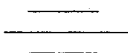


Ebből lett az „ügy”, a kivizsgálás a törvénytörtő eltávolítás visszavonása érdekében... Mindezek a határainkon túli és határainkon belüli esetek sajnálatos módon nem egyediek. És igazak akkor is, ha lehet ellenpéldákat állítani. És nem lehet vigasz, hogy más országokban, még a legfejlettebb demokráciákban is találni példát diszkriminációra. És hogy világhíreség a kiközösítettek, az áldozatok hibáztatása sanyarú sorsuk miatt. Amire a kiközösített kisebbségek egyre deviánsabb magatartással válaszolnak.

Az iskola, a pedagógus nem lehet képes a társadalmi problémák megoldására, de sokat tehet a problémák megoldása érdekében. Ha növekednek az etnikai feszültségek és az ebből fakadó devianciák, azokra még nagyobb türelemmel kell reagálni, hogy elősegítse az integrációt. És sohasem feledheti, hogy a gyermek nem tehet arról, milyen családba, milyen körülmények közé született, s a gyermeki jogok mindenképp felettié, azok maradéktalanul érvényesítése valamennyiünk elháríthatatlan kötelessége.



BÁCSI JÁNOS–KEREKES JUDIT–LÓDINÉ SZABÓ KATALIN–ZS. SEJTÉS GYÖRGYI
SZTE JGYTFK Gyakorló Általános Iskolája
Szeged

„BeszédMester” számítógéppel segített olvasástanítás és beszédterápia

1. Kutatásunk célja: a „BeszédMester”

A BeszédMester (továbbiakban BeMe) olyan gépi tanuláson alapuló szoftver, amely lehetővé teszi a részképességükben sérült és az olvasási nehézségekkel küzdő gyerekek beszédjavítás-terápiáját, illetve olvasástanítását.

Az alap kutatás során vizsgált modern gépi tanulási módszerek hatékonyságát beszédfelismerési alkalmazásban mérjük; az automatikus beszédfelismerést összekapcsoljuk a beszédjavítás-terápiával és az olvasástanítással. Célunk kettős: támogatni szeretnénk a normál hallású gyermekek beszédterápiáját, illetve olvasástanítását, valamint segíteni szeretnénk a hallássérültek érthető beszédének kialakítását.

Egyfajta fonológiai tudatosságot próbálunk kialakítani a gyermekekben, amely valós idejű, gépi fonéma-felismerésen és fonémaazonosításon, illetve a felismerés eredményének vizualizációján és visszacsatolásán alapszik.

A szoftver két részből áll: olvasástanítás és beszédterápia. Jelen cikkben az olvasástanítást fejtjük ki részletesebben.

1.1. A hallássérült gyermekek beszédterápiájának fejlesztése

A hallássérültek tanulási nehézségét az okozza, hogy hiányzik az auditív kontroll, így többségük beszéde halló embertársaik számára nehezen vagy alig érthető. Az auditív kontrollt a valós idejű vizuális visszacsatolással váltjuk ki. A szoftver a beszéd hangjait az elhangzás pillanatában vizuális képpé alakítja, így a fonémák megjelenítésével a hallássérültek a csökkent auditív kontroll, valamint az artikulációs mozdulatok vizuális megfi-

gyelése mellett nyelvi szintű vizuális információt is kapnak, amely hatékonyabbá teszi a kiejtéstanulást.

1.2. Az olvasástanítás

A normál hallású gyermekek esetében a beszédhibák javítására, diszlexia-, diszgráfia-terápiára, illetve az olvasási nehézségekkel küzdő, de egészséges kisiskolások olvasástanítására alkalmazzuk a BeMc-t.

A szoftver a hangképzési hibák javításakor kiválthatja a „tükörmódszert” a lopoté-diában; a diszlexiaterápiában akkor segíthet, ha a diszlexia oka a vizuális észlelés elégtelensége vagy helytelensége.

A géppel segített olvasástanítás alapvető célja, hogy a gyermek minél könnyebben és gyorsabban megtanulja a fonéma-graféma, illetve a graféma-fonéma megfeleltetéseket.

