

A vizuális beszédértékelés alkalmazásának történeti áttekintése a hallássérültek beszédoktatásában

HÉGELY GÁBOR ÉS KOCSOR ANDRÁS

MTA-SZTE Mesterséges Intelligencia Kutatócsoport
SZEGED, Aradi vértanúk tere 1., H-6720

kocsor@inf.u-szeged.hu

With the financial support of the Hungarian Ministry of Education (IKTA Grant No. 00051/2001), a new computer-aided software package, called SpeechMaster was developed for speech impediment therapy. The program – for therapy use – is freely available on the Internet at www.inf.u-szeged.hu. The core of it is a real-time phoneme recognizer which converts sounds into easily understood graphical elements such as flickering graphemes. This effect efficiently supports the development of correct articulation for the hearing impaired. Since the application of visual speech techniques is an old notion, we also decided to review the history of such developments. After the review section we highlight the significance of our result, which may prove useful in the future.

1. Bevezetés

A siketek oktatásának történetében sok olyan eszköz és eljárás látott napvilágot, amely több-kevesebb eredménnyel mind a siketek kiejtésének javítására törekedett. A szurdopedagógusok mindig is érezték, hogy a célcsoport kiejtési problémáit, ha gyökeresen nem is oldaná meg, lényegesen csökkentené egy olyan megoldás, amellyel a siket saját kiejtett hangját ellenőrizhetné, és kiejtését önmaga javíthatná. Ennek legeredetibb formája, hogy a hallássérült hallását javítjuk olyan mértékben, hogy képes legyen a beszédhangokat felfogni. Ehhez kialakultak a hallást támogató, a hallás minőségét javító eszközök. Kezdetben volt a hallócső. Napjainkban pedig egyénre szabható, nagy teljesítményű digitális hallókészülékek állnak rendelkezésre.

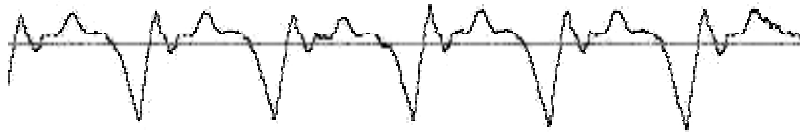
Igen súlyos halláskárosodás esetében sikeres cochlea implantációs műtéttel ma már ki lehet alakítani mesterséges hallást, aminek birtokában egy tanulási periódust követően tökéletesen elsajátítható a hangos beszéd.

Amikor e technikák alkalmazására nincs lehetőség, akkor a beszélőszervek izmainak mozgását a siketeknek látás útján kell megtanulni. Azonnal adódik az ötlet, hogy tegyük a siketek számára láthatóvá a hangzókat oly módon, hogy a szemükkel ellenőrizhessék saját kiejtésüket hasonlóan ahhoz, ahogyan mások a hallásukkal képesek erre. A „látható beszéd” technikájának megoldása manapság is intenzíven kutatott területnek számít, ahol a módszereket a megjelenítés szempontjából három generációba sorolhatjuk.

1.1 Első generációs megoldások: az oszcillogram használata

A beszédjel hosszú távú hullámforma jellegű információit jelenítik meg egyszerűbb műszerek, például oszcilloszkópok segítségével. Ezek az eljárások a szurdopedagógia gyakorlatában nem tudtak igazán elterjedni, mert a tanuló által ejtett beszédhang oszcillogramja bonyolult ábra formájában jelent meg (1. ábra), és hamarosan be kellett látni, hogy ezeknek a rezgéseképeknek a finom eltérései -

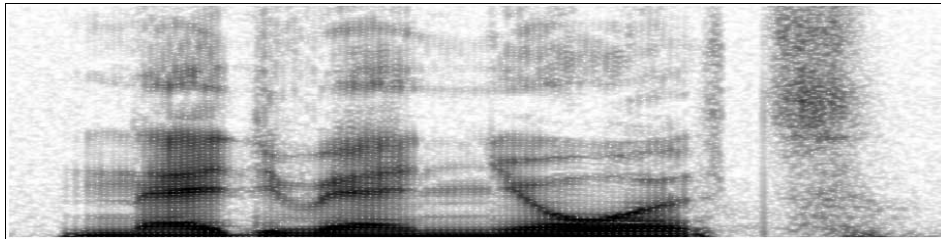
a látszólagos hasonlóság ellenére is - nagyon kevés rendszerességet mutatnak. Továbbá e képeknek az artikuláció pillanatában történő kiértékelése az ember számára megoldhatatlan feladatnak bizonyult, mivel szemünkkel csak gyengén tudunk megkülönböztetni nagyon nagy hangkülönbségnek megfelelő hullámforma-változásokat.



1. ábra
Egy felnőtt férfi által ejtett „a” hang oszcillogramja.

1.2 Második generációs törekvések: a spektrogram megjelenítése

Az előző generációhoz képest annyiban történt előrelépés, hogy a beszéd analízisében a spektrumszámítás módszerét követve információredukcióra van lehetőség, ugyanis a spektrális elemzésnek alávetett beszédhangból kinyerhetők a felismerés szempontjából releváns információk (Gósy, 1982, Gordos, Takács, 1983). Habár a spektrumképek a fonémák (Hégely, Sára, 1971, Hégely 1999) jellegzetes karakterisztikáját jól mutatják (2. ábra), a képek kiejtés pillanatában hallásunkhoz hasonlóan történő kiértékelésére az ember számára itt sincs lehetőség. A gyakorló személy csak a kiejtés befejeztével a saját és egy etalonnak számító spektrumkép hasonlósága alapján tud következtetni kiejtésének helyességére.



