix nyomtató elvű domborítók, vagy egyszerűbben nevezve, braille nyomtatók szolgálnak. Először az átalakított braille írógépek voltak használatban számítógépes perifériaként (ezeket leginkább a telexgépekkel rokoníthatnánk), később azonban célfejlesztések eredményei terjedtek el felhasználók körében.

Legelső fejlesztésemben én is a soft braille kijelzés útján indultam el 1978-ban. A CÉDRUS (Conversational Editor and Remote User Support) kétféle terminál rendszerében a TPA/70 számítógép szolgált ki a VT340 típusú terminálokat. Vezetéssel ebbe a terminálba építettünk be egy olyan kártyát, mely a VT340 felfrissítő memóriájából mindig egy karaktert kijelzett 6 elektromágneses tűskén. A kártya tartalmazott egy fix, szűkített nagybetűs ascii-to-braille konverziót.

A TPA/70 terminál kezelésébe beépítettem egy lassító rutint, mely a kijelzés menet közbeni olvasását szolgálta volna. Ennek a megoldásnak a használhatóságáról a 2.1. fejezetben írok bővebben. Már ezt a rendszert is kiegészítettem olyan számítógépes kommunikációs lehetőséggel, mely alapvető célkitűzése kell hogy legyen egy segédeszköznek, s mely a további fejlesztéseket is elősegítette.

## 1.2. SZINTETIKUS BESZÉDDEL KIJELEZŐ SEGÉDESZKÖZÖK

A számítógépes kijelzők másik nagy csoportját a szintetikus beszéddel működő kijelzők alkotják. További két csoportra oszthatjuk ezeket a beszélő eszközöket a szintetizátor szótára alapján. A beszéd szintézis lehet kötött szótáras, vagy kötetlen szótáras rendszerű.

Céljaim megfogalmazása idején (a 80-as évek elején) cél segédeszközökben (pl. beszélő óra) működtek különböző nyelveken kötött szótáras beszéd szintetizátoros rendszerek. Ezek hatására készítettem el az SDK-85 (System Design Kit for I8085) alapján egy programozási segédeszközt.

Az SDK-85 az Intel 8085 mikroprocesszoros fejlesztő rendszere, melynek hexadecimális monitora, lépésenkénti program belövést, hardver töréspontot és egyéb elemi fejlesztési lehetőséget nyújtott látók számára. Ezt a hardvert egészítettem ki első lépésben egy soros kommunikációs interfésszel, további program memóriával és egy Digitalker nevű fix szavas, amerikai nyelvű beszéd szintetizátorral.

A fejlesztő rendszer monitorát úgy alakítottam át, hogy minden funkciót, melyet látó el tudott végezni, vak is megtehesse. Már ebben a rendszerben is ügyeltem a beszéd megszakíthatóságára, melyet itt csak egy külön interrupt szint felhasználásával tudtam elérni. A fix szavas rendszer, az SDK-85 funkcióinak használatát maradéktalanul lehetővé tette módosított monitorommal vakok számára is.

A 80-as évek elején kötetlen szavas rendszerek csak angol nyelvterületen készültek. Ezek ún. fonéma szintetizátorral működtek (SCO01). A fonéma szintetizátor beszédminősége gyenge, mert a hangkapcsolatokat, az elő és utóhatásokat nem tudja figyelembe venni. Mégis nagy előnyük, hogy kevés gyakorlással a hangok jól megtanulhatók, és így a munka során jól felismerhetők.

Magyar nyelv céljára az angol fonéma szintetizátor gyakorlatilag használhatatlan, mert a magyar magánhangzók és egyes mássalhangzók idegenül hatnak az angol fonémákkal (pl. hehezett p, t). A fonéma szintetizátor kétségtelen előnye a jobb ár/sebesség mutatójában van a soft braille kijelzőkkel szemben.

Később, angol nyelvterületeken, fokozatosan elterjedtek a formáns szintetizátorok. Ezek hangzása már természetesebb, hosszabb hallgatás után sem fárasztó használatuk. Az első többnyelvű formáns elvű rendszer, melyet segédeszközökben használtak a svéd Infovox volt. Magyar nyelvre csak ilyen formáns szintetizátor használata jöhetett szóba.

Az MTA Nyelvtudományi intézetében kifejlesztettek kötetlen szótáras, digitálisan vezérelt analóg szűrőkkel működő beszéd szintetizátoros rendszert. A program Fortran nyelven overlay szerkezettel betöltötte a teljes PDP11 típusú kisszámítógépet. A 80-as évek elején kezdtek meg az áttérést a digitálisan vezérelt, 4 digitális szűrőt tartalmazó integrált áramkörre (MEA-8000). A Budapesti Műszaki Egyetemen Dr. Gordos Géza vezetésével ekkor még csak kötetlen szótáras rendszerek fejlesztésére vállalkoztak. Akkori publikált nézetük szerint full real-time text-to-speech rendszer mikroprocesszoron nem, csak közepes és nagyobb teljesítményű kisszámítógépen valósítható meg.

Ez volt az egyik oka annak, hogy saját magam kezdtem beszédtechnológiai kutatásba és fejlesztésbe. A másik fő oka a külön text-to-speech rendszer kialakításának az a meggyőződés volt, hogy segédeszközbe speciális beszédtechnológiát kell beépíteni. E meggyőződés az évek során csak tovább erősödött.

Az SDK-85 fix szavas, angolul beszélő fejlesztő rendszeremet kiegészítettem braille klaviatúrával, a szövegek bevitele céljából (a soros interfész mellett), valamint a MEA-8000 áramkörrel. Párhuzamos fejlesztés céljából készítettem C64-es mikroszámítógépes MEA-8000-es fejlesztő rendszert is. Ezzel főleg a magyar formáns paraméterek kialakításának feladatát céloztuk meg. Ebben Kiss Gábor volt segítségemre.

A kezdeti, segédeszközökben alkalmazott text-to-speech rendszerek közös jellemzője volt, hogy külön processzoron futott a szintetizátor program, vagy annak adatbázisa dedikált, külön tárolóban volt. Ennek oka a nagy program méret és a beszéd prototípusok nagy terjedelme. Ezt a szeparáltságot mindenképpen el akartam kerülni az ár/teljesítmény viszony megjavítása és a szoros real-time kapcsolat érdekében.

## 2. fejezet. KÖVETELMÉNYEK, CÉLKITŰZÉSEK

### 2.1. LÁTÁSSÉRÜLTEK KOGNITÍV FUNKCIÓI TAKTILIS ÉS BESZÉDKIJELZŐKKEL

Amikor vakoknak és gyengénlátóknak készítünk segédeszközöket, akkor tekintettel kell lennünk ezeknek az embereknek bizonyos pszichés problémáira is. Ez természetesen nem jelenti azt, hogy ők más emberek, különleges csodabogarak! Feltétlenül hangsúlyoznunk kell azt, hogy inkább általános emberi tanulási és felhasználási problémákat kell látássérültek esetében is megoldani, nem pedig valami egészen speciálisakat.

Ellentétben a hiedelmekkel, vak emberek hallása és tapintása fiziológiailag nem jobb mint a látóké. Ezen érzékszervek funkciói viszont mások a megismerés folyamatában. A látó tapintással csak kiegészítő információkat szerez, a vak haptikus érzékelésének alapvető kognitív funkciói vannak.

Felvetődik a kérdés, hogy mi fontosabb a látássérült számára: a hallás vagy a tapintás? Milyen képet lehet alkotni egy szöveg tüköréről tapintás és hallás útján? Egy elvont logikai program struktúra megértéséhez hogyan segíthet egy tapintható blokkdiagram?

A szövegszerkesztőkkel létrehozott mértani alakzatok viszonylag egyszerűek, így ezekről akár hallás útján is kialakíthatók megfelelő képzetek! Ehhez azonban a mesterséges beszédet előállító rendszerbe megfelelő támogatást biztosító üzemmódokat kell beépíteni. Monoton intonálással jelölhetjük a sorok végét, és a szünetek hossza utalhat a szavak egymás közötti távolságára.

A hangmagasság és a szünetek hossza analóg értékek, melyeket az ember nagyon jól tud kezelni. Egyéb hang effektusok (furcsa beszéd, nagyobb hangerő) kiegészíthetik, gazdagíthatják a kétdimenziós formák felismeréséhez a jelzéseket. A braille hardcopy haptikus érzékelése viszont elengedhetetlen a hallás útján végzett interaktív munka teljessé tételéhez.

Dr. Pálhegyi Ferenc kísérleti vizsgálatait és kutatásait alapján az állapítható meg, hogy a vak emberek tapintás útján szerzett formaérzékelésében a legnagyobb gondot a globális formák érzékelése jelenti. Más kísérletek alapján is kiderült, hogy vak, gyengénlátó és látó vizsgálati személy által végzett tapintási tesztekben, különösen két kézzel történő vizsgálódáskor, a látók és gyengénlátók ill. a később megvakultak előnyösebb helyzetben vannak a született vakokkal szemben. (Természetesen minden vizsgálati személy szeme a kísérlet alatt be volt kötve).

A globális formák felismeréséhez a beszéd nagyon jól használható. Ki lehet mondani a beszéd szintetizátorral, hogy a fájl melyik részén vagyunk éppen, hanyadik sor, hanyadik oszlopában található a cursor, mennyivel beljebb íródtak a bekezdések. A vak emberek metrikus megismerési elvét ezek a közölt számértékek nagyon jól segítik.

A beszéd kijelzős segédeszközök használatakor megfigyeltem egy érdekes pszichikai jelenséget (megjegyzem, hogy ezt látóknál is tapasztaltam, amikor beszélő számítógépet adtam a kezükbe). Kezdetben a felhasználó mindig kivárja a mesterséges beszéd végét, addig nem nyúl a következő billentyűhöz, amíg az beszél. Látók esetében ez a korlát feszültséggel párosul, mert az ernyőről sokkal hamarabb leolvassa az üzeneteket. Vakok esetében hamarosan megjelenik az az igény, hogy gyorsabban dolgozhassanak, mint ahogy a gép beszél. A legfontosabb, hogy ezt az igényt minél teljesebb körűen kielégítsük a segédeszközben.

Tekintsünk egy másik kísérletsorozatot, melyben a közvetlen és megtartó emlékezet lélektanát vizsgálták. Ebben a vak kísérleti személyek rendre jobb eredményeket értek el, mint a látók és alig látók. Fontos eredménye a vizsgálatnak, hogy a megtartó emlékezetben a tapintásnak fontosabb szerep jut, mint a hallásnak.

Az egykarakteres braille kijelző és az egy síkírásos karaktert megjelenítő (Optacon) eszközökben külön gondot jelent, hogy a taktilo-motorikus észlelés tere be van szűkítve. Itt a haptikus érzékelés mozgás elve korlátozódik.

Hosszabb braille sor kijelzése valamit javít a helyzeten. Készültek olyan kísérleti berendezések az elmúlt években, ahol a braille sor kézzel elmozdítható fel és le irányban, így igyekeztek megközelíteni, hogy egy braille oldal olvasásának képzetét nyújthassák. A túl magas ár miatt egyedi kivitelben található teljes braille lapot kijelző táblák (Stuttgarteri egyetem 1985). Egykarakteres braille kijelző beszéddel kombinálva sorozatban is készült (Braillescope Belgium 1990).

Összefoglalva, mind a tapintásra, mind pedig a beszédkielzésre egyaránt szükség van, ezek mellett leginkább a látássérült emlékezetére számíthatunk. A fokozottabb térérzékelés szükségessége és a jobb ár/teljesítmény miatt, interaktív munka céljára a szintetikus beszéd mellett döntöttem, braille nyomtató kiegészítéssel. Egy központilag használt braille nyomtató több felhasználónak nyújtja előnyeit, míg a beszédkiemenettel ellátott számítógépes segédeszköz az egyén céljait szolgálja.

## 2.2. A BESZÉDÉRTHETŐSÉG ÉS TERMÉSZETESSÉG DILEMMÁJA SEGÉDESZKÖZÖKBEN

A folyamatos beszédelőállító rendszerek megítélésében két paraméter játszik fő szerepet: az egyik a beszéd érthetősége, a másik pedig a beszéd természetessége. Ha segédeszközbe szánt mesterséges beszédre van szó, akkor a beszéd érthetőségét kell előtérbe helyezni. E két paraméter kielégítéséhez sajnos sokszor egymásnak ellentmondó megoldásokon keresztül vezet az út.

Vizsgáljuk meg először a beszéd érthetőségének követelményét. Kutatásom tárgyát képező speciális ember-gép kapcsolati esetben nemcsak a szöveg, mondat és szó érthetőségének van kiemelt jelentősége, de a látszólag szükségtelen logatomban és hang érthetőségekre is meglehetősen nagy figyelmet kell fordítani. Gondoljunk csak arra, hogy a beszéd sok esetben az egyetlen támasz a szövegek szerkesztése, javítása közben.

A beszéd természetessége helyett inkább a mesterséges hang fárasztóságának fogalmát vezetném be. Hosszabb szövegek meghallgatásakor a természetesség előny, hiszen a szöveg érthetőségét is elősegíti a természetes intonálás és hangsúlyozás. (Itt egyelőre csak a szöveg szintaktikai vizsgálatára gondolok és nem annak szemantikai analízisére. Az adott alacsony kategóriás processzorok kapacitása mellett ilyen szemantikai elemzés real-time elvégzése egyelőre nem lehetséges.)

A beszéd érthetőségének és természetességének növelése sokszor ellentmondó követelményeket támaszt a rendszerek kialakítása során, de néha e két tulajdonság erősítheti is egymást. Elsőként nézzük a beszéd tempójának kérdését.

Ha a leghosszabb időtartamú szeletekből állítjuk össze a paramétereket, akkor tág határok között gyorsíthatjuk a beszédet. Ha a sebességet vezérlő paramétert a leggyorsabbra állítjuk, akkor a hangszeletek időtartamának legrövidebbre váltásával el lehetne érni még gyorsabb beszédtempót is. Ez segédeszközben nagyon fontos szempont. Ugyanakkor a természetesség csorbát szenved. A rövid hangszeletekkel finomabb mikrointonálási struktúra hozható létre.



A tempó és a mikrointonáció kérdései után következnek a hangok és logatomok érthetőségének problémái. Felvetődik a kérdés: mikor van szükség arra, hogy szavaknál kisebb egységek érthetősége számítsa? A szöveg szerkesztése közben olyan részei szólhatnak meg a szavaknak, melyek önmagukban értelmetlen hangsorok. Gépeléskor a billentyűk ehózására a hangalakok szolgálhatnak.

A billentyűk ehózása olyan fontos ember-gép kapcsolati követelmény, hogy megérdemli a kutató különleges figyelmét. Magyarországon én vezettem be azt a billentyű echózási módszert, melyet az imént a hangalakkal történő megjelenítésnek neveztem. Sokszor láthatunk még ma is, más, viszonylag fonetikus nyelveken beszélő segédeszközökben olyan megoldásokat, hogy a betű nevét mondják ki a megfelelő billentyű leütésére.

Ez azért hibás megoldás, mert a gyors gépelést akadályozza. Szinte kényszeríti a felhasználót arra, hogy minden alkalommal kivárja a betű megnevezését. Ez abban az esetben, ha a beszéd meg sem szakítható, kifejezetten idegesítő a munkában. Ha a billentyűk echózásának azt a módszerét választják, hogy csak a szavak végén szólal meg már a teljes szó, akkor elvész az egyidejűség, a folyamatos kontroll megnyugtató érzése. Bizonyos nyelveknél (pl. angol) persze az utólagos szó echó tűnik csak járható útnak ld. Braille 'n Speak.

Vegyük a magyar nyelv példáját, és tekintsük az eL, eM, eN és eF betű megnevezésekkel történő meghangosítását. Ha a felhasználó kellő gyorsasággal gépel, és a beszéd megszakíthatósága is adott, akkor csak az E hang hallatszik. Ez teljesen félrevezető.

Még továbbra is áll a kérdés, hogy milyen hangparamétereket szóaltassunk meg a text-to-speech programunkban? Egy lehetséges megoldás a szókezdő és szóvégi paraméterek összekapcsolásából adódik. Ehhez azonban ezeknek a szó eleji és szó végi hangoknak a megtervezésében az ilyen jellegű felhasználást figyelembe kell venni! A beszéd hangzásában itt is kompromisszumok megkötésére kényszerülünk. Egy másik megoldás a külön paraméterek tárolása a hangalakok echózásának céljára. Ez többlet memória használattal jár.

Fontos feladat a kis és nagybetűk megkülönböztetése. Gondoljunk csak a C nyelvben történő programozásra. Mivel a braille írás eléggé elhanyagolja a nagybetűk jelölését, ezért egy segédeszközben, mely a látók világával köti össze a vak felhasználót, erre a kérdésre nagy súlyt kell fektetni. Másoknál sok megoldást láthatunk a nagybetűk kijelzésére. Leggyakrabban kiegészítő hanghatást alkalmaznak (pl. J. Frank TASO). Én egy olyan, eddig nem alkalmazott, természetes megoldást választottam, mely plusz időt nem igényel, és mégis nagyon hatásos. A nagybetűk hangosan szólnak.