A számítógép felruházása valós idejű vizuális beszédkiértékelési lehetőségekkel a hallássérültek oktatásában pótolhatja a hiányzó auditív kontrollt, illetve az olvasástanításban hatékonyabbá teheti a beszédhangok és a grafémák helyes asszociációját. Mivel mind a beszédhibás, mind a hallássérült gyerekek kiejtéstanítása többnyire az olvasás tanulásával együtt történik, ezért a nyelv hangjait a betűkép, illetve ún. hívóképek formájában jelenítjük meg a képernyőn.

A biztos és sikeres olvasáselsajátítás egyik legfontosabb feltétele a kommunikációs készség fejlettsége, ezen belül kiemelt fontosságúak a beszédészlelési, beszédértési folyamatok, a beszédprodukció. Ehhez kapcsolódnak még olyan részterületek, mint a vizuális és verbális memória, ritmuskészség, vizuális és verbális szerialitás, vizuális differenciáló képesség, irányfelismerés.

A fejlett kommunikációs készségű gyermekek könnyebben és sikeresebben tudják az írott nyelvet elsajátítani, mint azok a társaik, akiknek a kommunikációja sérült (pl.: megkésett beszédfejlődés, egyes diszlexiák, beszédhibák stb.) Az olvasás tanulása nem az első osztályban kezdődik, hanem a vizuális és verbális kommunikáció kialakulásával egyidejűleg.

Az anyanyelv bizonyos típusú beszédhangjaira már a magzat is reagál az anyaméhben. A kommunikációs készség fejlesztése a fonológiai tudatosság megalapozásával kezdődik. Óvodás korban a gyermekek játszanak a nyelvvel: új értelmes vagy értelmetlen szavakat hoznak létre, rímelő szavakat keresnek, ritmusos mondókákat, versikéket mondogatnak. Ez nem más, mint a fonológiai tudatosság felé vezető út első lépése.

Adams (1990) a fonológiai tudat öt szintjét különbözteti meg:

1. A fonológiai egység nagyságának érzékelése (könnyebb a mondatokat szavakra bontani és a szavakat szótagokra, mint a szótagokat fonémákra);
2. A fonémák számának meghatározása egy szóban (könnyebb felbontani egy rövid szót, mint egy hosszú szót);
3. A fonéma helyzetének felismerése a szóban (a középhelyzetben lévő fonémák beazonosítása a legnehezebb, a szó utolsó hangjának beazonosítása könnyebb, legkönnyebbek a kezdő pozícióban álló hangok);
4. A szóban szereplő hangok fonológiai sajátosságainak felismerése (a folyamatosan hangzó fonémákat könnyebb felismerni, mint a csak nagyon röviden ejthető hangokat);
5. a fonológiai tudatosság különleges eseteinek beazonosítása (a rímelő és kezdő fonémák beazonosítása könnyebb, mint az összeolvasás vagy a szétbontás).

Ezek a szintek vagy legalábbis többségük, hatéves korra rutinszerűen működnek a legtöbb gyermekben.

A szintetikus olvasástanítási módszer (hangoztató-elemző-összetevő módszer) előkészítő szakasza ennek fejlesztését tűzi ki célul. Ebben a szakaszban kerül sor a gyermekek kommunikációs készségének megfigyelésére, fejlesztésére, a tiszta fonémahallás, -ejtés fejlesztésére, a hiányosságok pótlására, azaz a betűtanulás megalapozására. A betűtanítás szakaszában általában egyszerre történik a kis és a nagy nyomtatott és az írott kisbetűk tanítása. A graféma megtanulása, beazonosítása a hívókép nevének kezdő hangjának leválasztásával, majd értelmetlen és értelmes szótagok, rövid szavak olvasása következik. A készségfejlesztő szakaszban az olvasástechnika fejlesztése, az olvasás gyakorlása, a szövegértés kialakítása, megalapozása kerül előtérbe.

A módszer fontos sajátossága, hogy szótagolva tanít olvasni, megalapozva ezzel a helyesírás tanítását.