2. ábra
A „harminc” szóhoz tartozó spektrogram egy felnőtt férfi ejtésében

1.3 Harmadik generációs eredmények: az automatikus beszéd felismerés alkalmazása

Az eddigiekkel szemben nem közvetlenül a hang oszcillogramja, vagy spektruma jelenik meg a gyakorlás során, hanem valamilyen származtatott, az ember számára könnyebben kiértékelhető vizuális információ. A vizuális kijelzés az elhangzással azonos időben történik, általában egy igen egyszerű ábra, például a betű képeinek formájában. Mind a beszéd analízise, mind a beszédhangok azonos idejű és fonémaszintű feldolgozása önműködően megy végbe. (3. ábra)



3.ábra

Az „a” beszédhang bemozdításakor megjelenő graféma.

2. Történeti áttekintés

A siketek ezen problémái, és így a „látható” beszéd technikájának kidolgozása az egész emberiséget érintő nemzetközi feladat. A probléma megoldására különböző technikákat dolgoztak ki az elmúlt 100 év során (Fónagy, 1978, Hégely, 1980). Mivel a kutatások magyar, azon belül szegedi vonatkozása jelentős, összefoglalónk is a történet ezen szegmensével foglalkozik.

2.1 Bell kísérletei

A telefon feltalálójának életrajzi adataiból tudjuk, hogy siketnémák oktatásával is foglalkozott, felesége is hallássérült volt. Bell megpróbálta valamilyen módon láthatóvá tenni a hanghullámokat, amelyek mintául szolgáltak volna a süketeknek a beszédtanulásban.

Bell arra gondolt, hogy egy csőből kiáramló gáz mennyiségét hang hatására rezgő, rugalmas membránnal szabályozza. A gázt meggyújtva a lobogó láng képét egy forgó tükörrel folyamatos hanghullámgörbévé akarta alakítani. A nyelv egyes hangjairól így fényképfelvételeket készíthetett volna. A süket gyermek maga elé téve ezeket a fényképeket, a membránra beszélve addig próbálná a helyes kiejtést eltalálni, amíg a forgó tükörben éppen olyan hanggörbét nem látna, mint a fényképen. Ezáltal a süket gyermek úgy tanulhatna összehasonlítás útján, hogy nem a füle, hanem a szeme irányítaná.

Bell hamarosan rájött, az 1800-as évek végén ezt a problémát nem lehet megoldani. Ötletét több évvel később Franciaországban valósították meg az oszcilloszkóp segítségével. A Groningenben 1950-ben megtartott Süketnéma-oktatásügy Nemzetközi Kongresszusán beszámoltak arról, hogy a párizsi Institut National des Sourds-Muets már 1938 óta használja az oszcillográfot süketnémák kiejtésének javítására (Zemmach, 1956).

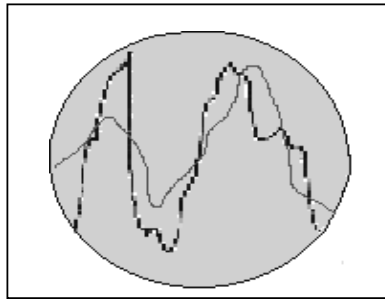
2.2 A soproni oszcillogramos kísérletek

Hazánkban, az ötvenes években a siketek soproni intézetében Tóth Béla alkalmazta elsőnek az oszcilloszkópot a hallássérültek beszédoktatásában. Figyelemre méltó, hogy a soproni intézet tanára kísérletei során már fonológiai

szemléletet követett. Tanulmányozta Laziczius Gyula műveit (Laziczius, 1944). Ismerte Tarnóczy Tamás 1941-ben közreadott „A magyar magánhangzók akusztikai szerkezete” című dolgozatát (Tarnóczy, 1941), amelyben a szerző leírja a magyar fonémák matematikai módszerrel feltárt akusztikai szerkezetét.

Ennek következtében ismertté váltak számára olyan fogalmak, mint a fonéma és variánsai, a formáns, a formáns energiája és frekvenciája. Tarnóczy tanulmányának igazolásaként Tóth Béla is azt tapasztalta, hogy a magánhangzók igen jellegzetes görbéi figyelhetők meg az oszcilloszkóp képernyőjén. A különböző személyek ejtésében annak ellenére, hogy az egyes fonémák eltérő képet adnak, mégis nagyfokú hasonlóság figyelhető meg.

Ezért gyermekek és felnőttek kiejtéséből vett nagyszámú hanggörbe segítségével a magyar nyelv fonémáinak tipikus oszcillogramjait igyekezett megrajzolni. Az így kialakított mintagörbék celluloid lemezen a képernyő elé helyezve mintául szolgáltak a siket gyermeknek.



4. ábra

*Az „ö” hang oszcillogramja. A minta és a képzés közben megfigyelhető rezgéskép.
A magyar nyelv fonémáinak Tóth Béla által kialakított rezgésképei.*

Tóth Béla kísérleteiről 1958-ban számolt be (Tóth, 1958). Az oszcilloszkóp alkalmazásának ötletét a Művelődési Minisztérium újításként fogadta el, és a hallássérültek iskoláit oszcilloszkópokkal szerelték fel. Hogyan is alkalmazták a műszert a kiejtés javításában? A gyermek és a nevelő elhelyezkedett a készülék előtt. A nevelő felszólította a siketnéma gyermeket a készüléken lévő mintahanggörbe utánzására. (5. ábra) A tanár hangoztatás közben természetesen állandóan utasította a gyermeket a klasszikus kiejtéstanítási módszerek ismeretében. Amennyiben a gyermek követte az utasításokat, elérhető volt, hogy a két görbe közeledett egymáshoz, és ezáltal a kiejtés észrevehetően javult.



