

PACZOLAY DÉNES, TÓTH LÁSZLÓ, KOCSOR ANDRÁS ÉS KERESKES JUDIT
MTA SZTE Mesterséges Intelligencia Tanszéki Kutatócsoport
pdenes@inf.u-szeged.hu, tothl@inf.u-szeged.hu, kocsor@inf.u-szeged.hu,
kerekesj@deltav.hu

Gépi tanulás alkalmazása egy fonológiai tudatosság-fejlesztő rendszerben **(A kutatás a KOMA XXIX/31 számú pályázat támogatásával folyik)**

The most important clue to the process of learning to read is the ability to separate and identify the phonemes and to associate these with its written forms. To also learn to read in a fruitful way young learners must be aware of the phonemes and must also be able to manipulate them. Many children with learning disabilities have deficiencies in their ability to process phonological information.

We constructed a computer-aided training software, which uses a very effective phoneme recognizer and provides a real-time visual feedback in the form of flickering letters on calling pictures. The brightness of the letters is proportional to the speech recognizer's output. We found we can make impressive progress in a very short training period, because the young people enjoy practicing with computers.

We collected and processed a special speech database and we also chose a good combination of the examined feature extraction techniques and machine learning methods to develop the best phoneme recognizer possible.

A számítógépek gyors fejlődésével és elterjedésével egyre nagyobb az igény oktatási feladatok automatizálására.

Jelen cikk egy olyan programrendszert mutat be, amely gépi tanuláson alapuló fonéma felismerőre építve, segíti a fonológiai tudatosság kialakítását, illetve növeli annak megbízhatóságát.

A gyermekek olvasástanulásánál az a legfontosabb momentum, hogy képesek legyenek megkülönböztetni és azonosítani a szavakban szereplő fonémákat, illetve összekapcsolni a fonémákat az írott formájukkal, a grafémákkal. Közismert, hogy az olvasási gondokkal küzdő gyermekeknek gyakran a fonológiai információk feldolgozásával van problémájuk. A fonológiai tudatosság fejlesztését nagymértékben elősegíti egy olyan számítógépes rendszer, amellyel a kisiskolások a fonémák és írásképek közötti asszociációs képességüket fejleszthetik. Egy ilyen program háttérében úgynevezett gépi tanuló algoritmusok állnak, így ezek ismertetésére cikkünkben is jelentős hangsúly esik.

A gépi tanulásra épülő fonémaosztályozás megvalósításának a menete a következő: első lépésben egy tanuló adatbázist kell elkészíteni. Jelen esetben ez beszédhangok rögzítését, illetve ezen belül az egyes fonémák kezdő és végpontjának bejegyzését jelenti. Második lépésben létre kell hozni egy modellt, amely segítségével leírhatók az osztályozásra váró fonémák jellemzői valós számok segítségével. A jellemzők előállításának egy lehetséges módja az, hogy a hangot szinuszos összetevők súlyozott összegére bontjuk és a kapott súlyokkal spektrális együtthatókat képezünk.