2. A konzorcium tagjai (1-3.) és feladataik

Az SZTE Informatikai Tanszékcsoportja alapkutatói és szoftverfejlesztési munkát végez beszédfelismerés és gépi tanulás témakörben.

Az SZTE JGYTF-i Kar Gyakorló Általános Iskolájának kutatócsoportja elkészített egy gyakorisági szótárt a hatévesek szóanyagából, amely az olvasástanítási rész szóadatbázisát alkotja. Az ország 14 gyakorlóiskolájában 500 gyerek hanganyagát rögzítettük a gépi feldolgozáshoz. Összeállítottuk a hívóképanyagot, kidolgoztuk a számítógéppel segített olvasástanítás metodikáját.

Az Óvoda, az Általános Iskola, a Diákotthon és a Gyermekotthon (Siketek Iskolája) a hallássérültek beszédkutatásának anyagából készített gyakorisági szótárt, hallássérült gyermekek hanganyagát rögzítette, didaktikát ír a BeMe végleges verziójához.

3. A gyakorisági szótárak létrehozása

A munkafolyamat első szakaszának részeredményeként létrejött egy gyakorisági szótár, amely relevánsan tükrözi azt a közel 2000 szót, amellyel egy 1. osztályos az olvasástanulás menetében a leggyakrabban találkozhat. A szótár létrehozásával olyan adatbázist kaptunk, amely több területen hasznosítható: beszédterápiák kidolgozásában, nyelvészeti vizsgálatokban, tankönyvírásban, a tanítók mindennapi munkájában és egy szemantikai hálósztár elkészítésében.

Célunk a szótár elkészítésével az volt, hogy empirikus adatok alapján megállapítsuk, melyek azok a leggyakoribb szavak, amelyekkel egy első osztályos az olvasástanulás menetében találkozhat. Ez a szóképlet adja azt az adatbázist, amelyet hatéves tanulók közreműködésével az ország tizenöt helyén rögzítettünk; majd a szavak szegmentálásával megfelelő mennyiségű és minőségű fonémakészletet fogunk kapni a szoftver hanganyagbázisának kifejlesztéséhez (lásd később részletesen).

A szótár elkészítésének menete a következő volt:

(Munkánkhoz az OM által 2001/2002-ben engedélyezett tankönyvek listája alapján 13 elsős olvasókönyv szóanyagát használtuk fel.)

1. Az olvasókönyvek teljes szóanyagát számítógépre vittük, ezután javítottuk a hibákat.

2. Az olvasókönyvek szóanyagát a Wordcounter programmal összesítettük, és betűrendbe soroltuk. (Létrejött az első abszolút gyakoriságot mutató szótár.)

3. Ezután töröltük azokat a grafémasorokat, amelyek nem elemei a magyar nyelv szókészletének. Erre azért volt szükség, mert bár az olvasástanítás menetében nélkülözhetetlenek az olyan grafémasorok, amelyek a betűkapcsolások (összeolvasás) tanításához szükségesek (pl. am, ib, pa stb.), de nem elemei szókészletünknek, ezért nem kerülhettek be a gyakorisági szótárba.

4. A szókészletből töröltük a nem tartalmazó szavakat. Azokat a morféákat tekintettük tartalmazó szavaknak, amelyek a jelentésnek mind a négy szintjén megfeleltethetők (lexikai, grammatikai, pragmatikai és stilisztikai). Töröltük azokat a szavakat, amelyek a mondatban vagy a szövegben anafora vagy kataforakötést kívánnak maguknak (az összes névmás, határozatlan számnevek és a nem tartalmazó határozószók), azaz antecedenst. Kihagytuk a félszabad morféákat, amelyek elsődleges funkciója valamilyen viszonyjelölés (névelők, névutók és segédigék), a kötőszókat, a módosítószókat és a tagadószókat, amelyek logikai formát vagy preszuppozíciót írnak elő. Technikai okokból töröltük azokat az igekötőket, melyek nem közvetlenül az ige előtt állnak, vagyis nincsenek az igével egybeírva. Töröltük a felelőszókat (igen, nem), mivel csak az eldöntendő kérdéssel együtt interpretálhatók, az indulatszavakat, mivel lexikai jelentésük függ a szöveggörnyezettől.