5. ábra

A saját kiejtés ellenőrzése a bemutatott minta és a saját produktum képeinek összehasonlításával történik.

A közleményben leírt és Bell elképzelésére emlékeztető gyakorlási mód a fonetizálásban alapesetnek tekinthető, amely napjainkig meghatározta mind a kutatások irányát mind az alkalmazás módszereit.

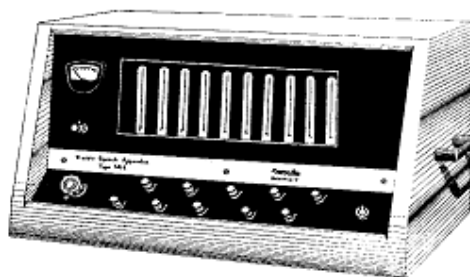
Az oszcilloszkóp alkalmazása a szurdopedagógia gyakorlatában nem tudott elterjedni. Egy-egy magánhangzó etalonjának számító rezgéskép és a beszéd közben rövid időre felvillanó oszcillogramok hasonlóságának mértékét nehezen lehetett figyelemmel követni. Ezeket a képi információkat artikuláció közben a gyermeknek kellett volna értékelni, ami számukra szinte megoldhatatlan feladatot jelentett. Az oszcillogramok ezért az önellenőrzéshez nem nyújthattak értékelhető információt.

2.3 A színeképelemzés módszerének alkalmazásai

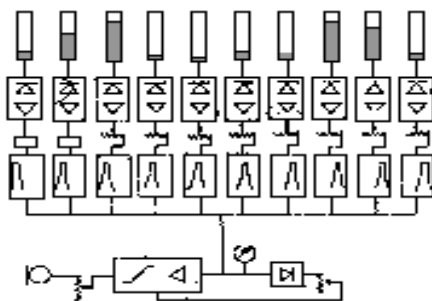
A fonetikai kutatások eredményeinek alkalmazásában új fejezet kezdődött, amikor szélesebb körben ismertté váltak a szonagrafikus vizsgálatok eredményei (Molnár, 1970, Hégely, 1975). A fülünk működésében is fellelhető spektrum módszerét kezdték alkalmazni. Lehetővé vált, hogy a beszédhangok hangszíneképét tegyék láthatóvá (Tarnóczy, 1966). A beszéd teljes információtartalmából így ki tudták emelni a hangzók azonosításához lényegesnek tűnő akusztikai információkat. Ezt követően jelentek meg a különböző „Visible Speech” készülékek és hangindikátorok, amelyeket kifejezetten a siketek és beszédhibások oktatásához fejlesztettek ki.

2.3.1 A KAMPLEX típusú visible speech apparátus

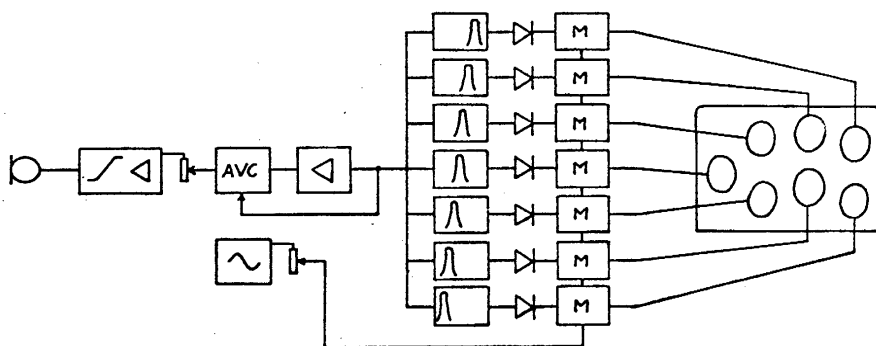
A hangzók spektrumképei először analóg jelként kerültek a képernyőre. A készülék hosszított ködfénylámpái folytonos fényerőváltozással jelezték a formánsok energiaszintjét a különböző frekvenciatartományokban.



6. ábra
A KAMPLEX típusú visible speech apparátus az 1960-as évekből.



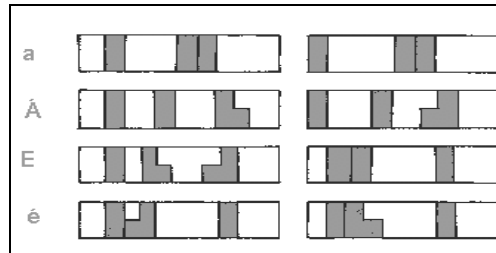
7. ábra
Az „é” hang spektrumképe a készülék képernyőjén.



8. ábra
A spektrumképek speciális megjelenítésével lehetőség kínálkozott a fonémák taktilis érzékeltesére is.