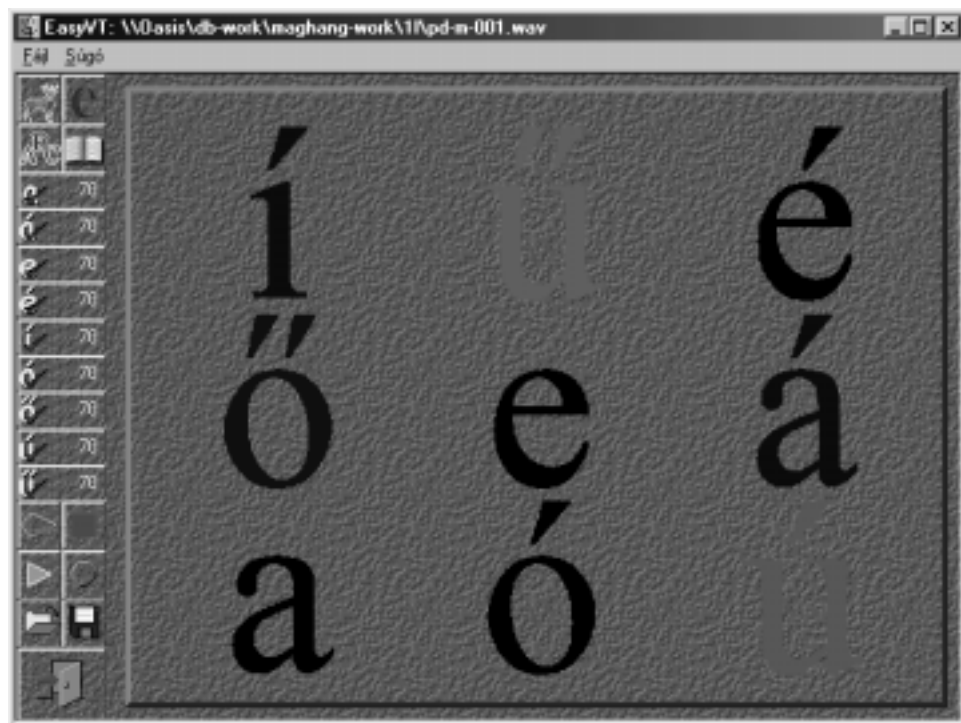
Végezetül felhasználunk egy gépi tanuló algoritmust is a jellemzőkkel adott fonémák modellezésére. Általában ez a három lépés nem csak lineárisan követi egymást, mert esetleg vissza kell lépni az előző szintre, akár az adatbázisgyűjtésig, ha nem értük el a kívánt eredményt. A továbbiakban az Easy Speech – Vowel Teacher rendszer rövid ismertetése után az adatbázis elkészítésének körülményeiről, majd a jellemzőki-nyerés technikájáról illetve a gépi tanulókról lesz szó. Végezetül az eredmények ismertetése és a konklúzió következik.

Az 'Easy Speech – Vowel Teacher' rendszer működése

Az 'Easy Speech – Vowel Teacher' rendszerben a fonológiai tudatosság fejlesztése a következő módon történik: a rendszer meghatározza, hogy a hang, amit a tanuló kiejt, mennyire hasonlít a rendszerrel tanított magánhangzókhoz, majd eredményképpen megjeleníti a grafémát, és/vagy a hozzátartozó hívóképet. A fonémákhoz rendelt grafémák színe, illetve láthatósága ad információt arról, hogy az adott hangot, mint egy-egy magánhangzót, mennyire ismerte fel a gépi tanuló algoritmus. Ha nem talál hasonlóságot, akkor a graféma nem látható, ha igen, akkor lényegesen kiemelkedik a háttérből. Jelenleg a program 9 magánhangzót tartalmaz, ugyanis a rövid-hosszú párokat nem különbözteti meg. Így a következő magánhangzókat ismeri: „a” „á” „e” „é” „í” „ó” „ő” „ú” „ű”.

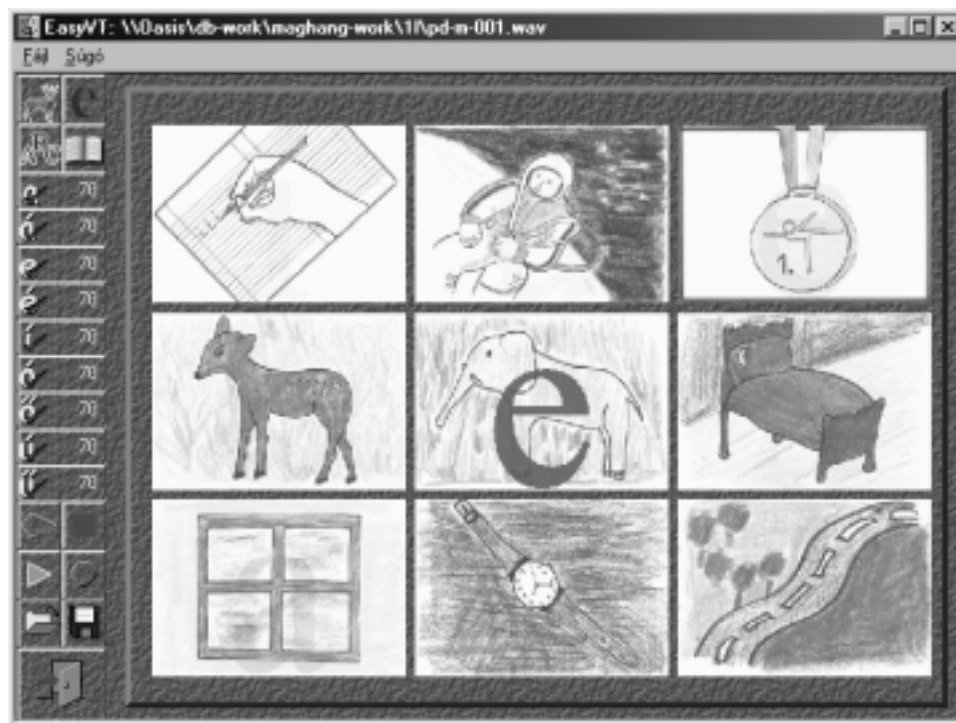
A jelenlegi programrendszer több adatbázist és több megjelenítési módot foglal magában. A rendszer legfontosabb alkalmazási területe a magánhangzók helyes kiejtésének, illetve az írott formájukkal való összekapcsolásának gyakoroltatása. Az első esetben első és második, míg a másik esetben 1-8. osztályos tanulók hangjára épül az adatbázis. A két eltérő feladathoz két-két megjelenítési mód tartozik.

