

A BESZÉDASSZISZTENS KONCEPCIÓ

¹Czap László – ²Pintér Judit Mária

Miskolci Egyetem

Automatizálási és Infokommunikációs Intézeti Tanszék

3515 Miskolc, Miskolc-Egyetemváros,

¹ egyetemi docens, e-mail: czap@uni-miskolc.hu

² PhD hallgató, e-mail: pinterjm@uni-miskolc.hu

Összefoglalás

A Debreceni és a Miskolci Egyetem közös projektjének célja a hallássérültek beszédmegértését és beszédtanulását elősegítő olyan összetett rendszer kifejlesztése, amely az eddig ismert módszereknél hatékonyabb támogatást tud nyújtani. A projekt előzménye a Debreceni Egyetemen kidolgozott audiovizuális transzkóder a hang vizualizálására, és a Miskolci Egyetemen kifejlesztett „beszélő fej”. A cikkben bemutatásra kerülnek az eddig elért eredmények, a keretrendszer és a projektben résztvevő kutatócsoportok munkái.

Kulcsszavak: beszédfelismerés, beszélő fej, hallássérültek beszélni tanítása

Abstract

A speech assistant system is developed at the University of Miskolc and at the University of Debrecen granted by the European Union. The project aims to help training of deaf and hearing impaired people to speak. The idea of the project has come from a three-dimensional head model for articulation presentation, called “talking head” developed by the University of Miskolc, and an audio-visual transcoder for sound visualization developed by the University of Debrecen.

Keywords: speech recognition, talking head, teaching of hearing impaired people

1. Bevezetés

Beszédjelek alatt érthetünk akusztikus vagy akár vizuális jeleket is. Akusztikus jelként definiálva a beszéd nem más, mint hangnyomás hullámok keltése, azaz beszédhangok kibocsátása. Agyi absztrakció során feleltetjük meg a hangokat fonémáknak, de a beszéd nem csupán fonémák sorozata, hanem fontos a hangsúlyozás a hanglejtés és számos más szuprasegmentális jellemző is [1,2]. Utóbbi esetben a vizuális jelek (gesztikuláció, arcminimika, szájmozgás stb.) feltérképezésével, megértésével, és rekonstruálásával az audiovizuális beszédfelismerés foglalkozik [3].

Az elmúlt évtizedekben a gépi beszédfeldolgozás dinamikusan fejlődött, egyre több alkalmazás jelenik meg. A beszédfelismerést igen sok területen és különböző céloknak megfelelően alkalmazzák. A hallássérültek beszélni tanítását és az önálló gyakorlás lehetőségének megteremtését elősegítő szoftverek megalkotásához elengedhetetlen mind az akusztikus mind pedig az audiovizuális beszédfeldolgozás alkalmazása.

2. A projekt kutatási koncepciójának ismertetése

Az internet lehetővé teszi a beszédasszisztens rendszernél olyan szolgáltatások betervezését, amelyek még a mai számítógépek közül is a nagy számítási kapacitású szerverek rendszerbe állítását igénylik. Ilyen a beszélő fej és a prozódiai jellemzők számítása.

A kutatás elméleti alapját vizsgálatok igazolták, hogy az agyban az akusztikus és vizuális modalitás integrálása optimális a maximális érthetőség elérésére. A hallássérültek esetében minél erősebb az akusztikus jel torzulása, annál inkább támaszkodnak a vizuális jelre. Várakozásunk szerint ez akkor is teljesül, ha a vizuális jel nem a beszélő képe, hanem a hang vizuális átirata. Ennek igazolása fMRI-vel az agyi plaszticitás bizonyításának egy újabb jelentős állomása lehet. Az alábbiakban a fejlesztésben résztvevő kutatócsoportok célkitűzéseit és feladatait ismertetjük.