Így a következő szófajokba tartozó szavak kerültek a gyakorisági szótárba: igék, főnevek, számnevek (határozott tő-, tört- és sorszámnevek), határozószók (valódi határozószók), főnévi igenevek, melléknévi igenevek (folyamatos és beálló), határozói igenevek.

Erre a felosztásra azért volt szükség, mert a BeszédMester szóanyagának létrehozásához csak olyan morféákat akartunk felhasználni, amelyeknek a hatévesek számára közvetlen lexikai jelentésük van.

5. Az azonos alakú szavakat egy szótári egységben szerepeltettük.

6. A toldalékos szavak besorolásakor a következő szempontokat vettük figyelembe:

a) Az igék rag és jel nélküli alakját, az Magyar értelmező kéziszótárnak megfelelően, önálló lexémaként kezeltük, valamint egy egységbe soroltuk azokat a képzett alakokat, amelyek nem változtatják meg az ige vonzatstruktúráját.

b) A ragos és jeles főnevek szintén önálló lexikális egységek, a képzős alakok új önálló lexémák.

A keresztnévek becézett alakjainál, ha egyértelműen el lehetett dönteni, hogy a becézett alak csak egy névhez tartozhat (pl. Misi → Mihály), akkor egy szótári egységgé vontuk őket össze, amennyiben egy becézett alak kettő vagy több keresztnév alakváltozata is lehet (pl. Marci → Márton, Marcell), akkor a becézett alakot önálló szótári egységként hagytuk meg, és nem vontuk össze egyik keresztnévvel sem.

c) A melléknévek alapfokban kerültek a szótárba, a képzett alakok minden esetben önálló szótári egységként szerepelnek.

d) A számneveket a melléknévhez hasonlóan soroltuk be.

e) A főnévi igenevek ragozott alakjait egy szótári egységként szerepeltettük.

f) A befejezett melléknévi igenév nem szerepel önálló szótári alakként, az alapigével vontuk össze, mivel alakilag megegyezik a múlt idejű igével. A folyamatos és a beálló melléknévi igenevet önálló szótári alakként szerepeltettük, toldalékos alakjaikat egy szótári egységként kezeltük.

g) A határozói igenevek önálló szótári egységet alkotnak.

7. Az így létrejött szóanyagot csökkenő gyakorisági rendbe soroltuk. Ha a gyakorisági mutató több szó esetén megegyezett, az azonos indexszám alá tartozó morféákat betűrendbe soroltuk. (Létrejött a második szótár.)

Bár terveink szerint a 2000 leggyakoribb szó lett volna a BeszédMester szoftver alapja, de ez 1953 szóra módosult, mivel a 2000. szó olyan gyakorisági pontot metsz (9), amely előtt és után is szerepelnek ugyanilyen gyakorisági indexszel szóalakok. Ezért úgy döntöttünk, hogy a munkánkhoz azokat a szavakat fogjuk felhasználni, amelyek legalább tízszer szerepelnek, vagyis gyakorisági indexük eléri vagy meghaladja a 10-et. E szavak száma 1953. (Létrejött a harmadik szótár.)

Az *első szótár* tartalmazza a 13 első osztályos olvasókönyv összes **27.293** grafémasorát csökkenő gyakoriság szerint.

A *második szótár* már csak az általunk felállított kritériumrendszer alapján tartalmaz minősített szavakat tartalmazza csökkenő gyakoriság szerint. Ebben a szótárban **12.226** szóalak található.

A *harmadik szótár* azt az **1953** szót tartalmazza csökkenő gyakoriság szerint, amelyek legalább 10-es gyakorisági mutatóval fordultak elő a második szótárban. Ez a szóanyag képezi a BeszédMester szoftver adatbázisának alapját.