Ez a látszólag egyszerű megoldás nagyon sok problémához vezet olyan formáns szintetizátor esetén, ahol nem normálható könnyen a jel/zaj viszony (MEA-8000). Itt zajos beszéddel kell fizetni ezért az ember-gép kapcsolati megoldásért. A PCF-8200-nál a DAC faktor ezt a problémát megoldja.

A logatom és szóérthetőség kérdéséhez tartozik a javító parancsok új echózási módja. A szövegszerkesztő használata közben a betű és szójavítások során részhangok, szavak szólalnak meg, melyek összeolvadva a javító hangokkal, végül a javítás befejezéseként helyes szó hangzást eredményeznek. Ez olyan természetes beszéd általi szövegszerkesztést tesz lehetővé, mely a felhasználók körében szinte nem is tudatosul, csak más, ilyen módszert nem alkalmazó eszközök használatakor tűnik fel hiánya. Ahhoz, hogy ezek a hang összeolvadások jó hatást érjenek el, a beszéd paramétereinek alapos „összecsiszolására” van szükség, nem beszélve a szintetizátor real-time programozásáról.

A szavak érthetőségének vizsgálata után elértünk a mondatok és szövegek érthetőségéhez ill. a beszéd fárasztóságához. A részletdús hanganyag nagyban javítja a beszéd természetességét, csökkenti a fárasztóságát. A formáns szerkezetben tárolt „mikrointonálást” a PCF-8200-nál érdemesebb megvalósítani, mert az 5 programozható szűrőt tartalmaz. Nem a frame-eket kell tekinteni a beszéd alapegységeinek, hanem teljes szakaszokat. Ennek a tömör tárolási módnak a megvalósítását az 5.1.2. pontban írtam le.

Az automatikus intonálás megvalósításánál újra szembe kerülnek a gyors érthetőség, az alkalmazhatóság követelményei a beszéd természetességének követelményével. Olyan intonálási görbéket kell alkalmazni, melyek már a mondat elején utalnak a mondat végi jelre. Hogy mégse legyen fárasztó a mesterséges beszéd használata, azt az intonálási görbék nagyobb számával lehet elérni.

A beszéd érthetőségét úgy lehet fokozni, hogy még karakterisztikusabbá formáljuk az egyes hangok paramétereit, vagy erősen artikulált beszéd mintákból indulunk ki. Előnyösebb segédeszközben a túl artikulált beszéd, mint a köznapi ejtés, ahol sok részletet elharapnak. Itt is győzzön inkább az érthetőség.

A jelöletlen hasonulásokat a köznapi nyelvben jelöljük. Segédeszközben mindenképpen szükséges, hogy az eredeti, a hasonulásban résztvevő hangok is meg tudjanak szólalni. Így egyértelműbb a hang alapján történő szövegszerkesztés.

### 2.3. ERGONÓMIAI ÉS BESZÉDTECHNOLÓGIAI KÖVETELMÉNYEK BESZÉDKIJELZŐS ESZKÖZÖKBEN

Ha eltekintenénk attól a tényről, hogy látássérülteknek kell hardver-szoftver eszközt készíteni, akkor elég jól megfogalmazhatók lennének azok az általános követelmények, melyeknek meg kellene felelni egy jól használható berendezésnek. Ezek a megállapítások szinte pontról pontra alkalmazhatók lennének a vakügyi segédeszközökre is. Ennek fényében két fő irányzat figyelhető meg a világban.

Az egyik szerint nem kell külön látássérültek számára berendezéseket készíteni, mivel ez az integrációjuk ellen is hatna. Ők kisebbségben vannak a társadalomban, ezért soha nem érhető el, hogy a látóknak fejlesztett minden program és berendezés vakos változata is elkészüljön. Csak olyan kiegészítő adapterre van szükség, mely valamennyire hozzáférhetővé teszi az információt vakoknak és gyengénlátóknak.

A másik irányzat szerint speciális berendezéseket kell fejleszteni és készíteni látássérülteknek. Még ha széria gépen is dolgoznak, külön számukra kifejlesztett programokat kell kezükbe adni. Az alkalmazói programokat célorientáltan kell megírni.

Én mindkét nézetet egyaránt osztom, bár különbséget kell tenni a felhasználók és a felhasználás körében. Ha kezdőknek adunk segédeszközt, akik kevés számítógépes ismerettel rendelkeznek, akkor célberendezés és célprogram szükséges. Ugyan ez igaz akkor, ha mindennapi felhasználásra, papír ceruza helyett használt berendezésről van szó.

Az első irányzat mellett kell érvelnem akkor, ha munkarehabilitációs célokat szolgál a segédeszköz. Az integrációt ebben az esetben a nem speciális gépeket és programokat használni tudó kiegészítés támogatja a legjobban. A beszédkimenet elé állított célok egyformák mindkét típusú berendezés esetében. Ezek a következők:

- kialakítani egy olyan hang adatbázist, mely alkalmas hallás útján történő interaktív szövegszerkesztésre és programozásra, de hosszabb használat mellett se legyen fárasztó.

- olyan magyar text-to-speech programot létrehozni, mely teljesen valós idejű, akár 8 bites mikroprocesszoron is működőképes, a gyors interaktív munkát nem akadályozza.

A megalkotott kötetlen szótáras mesterséges beszédet úgy kell beépíteni informatikai segédeszközökbe, hogy az mind gyakorlott, mind pedig kezdő felhasználó igényeit kielégítse. A kialakított segédeszköz:

- legyen olcsó, hordozható, lehetőleg hálózatifüggetlen is. Kövesse a felhasználója maximális munkatempóját.

- tartalmazzon olyan braille szövegszerkesztő lehetőséget is, melynek segítségével a látó és braille írást elég egyszer rögzíteni és kezelni egy közös „laboratóriumban”.

- a magyar braille kis és nagy rövidírást automatikus fordító program segítse.

- adjon lehetőséget a számítógépes kommunikációra (beszélő terminál program), mert a számítógépes hálózatok az oktatás és munkarehabilitáció egyik legfontosabb közege.

A fenti céloknak saját fejlesztésű, két különböző típusú formáns szintetizátorra kifejlesztett hang adatbázis, és két mikroprocesszorra készített magyar és német text-to-speech program felelt meg. A cél segédeszközök iránt támasztott követelményeknek a BraiLab és a BraiLab Plusz gépeim felelnek meg. Az általános gépek és programok használatát az IBM PC-re kifejlesztett BraiLab PC adapterem és szoftverem segíti a legjobban.

### 3. fejezet.

## A KUTATÓ-FEJLESZTŐ MUNKA ÉS MÓDSZEREI

A kötetlen szótáras mesterséges beszédelőállító rendszer hanganyagának kidolgozásában az ún. diádos módszert használtam és fejlesztettem tovább. Elsődleges szempontnak tartottam a beszéd érthetőségét, és kevésbé fontosnak a beszéd természetességét.

A BraiLab gépcsalád első és második tagjánál, melyek mikroprocesszorai 8 bitesek, s melyekben még a kis méret is elsődleges szempont volt, a tömör tárolás érdekében figyelembe vettem az egyes beszédhangok fonetikai tulajdonságainak rokon voltát. A beszédképző szervekben hasonlóan képzett mássalhangzókat csoportosítottam, hogy az általam kifejlesztett átlapolási előkeresési móddal a beszéd frame prototípusok még tömörebben legyenek tárolhatók.

Percepciós és statisztikai módszerrel ellenőriztem az érthetőséget, és ahol szükséges volt, több, a környező hangok hatását jobban kifejező beszéd frame-et tároltam. A MEA-8000 formáns szintetizátorra épült rendszereknél természetesen ejtett szavak szonogramjaiból indultam ki. A BraiLab PC-ben használt PCF-8200 esetében erősen artikuláltan ejtett szavakat analizáltam.

A magyar braille rövidírás fordító rendszerének, valamint a full text-to-speech kivétel-szótárának megtervezéséhez Jelinek Schneider, változó hosszúságú blokkok tömör kódolására kifejlesztett módszerét használtam. Segítségével vizsgáltam magyar nyelvi mintán az irodalmi nyelv entrópiáját, ami segítette a két kivételszótár optimális kialakításában.

Az ember-gép kapcsolati kérdések elemzésénél abból indultam ki, hogy a sérült ember épen maradt képességei területén erős kompenzáció alakul ki. Ezek közül vakoknál, a beszéd-kimenet esetében a hallást és az emlékezést kell kiemelni. A tanulás módját tekintve két alaptípust feltételeztem. Az egyik előre elsajátítja a parancsok összes lehetséges változatát, és csak ezután kezdi hallás útján szerezni a tapasztalatait. A másik, folyamatosan tanul, sokkal többet hibázva eközben. Mindkettőjük számára kielégítő módszert és rendszert kellett alkotni.

Az évek során a vakok oktatása alapján nagyon sok gyakorlati tapasztalatom gyűlt össze a segédeszköz legjobb ergonómiai kialakításához. A legcélravezetőbb tervezési módszernek a beleélés bizonyult. Az egyes programok használata közben az ernyőt nézve elképzeltem a látássérült felhasználó reakcióit, hibás lépéseinek lehetséges kijelzéseit, korrekcióit, és ezeket a megfigyeléseket visszacsatolva, a beszédkieljesítés és az operációs rendszerbe történő beépítés tökéletesítésére használtam fel.

Egy évtizedes kutató-fejlesztő munkám során, az általam kidolgozott módszereket alkalmazva, három BraiLab típusú számítógép készült sorozatban. Az első (1985-ben), a HomeLab típusú gép MEA-8000 szintetizátoros kiegészítő blokkal alkotja a BraiLab alapgépet. Ennek készült, egy gyártásba ugyan nem került soros vonallal történő kiegészítője (ld. 6.1.3.) is. Beégetett BASIC interpreteréhez, input/output rendszeréhez, Assembleréhez és monitorához illesztettem a beszédkieljesítést úgy, hogy a gépet vakon lehessen használni. Ez az alapgép főleg oktatási célokat szolgál.

1987-ben útmutatásaim alapján készült el a BraiLab Plusz nevű gép hardver kártyája. Mind a BraiLab alapgép, mind pedig a BraiLab Plusz Z80 mikroprocesszort és MEA-8000 formáns szintetizátort tartalmaz. Mindkettőn az általam, a beszélő SDK-85 segítségével kifejlesztett magyar nyelvű real-time full text-to-speech programrendszer működik. A BraiLab Plusz CP/M operációs rendszerébe olyan kiegészítéseket terveztem, melyek a fent említett célokat kielégítik. Beszélő szövegszerkesztőt, terminálkezelőt, adatbáziskezelőt alakítottam ki. Ez a gép már valós munkarehabilitációs elvárásoknak is megfelel.

A BraiLab Plusz-ra alapozva elkészítettem, a magyar braille rövidírást is támogató fordító és nyomtató rendszert. A programot számos braille sornyomtatóra és nyomdagépre is illesztettem. A kivételszótárba a magyar nyelv entrópia vizsgálatának eredményeit is beleépítettem.

A BraiLab PC rendszer 1990-ben készült el. Ennek hardver része egy IBM PC-re illesztett adapter, mely PCF-8200 típusú CMOS formáns beszédszintetizátort tartalmaz. Így az eszköz már nemcsak hordozható, de megfelelő PC-vel teljesen hálózatifüggetlen segédeszközként funkcionál.

Az új mesterséges beszédelőállító áramkör számára teljesen új elvű, kötetlen szótáras programrendszert készítettem. A nagyobb tárolókapacitás lehetővé tette a korábbi 4 Kbájtos beszédparaméterek 20 Kbájtosra növelését. Ezzel a beszéd minőségét lényegesen meg tudtam javítani. A hanganyagot a korábbi fejlesztő rendszerek kibővítésével fejlesztettem ki, és már erősen artikulált beszédmintákat alkalmaztam.

A full text-to-speech rendszert kivételszótárral kiegészítve beintegráltam az MSDOS operációs rendszerbe. Az ember-gép kapcsolati szempontokat össze kellett egyeztetnem az IBM PC gazdag szoftverellátottságával, így egy általános célú intelligens ernyő felolvasó rendszert készítettem. Ez a segédeszköz, bár nem célszámítógépen fut, könnyen használható látás-sérültek által.

## 4. fejezet. ÚJ MŰSZAKI TUDOMÁNYOS EREDMÉNYEK

1. Módszert dolgoztam ki, mely lehetővé tette, hogy elkészítsek Magyarországon elsőként egy teljesen real time full text-to-speech rendszert mikroprocesszorra, bebizonyítva ezzel, hogy ilyen rendszert nemcsak komolyabb teljesítményű számítógépen lehet megalkotni.
2. Kialakítottam egy olyan eljárást beszédparaméterek tárolása céljára, melynek segítségével nemcsak nagyon kis méretű beszéd frame prototípus tárat építettem, de igen jó minőségű, nagy részletességű adattárat is viszonylag kis helyen.
3. Kidolgoztam egy a magyar (és német) braille írást követő Metabrilie kódolási eljárást, mellyel a text-to-speech, és a braille nyomtató rendszert integráltan és hatékonyan működő egységben hoztam létre.
4. Kidolgoztam a hallás útján történő interaktív számítógépes munkavégzés módszereit látássérültek számára, és elkészítettem az első, teljesen on-line elvű beszéd echóval megvalósított mikroszámítógép alapú segédeszközt. Ez volt egyben Magyarország első beszélő személyi számítógépe.
5. A Recognita karakterfelismerő programmal elsőként valósítottam meg egy teljesen automatikus, síkírást felolvasó rendszert a BraiLab Basic számítógépem terminálemulátora segítségével.
6. Kidolgoztam és elkészítettem az első, és máig egyetlen automatikus braille rövidítő fordító és szerkesztő rendszert, mely az ASCII szöveget az általam bevezetett Metabrilie kódra fordítja, és így lehetővé teszi a braille írású nyomtatandó anyagok dokumentum szerkesztését mind vakok (hallás útján), mind pedig látók számára.
7. Elkészítettem egy bármilyen magyar éneket automatikusan eléneklő programot, mely a megadott szöveg, ritmus és Kodály Zoltán relatív szolmizációjával leírt dallam alapján működik.
8. Kidolgoztam egy kivétel szótár kezelési eljárást, melynek segítségével nemcsak a braille rövidírás kivételei kezelhetők effektíven, de a text-to-speech program kiejtésbeli kivételei is ezzel a módszerrel real time kezelhetők.

## 5. fejezet.

# A MEGVALÓSÍTOTT TEXT-TO-SPEECH RENDSZEREK

### 5.1. ÁTLAPOLÁSOS HANG ADATBÁZISOK

#### 5.1.1. A MEA-8000 HANG ADATBÁZISA

Ahhoz, hogy a text-to-speech rendszerekben bármilyen szöveg konvertálható legyen beszéddé, szükség van valamilyen hang alapegységek tárolására. A szomszédos hangok egymásra hatással vannak, ezért nem elegendő egy nyelv alap hangjait eltárolni. A leggyakoribb megoldás egy kétdimenziós alaphang mátrix összeállítása, amely minden hang-hang átmenetet úgynevezett diádot tartalmaz.

Én is ezt az utat választottam, de én nem fix számú hang alapelemből (frame-ből) indultam ki, melyeknek kódjait a diád mátrix tartalmazza, hanem nálam a mátrix egy összetett pointert tárol csupán. A kód tárolás azzal a hátránnyal jár, hogy a mátrix mérete megnő, mert legalább 6 kódszámot kell tartalmaznia minden diád számára. A másik hátrány abból származik, hogy viszonylag kevés frame áll rendelkezésre pl. 256, ha a kódszám egy bájtos, ami a beszéd finomságát korlátozza. A harmadik hátrány a nehéz kezelhetőség. Sokszor fonetikailag teljesen idegen hangátmenetekben kell azonos frame-eket használni, mely nehezen áttekinthető.

A pointeres megoldásom előnye az igen egyszerű kezelhetőségen túl kettős. Szükség esetén nagyon tömör adatbázis is létrehozható, mint ez a MEA-8000 első magyar adatbázisánál látható, de ha minden pointer külön beszéd szakaszokra mutat, akkor a hanganyag nagyon gazdag lehet, és nem korlátoz a frame kód kicsi értéke. (A frame kód nem könnyen növelhető, mivel kétbájtos frame kód érték használatakor a diád mátrix mérete

$$\text{hangszám} * \text{hangszám} * 2$$

értékkel nő, ami már igen tetemes. A pointeres megoldás teszi lehetővé az átlapolásos frame szakasz tárolást. Ez még a gazdag hanganyag kialakításnál is jelentős megtakarítással jár.