2.1. Debrecen 1. kutatócsoport (D1)

Az audiovizuális transzkódolás lényege a hangok absztrakt képi jelekké alakítása hangmagasság, a frekvencia-összetevők hangmagasság és hangerősség kódolt módon. A jelek négyzetek, melyek mérete hangerő-függő módon kisebb-nagyobb, a frekvencia-összetevő hangmagasságtól függően szín kódolt: a mély hangok a nagyobb hullámhosszúságú színek (vörös) a magas hangok a rövid hullámhosszúságú színek (kék). A beszéd szempontjából fontos hangtartomány a 125 Hz – 8000 Hz hangmagasság tartomány [4]. Ezt 20 fix frekvenciasávra bontjuk, amelyek megjelenítése a monitor képsíkjának 4x5 szektorában kerülnek megjelenítésre. A beszéd frekvencia komponense tehát nem csak szín kódolt, hanem pozíció kódolt is. Ez már nyújt akkora hangkép különbözőséget a szóképek között, hogy az egymáshoz hasonló hangalakú szavak megkülönböztetése lehetővé váljék. Korábbi tesztek alapján az ismertetett módon a 20 szektor oszlopos és mátrixos formában fog megjelenni.

A hangkártya mikrofonja érzékeli a 125 Hz és 8000 Hz közötti analóg hangokat, melyeket a számítógép digitalizál, majd Fourier transzformációval frekvencia specifikus jelekké alakít és szűr. A hang képi transzformációja valós idejű és így alkalmas a beszéd folyamatos leképezésére. Ez lehetővé teszi, hogy egyidejűleg a tanuló a számítógép monitorját is figyelje, amely a beszélőfej korrekt ajak-nyelv mozgásokkal demonstrálja a helyes kiejtést.

2.2. Debrecen 2. kutatócsoport (D2)

Kutatásunk során alapvető fontosságú az eredmények alátámasztása korszerű képalkotó diagnosztikai és számítógépes képfeldolgozó algoritmusok segítségével. Feltételezzük, hogy a kifejlesztett tanítóprogram alkalmazása az beszéddel összefüggésbe hozható agyi területeken plasztikus elváltozásokat okoz. Vizsgálataink három szinten történnek:

- Strukturális MR képalkotás és diffúziós tenzor képalkotás alkalmazása minden vizsgálat során. A diffúziós tenzor képadatok korszerű számítógépes analízise lehetővé teszi a szöveti anizotróp diffúzió karakterizációját a beszéddel és nyelvvel összefüggésbe hozható területeken, illetve strukturális agyi kapcsolatok indirekt, probablisztikus követését és vizualizációját is.
- Funkcionális MR vizsgálatok, az agyi aktiváció kimutatására, csoportok közötti aktivációs különbségek kimutatására, valamint a tanulási folyamat során észlelhető aktivitásváltozás időbeli követésére. Egy megfelelő aktivációs paradigmával készített fMRI vizsgálati anyag utólagos feldolgozásával kiemelhetők a feladat

végrehajtása során aktiválódott területek. Ezen aktivációs területek paramétereinek (lokalizáció, intenzitás, térfogat) populáció szerinti elemzése valamint a tanulási folyamattal mutatott korrelációjuknak vizsgálata a tanulással indukált funkcionális változások karakterizálását teszi lehetővé.

- A fMRI adatok elemzésével ún. funkcionális konnektivitás térképezhető fel, ami a strukturális konnektivitási adatokkal együtt az agyi hálózatok állapotáról, neuroplasztikus átrendeződéséről hasznos adatot szolgáltat. A funkcionális konnektivitás a spontán szinkronizációval kimutatható agyi nyugalmi hálózat tanulmányozására, e hálózat karakterisztikáinak csoportok közötti eltéréseinek kimutatására, valamint a tanulási folyamat során észlelhető regionális hálózati szerepkörök változásának időbeli követésére használható matematikai modell. A funkcionális konnektivitás és a strukturális konnektivitási adatok együttes elemzése az agyi hálózatok állapotáról, neuroplasztikus átrendeződéséről hasznos adatot szolgáltat.
- A strukturális és funkcionális konnektivitás adatokból az agyi hálózatok matematikai modellezését végezzük, amely globális és regionális hálózati tulajdonságok számolását is lehetővé teszi. Így vizsgálni lehet az egyes régiók kapcsolaterősségét, a hálózaton belüli információ továbbításban betöltött szerepét valamint az alhálózatok kialakulását és ezek laterizációját. A funkcionális hálózati modell kialakítása során a strukturális információval az fMRI-vel megbecsült kapcsolaterősség korrigálható, amivel az ún. effektív hálózati modell paraméterei számolhatók. Az ilyen módon vizsgált regionális hálózati tulajdonságok populációs szintű időbeli követésével a tanulási folyamat és az ehhez kapcsolható funkcionális/strukturális plaszticitás dinamikus hálózati paraméterekkel karakterizálható, azaz a változásokat kvantitatívan is jellemezni lehet.