Gyakorisági „szó-tár” készült a hallássérültek speciális iskolai oktatásában a kommunikáció fejlesztéséhez használatos anyagok alapján. A hallássérültek beszédfejlesztésének módszeres eljárásai során rögzített szövegek, a súlyos fokban hallássérültek speciális iskolai oktatásában használatos anyanyelvkönyvek olvasmányai, valamint az intenzív nyelvi fejlesztési szakaszra ajánlott tantervi szókészletek adták az anyagot.

3. Az adatbázis további felhasználhatósága

Munkánk célja az is, hogy a gépi tanulás modern, statisztikai módszereit gazdagítsuk, továbbfejlesszük. A speciális szempontok alapján kialakított, áttekinthetően megszerkesztett és fonetikailag feldolgozott szóadatbázis ad alapot a program háttérben működő valós idejű, gépi fonéma-felismerés kialakítására. Az adatbázis felépítése és tartalma alapján nyelvészeti és beszédtechnológiai kutatásokra is felhasználható a továbbiakban.

4.1. A szóadatbázis felépítése

A szóadatbázis 500 adatközlő 100-100 szószintű bemondását tartalmazza. A gyakorisági szótár alapján szerkesztettük meg az 500 szó bemondó lapot.

Ezek felépítése a következő:

- Az első rész a 14 db magyar magánhangzót tartalmazza.
- A második rész 40 szóból áll. Ezek olyan tartalmas szavak, amelyek betűrendben követik egymást. Az első és második rész minden adatközlőnél azonos.
- A harmadik rész 60 szava adatközlőnként változó. A leggyakrabban használt 1953 szót osztottuk el előfordulási gyakoriságuk alapján, tehát egy-egy gyakori szó több lapon is szerepelhet.

Mivel az adatközlők kisiskolások, a bemondásokat tartalmazó szólapok az átlagosnál nagyobb (36-os) betűmérettel készültek.

1. táblázat: Példa egy bemondólapra

1. rész: a, á, e, é, i, í, o, ó, ő, ö, u, ú, ü, ű
2. rész: alma, ág, béka, cica, csiga, dió, dzé betű, dzsungel, egér, ég, fészek, gólya, gyerek, hegy, iskola, ír, játék, kakas, lepke, lyuk, maci, nap, nyúl, oroszlán, óra, ördög, őz, pók, róka, süni, szőlő, tulipán, tyúk, uborka, út, üveg, úrhajós, virág, zokni, zsiráf
3. rész: jó, húz, nyit, fej, kérdez, három, marad, rétes, szám, aludni, meleg, ősz, madzag, haladva, válaszoló, autó, lyukas, betű, tűkőr, kíváncsi, áll, lúd, lel, Füles, néz, könyvtár, galamb, szó, farkas, végez, hoz, karika, kár, kis, alkot, téli, mély, mell, köszön, Imre, Gyuri, van, hall, fut, róka, pap, nagy, zaj, olvas, fehér, mondat, érdekes, fiú, ejt, kész, fa, ész, egy, mos, fűz

Az 500 kisiskolás 7000 magánhangzót és 50 000 szót olvasott a számítógépbe, amely összességében 57 000 bemondást jelent.

4.2. Az adatközlők kiválasztása

Adatközlőink 6-7 éves kisiskolások az ország 14 gyakorlóiskolájából. Az 500 gyerek (250 lány, 250 fiú) kiválasztásában fő szempontunk az volt, hogy minél közelebb legyenek az olvasástanulás kezdeti szakaszához, mivel a BeMe majdani felhasználói ebbe a korosztályba tartoznak.

4.3. A hangfelvételek helyszínei

Elvi kiindulásunk az volt, hogy úgy jelöljük ki a felvételek helyszínét, hogy az minél jobban megfeleljen a magyar nyelvjárási régióknak. Tettük ezt azért, hogy a BeMe szoftvert az ország bármely területén élő gyerek eredményesen tudja használni.

Készítettünk felvételeket lakótelepi, illetve belvárosi iskolában is.