Az első magyar hang adatbázisom, mely MEA-8000 típusú formáns szintetizátorra készült, normális hétköznapi ejtésű szavakon alapult. Kiválasztottam egy jellegzetes hangzású mássalhangzót, melyet szókezdő és szóvégi diádjaival együtt szintetizálva nagyon jól fel lehetett ismerni. A szonogramokból nyert formáns értékeket gyakran kellett sarkítani, természetes ellenes irányba, hogy a rokon hangok (pl. b, d, g) a fent leírt módon, könnyen megkülönböztethetők legyenek. Ez persze azt vontam maga után, hogy a hangzás sokszor természet ellenes lett, de kis tanulással viszont hallás útján biztos felismerést tett lehetővé.

A nagyfokú tömörítés érdekében a hangátmeneteket elnagyoltam. Nulla amplitúdójú, hosszan hangzó frame-eket alkalmaztam, és bíztam a MEA-8000 lineáris interpolálási funkciójában, mely a korábbi szűrőértékeket csak folyamatosan elsimítva változtatja meg. A beszéd természetessége ellen hatott a MEA-8000-nél a frame hosszak lineáris lépcsőzete. A PCF-8200-nál ez már logaritmikusra változott.

Az átlapolásnál a következő elveket érvényesítettem: a mássalhangzó-magánhangzó diádokban nem tároltam a mássalhangzó jellegzetes formánsait, csupán annak átmeneti részeit képző, gyakran nulla amplitúdójú frame-eket. Ezek a formáns értékekben és hosszban a különböző nagy osztályokba sorolt mássalhangzóknál eltértek egymástól, de végül a magánhangzó tiszta fázisában már teljesen azonosak voltak.

Az alábbi példákban a frame-ek beosztása a következő:

```

+-----+
|   FD  AMPL  PI   F1   B1   F2   B2   F3   B3   B4   |
+-----+

```

FD=frame duration msec; AMPL=amplitude; PI=pitch increment;

F1, F2, F3 = formáns frekvenciák; B1, B2, B3, B4 = formáns

sávszélességek. Az Fn Bn értékek Hz-ben adottak. PI=16=zaj.

```

+-----+ A k-a diád címkéje. 5 frame-et tartalmaz
| $K.A:   ;5 | az átmenet. Egy frame-el többet, mint a
+-----+ nagy többség.

```

```

+-----+
| #    16    0    0    784  125  1110  125  2400  309  125 |
+-----+

```

```

+-----+
| $SP.A:   ;4 |           |           |           |           |
| $A.A:    ;4 |           |           |           |           |
| $AA.A:   ;4 |           |           |           |           |
| $E.A:    ;4 |           |           |           |           |
| $EE.A:   ;4 |           |           |           |           |
| $I.A:    ;4 |           |           |           |           |
| $O.A:    ;4 |           |           |           |           |
| $OE.A:   ;4 |           |           |           |           |
| $U.A:    ;4 |           |           |           |           |
| $UE.A:   ;4 |           |           |           |           |
| $C.A:    ;4 |           |           |           |           |
| $CS.A:   ;4 |           |           |           |           |
| $J.A:    ;4 |           |           |           |           |
| $L.A:    ;4 |           |           |           |           |
| $M.A:    ;4 |           |           |           |           |
| $N.A:    ;4 |           |           |           |           |
| $P.A:    ;4 |           |           |           |           |
| $S.A:    ;4 |           |           |           |           |
| $SZ.A:   ;4 |           |           |           |           |
| $T.A:    ;4 |           |           |           |           |
- $V.A:   ;4 - Ez az első frame a négyből||-
| $Z.A:    ;4 |           |           |           |           |
| $ZS.A:   ;4 |           |           |           |           |
+-----+

```

Csak a k-a átmenethez tartozó frame.

A hangzás alapján összevont hangátmenetek címkéi. Mindegyik azonosan 4 frame-et tartalmaz.

```

+-----+
| #    32    0    0    554  50   988  50   2400  50   50 |
+-----+

```

```

+-----+
| $B.A:   ;3 |           |           |           |           |
| $F.A:   ;3 |           |           |           |           |
| $G.A:   ;3 |           |           |           |           |
| $GY.A:  ;3 |           |           |           |           |
- $H.A:   ;3 - Ezek az átmenetek csak 3 frame-ből
| $NY.A:  ;3 |           |           |           |           |
| $R.A:   ;3 |           |           |           |           |
| $TY.A:  ;3 |           |           |           |           |
+-----+

```

A d-a diád nincs semmivel összevonva, de az a tiszta fázisát meg kellett ismételni.

```

+-----+
| #    32   177  0    554  50   988  50   2400  50   50 |
| #    16   250  0    554  50   988  50   2400  125  50 |
| #    16   250  0    554  50   988  50   2400  50   50 |
+-----+

```

```

+-----+
| $D.A:   ;4 |           |           |           |           |
| #    16   125  0    391  125  1337  50   2842  125  125 |
| #    32   250  0    554  50   988  50   2400  125  50 |
| #    16   250  0    554  50   988  50   2400  50   50 |
| #    16   250  0    554  50   988  50   2400  50   50 |
+-----+

```



```

+-----+
| $A.K:      ;4 |
| $AA.K:     ;4 |
| $E.K:      ;4 | A magánhangzó-k és mássalhangzó-k átmenetek
| $EE.K:     ;4 | gyakorlatilag teljesen közösek.
| $I.K:      ;4 |
| $O.K:      ;4 |
| $OE.K:     ;4 |
| $U.K:      ;4 |
| $UE.K:     ;4 | +-----+
| $B.K:      ;4 |
| $C.K:      ;4 |
| $CS.K:     ;4 |
| $D.K:      ;4 |
| $F.K:      ;4 |
| $G.K:      ;4 |
| $GY.K:     ;4 |
| $H.K:      ;4 |
| $J.K:      ;4 |
| $K.K:      ;4 |
| $L.K:      ;4 |
| $M.K:      ;4 |
| $N.K:      ;4 |
| $NY.K:     ;4 |
| $P.K:      ;4 |
| $R.K:      ;4 |
| $S.K:      ;4 |
- $SZ.K:    ;4 - Csak a szókezdő k erőteljesebb a
- $T.K:      ;4 - hangalakkal történő echózás
| $TY.K:     ;4 | miatt is.
| $V.K:      ;4 |
| $Z.K:      ;4 |
- $ZS.K:    ;4 - A szóvégi k -val együtt alakul
| | | ki a hangalakkal történő echó|||
+-----+
| $SP.K:     ;5 | +-----+
+-----+

+-----+
| #   16   0   0   250  50  1179  50  2400  50  50 |
| #   16   0   16  1047 309  1428 125  2400 309 125 |
| #   32  31  16  1047 309  1428 125  1761 125 125 |
| #   16  31  16  1047 309  1428 125  1761 125 125 |
+-----+
| #   16  62  16  1047 309  1428 125  1761 125 125 |
+-----+
+-----+
| $K.SP:     ;2 | +-----+
| #   16  31  16  1047 309  1428 125  1761 125 125 | +--+
| #   16   0  16  1047 309  1428 125  2400 309 125 |
+-----+

```

A magánhangzó-mássalhangzó átmenetek tartalmazzák, sokszor csak egyetlen példányban, az adott mássalhangzóra jellemző hanganyagot. Ha az érthetőség megkövetelte, könnyen létre lehetett hozni új szakaszokat az adott diád számára. A példákban szereplő frame értékeket, címkéket egyszerű szövegszerkesztővel lehetett szerkeszteni, áttekinteni, rendezni. Az 5.2. pontban leírt fejlesztő rendszerek biztosították az automatikus adatbázis módosítást.

A leírtakból kitűnik, hogy a magyar nyelv jelöletlen hasonulásait a MEA-8000-es rendszerben nem vettem figyelembe az adatbázis kialakításakor. Ez a hallás útján történő szövegszerkesztést egyáltalán nem zavarta, sőt elősegítette. Ezért a PCF-8200-as hang adatbázisánál, ahol már tekintettel voltam a jelöletlen hasonulásokra, bevezettem egy speciális üzemmódot, hogy visszatérhessünk a korábbi jelöletlen kijetésre is.

A példából az is látszik, hogy nem használtam legrövidebb hosszúságú, 8 msec időtartamú frame-eket. Ennek az az oka, hogy a MEA-8000 szintetizátorban nem volt lehetőség a sebesség állítására. Ez pedig a vakügyi segédeszközökben elengedhetetlen. Ezért alkalmaztam, ugyancsak először Magyarországon, hogy a mesterséges beszéd sebességét a frame hosszak felezésével gyorsítottam. A lassítás a frame-ek hosszának duplázásával, vagy ha már erre nincs tovább mód, akkor maguknak a frame-eknek a duplázásával érhető el.

A lassítás fogalmával rokon módon kezeltem a hosszítást. A magánhangzókat úgy alakítottam ki, hogy a tiszta fázisukban szereplő utolsó két frame idejének duplázásával a hosszú párjuk érzetét adják. A mássalhangzókat hosszításuk szempontjából két csoportra osztottam. Az első csoportba tartoztak az ún. zárfelpattanásos hangok (c, cs, p, t, ty, k), míg a másik csoportba az összes többi (b, d, f, g, gy, h, j, l, m, n, ny, s, sz, z, zs). Az (r) kivételt képezett, és jobb híján, a magánhangzók szabálya szerint hosszabbodott.

A zárfelpattanásos mássalhangzók hosszítását a nulla amplitúdójú frame-ek lassításával értem el. A többi mássalhangzót úgy hosszítottam, hogy négy hangzó frame-jüket lassítottam. Látszik, hogy a lassítás és a hosszítás együttesen kezelődnek, ezzel is megrövidítve a feldolgozás időigényét.

Az első változatú magyar hang adatbázist később Olasz Gábor segítségével természetesebbé tettük úgy, hogy néhány átmenetnek külön frame sorozatot alakítottunk ki. Ugyancsak az ő segítségével dolgoztunk ki MEA-8000-re német nyelvű adatbázist is. Ebben a mássalhangzók hosszítása természetesen nem szerepelt.

Mind a magyar, mind pedig a német ének generálásához külön hang adatbázist kellett készítenem a MEA-8000 számára. Ennek részletesebb tárgyalása az 5.4. pontban található. A magyar hang adatbázis mérete 4 Kbájt(!) alatt maradt, a német e fölé nyúlt.

Az amplitúdók meghatározásánál külön figyelembe kellett venni azt a teljesen egyéni megoldásomat, hogy a nagybetűket hangosabban szintetizálom. A MEA-8000-nél sajnos ezért nagy árat kellett fizetni, mert így a digital analog konverter erős zajjal működik. Ez a PCF-8200-nál szerencsére kiküszöbölhető volt.

### *5.1.2. A PCF-8200 HANG ADATBÁZISA*

A PCF-8200 típusú formáns szintetizátor elődjéhez képest egy további formánssal gazdagodott, és mind az 5 szűrő frekvencia és sáv paraméterei szabályozhatók. A Philips cég ez újabb termékére már érdemesebb volt jó minőségű beszéd adatokat kifejleszteni. Ezért az átlapolásos tárolási módszeremet továbbfejleszttem. A MEA-8000-nél csak egy 8 frame-es szakasz tárolására volt lehetőség. Az újabban kialakított rendszeremben két szakasz 8 ill. 7 frame hosszúságú szűrő sorozat paramétert lehet egy-egy diád számára lefoglalni.

Szakítottam a korábbi frame hosszúságokkal is. A PCF-8200-as rendszeremben szinte kizárólag alap hosszúságú, azaz a legrövidebb frame-eket használom. Így egy hangátmenet tárolására 15, apró részleteket is tartalmazó frame-et használok. Ez igen gazdag, részletdús beszédet eredményez, mely még monotonon sem fárasztó, mivel szerintem a szó tágabb értelmében nem is monoton! (ld. még 5.4.).

A rendszer továbbra is átlapolásos maradt, így a duplára nőtt pointer mátrix és a hangzó anyag együttes hossza mégsem éri el a 20 Kbájtot. Az átlapolás technikáját itt inkább teljes szakaszokra, és nem a szakaszokon belül alkalmaztam.

A hangmintákat saját hangomból készítettem. A szavakat ill. hangzókapcsolatokat erősen artikulálva ejtettem. A szonogramok kialakításánál olyan időegységeket választottam, hogy egy-egy hangátmenetben minél több apró formáns részletet tudjak tárolni. Ez valamelyest a természetesség ellen hat, de a beszéd fárasztóságát nem befolyásolja (ld. 2.2.).

A beszéd sebességét így csak a PCF-8200-as általános beszédsebességi paramétereivel lehet szabályozni, de ez a gyakorlatban elegendőnek bizonyult.

A hangátmenetek szelelelésénél a következőképpen jártam el: a mássalhangzó-magánhangzó diádban tároltam a mássalhangzó nullától eltérő amplitúdójú részeit, a magánhangzóba áthajló speciális átmenetekkel együtt, de még a magánhangzó tiszta fázisát is. Ha erre nem volt elegendő a 15 legrövidebb hosszúságú frame, akkor összevontam néhány, a lineárisához közel változó frame-et hosszabb időtartamú frame-ekbe. A magánhangzók tiszta fázisa mindig külön szakaszba tartozik.

Az alábbi példákban a PCF-8200 frame-jeiben a számértékek csak kódok és nem Hz értékek vagy msec-ok.

```
+-----+
|  FD AMPL      PI    F1 B1  F2 B2  F3 B3  F4 B4  F5 B5  |
+-----+
```

FD = frame duration; AMPL = amplitude; PI = pitch increment;

F1, F2, F3, F4, F5 = formáns frekvenciák; B1, B2, B3, B4, B5

= sáv szélességek. PI = 16 = zaj.

```
+-----+
- $K.A: ;7 +| A k-a diád első szakasza, benne a k zárfel-
+-----+ pattanása és a (k-a)-ra jellemző rész. -
```

```
+-----+
| # 0 4 16 25 1 17 3 5 1 5 1 1 3 |
| # 0 7 16 23 1 19 2 4 1 4 3 1 3 | +-----+
| # 0 9 16 20 2 15 2 4 2 4 3 1 1 |
| # 0 11 0 18 3 14 4 4 2 4 2 0 1 |
| # 0 10 0 19 2 14 3 4 2 3 2 0 1 |
| # 0 10 0 20 2 13 2 4 2 3 3 0 2 | +-----+
| # 0 10 0 20 2 13 1 4 2 3 3 0 2 |
+-----+
```

```
+-----+
- $SP.A: ;4 +| A szókezdő (a) 2. szakasza. Az első szakasz
+-----+ itt nem látható.
```

```
+-----+
| $A.A: ;4 +- Az a-a átmenet csak egy szakaszból áll.
+-----+ Amplitúdói kiegyenlítettek.
```

```
+-----+
| $B.A: ;4 |
| $C.A: ;4 |
| $CS.A: ;4 |
| $D.A: ;4 |
| $F.A: ;4 | +-----+
| $G.A: ;4 |
| $GY.A: ;4 |
| $H.A: ;4 |
| $J.A: ;4 |
+-----+
```

```
| $K.A: ;4 +-- A k-a átmenet második szakasza.
```

```
+-----+
| $L.A: ;4 |
| $M.A: ;4 |
| $N.A: ;4 |
| $NY.A: ;4 |
+-----+
```

```

| $P.A: ;4 +-- A mássalhangzó-a átmenetek második----+
| $R.A: ;4 | szakasza mind összevonható. Ez az
| $S.A: ;4 | (a) hang tiszta fázisa. Az Fn Bn
| $SZ.A: ;4 | finom változásaiban látható a mikro-
| $T.A: ;4 | intonálás.
| $TY.A: ;4
| $V.A: ;4
| $Z.A: ;4
| $ZS.A: ;4
+-----+
+-----+
| # 0 11 0 21 3 13 3 5 3 4 3 1 1 |
| # 0 11 0 21 3 13 3 5 2 4 2 0 1 +-- Az (a) tiszta
| # 0 11 0 22 3 12 3 5 3 4 3 1 1 | fázisa.
| # 0 11 0 22 3 12 3 5 3 4 3 1 1 |
+-----+
+-----+
| $A.K. ;6 +|| Az a-k diád első szakasza. Tartalmazza az
+-----+ (a) változásait a
+-----+ (k)-ba haladva. Az
| # 0 9 0 19 3 13 5 5 2 4 2 1 1 | amplitúdó csökken
| # 0 6 0 15 2 13 4 6 1 3 1 1 1 |
| # 0 4 0 15 2 12 3 6 1 3 1 1 2 |
| # 0 1 0 17 3 12 3 5 1 3 2 1 1 +-----+
| # 0 1 0 18 3 11 2 5 1 4 2 1 1 |
| # 0 2 0 18 2 13 1 5 1 4 1 1 2 |
+-----+
+-----+
- $SP.K. ;3 +|| Itt csak a szókezdő (k) első szakasza van.
+-----+
| $A.K: ;3 +--- Az a-k diád második szakasza a (k) néma
+-----+ fázisával.
| $AA.K: ;3
| $E.K: ;3
| $EE.K: ;3
| $I.K: ;3
| $O.K: ;3
| $OE.K: ;3 +---- A magánhangzó-k és
| $U.K: ;3
| $UE.K: ;3
| $B.K: ;3 mássalhangzó-k átmenetek második
| $P.K: ;3 szakaszai mind összevonhatók.
| $D.K: ;3
| $T.K: ;3
| $CS.K: ;3
| $F.K: ;3
| $H.K: ;3
| $GY.K: ;3
| $TY.K: ;3
| $C.K: ;3
| $J.K: ;3
| $L.K: ;3
| $M.K: ;3 +-----+
| $N.K: ;3
| $NY.K: ;3
| $R.K: ;3
| $S.K: ;3
| $SZ.K: ;3
| $V.K: ;3
| $Z.K: ;3
| $ZS.K: ;3
+-----+
+-----+
| # 0 0 0 19 2 14 2 5 2 4 1 1 1 |
| # 0 0 0 19 2 14 2 5 2 4 1 1 1 |
| # 1 0 16 19 2 14 2 5 2 4 1 1 1 +-----+
+-----+