Képző és képfeldolgozó kutatási lépéseinket a következő csoportokon tervezzük elvégezni:

- jelelő gyermekekben, akik hangzó beszédre képtelenek, a jelbeszéddel (percepció, produkció) aktiválódó agyi területek kimutatása (látókéreg, mozgató kéreg), strukturális és funkcionális agyi hálózatok térképezése;
- jelelő gyermekek (12-14 éves korú) a project tanítóprogramjai segítségével kifejlesztett hangzóbeszéde során aktiválódó jelpálya kimutatása, strukturális és funkcionális konnektivitásában bekövetkező változások időbeli monitorozása;
- cochleáris implantáción átesett gyermekek beszédmozgató agyi vizsgálata.

A multimodális kommunikáció során a nem verbális csatornák (mimika, gesztusok, testtartás stb.) által közvetített információ tartalom egyes kutatók szerint meghaladja a beszédhangból nyert információ mennyiségét. Hallássérülteket segítő rendszernél elengedhetetlen a nem verbális kommunikációs elemek megjelenítése.

A gyakorló minták elfogadása megköveteli a fejmozgás, pislogás, szemöldök mozgás természetes megjelenítését. Hírolvasók és mindennapi helyzetet szimuláló felvételek elemzésével a szupraszegmentális jellemzők függvényében alapkutatásként megalkotható az artikulációt kísérő mozgások modellje. Szubjektív tesztekkel a valódi, véletlenszerű és a megalkotandó modellnek megfelelő minták tesztelésével a modell verifikálása elvégezhető.

2.3. Miskolc 1. kutatócsoport (M1)

A kutatásokkal érintett alkalmazások a hálózat nem szokványos használatát jelentik, a jellemzők gyökeresen eltérnek a böngészés-letöltés jellegű felhasználástól. Újszerű minőségi jellemzőket igényelnek a hálózattól, a hálózati erőforrásokat másként használják.

Az adott alkalmazásoknál lényeges a maximalizált csomagkésleltetés és maximalizált késleltetés ingadozás. Ez az alkalmazások "near real time" jellegéből fakad, a hálózati protokollok újragondolását is jelentheti pl. TCP - UDP. A hálózati erőforrások használata más szempontok szerint optimalizálható, mint a hagyományos felhasználásoknál.

2.4. Miskolc 2. kutatócsoport (M2)

Az artikuláció dinamikus leírása alap kutatás. A beszédet alkotó fonémák statikus jellemzőinek leírása megtalálható különböző hangalbumokban. A paraméterek dinamikus változására fellelhető adatok azonban csak az adott mintaszóra alkalmazhatók. A nem látható (vagy kis részben látható) beszédszervek mozgásáról kevés adatunk van. Léteznek nyelvmodellek, de magyar nyelvű, nyelvmozgásra vonatkozó adatbázis nincs. Az ajakformákat az audiovizuális beszédfelismeréshez létrehozott adatbázisunkból vettük, így a beszélő fej szájmozgása csak minimális korrekciót igényel [3,5]. A nyelvmozgás követéséhez igen nehéz adatokat szerezni, az élethű fejmodellhez elegendő volt a hangalbumokból vett mozgásfázisokon alapuló animáció [6,7]. A hangképzés bemutatásához a transzparens arcú fejmodell nyelvmozgásának pontosítása szükséges, önmegfigyeléssel és a szurdopedagógus-logopédus kutatók bevonásával.