2. táblázat: A felvételek helyszíne és ideje

Város	Az iskola neve	Nyelvjárási régió	Régióon belüli csoport
Szeged	Tabán Általános Iskola	Dél-alföldi	Szeged környéki
Szeged	SZTE JGYTFK Gyak. Ált. Isk.	Dél-alföldi	Szeged környéki
Győr	Apáczai Csere János Tanító- képző Főiskola Gyak. Ált. Isk.	Közép-dunántúli- kisalföldi	Észak-dunántúli
Pécs	Pécsi Tudományegyetem I.sz. Gyak. Ált. Isk.	Dél-dunántúli	Észak-baranyai
Szekszárd	PPF Gyak. Ált. Isk.	Dél-alföldi	Baja környéki
Kaposvár	K.Egyetem Cs.V. M. Ped. Főisk.Gyak. Ált. Isk.és Gimn.	Dél-dunántúli	Közép-somogyi
Baja	Eötvös J. Főiskola Gyak. Ált. Isk.	Dél-alföldi	Baja környéki
Eger	Eszterházy K. Főisk. 2.sz. Gyak. Ált. Isk.	Palóc	Déli tömb Eger vidéki
Nyíregyháza	Ny. F. Bt. és Műv. K. A. Cs. J. Gyak. Ált. Isk.	Északkeleti r.	Szabolcs-szatmári

Város	Az iskola neve	Nyelvjárási régió	Régió belüli csoport
Sárospatak	Miskolci Egyetem Com. Tanítóképző F.K.Á.J.	Palóc	Keleti tömb Keleti palóc
Debrecen	Kölcsey F. Ref. Tanítók. Főisk. Gyakorlóisk.	Tisza-Körös vidéki	Hajdú-bihari
Békéscsaba	Békéscsaba 10. sz. Általános Isk.	Dél-alföldi	Szeged környéki
Budapest	ELTE Tan- és Óv. F. K. Gyak. Ált. Isk.	*Budapest	-----
Kecskemét	K. Tanítók. Főisk. Kar Petőfi S. Gyak. Ált. Isk.	Dél-alföldi r.	Kiskunsági cs.

* Budapestet külön régióként kezeljük a település nagysága, lakóinak száma és összetétele miatt.

Az anyaggyűjtés professzionális volt, mert a terepmunkában (a felvételek készítésében) szakemberek: nyelvészek és fejlesztő pedagógusok vettek részt.

Az anyaggyűjtés egy-egy adatközlőnél 2-10 percet vett igénybe. Az időkülönbség egyéni sajátosságokból adódott (eltérő olvasási tempó, illetve készségszint).

Számolnunk kell azonban az aktív módszer azon hátrányával, hogy a felvételek során az interjúhelyzet megnöveli a gyerekek nyelvi önkontrollját, s a formális beszédhelyzet, a magyar köznyelv kívánalmi szerint olvassák a szavakat.

Meg kell említenünk azt a nyelvi sztereotípiát is, amely csak erős megszorítással igaz ugyan, hogy Magyarországon a falusiak beszélnek nyelvjárásban, a városiak pedig köznyelven. A gyakorlóiskolákat úgy választottuk ki, hogy az az ország minél több régiójának nyelvhasználatát tükrözze.

A regionális nyelvhasználat és az életkor azon összefüggését is szem előtt kell tartanunk, mely szerint az idősebbek jobban őrzik a nyelvjárási vonásokat, mint a hat év körüli kisiskolások.

4.4. Az adatházis hanganyaga

A BeMe hanganyaga korlátozottan érhető el, az adatházis struktúrájának specifikációja nyilvános. Ennek célja, hogy az adatházis felhasználható legyen más kutatásokban pl.: a gépi beszédfelismerés más alkalmazásaiban, és nyelvészeti (fonetikai, fonológiai, szociolingvisztikai) kutatásokban.

Az adatházisban megtalálható a digitálisan tárolt hanganyag. A hullámforma mellett a hozzájuk tartozó ortografikus karaktereket is rögzítettük. A teljes hanganyagot annotáltuk és szegmentáltuk.