2.3.2 A Danavox készülék

A gyakorlati alkalmazás során az is kiderült, hogy a formánsok folytonosan változó energiaszintjeinek képi ábrázolása, ha azt bizonyos módon egyszerűsítik (kvantálják), akkor ez további információredukálást eredményez, amely könnyíti a fonémák vizuális azonosítását az ember számára. Például elegendőnek bizonyult a formánsamplitudók két-három szinten történő kijelzése. A Danavox készülék képernyőjén a fonémák spektrumképei 10 oszlopból és 3 sorból álló mátrixmezőben jelennek meg. (9. ábra)



9. ábra

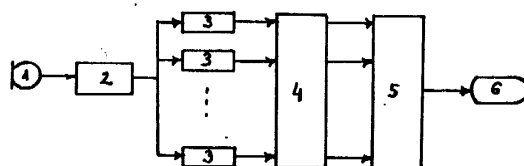
2.3.3 A KAMPLEX rendszer továbbfejlesztése Szegeden

1971-ben az Akusztikai Világkongresszus szegedi szekciójában Hégyes Gábor és a József Attila Tudományegyetem Kibernetikai Laboratóriumából Sára Attila olyan készüléket mutattak be, amely segítségével a hangzók spektrumképeit az eddigiektől eltérő módon lehetett ábrázolni (Hégyes, 1972, Tarnóczy, 1973, Hégyes, Sára, 1976, Subovits, 1977). A folyamatos beszéd frekvencia-spektrumra bontását a Kamplex cég egy kismértékben átalakított készülékével végezték. A készülék, 10 egymást féloktáv közönként követő frekvenciaszelektív, analóg kimenetét 10 komparátor áramkörre vezették, azaz voltaképpen digitalizálták. Az analóg digitálkonverter jelzőlámpái folyamatosan jelezték, hogy melyik frekvenciasávban érte el a formánsamplitúdó a beállított értéket. E lámpasorral párhuzamosan további 10 lámpa szolgált az akusztikai jelenségből kézi nyomógombbal vezérelt tetszőleges időpillanatban vett minta tárolására.

A tanuló előtt két sor parányi izzó volt. Az egyik sor az előzetesen rögzített mintát jelezte, a másik sort pedig a folyamatos beszéd működtette. Artikuláció közben arra kellett figyelni csupán, hogy a kijelölt hangzóra jellemző és előzően felgyújtott izzók párja gyulladjon ki. A két lámpasor egyidejű megfigyelésével az önellenőrzés valamelyest egyszerűbb lett.

2.4 Egy korai /magyar/ valós idejű beszédfelismerésen alapuló „látható beszéd” rendszer

A KAMPLEX rendszer szegedi továbbfejlesztése folytatódott. Voltaképpen ez volt az első harmadik generációs törekvés. A rendszert telefonkulcsokból és relékből felépített logikai mezővel bővítették, amely a digitalizált információk rögzítését és a folyamatos beszéd feldolgozását tette lehetővé. A fonémák spektrumképében megfigyelhető akusztikai megkülönböztető jegyek figyelembe vételével, a kapcsolók különböző állásával programozni lehetett az egyes magánhangzók felismeréséhez szükséges áramköri formákat.



10. ábra

A kísérleti berendezés blokkvázlata.

A mikrofonról érkező jelek (1) információtartalmát a dinamikakompresszor (2), a frekvenciaszelektív-erősítők (3) és az analóg–digitál átalakító (4) redukálja, illetve elvégzi a jelek átalakítását (digitalizálását). Ehhez csatlakozik a felismerő egység (5). A felismert hangzókat a kijelző egység (6) jeleníti meg. A vizsgált hangból a kellő időpillanatban kézi mintavételezéssel vett nagyszámú minta alapján kellett megállapítani hogy melyik frekvenciasávban jelentkeznek a leggyakrabban az egy-egy fonémához tartozó formánsok, azaz milyen lámpák villanak fel a leggyakrabban. Ezeket, a magánhangzókról nyert digitális információkat lehetett felhasználni a felismerést szolgáló logikai áramkörök kialakításához.

A fejlesztők soros kapcsolással háromállású kapcsolókat (telefonkulcsokat) telepítettek 125 Hz-től 8kHz-ig terjedő 10 frekvenciatartomány oszlopában, a magyar magánhangzónak megfelelően 9 sorban.

- A különböző magánhangzók leggyakrabban jelentkező felharmónikusaihoz tartozó telefonkulcsokat felső állásba lehetett állítani. Ha a kívánt fonéma formánsai álltak elő a kulcshoz tartozó frekvenciatartományban, akkor a kapcsoló a kérdéses fonéma sorában zárt állásba került (K-állás). Ez jelentette a felismeréshez szükséges releváns ún. „kell” értéket.

- A kulcs alsó állása azt jelentette, hogy a fonéma azonosításához nem szabad a kulcshoz tartozó frekvenciatartományban számottevő energiának jelentkezni, mert az más fonéma meghatározó eleme lehet. A jel hatására ekkor a kapcsolónak bontani kellett a vizsgált fonéma frekvenciatartományában az áramkört (T kapcsoló állás). Ez a kapcsolóállás szolgált a formánsátfedések kiküszöbölésére.

- A kulcs középső állásában a kapcsoló rövidzár állapotba került (R kapcsoló állás). Ekkor a kérdéses frekvenciatartomány viselkedése a vizsgált fonéma felismerése szempontjából közömbös volt. Ezzel ki lehetett szűrni a beszédhangok azonosításához nem lényeges, ún. redundáns információkat tekinthető felhangokat.

Hz	150	250	300	480	680	1000	1400	1900	2500	4000
Á	R	R	R	T	K	K	K	R	K	R
A	R	R	R	R	K	K	T	R	K	R
E	R	R	T	K	K	R	R	K	T	R
É	R	K	K	K	T	R	R	K	K	R

11. ábra

A kapcsolók három állásával kialakított felismerő minták.