A szájról olvasók érdekében a hivatásos jeltolmácsok közül némelyek nagyon kihangsúlyozva a megértés szempontjából lényeges beszédelemeket, az átlagos beszélőknél lényegesen intenzívebb artikulációval beszélnek. A tanítást ezért nem az átlagos beszélőkkel, hanem ilyen jeltolmácsokkal célszerű végezni. A koartikulációs hatások teljes körű leírása, az egyes jellemzők domináns jellegének meghatározása, a paraméterek közötti interpoláció szabályainak kidolgozása a projekt keretében végrehajtandó feladat.

A tervezett kliens-szerver kiszolgálással és Internet kapcsolattal a gyakorlásra bárhol, bármikor lehetőség nyílik. A számítógép (laptop, asztali számítógép, táblagép) és az okostelefon kijelzőjének eltérő mérete szükségessé teszi a minták megjelenítését különböző nagytításokban. Az okostelefonon lehet, hogy csak a száj közvetlen környezetét és mozgását célszerű ábrázolni, a számítógép monitorán bizonyára megfigyelhető az artikuláció minden részlete az egész fej megjelenítésével is. Tesztekkel vizsgálandó a legelőnyösebb nézet kiválasztása.

Az artikuláció mélységének skálázhatóvá tétele a jelnyelvi tolmács természetesnél hangsúlyosabb szájmozgásától az átlagos intenzitásig a beszélő fej egyik kialakítandó szolgáltatása. A minták legmegfelelőbb tempója is vizsgálandó, a gyakorláshoz a mindennapi életben megszokott, vagy annál valamivel lassúbb beszédtempó tesztekkel vizsgálható. A mintákat különböző beszédtempókkal elő kell állítani, az előrehaladás egyik mutatója a beszédprodukciónak sebessége. Egy beszédsegmentumhoz tartozó artikuláció és vizualizált hang részletes megfigyelését segíti a kijelzés utólagos kimerevítése. A beszédtempóhoz alkalmazkodnia kell az artikulációnak is, hiszen gyors beszédnél a jellemzők kevésbé közelítik meg névleges értéküket. Az interpolációs szabályoknak a beszédtempót is kezelniük kell. A visszajelzéshez a fejmodell érzelmi töltést is társítson az attraktivitás fokozása érdekében.

Audiovizuális transzkóder (AVT) szimultán alkalmazásával a hangzóbeszédet tanuló számára egyszerűsített önkontroll a szurdopedagógus számára pedig, a beszélőfej mellett (amely a beszédprodukciónak mozgás elemeinek vizualizálása), a beszéd komplex követését és a hibás rész(ek) kiragadásával annak könnyebb és látványos javítását szolgáltatja.

A szerver oldali alkalmazás az Internet kapcsolaton keresztül kiszolgálja a kliens oldali alkalmazásokat és elvégzi az eredmények regisztrálását, az említett kiértékeléseket, a soron következő gyakorló minta kiválasztását is. Az elfogadottságot erősíti a visszajelzés változatos tétele a szerveren tárolt üzenetek folyamatos frissítésével. A kliens oldali alkalmazást különböző operációs rendszerekre kell kifejleszteni (Windows, Linux, Android, Windows phone). A kliens oldali alkalmazás eltérő a célcsoport életkorának és állapotának megfelelően.

2.5. Miskolc 3. kutatócsoport (M3)

A tanulás során a referencia kiejtést a szerver vagy a tanár produkálja. A diák ezt igyekszik utánozni az ő aktuális bemondásával. Ezzel rokon probléma merül fel a beszéd gépi felismerésénél: Előre (modellezés segítségével) eltárolt, valóságos beszédből származó beszéd-részleteket (hang, hangátmenet, szó, stb) közül kell a felismerendő beszéd-részlethez leghasonlóbbat megtalálni, és ha a hasonlóság elég nagy, akkor a beszéd-részlet felismertnek tekinthető. A hallássérültek beszélni tanításánál a hasonlóság automatikus ellenőrzése és a visszajelzés generálása alapvető, amely megköveteli egy hasonlósági mérték kidolgozását. A hasonlósági mértéknek monoton összefüggésben kell lenni a hallássérült és halló bemondók által kiejtett hangok, hangkapcsolatok, szavak szubjektív (épen halló emberek által végzett) tesztek átlagos megítélésével (MOS = Mean Opinion Score). A különböző lényegkiemelési és távolság számítási módok elemzésével kidolgozható a szubjektív értékelésnek megfelelő hasonlósági mérték. Ez az alapja az előrehaladás értékelésének és a visszajelzés generálásának. Az értékelés nyilvánvalóan a korábbi eredményekkel összevetve alakítható ki, hiszen ugyanaz a kiejtés egyik tanulónál siker, a másikonál kudarc lehet. Az automatikus értékelés verifikálása érthetőség vizsgálattal történhet.