```

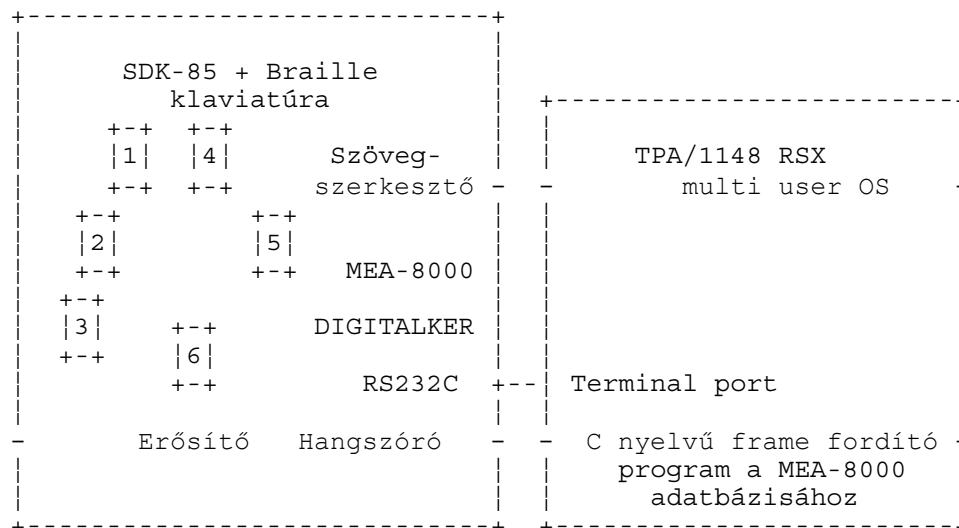
A magánhangzó mássalhangzó átmenetekben első szakaszban a magánhangzóból a mássalhangzóba áthajló részt tettem, majd a mássalhangzó bevezető (f, h, j, l, m, n, ny, s, sz, v, z, zs) vagy néma fázisát (b, c, cs, d, g, gy, k, p, r, t, ty). E második szakaszt itt is nagyon jól lehetett közösitni. A dz hang problémáját, a néma j-t, (pj) és a p-f átmeneteket stb. mind nagyon szépen meg lehetett oldani a jelöletlen hasonlóságokkal egyetemben. Azonban mindez nem segítette a hangok útján történő szövegszerkesztést.

Ennek a problémának a kiküszöbölésére bevezettem a részletező üzemmódot, melynek a hang adatbázis minőségi kialakítására nem volt negatív hatása (ld. 5.3.). A hangzók hosszítását is a diád mátrix segítségével oldottam meg. Így az r hang is több perdülettel hosszabbodik, és a magánhangzók és mássalhangzók hosszú változatai is egyéniek.

## 5.2. A HANG ADATBÁZISOK FEJLESZTŐ RENDSZEREI

A MEA-8000 adatbázisának első fejlesztő rendszere SDK-85 és PDP/11 gépeken futott. Az SDK-85-nek csak braille klaviatúrája volt, és a formáns szintetizátor mellett egy fix szókészletű rendszert (Digitalker) is tartalmazott. A két gép soros vonalon kommunikált. A PDP/11 gépen egy C nyelvű program kezelte a forrás formában megadott frame-eket és a hangátmenetek helyét jelző címkéket.

A szeleteket kézzel, szövegszerkesztővel kellett összeállítani. A C nyelvű fordítóprogram egy menetben alakította ki a text-to-speech program adatbázisának pointeres struktúráját. A kész adatbázist soros vonalon lehetett az SDK-85 alapú, kötetlen szótáras fejlesztő rendszerbe letölteni. Ehhez ki kellett fejlesztenem egy egyszerű, braille inputtal rendelkező, magyar szövegszerkesztő és soros vonali kezelő rendszert. Az első hang adatbázis kialakítását segítőként készítettem egy fejlesztő rendszert C64 számítógépre is (ld. F. melléklet).

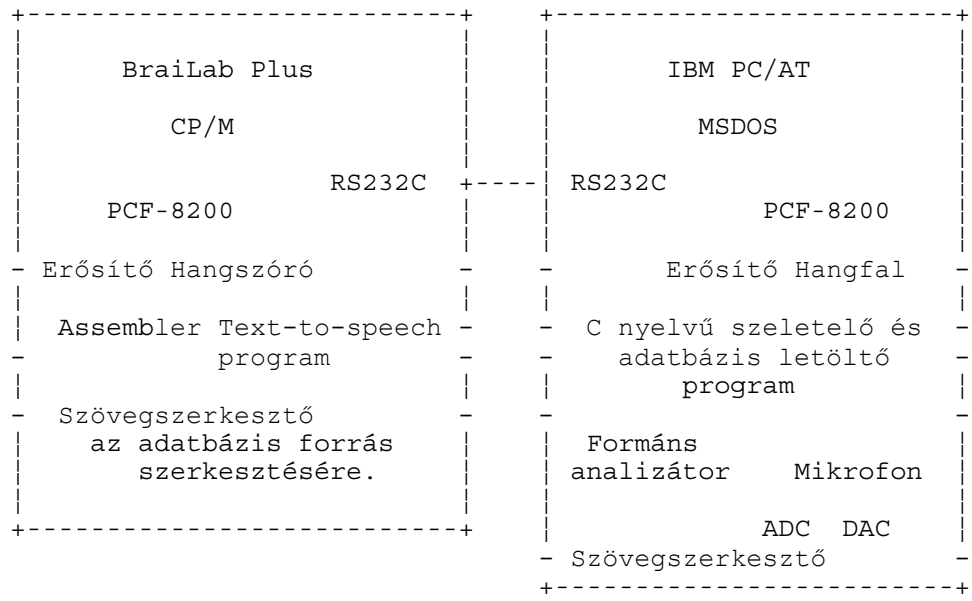


A MEA-8000 hang adatbázisának egyik fejlesztő rendszere.

Később a BraiLab Plusz létrehozásával a fejlesztő rendszer egy gépen, a CP/M operációs rendszer alatt üzemelt. A C nyelvű fejlesztő rendszer megfelelő rugalmassága tette lehetővé, hogy különböző gép összeállításokban más és más operációs rendszerekkel legyen használható a szeletelő fejlesztő rendszerem.

A PCF-8200-as adatbázis fejlesztő rendszer BraiLab Pluszon és IBM PC-n futott, újra kétféle rendszerre alkotva. Az újabb adatbázis kialakítása nagyobb terjedelme miatt nagyobb feladatokat rótt a fejlesztő rendszerre. Ebben már a szeletelést külön rutinok végezték a

szonogrammból kinyert formánsokból. Az egyes szeleteket szövegszerkesztővel lehetett egy nagy forrásszöveggé összeszerkeszteni, amit az új típusú fordító program konvertált kettős pontereket tartalmazó adatbázissá.



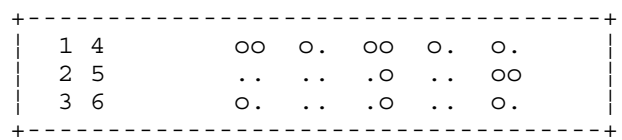
A PCF-8200 adatbázisának kétgépes fejlesztő rendszere.

### 5.3. ASCII -> METABRAILLE -> DIÁD KONVERTÁLÓ PROGRAM

A kötetlen szótáras beszédelőállító rendszer a hang adatbázist felhasználó konverziós programok összessége. Az én rendszerem programjainak újdonsága abban áll, hogy célzottan egy menetben képes konvertálni mind a bejövő ascii, mind pedig a braille kódokat. Az egy menetes működést az a felhasználási mód követelte meg, hogy segédeszközben a különféle billentyűk és billentyű kombinációk echózása közben megengedhetetlen a lassú reakció. Az alapvető billentyű echók esetében ez az érték 10-20 msec-nál nagyobb semmi esetre sem lehet.

A Metabrilie kód bevezetése előtt ki kell térjek a magyar és a német braille írás rövid ismertetésére. A braille írás, bár alapvetően minden nemzet a Luis Braille által megalkotott 6 pontos rendszerből indul ki, nem nemzetközi. Minden nyelvben igyekeztek a braille írással a hangok és nem a betűk jelölésére. Nagyobb braille írásos hagyományokkal rendelkező nemzetek bevezettek rövidírási rendszereket is (pl. angol, német), ahol gyakran hangcsoportokat is rövidítettek.

A magyar nyelvben, hasonlóan a némethez, már a rövidítetlen braille írás is betű csoportokat jelez egy braille jellel, mivel ezek gyakran egy hangot jelölnek, így helyes ezt az írást első szintű rövidírásnak nevezni. A braille írás pontjainak elhelyezkedése és számozása a következő:



A hat Braille írású pont elhelyezkedése.

A minta szöveg: „magyar”. Még az is jellemző, hogy a gy hang a d hang braille jeléből a 6. pont hozzáadásával képződik. Ebben is hangtani szabályokat követ (ld. d-j jelöletlen hasonulása gy hangra).

Az alábbi táblázat tartalmazza a magyar és a német braille írásban használatos speciális hangzó jelöléseket, amikor a betű csoportoknak egy braille jel felel meg.

GY	LY	NY	TY	CS
SZ	ZS			

A magyar braille írás első szintű hang rövidítései.

EI	EU	ÄU	AU	IE
CH	SCH	ST		

A német braille írás első szintű rövidítései.

A text-to-speech programban a billentyűk megnevezéseit, és a számok kimondására szolgáló szavakat és szó részleteket ilyen Metabraille-hang kódokban tároltam. (A számokat a program 9 jegyig olvassa össze). Ennek nagy előnye származott a gyors feldolgozás lehetőségében, mivel ez egy első szintű kivétel szótárt képezett. Ha ascii kódokból állítottam volna össze pl. a kilencszáz szót, akkor a kétjegyű mássalhangzók feldolgozásakor abból kilencs-záz lehetne.

A BraiLab Basic gépben, mert kevés volt a rendelkezésre álló memória, ezt a Metabraille kódú bevitelt csak kézzel lehetett megvalósítani. A BraiLab PC text-to-speech programjában már van egy második szintű kivétel szótár kezelés is.

A magyar nyelvű, MEA-8000-re készített programjaim, mivel ezek 8 bites gépeken futnak, egy menetben oldják meg az ascii Metabraille hangparaméter konverziót. Kivételt képez a makrointonálási szerkezet utólagos kialakítása a beszéd framekre. Ezekben a rendszerekben csak néhány nagyon egyszerű dallamot generálok intonálási egységenként.

A német nyelvű programban több menetet kellett beiktatni az ascii feldolgozásban is, ezért a német nyelvű szövegszerkesztés 8 bites gépen észrevehetően lassúbb lett. Az egyik vak felhasználó meg is jegyezte: „mintha süketen kellene dolgoznom vele”.

A magyar Metabraille hangkód konverziós fázisban külön típus kóddal jelöltem meg az egyes hangcsoportokat (magánhangzó, zárfelpattanásos mássalhangzó, egyéb mássalhangzó), hogy a hosszítást és lassítást könnyen el lehessen végezni. A PCF-8200-as rendszerben erre már nem volt szükség. Mindkét rendszerben a használt pointer táblázat, mely frame szakaszokra mutat, gyorsabb feldolgozást tesz lehetővé.

A programokat I8080, Z80 és a BraiLab PC esetében I8086 assembler nyelven írtam. A MEA-8000-re csak 8 bites program változatok készültek, míg a PCF-8200-ra mind 8, mind pedig 16 bites program forrás kód is készült. Adatbázis nélkül egyelőre, de elkészítettem a német text-to-speech programot is a PCF-8200-ra.

#### 5.4. MIKROINTONÁCIÓ, INTONÁCIÓ, ÉNEK

A beszéd természetességét és érthetőségét is javítják a mesterséges intonációk. Olasz Gábor azt az alaphang magasság változást nevezi mikrointonálásnak, mely a beszélő akaratától függetlenül jelenik meg a hangjában monotonon ejtett szavak esetében. A mikrointonáció lüktetést ad a mesterséges beszédben.

A MEA-8000-es rendszerben csak a német programba építettem kísérletképpen mikrointonálást alaphang magasság változtatással, a magyarba nem. A PCF-8200-as rendszerben másképpen jártam el. A „mikrointonációs” szerkezetet a magyar hang adatbázisban nem az alaphang változás biztosítja, hanem a formánsok frekvenciáinak és sáv szélességeinek természetes lüktetése.

A beszéd makrointonációs szerkezete a beszélő szándékát segíti kifejezni. Az első magyar text-to-speech rendszeremben először valósítottam meg automatikus dallamgenerálást a mondatjelek és bizonyos minimális szintaktikai vizsgálatok alapján. A Nyelvtudományi Intézet akkori rendszere ehhez külön speciális jeleket alkalmazott. Megjegyzem, hogy azóta ezen a területen Olasz Gábor nagyon szép új eredményeket ért el.

Az automatikus dallam generálásnak a segédeszközbe épített rendszereimben sajátos funkcióit valósítottam meg. A vak felhasználónak, mikor sorokat, vagy mondatokat hallgat szövegszerkesztés közben, az intonálás gyors jelzést ad arra, hogy milyen mondatjel van az adott mondatrész végén. Ezért arra törekedtem, hogy ez az intonációs jelzés akkor is egyértelmű legyen, ha a felhasználó korán megszakította az intonálási egység kimondását, mert már a következő sorra ugrott.

A MEA-8000-es rendszerben mindössze 4 alap intonálási típust valósítottam meg. A kijelentő mondathoz kettőt, a kérdő mondathoz kettőt, a felszólító mondathoz pedig a kérdő szavas kérdőmondat dallamát használtam. Az eldöntendő kérdő mondatok hossza alapján csak két altípust különböztettem meg. A kérdőszavakat és a határozott névelőket csak a mondat elején vizsgáltam egy nagyon egyszerű algoritmus szerint. Minden HO HÁ MI ME KI karakter párossal kezdődő mondatot kiegészítendő kérdő mondatnak tekintettem. Ezzel elég nagy százalékban jól működő rendszert sikerült alkotni igen kis helyen. Az első magyar text-to-speech programom hossza hang adatbázis nélkül mindössze 4 Kb-ot.

A PCF-8200-as rendszerben már különbséget tettem az egyes mondat típusok hossza között is, a szótagszám alapján. Mindkét rendszerben szó vektorban tárolom a szókezdő frame-ek helyét, de a szó intonálást még egyik rendszerben sem valósítottam meg. Ennek segédeszközökben történő felhasználás esetén nincs nagy jelentősége.

Az automatikus ének generálásnak legfőbb felhasználási területe a braille kottairás lehet. Interaktív munkában a hangmagassággal jelzett hangjegyzírás természetes, gyors munkát tehet lehetővé. Ez az alkalmazás egyelőre nem készült el. Maga az automatikus ének generálás viszont igen. Ehhez a MEA-8000-es magyar és német hang adatbázisok külön változatát kellett létrehoznom.

Egy egyszerű ének, kezdő alaphangmagasságok megadásával képezhető. Hajlításokhoz, és pontosabb ritmus létrehozásához már alaphangsík inkrementumokkal kellene dolgozni. A kezdeti alaphangmagasság változtatásával csak akkor lehet jól éneket előállítani, ha a text-to-speech program gyorsan lefut, mert különben a feldolgozási idő torzítja a ritmust. A MEA-8000-es hang adatbázisokat (magyar és német) úgy módosítottam, hogy minden szókezdő és szóvégi diád legutolsó frame-je is minimális hosszúságú legyen.



A másik alapvető módosítás, melyre szükség volt az ének generálásához a hang adatbázisban, a magánhangzó-magánhangzó kapcsolatok azonos amplitúdóra hozása. Abból a célból, hogy a magánhangzók egymástól jól elkülönüljenek, ezek közötti átmenetek nulla amplitúdójú frame-eket tartalmaztak, az éneknél viszont, különösen az azonos magánhangzók egymáshoz kapcsolódásában a diád mátrix pointereit kellett figyelembe venni, és az így nyert frame-ekben az amplitúdókat ki kellett egyenlíteni.

A PCF-8200-as hang adatbázis már minden tekintetben alkalmas volt az ének létrehozására is. A magánhangzón belüli „mikrointonációs” szerkezetem pedig a kitartott énekhangot is gazdagítja.