Ha egy fonémára jellemző spektrumkép-karakterisztika jelentkezett, akkor teljesült a fonémának az „És” kapcsolókkal kialakított logikai definíciója. Ekkor

a fonémának megfelelő sorban a különböző frekvenciasávok oszlopaiban jelentkező információk egyetlen jelet adtak, tehát a felismert fonémának megfelelő sorban villant fel a lámpa. Ezáltal a folyamatos beszédben az elhangzással azonos időben is lehetett jelezni a helyesen hangzó bármelyik magánhangzót. Nem kellett a hangzókat külön-külön bekapcsolni. A folyamatos beszédben felismert fonéma kijelzése ezentúl akár szegmensekből felépített betűkép is lehetett (Bolla, 1982).

2.5 A szegedi fejlesztések folytatása

A számítógépek megjelenésével szinte önmagától adódott a gondolat, hogy a fonémákat jellemző kézi mintavételezéssel gyűjtött akusztikai információkat a 9 magyar magánhangzónak megfelelő csoportosításban digitálisan tároljuk (Kassai, 1998). Ilyen módon az eddigieknél lényegesen több információval ún. elfogadási sávokat lehetett egy-egy fonéma felismeréséhez rögzíteni, szemben a felismerő kapcsolómező programozásával, ahol alig pár mintát tudtunk csak figyelembe venni. A program által vezérelt, és egy-egy tanuló kiejtéséhez igazodó felismerés eredménye lehetett a fonetikai tükörben megjelenő betű. Ennek az elgondolásnak egy működő változatát a Művelődési Minisztérium igen jelentős anyagi támogatásával sikerült megalkotni.

A kézi mintavételezéshez 10 kHz frekvenciasávban, 12 frekvenciatartományban dolgozó meredek sávszűrők és 30 dB dinamikasávban, 10 komparálási szinten megjelenő információsorozatot lehetett rögzíteni. A készülék képernyőjén egy-egy elhangzó fonémáról egy 10-szer 12-es mátrix celláiban kaptunk információt. Minden frekvenciasávban, három oszlopban ledeket telepítettünk. A zöld oszlopsor a folyamatos elemzést mutatta, a piros sorban gombnyomásra lehetett mintákat gyűjteni. A sárga izzósorra lehetett megjeleníteni a korábban tárolt mintákat. A spektrumképekből gyűjtött és tárolt minták alapján a mintaillesztést is számítógéppel lehetett elvégezni, és így a memóriában személyenként tárolt nagyobb számú adat felhasználásával lehetett a magánhangzók önműködő felismerését megvalósítani. Tökéletes kiejtés esetén a magánhangzó betűképe egy félig áteresztő tükör mezőjében tűnt fel.

Az elért eredményt elismerve az Országos Találmányi Hivatal 1984-ben mintaoltalmat adott a működő készülékre (Hégely, 1985).

A fejlesztőmunkában azonban továbbra is súlyos gondok jelentkeztek:

- Bizonytalanul működött a sávszűrőkből felépített analóg üzemmódú hanganalizátor,
- Továbbra is kevésnek bizonyult az a felhasználható információ, amelyet a hangzók formánsstruktúrájáról a felismerő mátrixon meg lehetett jeleníteni.
- Nem vált be az az elgondolás sem, hogy a folyamatos beszédből kézi mintavételezéssel, gombnyomással gyűjtött mintákkal ki lehet alakítani a felismerésben jól alkalmazható etalonokat. A magánhangzók felismeréséhez ugyanis, a spektrumok 30-90 mill. sec alatt lecsengő, állandósult szakaszaiból

kellett volna információkat gyűjteni. Ilyen rövid idő alatt ezt kézi mintavételezéssel természetesen nem lehetett megvalósítani.

- Ahhoz, hogy a fonémaszintű és azonos idejű felismerés biztonságosan működjön az adatbázis kialakításához és működtetéséhez más megoldás kidolgozására van szükség. A hangzó beszédből valós időben vett nagyszámú minta gyűjtéséhez, és feldolgozásukhoz korszerű matematikai eljárásokra és számítástechnikai módszerekre lett volna szükség.

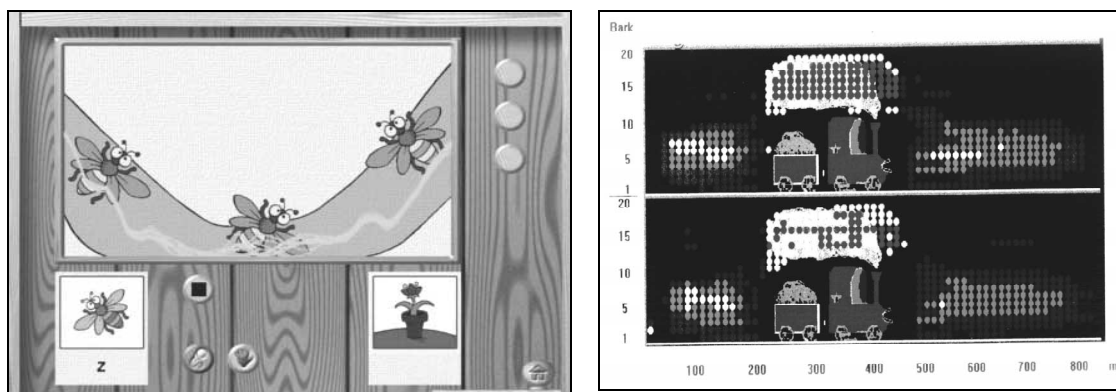
3. A mesterséges intelligencia fejlődésének hatása a látható beszédet megvalósító rendszerekre: az új generáció

A számítógépek rohamos fejlődése közvetlenül hatott a mesterséges intelligencia és azon belül is a gépi tanulás fejlődésére. Sokkal komplexebb matematikai algoritmusok kifejlesztésére és alkalmazására nyílt lehetőség (Tóth, 1958, Hégyel, et al., 1984). Ennek következményeként időközben a látható beszéd megvalósítására nagy számú intelligens algoritmus vált alkalmassá.