2.6. Miskolc 4. kutatócsoport (M4)

A technika az utóbbi években rengeteget fejlődött, és számos korábban technikailag nem megoldható probléma, például az interneten keresztüli használat, vagy a prozódia jobb megjelenítése ma már megoldható. A prozódia elemzése és vizualizálása a helyes prozódia elsajátításának nélkülözhetetlen feltétele. A szuprasegmentális jellemzők megfelelő kialakítása az érthetőség és természetesség fontos mutatója. Hallássérültek beszédének egyik legszembevetőbb jellegzetessége a prozódia rossz használata. A prozódiai paraméterek a hangsúly, hanglejtés, ritmus megfelelő lényegkiemelése, normalizálása, időbeli vetemítése kritikus része az oktatásra alkalmas megjelenítésnek [8,9]. A referencia kiejtés és az aktuális bemondás hasonlóságának, eltérésének vizuális megítélése még a kisgyermek számára is könnyen kell, hogy menjen. Ezt a vizuális értékelést automatikus válaszadással is tervezük erősíteni, de az automatikus hasonlóság mérése, itt más, mint az M3-ban felvázoltak. A feladat komplexitása és a folyamatok célszerű allokálása a prozódiai elemzés szerver oldali megvalósítását diktálja, a megjelenítés kliens oldali funkció, az aktuális bemondáshoz tartozó digitális hangminta szerverbe juttatása és az eredmény közvetítése a kliens felé Interneten zajlik.

2.7. Miskolc 5. kutatócsoport (M5)

A gyakorló minták kijelölése a halláskárosodás jellege szerint a kutatásban résztvevő összes szereplő együttműködését igényli. A kliens oldali megjelenésnek a tanuló életkora szerint eltérőnek kell lennie. Egy játékos környezet a kisgyerekeknek vonzó lehet, az idősebbeknek neveléses. A nyelvi és értelmi fejlődés befolyásolja a minták kijelölését is. A hangok, hangkapcsolatok gyakorlása megköveteli az absztrakció és a hangképző szervek koordinált mozgásának képességét. A projekt eredményeképpen létrejövő alkalmazás használata módszertanának kidolgozása és betanítása az iskolai tanárok részére a projekt immanens része. A pedagógusok és a tanulók közösséggé szervezése a tapasztalatok hasznosítása végett az egyik legfontosabb cél.

3. A beszédasszisztens rendszer

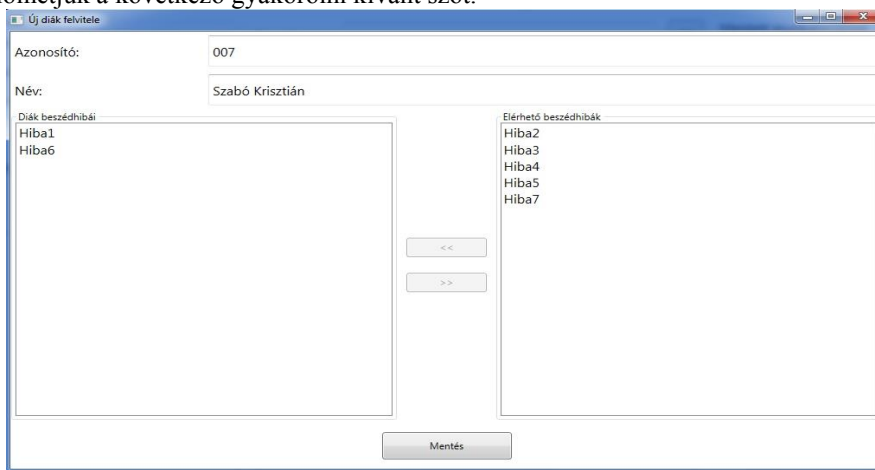
A rendszer tesztelése 2013. szeptemberben kezdődött 14 pedagógus részvételével és eltérő korosztályú és fejlettségi szinten álló gyerekekkel. A rendszer felhasználásának módszertana még nem egységes, a pedagógusok szabadkezet kaptak annak tanórán belüli alkalmazására, hogy tapasztalatokat gyűjtsenek és ajánlást tegyenek a rendszer továbbfejlesztésére.