Az első esetben, amikor a kiejtett magánhangzók tisztaságát gyakoroltatjuk, a felismerés eredményét színekkel jelezzük. A programban beállíthatunk minden egyes fonémához egy-egy 'elfogadási' határt: ha a határ alatt vagyunk, akkor az adott magánhangzó színe feketéből zöldbe, míg felette zöldből élénk piros színbe megy át. A következő ábra egy olyan esetet szemléltet, amikor az ejtett hang nem tisztán [ű], hanem az [ú] hangot is fellelhetjük benne.



1. ábra

A második esetben, amikor a kiejtett magánhangzóbetűk írott formájukhoz csatolását gyakoroltatjuk, akkor a megfelelő hívóképen egy-egy piros színű magánhangzóbetű jelenik meg. A betű láthatósága arányos a hang felismerésével. Amennyiben köztes hangot ejtünk halványan több helyen is látható az írott magánhangzó.



2.ábra

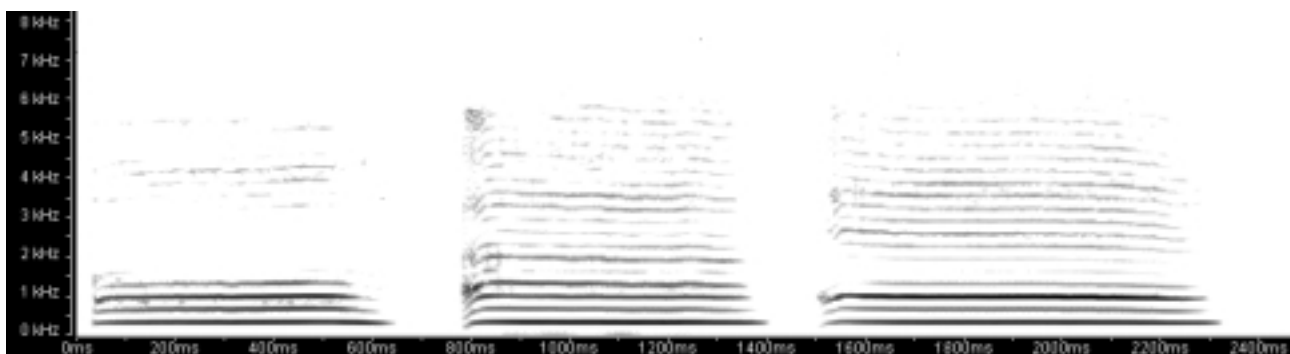
Mi áll az 'EasySpeech - VowelTeacher' rendszer háttérében

A beszédkorpusz

A tanulmányhoz használt korpusz 50 elsős és 10 másodikos fiú, és ugyanennyi lány tanuló kitartott magánhangzó felvételeire épül. Minden tanuló a fent említett 9 magánhangzót mondta be, amelyek mindegyike egy-egy példát szolgáltatott. Az adatbázisba került a magánhangzókat elválasztó 8 szünet is. Néhány bemondásból lemaradt egy-két hangzó, így végül a 2040 példa helyett 2007 fonéma-példányból áll az adatbázis. A bemondott mintákat két részre bontottuk: a gépi tanuló algoritmusok tanításhoz 1514, a tesztelésükhöz 493 elemet tartalmazott az adatbázis.