Ahogy korábban láttuk, az oszcillogram és a spektrogram valós idejű direkt emberi kiértékelésére nincs lehetőség, ezért a számítógépnek kell az ember számára gyorsan feldolgozható vizuális formában megvalósítani a kiejtés helyességének kontrollálási lehetőségét. Ilyen, a kiértékelést segítő, megoldási módszerekre már láttunk példát az előző szakaszban a KAMPLEX rendszer szegedi továbbfejlesztésénél és az azt követő korai számítógépes megoldásnál, azonban e rendszerek hatékonysága a felhasználás szempontjából nem volt kielégítő. A gyorsabb számítógépek megjelenésével újabb megoldások születtek (Hégyel, 1999, Hégyel, 2001, Vicsi, 2002, Kocsor, Tóth, 2004).

3.1 A VARÁZSDOBOZ

Vicsi Klára vezetésével a BME TTT Beszédkusztikai Laboratóriumában kidolgozták a VARÁZSDOBOZ csomagot (Vicsi, 2002), amely tartalmaz valós idejű kiértékelési lehetőséget adó beszédhang-korrekcióra, illetve nem valós idejű kiértékelésre használható szókiejtés-korrekcióra alkalmas modulokat. Az első esetben a spektrum idő szerinti görbét kell egy adott sávban tartani (A), míg a második esetben játékos grafikai elemekkel ellátott spektrogramkép segíti a kimondott szó összehasonlítását egy etalonnak tekinthető bemondással (B).



(A)

(B)

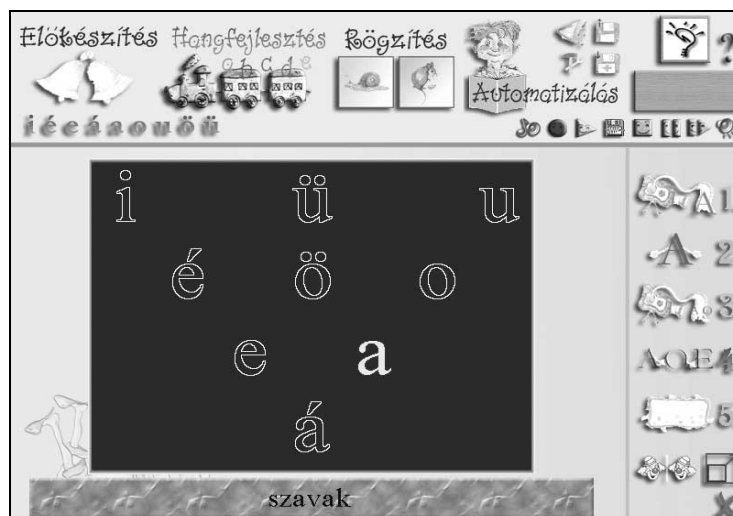
12. ábra

3.2 A BESZÉDMESTER

Egy másik beszédjavítás-terápiai technológia kifejlesztése a Szegedi Tudományegyetemen készült a jelen publikáció szerzőinek elképzelése alapján. A fejlesztés egy IKTA projekt formájában történt együttműködve a kaposvári Óvoda Általános Iskola Diákotthon és Gyermekotthon (Siketek Iskolája) munkatársaival. Informatikusok, modellalkotó matematikusok, fejlesztő pedagógusok és szurdopedagógusok közösen láttak hozzá a BESZÉDMESTER szoftver beszédjavítás-terápiai moduljának kidolgozásához. Ebben a cikkben a kapcsolódó kutatások és a rendszer részletes leírását mellőzzük, hiszen az számos helyen megtalálható (Paczolay, et al., 2002, Paczolay, et al., 2004, Sejtes, Kocsor, 2004, Kocsor, Tóth, 2004), azonban a beszédjavítás-terápiára alkalmas újszerű vizuális beszédkiértékelési eredmények leglényegesebb momentumait áttekintjük.

A BESZÉDMESTER rendszerben a vizuális kijelzés mindig az elhangzással azonos időben történik, általában igen egyszerű ábra, például a betű képe formájában. Mind a beszéd analízise, mind a beszédhangok azonos idejű és fonémaszintű feldolgozása önműködően megy végbe. Tehát ez a rendszer nem a mikrofon előtt közvetlenül elhangzó egyedi beszédhangot jeleníti meg, hanem azt, amit „megért”, azaz a hanghoz társított fonémát.