A rendszer a beszédasszisztens elnevezést kapta. Bejelentkezés után lehetőségük van a pedagógusoknak, hogy kiválasszák, melyik diákkal szeretnének foglalkozni és milyen szavakat akarnak gyakoroltatni. A 2. ábrán látható felületen kell az új diákokat felvenni és egyedi azonosítóval ellátni. A jövőben lehetőségünk lesz az egyes diákokhoz tipikus beszédhibákat hozzárendelni, és így célirányosabban megtervezni a gyakorlást. Már korábban elmentett munkaterületek is újból meghívhatók. Az 1. ábrán látható a kezelőfelület, amin ha kiválasztottuk a diákot (ellenkező esetben a rendszer nem hagy továbblépni), a gyakoroltatni kívánt szavakat és az aktuális szót, továbbléphetünk a gyakorlásra (3. ábra). (A *Súgó* tartalmaz egy rövid ismertetőt a rendszerről, és a lényegesebb részeiről.)

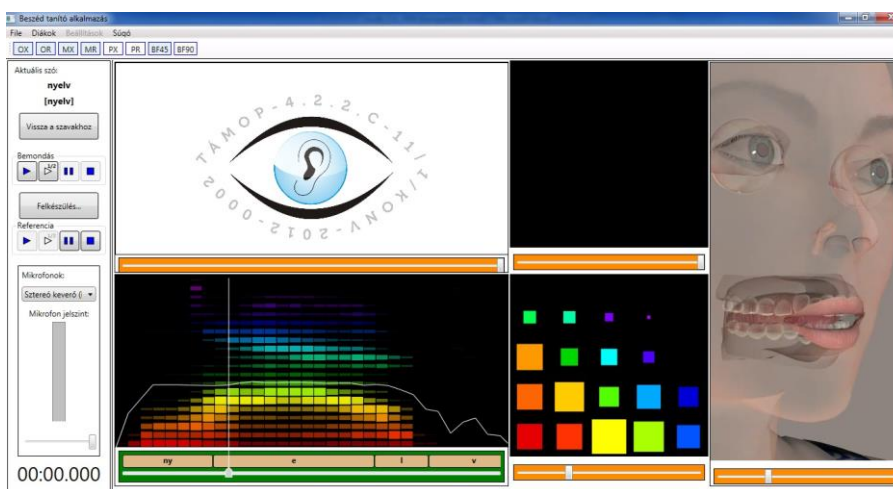


1. ábra. A beszédasszisztens kezdő felülete

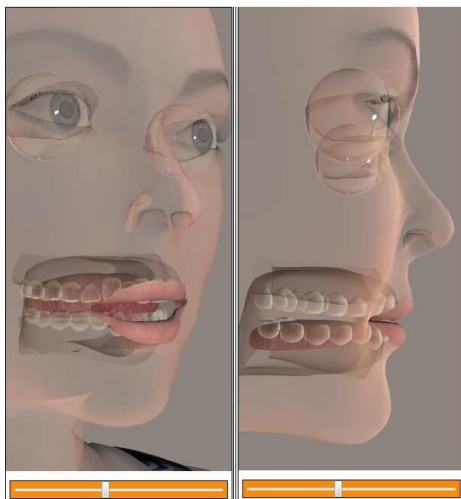
A gyakorló felületen az aktuálisan kiválasztott szónak a referencia bemondását több formában is megjelenik, balról jobbra haladva látható a referencia bemondás oszlopos és mátrixos megjelenítése és a felület jobb szélén a beszélő fej. A beszélő fej transzparens arccal (2.5 fejezet; 4. ábra) két eltérő szögben is megjeleníti a referencia bemondást, amit a kezelőfelület tetején található menüsorban választhatunk ki. Az aktuális bemondást, amit az éppen gyakorló személy rögzít az 5. ábra pirossal bekeretezett részében láthatjuk. Az aktuális és a referencia bemondás normál és fél sebességgel is lejátszható a nyomon követhetőségért. A csúszkák segítségével pedig szinkronizáltan tetszőleges pozícióba állíthatóak a megjelenítések, jobban megfigyelhető az egyes hangoknál