4.5. A hanganyag annotálása

Az annotálás azt jelenti, hogy minden hangfájl mellé egy címkefájl került, amely a hangfájl paramétereivel és tartalmával kapcsolatos változókat tartalmazza. A BeMe hanganyagának annotálásánál informatikai, nyelvi és társadalmi változókat vettünk figyelembe. Ezeket a SpeechDat-E és a The Budapest Sociolinguistic Interview alapján dolgoztuk ki. A SpeechDat személyfüggetlen beszédfelismerők betanítására alkalmas adatházisok létrehozásának szabályrendszere, melyet az Európai Közösség által létrehozott szakértői bizottság állított össze. A Budapesti Szociolingvisztikai Interjú az MTA Nyelvtudományi Intézetében készült magyar nyelvi szociolingvisztikai kutatás angol nyelvű leírással.

Az általánosan elfogadott szempontokon túl specifikus mutatókat is felhasználtunk.

Az elhangzott szövegek lejegyzésére a nyelvészetben általában az IPA szimbólum-rendszert használják. Ez a jelölésrendszer nem illeszkedik a számítógép billentyűzetéhez, ezért a nemzetközi SAMPA jelölésrendszerrel dolgoztunk.

5. Összegzés

A BeszédMester tehát olyan gépi tanuláson alapuló szoftver, amely hívóképek segítségével összekapcsolja a fonéma-grafémapirokat úgy, hogy az alkalmas legyen beszédjavítás-terápiára és az olvasástanításra. A hangosolvasás technikájának elsajátítása során a gyermek a szövegszónyi hangcsoportot hangokból, szótagokból tanulja meg előállítani úgy, hogy a szónál kisebb egységeket, a betűket megfelelteti az akusztikus jeleknek, a hängoknak.

Jelenleg a program tesztelése folyik. Az előzetes eredmények azt mutatják, hogy a gyerekek szívesen dolgoznak a szoftverrel, hiszen a motiváció nyilvánvaló; a gyerekek akár egyedül is képesek a számítógépes programmal „játszani”, illetve játszva tanulni abban a korban és annyi idő alatt, amely számukra a legalkalmasabb.

IRODALOM

1. Adams, M. I.: *Beginning to read: Thinking and learning about print*. MIT Press, Cambridge, 1990.
2. Bácsi János – Kerekes Judit: Az első osztályos olvasókönyvek szóanyagából készült gyakorisági szótár „Van szó” Módszertani Közlemények. 2003. II. szám 52-57.
3. Benkő László: *Az írói szótár (1979)* Akadémiai Kiadó, Budapest.
4. Bishop, C. M. (1995) *Neural Networks for Pattern recognition*. Oxford: Oxford University Press.
5. Kiss, J.: *Magyar dialektológia*, Osiris Kiadó, Budapest, 2001.
6. Kontra, M., Váradi, T., *The Budapest Sociolinguistic Interview: Version 3*. Linguistics Institute, Hungarian Academy of Science, Budapest, 1997.
7. 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.
8. Pollak, P., Cernocky, J., Boudy, J., Choukri, K., Heuvel, H., Vicsi, K., Virag, A., Siemund, R., Majewski, W., Sadowski, J., Staroniewicz, P., Tropf, H., Kochanina, J., Ostrouchov, A., Rusko, M., Trnka, M.: *SpeechDat(E) –Eastern European Telephone Speech Databases*. *Proceeding LREC’ 2000 Satellite workshop XLDB – Very large Telephone Speech Databases*, 29 May 2000, Athens
9. *Version 3*. Linguistics Institute, Hungarian Academy of Science, Budapest, 1997.
10. Vicsi, K., Tóth, L., Kocsor, A., Gordos, G., Csirik, J.: *MTBA-Magyar nyelvű telefonbeszéd-adatbázis*, *Híradástechnika*, LVII. 2002/8, Budapest, pp. 35-43.
11. Wels, J. at all: *Standard Computer-Compatible Transcription*. *Esprit Project 2589 (SAM)*, Doc.no. SAM-UCL-037. London: Phonetics and Linguistics Dep., UCL, 1992.