A programok, melyek generálják az éneket a MEA-8000-re assembler nyelven, a PCF-8200-ra C nyelven íródtak. A C nyelvű program egy fordító, mely a forrás szövegből egy escape szekvenciákat tartalmazó fájlt készít. Ez a fájl a szokásos módon a COM4 virtuális periférián keresztül megszólaltatható (ld. 6.3.2.). A forrásnyelvi szövegre álljon itt egy minta:

```
+-----+
|
| re,éééérik,La,a,re,szóó,mi,lóó
| do,haaaajlik,re,a,mi veee,re,szóó
| do,bo,Ti,dor,La,aaa lee,re,ve,La,leeee
| So,kééééet sze,Re,géeny,So,leeee,La,gééeny
| Fa,száánta,So,ni,La,mee,So,nnee
| Fa,de,Mi,nin,Re,cseen ke,So,nye,Re,ree
|
| re,vaaaaaan ve,La,res,re,haaaagy,mi,maaaa
| do,aaaaaa ta,re,risz,mi,nyáá,rebaaa
| do,ke,Ti,se,La,rúúú maa,re,gá,La,baaa
| So,szoool ga,Re,le,So,gééeny,La,neek
| Fa,heeej de,So,sze,La,gééeny,So,neek
| Fa,de,Mi,ke,Re,véés va,So,cso,Re,raa
|
+-----+
```

Az automatikus ének generálás input-ja.

Kodály Zoltán relatív szolmizálási jelölését követve, a szövegtől vesszőkkel elválasztva a hangmagasságot jelölő nevek láthatók. Ha a hang magassága változik, újabb vesszővel újabb szolmizációs név következik. A nagybetűvel kezdett nevek az egy oktávval mélyebb hangokat jelölik. A ritmust a magánhangzók megfelelő számával lehet szabályozni. A szüneteket a szóközök száma szabályozza. A BraiLab PC rendszerben az énekhez ki kell választani a szóközöket nem tömörítő üzemmódot.

Az újszerű „mikrointonálás” és az énekhez létrehozott speciális hang adatbázisok is bizonyítják az átlapolásos tárolási technikám nagyfokú rugalmasságát. Frame kódolással ennyi féle adatbázis és ilyen finom részleteket is tartalmazó hang anyag létrehozása nem lett volna lehetséges.

## 6. fejezet. A BRAILAB SEGÉDESZKÖZÖK

### 6.1. A BRAILAB BASIC

#### 6.1.1. A BRAILAB BASIC SEGÉDESZKÖZ FELÉPÍTÉSE

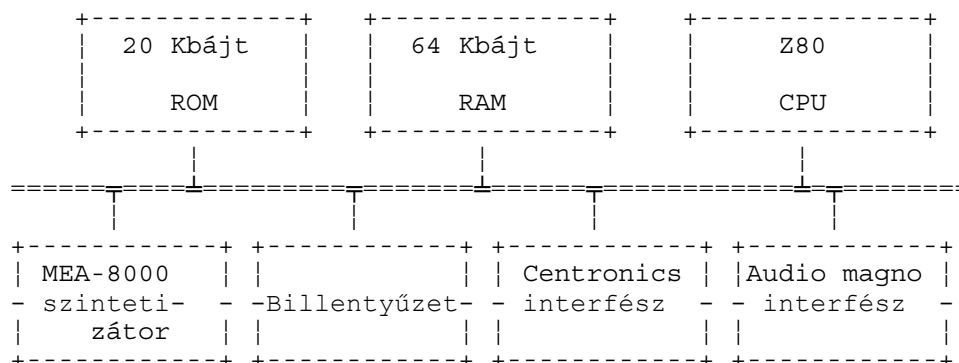
A BraiLab Basic segédeszköz alapja a Lukács József és Endre készítette HomeLab gép. Ez Z80 mikroprocesszor alapú, moduláris felépítésű eszköz. A BraiLab használathoz 64 Kbájt dinamikus RAM memóriát tartalmazott és 5 darab 4 Kbájt méretű EPROM-ot. A HomeLab gép elektronikája egy nyomtatott kártyára épült, melyet a billentyű kártya egészített ki. A BraiLab változathoz a gumi érintkezős billentyűket mechanikusakra kellett lecserélni. Néhány speciális billentyűt tapintással érzékelhetővé tettem.

A video memóriát és a karakter generátort a Lukács testvérek külön 2 Kbájtos RAM-ban ill. ROM-ban valósították meg. Az ernyő 25 sor és 32 vagy 64 karakter széles volt. A BraiLab változat csak ez utóbbit használta. A video kimenet fontos segítséget jelentett a gép oktatásában. A Z80PIO áramkör a centronics interfész céljait szolgálta. Az egyetlen háttértár lehetőség az audio magnós kimenet volt.

A HomeLab gép speciális memory management megoldása tette igazán lehetővé a hardver és a szoftver segédeszközzé alakítását. A video memória és a mátrix szerűen kialakított billentyűzetet memória referenciás utasításokkal lehetett elérni. Különböző ROM és RAM területeket lehetett ki és belapozni speciális input/output címekre kiadott utasításokkal.

A MEA-8000 formáns szintetizátor kiegészítő kártyára került egy használaton kívüli EPROM tokba helyezve. A szűrő és erősítő áramkörök után a hang kijelzés közösitve lett a HomeLab programozható hangkimenetével.

Az egész gép fröccsöntött műanyag dobozban nyert elhelyezést. A gép kivezetései: külső Z80 busz csatlakozó, Centronics kivezetés, magnó és tápfeszültség ki/bemenetek. A dobozban volt elhelyezve a hangszóró, és a hangszórót leválasztó fejhallgató kimenet is. Ez utóbbinak a csoportos oktatásban volt nagy szerepe.



A BraiLab Basic felépítése.

### 6.1.2. A TEXT-TO-SPEECH RENDSZER BEÉPÍTÉSE A BRAILAB BASIC-BE

A HomeLab gép szoftvere Basic interpreterból, interaktív Assembler-monitorból és az input/output rendszerből állt. A billentyűzet megnyomását különböző hangmagasságok kísérték. Az ernyő input és output berendezés is volt egyszerre. Ahhoz, hogy a közel 8 Kbájtos text-to-speech rendszert be lehessen integrálni a számítógépbe, az egyetlen 4 Kbájtos bővítő EPROM hely kevés volt, ezért a 4 Kbájtos duplapontos Basic aritmetikai lehetőséget ki kellett venni a gépből.

Az ernyő és a billentyű kezelést két ponton kellett módosítanom. Az egyik pont az ernyőre való kiírás, melyen keresztül az input/output rendszer segítségével a Basic interpreter normál és hibaüzeneteket írt ki. Ha a felhasználó közvetlen címezte meg az ernyőt, akkor ezek az üzenetek elkerülték a meghangosítást. Az ernyőre való íráskor a képváltáshoz történő szinkronizációt a program biztosította.

A másik módosítási pont egy olyan összetett rutinban volt, mely a billentyű scan-elést, screen editor szerű ernyő input/output kezelést és a kurzor villogtatást együttesen kezelte. Ha nem kontroll jellegű karaktereket gépelt a felhasználó, akkor ezeket a hangalak alapján, a text-to-speech program meghívásával meghangosítottam. A hangalak a megfelelő hang szókezdő és szóvégi diádijából alakult ki, vagy írásjelek és számjegyek esetén azok megnevezéséből a text-to-speech belső Metabrilie kivétel szótárából.

A kontroll jellegű billentyűk megnyomásakor először az ernyőre echóztam a megfelelő funkciót, pl. jobbra, balra, le fel, ernyőtörlés, és utána az ernyőről echóztam a kurzor pozícióját. Nagyon fontos itt megjegyezni, hogy nem a megfelelő funkciót mondtam ki (jobbra, balra stb.), mivel az nem segíti a hallás útján történő tájékozódást az ernyő síkján.

A törlő funkció megvalósításakor is az éppen a törölt karakter helyére belépő karaktert szólaltattam meg, és nem magáról a törlési funkcióról adtam tájékoztatást. Megneveztem az üres karaktert is, melynek funkciója nagyon eltért a szóközétől. Az ernyőről a sor beolvasása (két üres karakter közé eső rész), video memóriából történő kiolvasással valósult meg. Ekkor a teljes sort összeolvasva, intonálási egységekre bontva kimondattam a text-to-speech rendszerrel. Ez egyben egy végső ellenőrzést is lehetővé tett a Basic sorok begépelésekor. A kimondott sor azt demonstrálta, hogy a Basic interpreter az ernyőn lévő sort a program sorai közé felvette.

Az ernyőre kiküldött karaktereket is intonálási egységekre bontottam és azokat is, beleértve a számokat is, összeolvasva mondtam ki a kötetlen szótáras beszéd előállító rendszeremmel. A HomeLab Basic jelzéseit és hibaüzeneteit a jobb hangzás céljából kijavítottam, egységes mondatjelekkel láttam el. Így sikerült természetesebbé tennem a beszéd útján történő párbeszéd programozást.

A mesterséges beszéd paraméterei Basic-ből a peek és poke utasításokkal érhetőek el. Ezek segítségével programból egyszerűen változtatható a hangmagasság és a beszédsebesség. A programozható digitális fütty és az írásjelekkel is könnyen vezérelhető mesterséges beszéd kiváló lehetőségeket teremtett az oktató és játékprogramok létrehozói számára.

A HomeLab gép alapvetően megszakítás nélkül üzemelt, ezért a sorbevevő rutinban, valamint a text-to-speech programban folyamatosan figyelni kellett az újabb billentyű lenyomását. Ha ez bekövetkezett, akkor a formáns szintetizátort le kellett állítani, az újabb billentyűhöz a text-to-speech programot újra le kellett futtatni, és a megszakító billentyű megnevezésével jelezni, hogy milyen billentyű lett megnyomva. A programmegszakítás hiánya miatt itt olyan kevés idő maradt a real-time text-to-speech futása mellett, hogy a gyors billentyűzés követése

közben nem maradt idő a beszéd ideje alatti kurzor villogtatására és a beszéd szintetizátor korrekt leállítására. Ez utóbbi néha hibás utózengekhez vezetett.

Az ernyőn való tájékozódást RAM-ban futó kiegészítő program (BraiLab overlay) segítette. Ezt magnóról lehetett beolvasni Basic területen kívüli helyre. A programmal off-line ernyő-olvasó funkciókat valósítottam meg (hangalak helyett névvel történő betűzés, kurzor helyének kimondása). A BraiLab vakon történő használatához oktatási segédletet írtam, mely a mellékletben található.

A fent leírtak szerint úgy sikerült a HomeLab-ot módosítanom, hogy annak minden funkcióját vakon lehetett használni, a korábban írt minden program gyakorlatilag módosítás nélkül futtatható volt. Kidolgoztam egy olyan párbeszédés Basic programozási technikát, melynek segítségével a programozásban gyakorlatlan vak felhasználók is szinte azonnal képesek voltak a BraiLabot használni. Később már maguk a felhasználók „vakosították” a korábbi HomeLab-os játékprogramokat, vagy írtak saját tervezésűeket, hogy számítástechnikában kezdő sors-társaikkal megismertethessék a segédeszköz lehetőségeit.

A lavinaszerű sikert az alapozta meg, hogy a BraiLab Basic egy speciális, bolondbiztos cél segédeszköz. Annak használatához nem volt szükség külön parancsokat megtanulni, mindössze a személyi számítógép funkcióit kellett megismerni. A mesterséges beszéd mindig logikus, gyors visszajelzést biztosított a látássérült felhasználóknak.

### *6.1.3. AZ ELSŐ MAGYAR OLVASÓGÉP BELSŐ FELÉPÍTÉSE*

A BraiLab Basic-nek készítettem egy sorozatban nem gyártott változatát, mely a MEA-8000-en kívül soros vonali interfésszel is rendelkezett. A HomeLab input/output rendszerét módosítottam úgy, hogy beépítettem egy hardver flag-et annak jelzésére, hogy melyik memory management lapon fut éppen a program. Ezzel sikerült hardver interruptos programot is írnom.

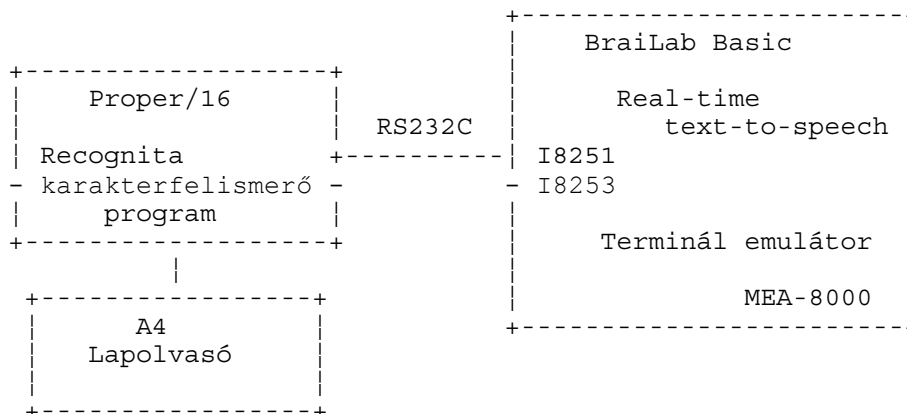
A Basic interpreter helyett ROM-ba készítettem el a soros vonalat interruptosan kezelő terminál programot. A soros vonalat I8251-gyel, a baud rate generátort I8253-mal valósítottam meg.

A soros vonalról érkező karaktereket silóban gyűjtöttem interrupt rutinnal, mely a termelő folyamatot valósította meg. Ha a karaktereket tároló buffer mérete egy adott érték alá csökkent, akkor XOFF karaktert küldtem ki a terminálon. Ha a bufferben hiszterézis függvény alapján kellő bájtt szabadult fel, akkor XON jelzést küldtem ki a soros vonalon.

A fogyasztó folyamatot a mesterséges beszéd jelentette. A bufferből a karaktereket intonálási szakaszokra osztottam, és kimondattam a BraiLab Basic text-to-speech rendszerével. A fogyasztó folyamatot a termelő folyamat interruptja a karakter beérkezésekor megszakította.

Ezt a beszélő terminál rendszert kapcsoltuk össze a Számítástechnikai Kutató Intézet által kifejlesztett Recognita nevű optikai karakterfelismerő rendszerével. A felismerő program Proper 16 számítógépen futott, és a szabványos RS232C interfészén küldte ki a felismert magyar 8 bites ascii karaktereket az első magyar beszélő személyi számítógépnek, a BraiLab Basic-nek.

Ezt az első magyar olvasógépet 1986-ban állítottuk ki a „Magyarok szerepe a világ természet-tudományos és műszaki haladásában” című tudományos találkozón, mint meghívott kiállítók.



Az első magyar olvasógép konfigurációja.

## 6.2. A BRAILAB PLUS

### 6.2.1. A BRAILAB PLUS SEGÉDESZKÖZ FELÉPÍTÉSE

A BraiLab Plus számítógép hardver és CP/M-es alapszoftverét is Lukács József és Endre tervezték, de már figyelembe vették mindazokat a követelményeket, melyeket kidolgoztam ahhoz, hogy a személyi számítógép látássérültek segédeszköze lehessen.

Először a BraiLab Basic gép buszkivezetéséhez készült el floppy diszk interfész. A CP/M-es operációs rendszer első változatába is beépítettük a MEA-8000-es text-to-speech rendszert, de ez a változat nem került sorozat gyártásba. A BraiLab Plus gépet újra terveztük egy nyomtatott áramkörre.

A segédeszköz két modulból állt. Hasonlóan a BraiLab Basic géphez, fröccsöntött műanyag dobozban található az elektronikát tartalmazó nyomtatott áramkör, és egy külön kártyán a kiegészített mátrix-os mechanikus kapcsolós billentyűzet. Külön alumínium dobozban nyert elhelyezést a BraiLab Plus hálózati tápegysége, együtt a hajlékony lemezes meghajtóval. Display nélkül a rendszer 5 Kg alatti, hordozható, bár nem hálózat független kivitelű.

A kétoldalú nyomtatott áramköri lemezen a Z80 mikroprocesszoron kívül 256 Kbájt dinamikus RAM, négy toknyi 16 Kbájtos EPROM, 2 darab Z80PIO, 4 Kbájt statikus RAM, karakter generátor EPROM, WD2793 floppy diszk controller, I8251 PUSART és a MEA-8000 beszédszintetizátor található.