Jellemzőkinyerés fonémákból

A felismerő tanítása a fonémákból kinyert jellemzők segítségével történik. A fonémák azon jellemzőinek megtalálása, amelyek segítségével valóban meg lehet különböztetni őket, nagyon nehéz feladat, és máig is nyitott kérdés [4]. A fonetikában használt spektrogramokat – szinuszos összetevők erősségének időbeli grafikonját – vizsgálva jól látható vízszintes vonalakat (azaz időben állandó frekvencia összetétel) találhatunk a magánhangzók kiejtésekor. A maximumhelyek a formánsok, melyeknek a helye más-más hangú beszélő esetén kissé máshol van, de relatív helyük és erősségük hasonló egy-egy magánhangzó esetén. Kitartott magánhangzók esetén a formánsok azok a jellemzők, amelyek segítségével a magánhangzók karakterizálhatók.



3. sz. ábra Az (a), (á) és (e) kitartott fonéma spektrogramja

A felismeréshez egyidejűleg több, egy-egy számértékkel megadható jellemezőt használunk fel, ezeket jellemzőkészletnek nevezhetjük. A jellemzőkészletek mindegyike a hangok energiaspektrumát írja le más-más módon. A tanulmányban négy készletet próbáltunk ki:

- Az első jellemzőkészlet ('FBLE') a frekvenciatartományt nemlineárisan, Bark skála szerint felosztja 24-sávra, majd itt súlyozott összegét képezi a szinuszos összetevők erősségeinek. Így tehát ha egy sávba formáns esik, akkor ott nagyobb értéket kapunk. A sávok (szűrők) átfedők és különböző szélességűek, mert ily módon szimulálják a hallás kritikus sávjait.
- A spektrum pillanatnyi értékéből az ember sem képes a fonémák pontos azonosítására. A második jellemzőkészlet esetén ezért további két-két szomszédos oszlopot (időintervallumot) is figyelembe vesz a tanulóalgoritmus (így 5x24 jellemzőt használva).

- Másik lehetőség, hogy a változást a frekvenciasávokban is figyelembe vesszük. Ezt a 'Deriv' jellemzőkészlet tartalmazza. Itt a 24 sávban számolt energiák mellett a 24 időbeli változást is hozzávesszük a készlethez.
- A negyedik jellemzőkészlet ('Grav') megkísérli megbecsülni az egyes formánsok helyét és szélességét. Ehhez négy sávra osztjuk a frekvenciatartományt, majd minden sávban kiszámoljuk a súlypontot, és meghatározzuk, hogy mennyire összpontosul az energia ebben a pontban.

A sávok száma mellett a jellemzőkinyerésnél a spektrogram paramétereit is lehet változtatni. A tesztheinkben egy nagy felbontású részletes, szélessávú spektrogramot használtunk.

Fonémafelismerés gépi tanulással

A jellemzők keresésével párhuzamosan meg kell találni a célnak leginkább megfelelő gépi tanuló algoritmust, amely megbízható és valós idejű, azaz a fonéma kiejtése közben azonnal, késleltetés nélkül szolgáltat eredményt.

A tanulóalgoritmus feladata, hogy a jellemzők alapján felismerje a magánhangzókat, és az ezáltal megadott kategóriákba sorolja őket. Várhatóan az egy csoportba tartozó hangokat hasonló számértékek jellemzik, azonban a hasonlóság nem egyértelműen definiált, és nem is teljesen különülnek el ezek a csoportok, így az osztályozás feladata nem egyszerű. Jelen tanulmány elkészítésekor több ismert tanulórendszert próbáltunk ki különféle paraméterekkel, és a tanulás eredményét összehasonlítottuk a tesztadatbázison. A klasszifikációs problémák megoldására szolgáló statisztikai tanuló algoritmusokat vagy diszkriminatív, vagy generatív tanulóknak nevezik, attól függően, hogy mit modelleznek. A diszkriminatív modellek minden osztályt közös jellemzőcsoporttal írnak le, és arra törekednek, hogy egy-egy osztályt rendre elkülönítsenek a többitől. Ezt egy elválasztó függvény paramétereinek megfelelő beállításával érik el, vagy az osztályokat elemekkel és távolságfüggvény definiálásával reprezentálják. Ilyen diszkriminatív tanuló algoritmusok például a *C4.5*, *OC1* [9] vagy a *Mesterséges Neuronhálózatok* (Artificial Neural Network, ANN) [3]