a nyelvállás valamint az oszlopos és négyzetes transzformációjuk. (5. ábra). Az egyes komponensek elhelyezkedése és mérete egyelőre kötött, de tetszőlegesen kiválaszthatjuk, melyek azok, amiket a gyakorlás során akarunk használni. A vissza gombbal pedig visszatérhetünk a kezdő felületre, ahol kijelölhetjük a következő gyakorolni kívánt szót.



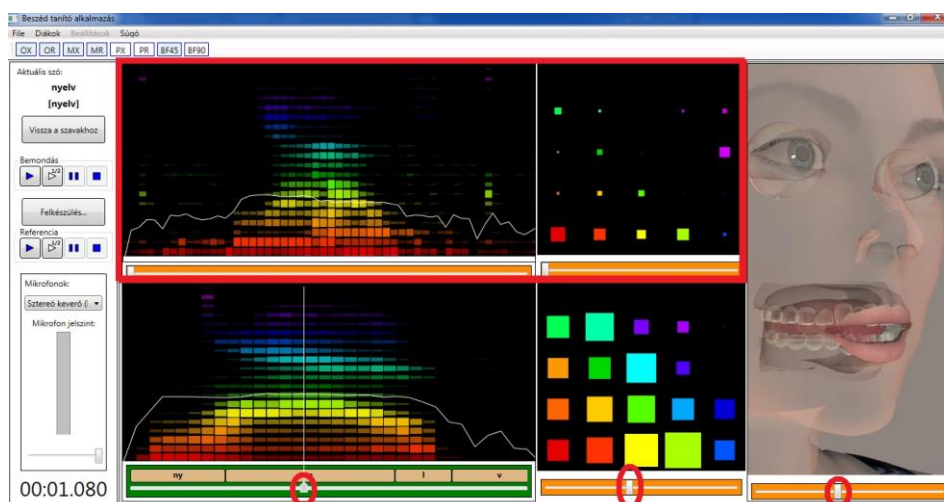
2. ábra Új diák felvétele



3. ábra A beszédasszisztens gyakorló felülete (referencia bemondás)



4. ábra A beszélő fej 45° és 90° szögben



5. ábra A beszédasszisztens gyakorló felülete (piros keretben az aktuális bemondás látható)

3.1. Beszédadatbázis

A rendszer tesztelésének 2013. szeptemberi indulásánál a rendszer 640 szót tartalmaz (referencia bemondást, amit a már elkészült 3000 szavas adatbázisunkból választottunk ki a szurdopedagógusok javaslati szerint), amelyeket a programban résztvevő pedagógusok válogattak össze az eddigi tapasztalataik alapján. A szó adatbázis alapját egy korábban szintén a pedagógusok által összegyűjtött 3000 szavas adatbázis képezte, amit az M5 kutatócsoport tagjai rendszereztek több szempont alapján:

- szófaji besorolás (ige, főnév, melléknév, számnév, névelő, névmás, kötőszó, határozószó);

- témaköri besorolás (család, iskola, szórakozás stb.);
- szótagszám alapján;
- hangok száma alapján;
- magánhangzó – mássalhangzó képlet alapján;
- külön jelölést kaptak a szavak az alábbi esetekben:
 - szavak belsejében előforduló, valószínűleg többet gyakorlandó mássalhangzó-torlódások;
 - szóvégi torlódások;
 - leírástól erősen eltérő ejtémód;
 - ha egy szó azonos alakú és két teljesen más jelentése van;
 - gyakorlást segítő minimál-párok felsorolása.