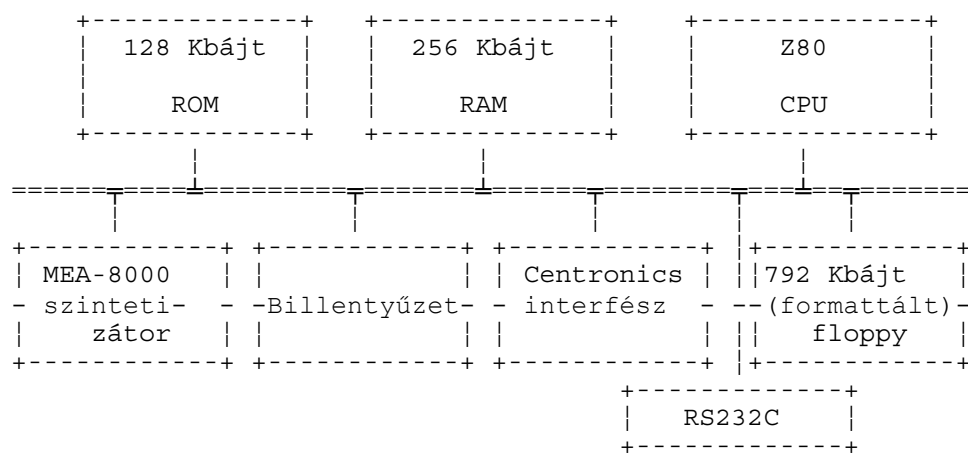
A 256 Kbájt RAM-ból 180 Kbájt fix RAM diszkként kezelődik a CP/M operációs rendszer alatt. A memory management támogatja a dinamikus RAM fix diszk használatát, biztosítja az operációs rendszer megfelelő címtartományát, és lapoz közöttük és az EPROM-ok között (ebből végül csak egyet használt ki a BraiLab Plus). A két darab Z80PIO végzi el a centronics interfészelt, a rendszer órajel interrupt kezelését és az I8251-gyel és HD-4702-vel megvalósított soros vonali interfész baud rate programozását.

A megnövelt méretű 4 Kbájtos memória 25-ször 80 karakteres ernyőt tett lehetővé. A billentyűzetet kiegészítettük a BraiLab Basic-hez képest további 5 billentyűvel. Négyet kereszt alakban külön tapinthatóan helyeztem el, evvel is elősegítve az ernyőn való tájékozódás biztonságát. Az optifikált ernyő síkjának megfelelően pozícionált nyilak megnyomásában a vak felhasználó a legritkábban téved. (Ezért sem kell a nyilak neveit echózni).

Az ötödik billentyű, melyet fontosnak tartottam bevezetni, az Escape volt. Ezt a klaviatúra jobb szélén helyeztem el, mert így minden hibás programkezelés után az Escape gomb lenyomásával egyértelműen vissza lehetett lépni a hibás tevékenységből. Tapintáshoz kiemelni mindössze 3 billentyűt kellett.

A MEA-8000 formáns szintetizátor hangkimenete a BraiLab Plus-ban is közös hangszóróba csatlakozik a programozható hangmagasság (fütyt) kimenettel. A nyomtatott áramkör gépi tervezésekor igyekeztem elkülöníteni az analóg hangrészeket a gép digitális zajától.

A szabványos soros, párhuzamos, valamint a floppy és tápfeszültség kivezetések mellett a BraiLab Plus-nak is van fekete-fehér video display kimenete és fejhallgató csatlakozója. A gép két moduljának összekapcsolásakor csak egy kombinált hálózati-floppy csatlakozót kell a felhasználónak összeilleszteni. A kicsatlakozások bolondbiztosak, egymással össze nem keverhetők.



A BraiLab Plus felépítése.

### 6.2.2. A TEXT-TO-SPEECH RENDSZER BEÉPÍTÉSE A BRAILAB PLUS-BA

A BraiLab Plus CP/M rendszere ROM-ból fut. Ezzel Lukács Endre elérte, hogy az operációs rendszer alatt futó programok megszakíthatók, és bizonyos funkciók (pl. DIR, TYPE stb.) végrehajthatók a megszakítás ideje alatt. Az ernyőn egy ablak nyílik, melyről a vak felhasználó szóban figyelmeztetést kap. Az operációs rendszer prompt-ja az ablakban más intonálási karaktert tartalmaz, így ez is figyelmeztető jelzés a gyengénlátó felhasználónak.

Az „ablakban” adhatók ki az egyes üzemmód váltások is, melyek menüszerűen vannak felsorolva. A látássérült felhasználónak csak a fel és le nyilakkal kell kiválasztani a megfelelő üzemmód váltó parancsot, mert a parancsok echója egyértelmű tájékoztatást ad számára. A kocs vissza billentyű szolgál a váltásra, amikor még egyszer hangos visszajelzést kap a kért parancsról.

A BraiLab Plus beszélő rendszere WordStar elvű. Ez azt jelenti, hogy a screen editoros CP/M sorbeevő program minden alapvető funkciója megegyezik a WordStar szövegszerkesztő funkcióival. A karakterenkénti, szavankénti léptetést a nyilak és a váltós párjuk aktivizálja. Az F1 funkció billentyű a balra és a jobbra nyíllal együtt a sor széleire vezérli a kurzort. Mindezek a funkciók azonosak nemcsak a WordStar-ban és az operációs rendszerben on-line-ban, de az off-line működésben is.

Az off-line bevezetésére, hasonlóan a BraiLab Basic overlay programjához, azért volt szükség, hogy betűzve ismételve az ernyőről minden leolvasható legyen abszolút érthetőségi biztonsággal. Az off-line másik funkciója abban áll, hogy a CP/M alatt futó nem „vakosított”

programokat is lehessen vakon használni. A BraiLab Plus rendszerben bevezettem fecseg és nem fecseg üzemmódokat. A vakok segédeszköze céljára átírt programok speciális karakter szekvenciákkal (%` ill. %@) be- ill. kikapcsolhatják a beszédet, mely a standard output-on keresztül az ernyőre kerül.

Ezzel a megoldással a programok nagy része könnyen volt módosítható úgy, hogy a megfelelő stringek átírásával a legfontosabb üzenetek hangossá váltak. A nem fecseg üzemmódban csak ezek az üzenetek hallatszanak. A fecseg üzemmód az ismeretlen, nem „vakos” programok futtatására szolgál.

A WordStar egy adott változatát úgy módosítottam debugot és batch programot használva, hogy a batch program a futtatható WordStar üzeneteit átírta, így az eredeti szoftver használhatóvá vált vakon. DbaseII felhasználói programot is írtam, többfunkciós beszélő telefonkönyv céljára. Ebben is a beszéd ki- és bekapcsoló szekvenciáit kellett alkalmaznom.

A Kermit szabad szoftvert forrásnyelven írtam át. Magyarul és németül (és angolul is) vezérelhető parancsokat és help rendszert építettem bele. A Kermit-re épülő kommunikációs rendszer több vak felhasználónak jelentett segédeszköz lehetőséget IBM PC-n folytatott munkájukban is addig, míg a BraiLab PC el nem készült.

A beszéd beillesztése a rendszerbe nem volt interruptos, de mivel az egyéb funkciókat 50 Hz-es óra interrupt vezérli, az egyes billentyűk meghangosítására már elegendő idő állt rendelkezésre. Ezért sikerült a billentyűk okozta beszédleállítást úgy megoldani, hogy nulla amplitúdójú frame-eket is volt idő kiküldeni a MEA-8000-re, ami így utózengezés nélkül áll le minden esetben.

A BraiLab Plus-ba többlépcsős némítási rendszert terveztem. Ez azt jelenti, hogy az egyes billentyűk nemcsak egy intonálási egységet némítanak, ha beszéd közben lenyomják őket, hanem pl. a sorvég jel operációs rendszer szinten némítja a teljes parancs output-ját. Ugyancsak több intonálási egység némítható a váltó billentyűkkel. Ezek mindaddig hatásosak, míg nyomva tartva.

A kontroll jellegű karakterek echózását (jobbra, balra, szó jobbra, szó balra, lapozás előre hátra stb.) külön szoftver időzítő számlálók vezérlik, hogy a szövegszerkesztő futási idejét ki lehessen várni. Az ernyőről már csak a megfelelő karakter, szó vagy sor olvasódik fel a kontroll karakter lenyomása után, a megfelelő időpillanatban.

Külön ki kell emelnem a Del karakter echózását. Ez kitörli visszafele haladva a szó utolsó karakterét, és úgy jelzi hangosan, hogy a szót kimondja addig, ameddig az éppen látható. A szavak törlésekor a törölt szó helyébe belépő szó hallatszik. Sorok törlésének echója a szavak törlésének echójával rokon.

A BraiLab Plus-ban bevezettem a szóköz echóját. A BraiLab Basic-ben ez csak beszéd szünetként jelentkezett. Ott nem okozott nagyon sok gondot, hogy a szóköz külön nem hallatszott, mert a Basic nyelvben a szóközők száma nem számít. Szövegszerkesztésnél azonban szükség volt a szóközők pontos leszámolására, amihez a beszéd szintetizátort hívtam segítségül. A szóköz echójára a következő frame-et szólaltatom meg:

```
+-----+
|FD=8  AMPL=11  PI=0  F1=150  F2=440  F3=1179  B1=B2=B3=B4=726|
+-----+
```

A szóköz hangja a MEA-8000 szintetizátoron.

Az autorepeat funkciót a BraiLab Plus-ban, hasonlóan a BraiLab Basic-hez, úgy oldottam meg, hogy annak ismétlési ideje nem fix, mint általában, hanem a kimondott hangalaktól függ. Addig tart, míg az adott billentyű echója hangzik. A kezdeti idő persze néma, és hosszabb, nehogy véletlenül több karaktert gépeljen be a vak felhasználó.

Egy másik billentyűzést elősegítő funkció az egyes váltók zárolási lehetősége. A bal és jobb nagybetű váltók külön zárolhatók a következő karakter leütéséig. Ugyan így használhatók kívánságra a kontroll és az F1 vezérlő billentyűk. Ezt nemcsak azok a felhasználók veszik igénybe, akik többszörös fogyatékoságuk miatt csak egy kézzel tudnak gépelni, de az ép kezű látássérültek szinte mind bekapcsolva használják az F1 váltót (van olyan, aki a kontrollt is), nagymértékben biztonságossá téve ezzel a felhasználói programok vezérlését.

A BraiLab Plus operációs rendszerében az F1 billentyű kombinációkhoz billentyű sorozatok rendelhetők CP/M szinten. Ezek könnyen előhívhatók az egyes felhasználói programokban úgy, hogy echójuk szintetizált beszéddel egyértelmű.

A szabványos soros port a BraiLab Plus beépített lehetősége. Kezelése interruptos, nagyon hasonló, silós megoldású, mint a BraiLab Basic minta példányának terminál emulátorában. A különbség abban áll, hogy a szoftver handshake mellett programozható a soros interfész működése hardver handshake-re is, mind adás, mind pedig vételi oldalon. A beszélő terminál és file átviteli program ezeket a lehetőségeket ki is használja.

### 6.2.3. A BRAILAB PLUS BRAILLE PUBLISHING RENDSZERE

A BraiLab Plus billentyűzete használható braille üzemmódban is. Ez azt jelenti, hogy az eddig lenyomásra megszólaló billentyűk felengedésre szólalnak meg a megfelelő magyar braille pont kombinációra. A braille-1 üzemmódban a következő billentyű hozzárendelések élnek:

F	-	1	4	-	J
D	-	2	5	-	K
S	-	3	6	-	L

A BraiLab Plus Braille Billentyűinek elhelyezkedése.

Braille-2, azaz egykezes változatban ezek a billentyűk közelebb kerülnek egymáshoz. A braille pont kombináció mindaddig nem végleges, míg legalább egy, az adott billentyűk közül le van nyomva. A szóköz billentyű azon kívül, hogy megőrizte eredeti funkcióját, a többi billentyűvel kombinálva kontroll funkciót lát el. Az É billentyű tölti be braille-1 módban az F1 szerepét.

Braille billentyűzés esetén természetesen lehet bevinni a gépbe CS, LY, GY stb. betűket. Ezek ábrázolására szolgál a Metabrilie kódolás. Az a szó, hogy magyar, metabrilie-ben így néz ki: maGar. Ennek a kódolásnak egyik nagy előnye mutatkozik meg abban, hogy egy braille cellának egy pozíció felel itt is meg. Az alábbi mondat látó és síkírásban példázza a metabrilie kódolás lényegét:

Magyar Vakok és Gyengénlátók Országos Szövetsége 1992.
maGar vakok és Gengénlátók orSágos Sövetsége D1992.

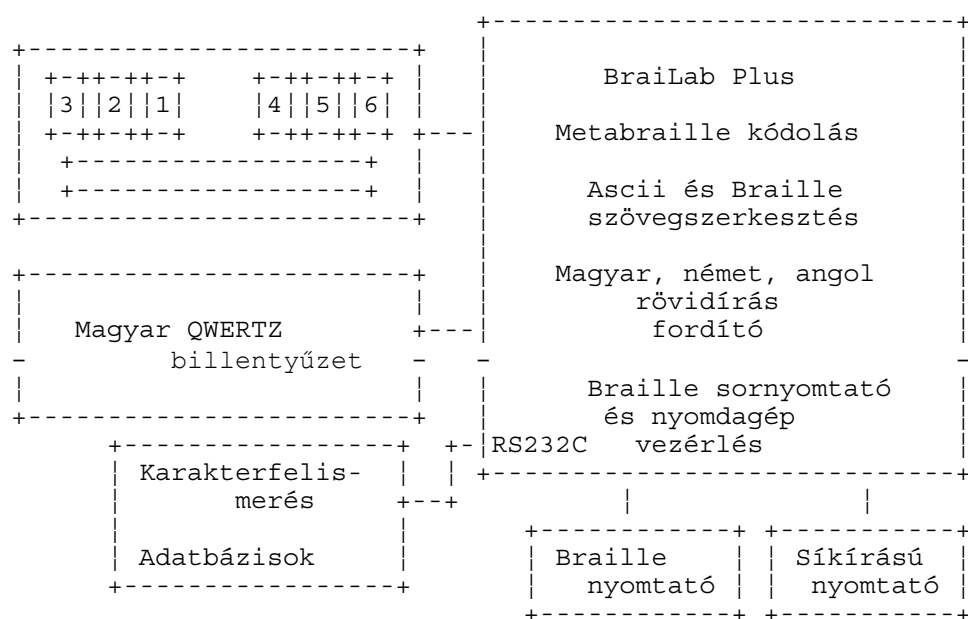
Metabrilie és síkírású kódolás a BraiLab Plus-ban.



Mind gépeléskor, mind pedig nagyobb egységek összeolvasása esetén a BraiLab Plus-ban braille üzemmódban helyes és teljes a beszéddel történő visszajelzés. A D metabraille kód a számjel jelzésére szolgál (hasonló a számjel használata a braille-ben mint az 5 bites telex kódban).

A Braille szerkesztő rendszerhez fordító programokat készítettem a magyar teljes írás, a 44-es és a legújabban kialakuló 77-es rövidírás konvertálására. Készítettem fordító programot német teljes és angol rövidírási rendszerre is. A fordító programok ascii-ből konvertálnak metabraille-re.

Metabraille-ben azután könnyen használható a beszélő WordStar szövegszerkesztő program a későbbi braille szövegek formázására, szerkesztésére. Négy féle braille nyomtatóra készítettem metabraille kódból braille domborítást végző programokat. A nyomdai eljáráshoz fém klisé lemezeket használnak a Magyar Vakok és Gyengénlátók Országos Szövetségében. Ennek a fémnyomó gépnek a vezérlését is az én programommal használják.



A BraiLab Plus Braille Publishing rendszere.

A nyomdai munka, így a következő lépésekből áll: az anyagot látó vagy braille írásban rögzítik (esetleg optikai karakter felismerővel beolvassák, majd a hibákat kijavítják), metabraille-be konvertálják. Ha kell rövidírási fordítóval automatikus nyelvi rövidítéseket végeznek megadott szabályok és kivételek alapján. A szerkesztéseket metabraille kódban végzik, majd braille nyomtatón „kefelenyomatot” készítenek. Az újabb szövegszerkesztések után elkészítik a fémlapokat a préseléshez.

Braille nyomtatási rendszerem nagy előnye, hogy annak minden lépése végezhető akár vak, akár látó által. Mind a kettőre van is gyakorlati példa (a szövetségben vak a nyomdász, az általános iskolában látó tanár is készíti a braille segéd anyagokat). Az anyagokat csak egyszer kell a BraiLab-on rögzíteni, utána előállítható belőle a látó és a braille anyag is. A „Braille Laboratóriumban” így találkozik a két kódolási rendszer Metabraille formában.