A Bayes-formulát felhasználásával, a C osztály feltételes valószínűsége egy x minta esetén, megadható a következőképpen:

$$P(C|x) = \frac{P(x|C)P(C)}{P(x)} \quad (1)$$

Így ahelyett, hogy közvetlenül a $P(C|x)$ valószínűséget modelleznénk, mint ahogy azt a diszkriminatív tanulók teszik, másik lehetőségként becsülhetjük az osztályok feltételes valószínűségét, azaz $P(x|C)$ -t minden egyes osztályra külön-külön. Ez az

úgynevezett generatív megközelítés. Ilyen generatív tanuló algoritmus a Rejtett Markov Modell (Hidden Markov Model, HMM), vagy a Kevert Gauss Modell (Gaussian Mixture Model, GMM).

HMM

Ez a módszer a fonémafelismerési feladatra leggyakrabban használt modell, azonban korábbi tesztjeink alapján a gépi tanulás hatékony módszerei elérik, sőt megelőzik az itt kapott HMM eredményeket

GMM

A Kevert Gauss Modell az osztályok feltételes valószínűségét $P(x|C)$ -t közelíti rögzített számú Gauss- vagy más néven haranggörbe segítségével. Ez a görbe sok természetben előforduló valószínűségi eloszlást ír le jól, így várható, hogy néhány Gauss-görbe összegével megfelelő pontossággal közelíthető a tanuló példák eloszlása. Ez után közvetlenül számolható, hogy egy elem milyen valószínűséggel esik egy adott osztályba. A tanítás a Gauss-görbe paramétereinek megtalálását jelenti.

C4.5

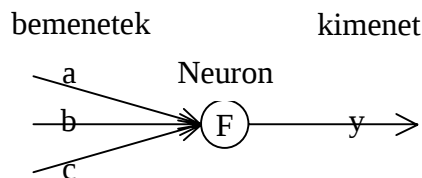
A C4.5 tanuló algoritmus egy úgynevezett döntési fát épít fel, és ennek segítségével sorolja az objektumokat osztályokba. A C4.5 döntési fája a következőképpen képzelhető el: az algoritmus minden lépésben kiválaszt egy tulajdonságot, és ennek segítségével kettéosztja az elemeket aszerint, hogy a kiválasztott tulajdonsággal rendelkeznek-e vagy sem. Ha a két részre osztott halmazban még osztályozásukat tekintve keverednek az elemek, akkor újabb tulajdonságot választ, és ennek segítségével ismét felbontja valamelyik halmazt két további részre. A döntési fa struktúrája írja le a kérdések egymásutánosságát. Egy csoport felbontása akkor jó, ha az utána keletkezett csoportokban már kevésbé keverednek a különböző osztályba tartozó elemek. A döntési fa hatékony tudásreprezentáció, ha kevés vizsgálattal eljutunk egy olyan csoporthoz, ahol már csak csupa azonos elem fordul elő. A mi esetünkben a fonémaosztályozásnál a jellemzőket valós számok képezik. A döntést, amellyel a példákat két csoportra bontjuk, kiválasztott jellemzőkomponens alapján hozzuk meg, ha ez a jellemző egy gondosan megválasztott küszöbértéket elér, akkor az adott egyedet egyik, különben a másik csoportba soroljuk.

OC1

Az OC1 algoritmus a C4.5 tanuló algoritmus egyik kibővítése. Szintén egy döntési fát hoz létre, de az elemeket felosztó döntést nem egyetlen tulajdonság alapján hozzuk meg, hanem több tulajdonságot súlyozottan vesszünk figyelembe. Ez az algoritmus bonyolultabb, de általában hatékonyabb, és kisebb lesz a döntési fa mérete is. (Kevesebb csoportokra vágást tartalmaz.)