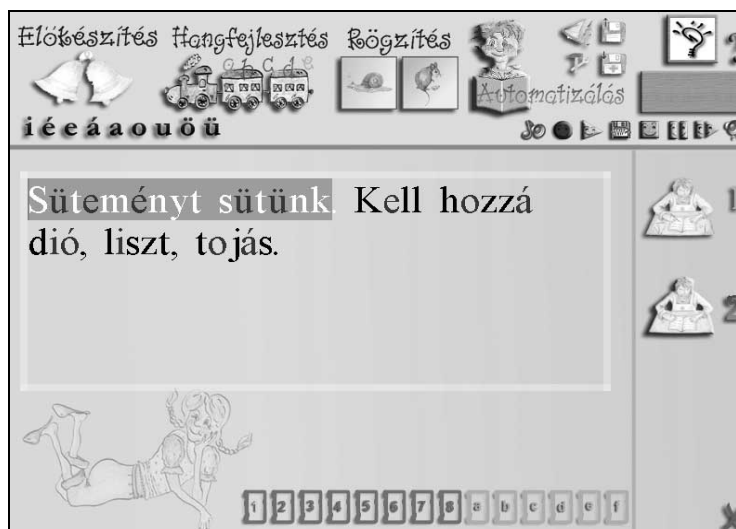
3.2.1 Valós idejű magánhangzó-felismerés



13. ábra

A képernyőn megjelennek a kiválasztott magánhangzók körvonalai. Hangképzés közben egyszerre több betűkép is világíthat, a betűk fényereje a megfelelés mértékével arányos.

3.2.2 Szöveg olvasásának gyakorlása valós idejű kontrollal



14. ábra

A gyakorlás folyamán a teljes szöveget kisebb egységekben olvashatjuk be a mikrofonba. Az egységeket az írásjelek definiálják. Az aktuális részt rózsaszín kiemelés jelzi, ezen felül a figyelt betűk pirossal jelennek meg.(13. ábra) A hangfelvétel elindítása után a szöveg beolvasásával párhuzamosan – a felismerésnek megfelelően – halad a kurzor (függőleges fekete vonal) a szövegben. Ha olvasás közben a figyelt fonéma megfelelése elér egy előírt szintet, akkor a betű piros kijelölése megszűnik. Adott részt akkor teljesítettünk, ha minden figyelt fonéma kijelölése megszűnik, ekkor automatikusan a következő részre lépünk.

3.2.3 Új helyzet a fonetizálásban: a fonetikai tükröt felváltja a WEBKAMERA



15. ábra

A „Beszédmester” képernyőjén a kiválasztott magánhangzók körvonalain túl megjelenik a webkamera által közvetített kép is. A betűk fényereje a megfelelés mértékével arányos. Követhetők a fonológiai (az elvárt magánhangzó helyett jól érthető másik magánhangzó képzése) és a képzési hibák (magánhangzó ejtése közelít egy másik magánhangzóéhoz), amelyek segítséget jelenthetnek a logopédusoknak.

A kamera megfelelő beállításával tökéletesen lehet szemléltetni az egyes magánhangzók artikulációs mozgulatait. A bemutatott szájrólolvasási kép, hallás nélkül azonban csak megközelítő információt ad arról, hogy a képzett hang mennyire felel meg a magyar nyelv ejtési szabályainak.

A korrekt artikuláció létrejöttéhez, ezért jelezni kell, hogy a tanár és a tanuló szájrólolvasási képéhez az elhangzás pillanatában milyen fonéma tartozik.(14. ábra) Ez valósul meg a felvillanó betűképek segítségével. Foglalkozás közben, amikor a tanár a kamera előtt mikrofonba beszél, akkor a szájrólolvasási képen megjelenő betűképpel nem a saját kiejtésével adott egyedi mintát mutatja be, hanem a gyakorlandó fonémát jelöli meg. Annak eldöntése pedig, hogy a tanuló részéről a fonémának megfelelő hang képzése mennyire volt sikeres, szintén nem a tanár feladata, mert az artikuláció közben feltűnő betű segítségével ez is a program felhasználásával történik.

A tanár közreműködése ezen túl csupán arra szorítkozik, hogy a szájrólolvasási minta bemutatásával segítse a tanulót a program által közvetített nyelvi minta követésében, alkalmazva a hagyományos fonetizálás eddig is jól bevált kiejtéstani módszereit.

A tanuló helyzete is lényegesen megváltozik. Gyakorlás közben az elhangzó beszédhangok spektrumképei helyett az egyedi hangzók fonémacsoportjának megfelelő betűjelek láthatók, amelyek artikuláció közben is jól követhetők. Az önellenőrzés nem a tanár által bemutatott minta és a saját produktum közvetlen összehasonlításával történik, mert a program a tanuló ejtésében is azt jelzi, hogy az mennyire felelt meg a magyar nyelv ejtési szabályainak. Tehát mind a helyes

ejtést, mind a megközelítés mértékét a program fogja jelezni az artikulációval gyakorlatilag azonos időben.

Az eddigi elért eredmények arra utalnak, hogy önműködő felismeréssel fonémahallást tudunk biztosítani a hallásukban gátolt egyének számára. Ennek segítségével szemükkel valóban ellenőrizni tudják kiejtésüket, ahogyan mi a fülünkkel vagyunk erre képesek. Megfelelő előkészítést követően a tanuló a beszédhangok artikulációját nyelvi akusztikai ellenőrzéshez hasonló helyzetben önállóan is tudja gyakorolni. Súlyos halláskárosodás esetén tehát a hallás nélkül is érthető beszédartikuláció kialakítását támogatjuk, nagyothallók és CI (Cohlea implantációs) műtéten átesett egyéneknél pedig meggyorsíthatjuk a fonémahallás kialakulást

Joggal reméljük, hogy a fonémák önműködő felismerésének alkalmazásával lényegesen megváltozik a hallássérültek kiejtési problémáinak megoldása.