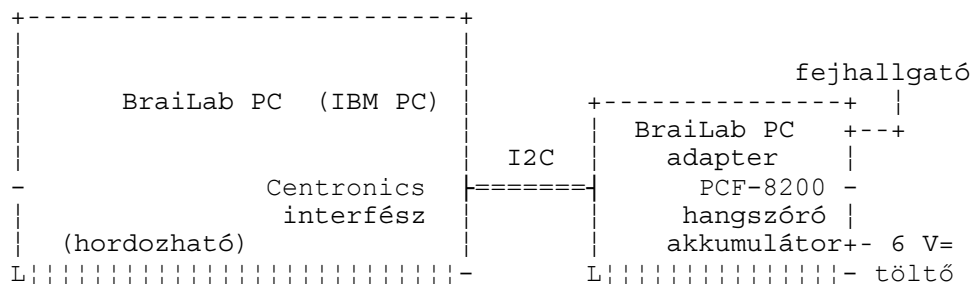
### 6.3. A BRAILAB PC

#### 6.3.1. A BRAILAB PC SEGÉDESZKÖZ FELÉPÍTÉSE

A BraiLab PC tekinthető a legkevésbé cél segédeszköznek. A hardver mindössze egy jó minőségű hangszóróba épített PCF-8200 típusú beszéd szintetizátor áramkörből áll végerősítővel, szűrővel kiegészítve. A szintetizátor áramkör az I2C interfészen keresztül csatlakozik az IBM PC centronics portjára úgy, hogy annak csak néhány adatvezetékét használja. A PC felőli csatlakozó átmenő érintkezős, így csatlakoztatható ugyan erre a portra egy másik eszköz is. Erre példa egy teletext olvasó elektronika, melynek segítségével vakok is hozzáférhetnek a képűrség információihoz. Némi módosítással az átmenő csatlakozóra nyomtató is dugható.

A BraiLab PC beszélő adapter fogyasztása olyan csekély, hogy a hangszóróba akkumulátor is építhető, ekkor a hálózati adapter töltőként is üzemel. A hordozható beszélő adapterem bármilyen, hordozható IBM PC kompatibilis számítógéphez kapcsolható, és ezzel kialakítható hálózattól független üzem is. Ennek a segédeszköz használhatósága szempontjából igen nagy jelentősége van.

A BraiLab PC adapter hardverének és az I2C szoftver illesztésének kialakításában Lukács József és Endre voltak a segítségemre. Az I2C interfész előnye a könnyebb gyárthatóságban van, hátránya viszont az időigényesebb programozás. Készült a beszélő adapternek szabályos centronics változata is.



A hordozható BraiLab PC segédeszköz felépítése.

#### 6.3.2. AZ ÚJ TEXT-TO-SPEECH RENDSZER BEÉPÍTÉSE AZ MSDOS OPERÁCIÓS RENDSZERBE

A BraiLab PC nem bolondbiztos segédeszköz. A PCF-8200 alapú, új, kötetlen szótáras beszédelőállító rendszerem segédeszközbe építéskor a legfontosabb célkitűzésem az volt, hogy az IBM PC lehetőleg minden funkciója használható legyen hallás útján. A grafikus funkciók kivételével ezt a célt sikerült is megvalósítani. Még a színek és ablakok is érzékelhetők beszéd útján bizonyos kényelmi szintig.

A BraiLab PC beszélő szoftvere az MSDOS interfészek nélkül 20 Kb-ot hang adatbázissal, 4 Kb-ot külső kivételiszótárral együtt 45 Kb-ot. Operációs rendszerbe illetve intelligens ernyőfelolvasó rendszerrel együtt 53 Kb-ot. Ez a program tárrezidens, ezért a rezidens részeket assembler nyelven írtam. A rezidens résszel kapcsolatot tartó részeket C nyelven programoztam.

A rezidens program több szinten kapcsolódik az MSDOS operációs rendszerhez. Legfelső fájl szinten úgy, hogy a COM4 virtuális periférián keresztül irányíthatók át az üzenetek CWI kódkészletben a mesterséges beszéd számára. BIOS szinten a legfontosabb ernyőre író rutinok adatai irányíthatók át egyenként vagy kombinálva a text-to-speech programhoz megszólaltatás céljából. A BraiLab Plus-nál ez a fecseg üzemmódhoz hasonlítható. A BraiLab PC rezidens programnak saját billentyű kezelő interrupt programja van, melynek segítségével a billentyű

echózás tartható kézben. Az ernyőre direktben kiírt adatokat a program off-line rendszere segítségével lehet szintetizálva kijelezni.

A BraiLab PC program ablakozás elvű. Ez azt jelenti, hogy mind on-line-ban, mind pedig off-line-ban egy aktuális ablakban történik az echózás. Az ablakok részben fixen letölthetően jelölődnek ki, részben automatikusan képződnek. A karakterek, szavak és sorok meghangosítása szövegszerkesztés közben, parancs sor beírásakor és off-line-ban hasonlóan működik, mint a BraiLab Plus-ban.

Az ablakok kijelölésével, melyet segédprogrammal végezhet a látássérült felhasználó, az ernyőn való tájékozódást lehet elősegíteni. Ha az egyes alakzatokban a vak felhasználó off-line módban olvassa le az ernyőt, akkor az egyes kurzor pozíciókat a rendszer eltárolja, hogy később könnyen vissza lehessen találni a korábbi ablakok pozícióiba.

Az üzemmód váltások, beleértve az összes beszédparaméter állítási lehetőségeket, escape szekvenciák segítségével történnek. Az üzemmódok közül fontos kiemelni a sorvég be/ki, szóköz be/ki, és a részletez be/ki szekvenciákat. Ha egy fájl folyamatos felolvasása történik pl. a

COPY fájl.név COM4

paranccsal, és ha a szöveg formátumára is kíváncsiak vagyunk, akkor a sorvégeket figyelembe kell venni, mint intonálási határokat, valamint a szóközöket is ki kell mondatnunk (sorvég-be, szóköz-be). A betűhibákra koncentrálhatunk, ha a jelöletlen hasonulásokat nem mondatjuk ki (részletez-be). A szemigrafikus jelek kimondását tilthatjuk le a kód-ki szekvenciával.

Az off-line üzemben nemcsak a karakterek betűzése történik, hanem az ernyő szín attribútumairól is információ kapható. A színek és az ablakok automatikus felismerése egy intelligens ernyőfelolvasó rendszer részei lesznek (fejlesztés alatt). Az ablakok és a kurzor pozícióiról szóbeli koordináta kijelzés kérhető.

A beszéd szintetizátor programozása nem interruptos. Ennek egyik legfőbb oka a programból történő I2C interfész kezelés. A billentyűk kezelése viszont megszakítható, és egy újabb billentyű megnyomása meg is szakítja a beszédet. A PCF-8200-as áramkörrel meg lehetett csinálni, hogy nem állítom le a beszédet megszakításakor, hanem azonnal az újabb szöveg frame-jeit adom a szintetizátor számára. Ehhez feltétlenül szükséges az igen gyors real time text-to-speech rendszer megléte.

Javítás és gyors gépelés közben a finom részletekig kidolgozott beszéd paraméterek összeolvadnak, és a magyar nyelv fonetikusságából adódóan a szavakat a gépeléssel teljesen egyidőben hallhatjuk. A szöveg szerkesztése ezáltal abszolút természetességgel végezhető hallás útján. A szövegszerkesztő javító, vezérlő gombjainak hatására a vak felhasználó mindig a szükséges információt hallja, magának a segédeszköznek a kezelésével igen kevés a gondja.

Ahhoz, hogy a billentyűzés közben ne késsen a hangos echó, arra is figyelemmel kellett lennem, hogy a hosszú magyar zárpfelpattanások idejét szókezdő pozícióban kiszűrjem (pl. kettő). Normális megszólaltatáskor persze, pl. fájl felolvasásakor) a zárpfelpattanások szókezdő néma fázisának ideje nem csökkenthető le, mert ez az idő nagyon jellemző az egyes mássalhangzókra.

A BraiLab PC autorepeat funkciója eltér az előző BraiLab-okétól. A billentyűzetekben helyben programozhatók az autorepeat idők (IBM PC/AT), ezért az automatikusan ismételt billentyűk megszakítják a beszédet. A beszéd sebességével és az időzítések állításával megválasztható a felismeréshez szükséges elégséges idő. A billentyűk kiosztása vakon is átdefiniálható segédprogrammal. Az egykezes gépelést a BraiLab Plus-hoz hasonlóan oldottam meg, kivéve az Alt gombot, mely folyamatosan zárható. Így bevihető minden kód a kiegészítő billentyűkről, ill. az operációs rendszer újra tölthető egy kézzel is.

## 7. fejezet. A BRAILLE RÖVIDÍRÁS ÉS A TEXT-TO-SPEECH KIVÉTELSZÓTÁRA

Molnár Pál egyetemi szakdolgozatának eredményeit feldolgozva kutattam a magyar nyelv entrópiáját. Ez két dolog miatt volt szükséges. Egyrészt a PCF-8200 alapú text-to-speech rendszert ki akartam egészíteni beszéd frame kivétel szótárral (ennek programozási munkái elkészültek), másrészt Magyarországon kialakulóban van a korábbiakhoz képest több rövidítést tartalmazó braille rövidírási szabvány.

Mindkét feladat számára fontos tudni a magyar nyelv információ tartalmát. Ki kell tudni választani azokat a szavakat, szókezdeteket és szóvégződéseket, melyeket érdemes külön frame-ekkel szintetizálni, ill. érdemes külön braille pont kombinációkkal rövidíteni.

Jelinek Schneider módszere változó hosszúságú kódszavait úgy választja ki egy adott adattömegeből, hogy a kapott kódszótár információ tartalma megegyezik a kiindulási szöveg információ tartalmával. A kódszavakról elmondható, hogy azok az egyértelmű kódolás miatt rendelkeznek a teljesség és a prefix tulajdonságokkal. A teljesség azt jelenti, hogy minden forrás string-nek van olyan kezdőszelete, mely már szerepel a kódszótárban. A prefix tulajdonság pedig azt jelenti, hogy bármely kezdőszelet csak egyféleképpen kódolható.

Egy szöveg információ tartalmának becsléséül a Jelinek - Schneider által bizonyított tétel szolgál, mely szerint:

$$\frac{H(\text{szöveg})}{\log_2(\text{abc-elemszáma})} = \frac{H(\text{kódszótár})}{\log_2(\text{szótár-elemszáma})}$$

Ahol a  $H(\text{szöveg})$ ,  $H(\text{kódszótár})$  az elemzett szöveg ill. a kódszótár entrópiáját jelöli.

A magyar irodalmi nyelv egy mintáján lefuttatva a változó hosszúságú blokkos kódolási eljárást, különböző hosszúságú kódszavakat kaptam. A rekurzív program hosszabb idejű futásával (5 óra VAX 8550-es gépen) már több teljes szót is tartalmazó kódszót is generáltam. Ezeknek a vizsgálata sokkal egyszerűbb volt, mint a több száz kilobájtos minta szövegé.

A braille rövidírást illetően érdekes eredményt adott a szóvégzódések analízise. Kiderült, hogy a 77 rövidítést tartalmazó hármas szintű magyar braille rövidírásban a „szor”, „szer”, „ször” rövidítése helyett sokkal előnyösebb lenne a „nak”, „nek” végzódéseket rövidíteni. A szabványosítás folyamata még ma is tart, így fontos kiemelni a braille publishing rendszerem visszacsatoló hatását is ebben a folyamatban. A teljes szavak rövidítései közé az entrópia vizsgálataim alapján mindenképpen felvenni javasoltam a „nemzet”, „mert” és „ott” szavakat.

A braille rövidírási fordító programjaim valamint a BraiLab PC-ben működő külső kivétel szótár kezelése céljából kialakítottam rövidítéseket leíró szabály rendszert. A külső kivétel szótár szerepe a text-to-speech rendszerben kettős; egyrészt a leggyakoribb rövidítéseket és idegen szavakat mondatja ki magyarul, másrészt a kettős hangzók összevonásának szabályait korrigálja. Ez utóbbira lássuk a következő két példát:

vízszint=víz-szint

kilencszáz=kilenc-száz

Mind a két példában az egyenlőség jel bal oldalán álló string összetett szavakban is állhat, így elegendő a szótárban csak ilyen rövid alakot tárolni. A következő sorban arra láthatunk példát, hogy a helyes ejtés (ss) céljából a tömörség érdekében a belső, metabraille kivétel szótárra történik hivatkozás:

kisebb=<

A rövidítéseket leíró szabályok (ld. mellékletben) illusztrálásához vegyük a magyar Eötvös családi név helyes kiejtése céljából beírt sort:

!eö-=ö

Ennek értelmezése, ejts minden szókezdő eö stringet ö-nek, ha még követi valami. A BraiLab PC helyes kiejtéséhez a következő kivételeket kellett bejegyezni:

!brai-=bráj

!pc.=pécé

Itt is a szókezdő brai stringet kell megfelelően ejteni, ami persze a braille szóra is érvényes. A PC stringet csak különálló szóban kell „pécé”-nek ejteni, pl. a „hapci” szóban nem. A kivétel szótár jelenleg 4 Kb-át mérete még szükség szerint növelhető, mert a real-time feldolgozás indexelt kezelés miatt nagyon gyors. A felhasználó saját maga is kibővítheti, átírhatja a kivétel szótárt a saját szövegeinek és céljának megfelelően. Kiegészítő programom segítségével bármikor, a teljes rendszer letöltése nélkül módosíthatja a felhasználó az éppen aktuális kivétel szótárt, akár a beszélő program futása közben is.

## **8. fejezet.**

### **A BRAILAB GÉPEK ALKALMAZÁSAI**

A HomeLab számítógép BraiLab alapgéppé alakításához szükséges modul gyártási jogát a Hardszoft Gmk ill. a dombóvári Color ipari szövetkezet kapta meg. Együttesen mintegy 400 BraiLab alapgépet értékesítettek vakok és gyengénlátók számára.

A BraiLab Plusz gépek gyártását a KFKI végezte, és 60 darabot értékesítettek belőle. A BraiLab PC adaptereket a KFKI Direkt Kft. gyártja, és eddig 70 darabot adott el elsősorban látássérülteknek.

Szűcs László egyéni vállalkozó, engedéllyemmel sorról sorra, átvette a MEA-8000-re írt text-to-speech programomat, ill. változtatás nélkül használja beszéd adatbázisomat a PC Voice nevű termékében, mely egy PC-be dugható kártya. Mivel az elért tudományos eredményeim az ő eszközében is hasznosultak a gyakorlatban, így együttesen számolva 550 berendezésem szolgálja a magyar vak és gyengénlátó felhasználókat. Ez a magyar vakok számát tekintve világviszonylatban is kimagaslóan nagy elterjedtséget jelent.

Vak felhasználók állítása szerint, a BraiLab megjelenése forradalmi változást jelentett az aktív munkát végző és az oktatásban résztvevő sorstársaik életében. Az alábbiakban tételesen felsorolom, milyen felhasználóknál, és hogyan hasznosul az az 550 segédeszköz, melyek egy évtizedes kutatás-fejlesztési munkám gyakorlati eredményeit demonstrálják.

- A Vakok Általános iskolája, a BraiLab alapgépet iskola-számítógépének tekinti. Oktató-programokat készítettek szinte minden tantárgyhoz, és a tantervükbe is beillesztették a többi BraiLab géptípus oktatásával együtt. Braille írású tananyagaikat a BraiLab Plusz rövidírási rendszerével készítik.

- A közép fokú iskolákban megjelent BraiLab Pluszok hatására egyre több vak diák jelentkezett különféle egyetemekre, legtöbbször az ELTE TTK programozói szakára. Sokan, pl. a Jogi karról kikerülve BraiLab Plusz segédeszközük segítségével könnyebben tudtak munkába állni.

- Az ELTE TTK-n folyó vakok programozói oktatásába a BraiLab Pluszok és BraiLab PC-k szervesen beépültek. A braille nyomtatót itt is a BraiLab Plusz vezérli. A szegedi JATE-n a PC voice-t és a BraiLab PC-t használják.

- A Magyar Vakok és Gyengénlátók Országos Szövetségében, valamint a Vakok Állami Intézetében a BraiLab gépeimet használják munkarehabilitációs tevékenységükben. A braille nyomdában rövidírási rendszerrel készítik a nyomdai termékeik legnagyobb részét. Nyilvántartási rendszereikben is a BraiLab gépcsalád tagjait hasznosítják.

Összefoglalva, a BraiLab gépek, a beszélő segédeszközök tekintetében szinte kizárólagos, de az egyéb taktilis számítógépes segédeszközöket is beszámítva döntő részét képezik a magyar vakok oktatásában felhasznált segédeszközöknek, az általános iskolától egészen az egyetemekig, valamint a felnőtt korban látásukat veszített vakok elemi rehabilitációs oktatásáig.

A BraiLab gépcsalád tagjai új munkalehetőségeket teremtettek látássérült embereknek, és az oktatás minden szintjére beépültek. A magyar vakok számítógépes braille nyomtatása kizárólag a BraiLab Plusz gépeimre épült. Munkám eredményeképpen elmondható, hogy Magyarországon a vakok és gyengénlátók számítógépes segédeszköz használata és braille nyomtatási technikája semmiben sem marad el az iparilag legfejlettebb országokétól.

## VÁLASZ

### Gordos Géza, a műszaki tudomány kandidátusa bírálatára

Először is meg szeretném köszönni bírálónak, hogy sok elfoglaltsága mellett ilyen részletes bírálatot írt. Egyaránt köszönöm a pozitív és a negatív megjegyzéseket. Nagyon sokat tanultam belőlük.