3.2. Az előrehaladás mérése

Az előrehaladás méréséhez a projektben résztvevő gyerekektől (a kontroll csoportban levőktől is) hangmintákat rögzítettünk 2013. szeptemberben (60 szót és 30 mondatot) és minden félév elején meg fogjuk ismételni.

Az adott iskola mintáit a másik két iskola pedagógusai minősítik összehasonlítva a korábbi mintákkal az alábbi, szurdopedagógusok által korábban egy másik részfeladat céljából kidolgozott skála alapján:

- **Érthetetlen (1):** az artikuláció teljesen torz; felismerhetetlenek a magán-és mássalhangzók; a szótagszám visszaadása sem megfelelő vagy nem kivehető; a levegővétel, a levegővel való gazdálkodás helytelen; rossz a tempó, a ritmus; dalamtalan, dinamikátlan vagy túl feszített a hangadás.
- **Nehezen érthető (2):** súlyos torzítások, hangelhagyások, hangcserék; csak a magánhangzók egy része kivehető; a légzés elégtelensége miatt létrejövő torzítások, pl. túl levegős vagy fojtott; eltérő, zavaró hangszín, ritmus, tempó jellemzi.
- **Közepesen érthető (3):** a magánhangzók ejtése helyes, a szótagszám megfelelő; súlyos beszédhibák előfordulhatnak pl. diszlália, orrhangzósság, fejhangzósság, stb.
- Prozódiai elégtelenségek
- **Jól érthető (4):** csekély mértékű beszédhibák; enyhe prozódiai elégtelenségek.
- **Hallók beszédével azonos szinten érthető (5):** legfeljebb 1-2 hanghiba fordulhat elő.

4. Összefoglalás

A „hallássérült” kifejezés gyűjtőfogalom: siketeket, nagyothallókat egyaránt magába foglal, függetlenül hallássérülésük fokától, súlyosságától, hallásállapotuk milyenségétől. Különböző fokozatok vannak a siketség és a nagyothallás között. A hallássérültek beszédmegértését és beszédtanulását elősegítő összetett rendszer lehetővé teszi az önálló gyakorlást és a hosszú távú fejlődés nyomon követését. Az eddig elért eredmények integrálásával és az eddigi kutatások során felhalmozott tudás egyesítésével, valamint további alapkutatási és alkalmazott kutatási feladatok megoldásával olyan alkalmazás kifejlesztésére nyílik lehetőség, amely a hallássérültek beszéd megértését és beszélni tanítását az eddig ismert módszereknél hatékonyabban tudja támogatni.

5. Köszönetnyilvánítás

A bemutatott kutató munka a TÁMOP-4.2.2.C-11/1/KONV-2012-0002 jelű projekt részeként az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósult meg.

6. Irodalom

- [1] Nooteboom, Sieb: The prosody of speech: Melody and rhythm. In: Hardcastle-Laver (eds): 640-674. 1999.
- [2] Gósy Mária: Fonetika, a beszéd tudománya. Osiris, Budapest, 2004. pp.182-243.
- [3] Kassai Ilona: Fonetika. Nemzeti Tankönyvkiadó, Budapest, 1998.
- [4] Czap L.: *'Audio-Visual speech recognition and synthesis'*. PhD Thesis, Budapest University of Technology and Economics. (2004)
- [5] Massaro D. W.; Light J.: *'Using visible speech to train perception and production of speech for individuals with hearing loss'*. Journal of Speech, Language, and Hearing Research Vol. 47; pp. 304-320 (2004)
- [6] Bolla K.: *'A Phonetic Conspectus of Hungarian'*. Tankönyvkiadó., Budapest. (1995)
- [7] Molnár J.: *'The Map of Hungarian Sounds'*. Tankönyvkiadó, Budapest. (1986)
- [8] Waibel, Alex: Prosody and Speech Recognition. Pitman, London, UK. 1988.
- [9] Gordos Géza, Takács György: Digitális beszédfeldolgozás. Műszaki Könyvkiadó, Budapest, 1983.