Irodalom

- Tarnóczy Tamás** (1941) *A magyar magánhangzók akusztikai szerkezete*. Budapest: Pázmány Péter Tudományegyetem.
- Laziczius Gyula** (1944) *Fonetika*. Budapest
- Süketnéma Oktatásügy Nemzetközi Kongresszusának kiadványa*, Hollandia
- Zemanch, H.** (1956) *Információelmélet*. Budapest: Műszaki Könyvkiadó.
- Tóth Béla** (1958) *Beszámoló a hangfrekvenciái oszcillográfoknak a süketnémák kiejtéstani tanításában való alkalmazására vonatkozó kísérletekről*. Gyógypedagógia (1)
- Magdics Klára** (1965) *Magyar Beszédhangok akusztikai szerkezete, Nyelvtudományi Értekezések*. Budapest: Akadémiai Kiadó.
- Tarnóczy Tamás** (1966) *Megoldható-e az önműködő beszéd felismerés kizárólag elemzés alapján*. Fizikai Szemle, 5.
- Molnár József** (1970) *A magyar beszédanyagok atlasza*. Budapest: Tankönyvkiadó.
- Hégely Gábor, dr. Sára Attila** (1971) *Visible Speech készülék alkalmazása a siketek kiejtéstani tanításában*. VII. Nemzetközi Akusztikai Kongresszus Szegedi Beszédszimpoziuma (Előadás és bemutató)
- Hégely Gábor** (1972) *Rutinvizsgálati mód siketek kiejtési állapotának objektív felméréshez*. Magyar Gyógypedagógusok Egyesülete, Vác
- Tarnóczy Tamás** (1973) *A magánhangzók vizsgálatának és felismerésének problémái*. Kép- és Hangtechnika XIX.
- Hégely Gábor** (1975) *Beszámoló egy Besztercebányán szervezett kiállításról*. (WS kísérlet bemutatása) Budapest: Gyógypedagógiai Szemle
- Hégely Gábor, dr. Sára Attila** (1976) *Egy valós-idejű magánhangzó-felismerő megjelenítő készülék alkalmazása a siketek artikulációs oktatásában*. Budapest: Magyar Nyelvtudományi Társaság.
- Subovits István** (1977) *Bevezetés a beszédakusztikába*. Budapest: Tankönyvkiadó.
- Fónagy Iván** (1978) *Utószó: (Laziczius Gyula: Fonetika)*. Budapest: Tankönyvkiadó. pp 189-206.
- Hégely Gábor** (1980) *A fonetikai kutatások eredményeinek alkalmazása a siketek oktatásában*. MTA Fonetikai Munkabizottság (Előadás)
- Bolla Kálmán** (1982) *A magyar beszéd akusztikai szerkezetének analízise és szintézise*. Magyar Fonetikai Füzetek 10, pp7-21.
- Gósy Mária** (1982) *A szegmentális hangszerkezet percepciójáról*. Magyar Fonetikai Füzetek. Budapest: MTA Nyelvtudományi Intézet.
- Gordos Géza, Takács György** (1983) *Digitális beszédfeldolgozás*. Budapest: Műszaki Könyvkiadó.
- Hégely Gábor, Winkler Tibor és társai** (1984) *Magánhangzó analízátor leírása Országos Találmányi Hivatal 81 090 szám alatt lajstromozott mintaoltalom*.

- Hégely Gábor** (1985) *Látható beszéd – kiejtésitanítás Kommunikáció – nehezített társadalmi beilleszkedés* Tankönyvkiadó
- Kassai Ilona** (1998) *Fonetika*. Budapest: Nemzeti Tankönyvkiadó.
- Hégely Gábor** (1999) *A siketek artikulációs oktatásában alkalmazható magánhangzó –felismerést megvalósító szoftver kifejlesztése*. KOMA XXIII. pályázat,
- Hégely Gábor** (2001) *Súlyosan hallássérültek artikulációjának rehabilitációjában alkalmazott – a magánhangzók azonos idejű felismerését megvalósító – rendszer*. (Előadás és bemutató) MAGyE Hallásfogyatékoság-ügyi Szakosztály 2001.06.22.
- Vicsi K.** (2002) *VARÁZSDOBOZ. Audiovizuális számítógépes beszédfejlesztő program beszédhibás gyerekek részére*. Démoszthenész Hírmondó 13. sz. pp. 8-16
- Paczolay Dénes, Tóth László, Kocsor András és Kerekes Judit** (2002) *Gépi tanulás alkalmazása egy fonológiai tudatosság – fejlesztő rendszerben*. Alkalmazott Nyelvtudomány II. évfolyam 2. szám 55-67.
- Paczolay Dénes, Kocsor András, Sejtes Györgyi, Hégely Gábor** (2004) *A „Beszédmester” csomag bemutatása, informatikai és nyelvi aspektusok*. (Alkalmazott Nyelvtudomány, megjelenés alatt)
- Sejtes Györgyi, Kocsor András** (2004) *A „Beszédmester” szoftver adatbázis-specifikációja*. (Alkalmazott Nyelvtudomány, megjelenés alatt)
- Kocsor András, Tóth László** (2004) *Kernel-Based Feature Extraction with a Speech Technology Application*. (IEEE Transaction on Signal Processing, megjelenés alatt).