Munkám „A BraiLab beszélő számítógépcsald” műszaki alkotás, tehát a minősítés alapja maga az alkotás. Az alkotás ismertetését a műszaki leírás és azok mellékletei segítik elő. Bírálóm kifogásolta, hogy a leírás nem tartalmazza a téma tudományos helyzetképét és a rokon eredmények pontos bemutatását. A leírásom első fejezete mellett a D, E és I mellékleteim is tartalmaznak a vakoknak szánt segédeszközökről áttekintést. Beszédtechnológiai helyzetkép leírásom megmagyarázza, hogy miért kellett egy kitaposatlan ösvényen elindulnom, azaz speciális beszéd kimenetet készíteni vakokat szolgáló segédeszközben.

Az irodalmi hivatkozásokat érintő bírálatot elfogadom. Formailag nem volt számomra világos, hogy egy műszaki leírásban is benne kell-e lennie vagy sem. Szeretnék utalni a mellékelt cikkeimben lévő hivatkozásokra. Például a G mellékletben 15 hivatkozás található. E válaszomhoz is mellékelek fontos irodalmi hivatkozásokat. A formáns szintézis tekintetében cikkeimben és e válaszomban inkább egy átfogó műre hagyatkozom, melyet Olasz Gábor írt [1]. Úgy érzem, hogy nála nem írhattam volna jobb történeti áttekintést. Egy vakok számára készült műszaki tudományos alkotásról szóló leírásban a beszéd-szintézisről csak mint egy alkotó elemről kellett hogy írjak.

Nem formáns elvű beszéd-szintézis, a nagy tárigény miatt, alkotásomban nem jöhetett szóba, ezért nem tekintettem át a hullámforma kódolással elért hazai és külföldi eredményeket. Bírálómmal egyetértően jegyzem meg, hogy az hullámforma kódolással jobb beszédminőség érhető el, mint formáns szintézissel ld. [2] a LIAWOX-ról.

A beszédminőség értékelésében a hangsúlyt a beszéd érthetőségére helyeztem. Erre utaltam a leírásom 2.2. fejezetében (ld. a fejezet első bekezdését), valamint az L mellékletben (2. oldal teteje). Az érthetőségi vizsgálatok néhány speciális eredménye az 1. számú mellékletben található. A beszéd természetességi mutatóit kevésbé tartottam fontosnak, ezért nem tértem ki részletesebben erre a kérdésre a leírásban és a mellékelt cikkeimben. Helyette viszont a vak felhasználó kognitív terhelésének csökkentését tűztem ki célul a PCF-8200 alapú második text-to-speech rendszeremben. A több mint 150 BraiLab PC felhasználó folyamatosan küldi pozitív szubjektív visszajelzéseit, melyből arra a következtetésre juthatunk, hogy célokat elértem.

Köszönöm, hogy bírálóm felhívta leírásom hiányosságára a figyelmemet az üzembiztonság tekintetében. Ezt a kérdést ketté kell választani. Az egyik oldalon vizsgálendő a szoftver komponens, míg a másikon a hardver komponens megbízhatósága. A szoftver megbízhatóságát határesetekre és közbenső esetekre alaposan bevizsgáltam minden új változat kiadása előtt. A szoftver megbízhatóságára soha semmilyen kifogás nem érkezett. A hardver esetén is külön kell vizsgálnunk a BraiLab Basic, BraiLab Plus gépeket, mint amelyek több hardver elemet tartalmaztak, valamint az akkori helyzet miatt több keleti gyártmányú elemmel rendelkeztek, és a BraiLab PC hardverét, melynek felület szerelt nyomtatott áramköre mindössze 12 négyzetcentiméter.

A BraiLab Basic megbízhatósága a vele egy kategóriába eső Sinclair gépekkel kb. azonos volt. A BraiLab Plus ugyancsak a CP/M kategóriájú gépekkel vette fel a versenyt az üzemi-képesség tekintetében. Erre utal ezen gépek szerviz statisztikája. A különbség persze az volt, hogy azokat az eszközöket a magyar vakok nem tudták használni. A BraiLab PC hardverére a nagy gyártási szám ellenére soha nem érkezett semmiféle kifogás.

Bírálom kitért a szintetizátort vezérlő parancsok, vagy ahogy a leírásomban írtam, a frame-ek tárolásának redundancia kérdésére. Valóban nagyon fontos, hogy elkerüljük a redundanciát. Szeretném meggyőzni bírálómat és a tisztelt bizottságot is, hogy az átlapolásos technikám szinte redundancia nélküli tárolást tesz lehetővé. A MEA-8000 alapú rendszeremben 216 frame-t használtam, melyek közül 18-szor mutatott a diád mutatóm már egyszer használt frame-re. Ez 8.3%. A PCF-8200 alapú text-to-speech programomban összesen 3242 beszéd frame-t tároltam, s közülük mindössze 42-szer mutatok már használt frame-re. Ez csak 1.3%-os redundancia! Ezeket az eredményeket a 2. számú melléklet alapján kaptam.

A kisebb tárméret 4 Kbájtos és a nagyobb 20 Kbájtos értékek természetesen a diád-matrix-al együtt értendők ld. műszaki leírásom 30. old. második bekezdés ill. a 39. oldal harmadik bekezdés alján és 43. oldal ábra alatt.

A redundancia és az összetett pointerok kérdésében érdemes kitérni a HUNGAROVOX, SCRIPTOVOX és a MULTIVOX text-to-speech rendszerek tárolási módjára ill. azok összevetésére a BraiLab-bal. A HUNGAROVOX-ban 370 hangszületet (frame-t) használtak fel összesen 2457 helyen. ld. [1] 291-292 oldalak. Ebből világosan látszik, hogy elvben sem lehet olyan nagy részletességgel tárolni paramétereket mint a BraiLab PC 3200 különböző, sehol sem ismétlődő frame-jeivel.

Felmerülhet a kérdés, hogy ha annyira triviális lenne az átlapolásos tárolási mód és az összetett pointerok használata által megnyerhető memória terület és gazdag hanganyag tárolási lehetőség együttes előnye, akkor miért nem használták korábban mások ezt a technikát.

A MEA-8000-el működő SCRIPTOVOX ld. [3], valamint a PCF-8200 alapú MULTIVOX rendszerben is mindössze 255 különböző hangszületet alkalmaztak az alkotók ld. [4]. A BraiLab PC-ben általam használt 3200 különböző frame-et összehasonlítva a 255 ill. 370-nel szeretném kihangsúlyozni a gazdag hangzás elvi lehetőségét. Csak ilyen sok frame tárolásával lehet gazdag hanganyagot előállítani. Ezeket az apró formáns változásokat az előadásom alatt is fogom demonstrálni.

Nem szeretném természetesen ezzel a HUNGAROVOX, SCRIPTOVOX és MULTIVOX rendszereket negatívan jellemezni. A HUNGAROVOX egyértelműen történelmi jelentőségű alkotás, a legelső magyar formáns text-to-speech rendszer. A BraiLab gépeimben használt kötetlen szótáras rendszerek legfontosabb tulajdonsága a tömörség mellett azon képessége, hogy vakok segédeszközeibe történő alkalmazásra lettek speciálisan megalkotva. Uthalatok itt a hangalak echo céljaira kialakított diádokra, a néma fázisok kihagyására a billentyű echo idején, a mondatjelek mihamarabbi jelzésére az intonálással, a teljesen azonos idejű megszólaltatásra, a Metabrilille kódolásra, a hangos nagybetűjelzésre, az ernyő attribútumok hangzással történő kiemelésére, a sorvégelek intonálással történő kimondására, a szóközök speciális szintetizált hangjára, a felolvasott szóközök szünetekkel történő jelzésére (ezen lehetőségek kikapcsolására), az igen gyors beszédmegszakítási lehetőségre, a különböző némítási funkciókra vagy a kétszintű kivételszótár kezelésre. Ezek nélkül nem lehetett jó vakügyi segédeszközt készíteni.

Bírálom 5.6. pontjában hiányolja a percepció vizsgálati eredményeimet. A bírálatot köszönettel elfogadom. Szeretném jelezni, hogy a speciális percepció vizsgálataim eredményeit az 1. számú mellékletem tartalmazza. A valós és gépi ejtés összehasonlítására előadásomban fogok



kitérni. Ugyancsak az előadáson, de azon kívül is lehetővé teszem a kötetlen szótáras szintetizátoraim beszédhangjának megismerését. A védésem színhelyén 9 óra 30 perctől, de kérésre máskor is, megtekinthetők, meghallgathatók az eszközeim.

A MEA-8000 hangadatbázisról a leírásban kifejtettem, hogy az elsődleges szempont az igen kicsi tárméret volt azért, hogy a számítógép kapacitása megmaradjon a felhasználó számára. Ez természetesen kihatott a beszéd minőségére, ami mindezek ellenére mégis igen jól használható volt vakok segédeszközében.

A GY-A, TY-A hangátmenetek hosszának számolásakor bírálónak bizonyára elkerülte figyelmét az, amit a MEA-8000 adatbázisról írtam műszaki leírásom 25. oldalán, mely szerint „a mássalhangzó-magánhangzó diádokban nem tároltam a mássalhangzó jellegzetes formánsait, csupán annak átmeneti részeit...”, ezért jutott bírálatának 5.8. pontjában arra a következtetésre, hogy e hangátmeneteim rövidebbek az átlagosnál. A szókezdő GY és GY-A diádok együttes hossza a MEA-8000 esetén nálam 152 msec.

A PCF-8200 adatbázisban kifogásolt K hangba történő átmenetek összevonása szerintem jogos, mivel itt csupán a második szakasról van szó! Ez a K hang zárpfelpattanás előtti néma fázisa. Erre utal az utolsó néma frame zaj paramétere, mely előkészíti a zárpfelpattanás hangzó zaját. Bírálóm ezen megállapításának talán az lehetett az oka, hogy az első szakaszok a 32. oldalon nem látszanak. Ott csupán az A-K átmenet első szakasza van feltüntetve. (ld. pont ill. kettőspont a címke után). Az 5.1.2. fejezetem írja le a PCF-8200 adatbázisában alkalmazott új átlapolásos módszeremet: ha a címkét pont zárja le, akkor az az első szakaszra utal, ha kettőspont, akkor a másodikra.

Elfogadom bírálóm megjegyzését a „jellegzetes hangzású” ill. „sarkítani” meghatározásaimat illetően. Jellegzetes hangzású G hang alatt azt értem, ahogy a G hangot próbáljuk kiejteni magánhangzó kíséret nélkül. Ez a GŐ hangsorban ejtett G hanghoz áll a legközelebb. A sarkítás alatt azt értem, hogy olyan intenzitással hozom létre például a szintetizált szóköz-G-szóköz jelsorozatnak megfelelő hanghatást, mely köznapiban nem hangzik el sem szókezdő, sem pedig szóvégi állapotban. Az „artikulált” ejtést az előadásom alatt demonstrációval szeretném megmagyarázni.

Az intonáció megvalósítása ugyan kötelező már az 1980-as években, de senki nem gondolt arra, hogy a mondatvégi írásjeleket hogyan lehet a mondatintonációval vakoknak effektíven jelezni. Ezt bírálóm az 5.15. pontban pozitívan is értékelte. A beszéddallam automatikus előállításáról először 1986-ban publikál Olasz Gábor [5] A HUNGAROVOK rendszerben az írásjel utáni számmal kellett jelezni a kívánt dallamformát ld. [1] 303-305. oldalak. Az általam bevezetett Ho Há Mi Me Ki szabály nagyon tömör, automatikus feldolgozást adott az esetek nagy többségére már 1985 szeptemberében (ld. 194 639 lajstromszámú szabadalom „intonálási egységek” „Eljárásra és berendezés hang szintetizálására előnyösen beszéd és/vagy ének szótár nélküli, a vezérléssel teljes egészében azonos idejű előállítására” A bejelentés napja: Budapest, 1985.09.18., „Útmutató a Braille-Lab beszélő Basic-jének használatához” 20. old. Budapest 1986 február.)

A „mikrointonációval” kapcsolatban fontosnak tartom kiemelni, hogy előttem senki nem tárolt még ilyen részletességgel formáns frekvencia és sáv szélesség stb. paramétereket diád táblázatokban a magyar nyelv szintetizálására. Ez adja a gazdag hangzását a PCF-8200 alapú rendszeremnek. Megjegyzem, hogy a legtöbb formáns szintetizátor rendszer a 4. és 5. formáns paraméterek tárolásával nem is foglalkozik.

A BASICVOX-ra történő közvetett utalásom megtalálható az L melléklet 3. oldalán. Ott a rendszer nevét nem említettem.

Bírálom a kivételszótárak új megadási módszerét „érdekesnek” tartotta, a jó használhatóságát a felhasználók tapasztalatai alapján véli bizonyítottnak. Örülök, hogy bírálóm fontosnak tartja a felhasználói tapasztalatok visszacsatolását. Szeretnék beszámolni arról, hogy az ember-gép kapcsolati szempontból igen egyszerű, de hatékony módszerem a felhasználóktól érkezett visszacsatolások alapján nagyon sikeresnek mondható. Igen széles körben használják vak emberek a legkülönbözőbb szövegek kényelmes felolvasatása céljára.

Az általános észrevételekre reflektálva ki szeretném egészíteni az új műszaki tudományos eredmények című fejezetem első pontját a következőkkel: 1984-ben nem volt számomra elérhető olyan magyar text-to-speech rendszer, melynek mérete kisebb lett volna 8 Kbájnál, és melyet vakok segédeszközében használhattam volna.

Nagyon köszönöm, amit bírálóm a műszaki alkotás szellemében írt. A 2.1. pontot bírálóm az összefoglalójában kiemelte „Alkotás-elemnek tekintem a komplex (taktilis és beszélő) segédeszköz felvetését és kialakítása általános szempontjainak megfogalmazását” Ugyancsak kiemelte bírálóm bírálatának 5.1. pontját: „Az (5.4. szerinti fenntartással) jelentős alkotásnak tekintem a MEA-8000 és a PCF-8200 formáns szintetizátorok segítségével megvalósított BraiLab, BraiLab Plusz és BraiLab PC beszélő számítógépeket és az ezek létrehozását segítő fejlesztő rendszereket.” Az 5.4. pont bírálatára a redundancia-mentességet érintette. Erre úgy érzem kielégítő választ tudott nyújtani a 2. mellékletem. Kiemelte bírálóm bírálatának 6.1. pontját is, mely így hangzik: „Önálló alkotásnak tekintem az alapvető vezérlési (írás, törlés, hibaüzenetek, szóköz, stb.) és bonyolult (pl. szövegszerkesztés) számítógépi funkciók kiegészítését alaposan megtervezett hang- illetve beszéd effektusokkal”. Bírálóm következő, általa kiemelt pontja a 6.6., mely így hangzik: „Alkotás-elemnek tekintem a TTS rendszer beépítését a PC típusú gépcsalád MSDOS rendszerébe” Bírálóm ezen pozitív értékelései olyan alkotásokat és alkotás-részeket emelnek ki, melyek új műszaki tudományos eredmények. Ezek vagy a magyar nyelv specifikumait a vakok segédeszközei céljára figyelembe vevő egyedi alkotások, vagy még külföldön is elfogadott új tudományos eredmények (pl. hangalak echó). Ezek talán elegendő érvet szolgáltathatnak az összefoglalóm 4. pontjának elfogadásához.

Bírálom ugyancsak dicsérően emelte ki bírálatának 5.13. pontját. „Az alkotás fontos elemének tekintem azt, hogy egy menetben kezeli az ASCII és a Braille kódok konverzióját, és hogy a Metabrilie-ből történő konverzió tekintettel van a beszéd-sebesség változtatására.” A Metabrilie kódolás és annak alkalmazása, melyet bírálóm ugyancsak „Értékes alkalmazás-elemnek” tekint, egy eddig mások által, sem belföldön sem külföldön nem használt új módszer. Remélem, hogy ezek alapján elfogadható bírálóm számára az összefoglalóm 3. pontja.

Bírálom 5.2. pontjában pozitívan ír az átlapolásos szintetizátor parancs-sorozat tárolásomról. Szeretném remélni, hogy a redundancia számításaim segítik abban bírálómat, hogy összefoglalóm 2. pontját is el tudja fogadni.

Elfogadom bírálóm kritikáját a pontatlan idézetek és hiányos utalások tekintetében. Ezek listáját bírálóm kérésére a 3. mellékletemben adom meg.

A szabadalmi és egyéb társszerzős cikkek részarányával kapcsolatban utalni szeretnék a TMB számára mellékelte írásbeli nyilatkozatra, melyben a társszerzők kijelentik, hogy a „BraiLab beszélő számítógépcsalád” az én alkotásom.

Befejezésül még egyszer megköszönöm bírálómnak alapos bírálatát, valamint azt, hogy elfoglaltsága ellenére mégis vállalta a bírálat elkészítését.

*Budapest, 1993 október 8.*

Arató András

## Mellékletek\*

1. Percepció vizsgálati eredmények
2. Redundancia számítások
3. Irodalmi hivatkozások listája

---

\* A mellékletek az elektronikus változatból hiányoznak *(a szerk.)*