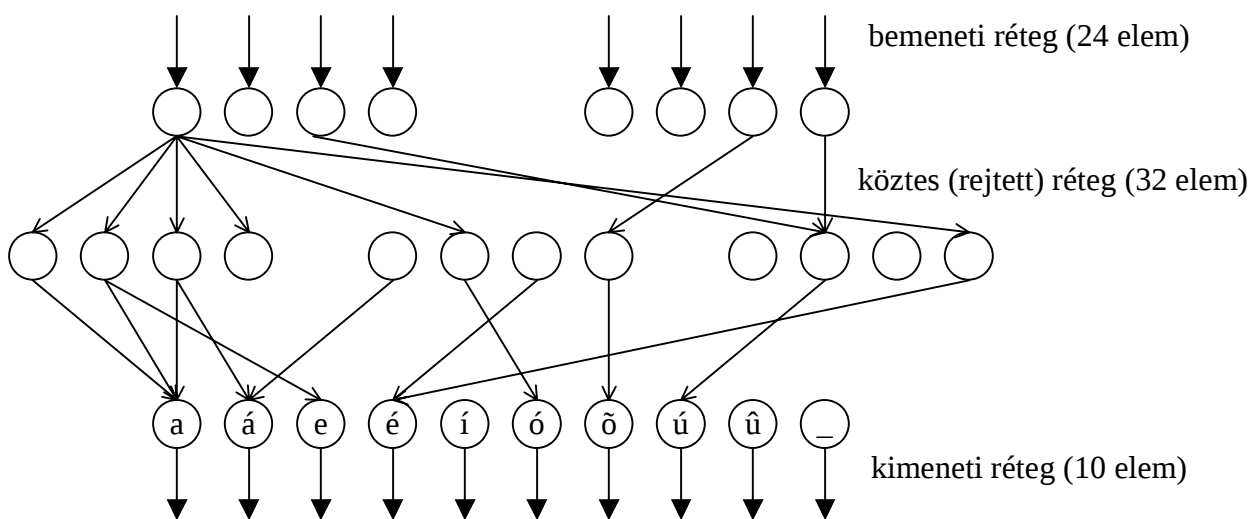
ANN

A mesterséges neuronháló egy olyan gépi tanuló algoritmus, amely tervezésekor a megalkotói a természetben megtalálható neuronhálóak egy egyszerűsített modelljéből indultak ki. A neuronháló egy építőelemét a következőképpen lehet elképzelni:



4. sz. ábra

Itt a bemenetek mindegyikéhez tartozik egy súly-érték, ami igazából a kapcsolat erősségét szimbolizálja. Jelen esetben a jelek és a súlyok (a,b,c) természetesen számok. A neuron kimenetére eső jel egy F függvény segítségével a súlyozott bemenetekből származik. Gépi tanuláskor általában rétegzett hálót használunk, ami azt jelenti, hogy neuronokból rétegekbe rendezett hálót építünk fel úgy, hogy az egy rétegben belüli csomópontok között nincs él, csak a szomszédos rétegek között. Az első rétegben található neuronok csak kimenetükkel csatlakoznak más neuronokhoz, a bemeneti értékeiket szabadon meghatározhatjuk. A speciális helyzet miatt ezt a réteget nevezzük bemeneti rétegnek. Az utolsó rétegben elhelyezkedő neuronoknak csak a bemenetéhez csatlakozik neuron, így a kimenetén megjelenő értékeket szabadon leolvashatjuk. Ez a réteg a neuronháló kimeneti rétegeként ismert.



5. sz. ábra Az egyik tesztben használt 3 rétegű (24-32-10) neuronháló vázlata

A neuronháló működése vázlatosan az esetünkben a következő: a bemeneti rétegre írjuk a jellemzőket sorban, majd a kimeneten megjelenő számok adják meg, hogy az adott minta mennyire hasonlít a tanult magánhangzókhoz. A neuronháló tanítása az élekre írt súlyok, és a neuronokra jellemző határok megfelelő beállítását jelenti a tanulópéldák alapján.

SVM

A Support Vector Machine [10] egy speciális típusú függvény (hipersík) segítségével osztja két részre a példákat. Egy ilyen függvény azt adja meg, hogy egy elem mennyire tartozik egy adott osztályba. Ha több osztály van, akkor több vágó függvény segítségével lehet a klasszifikációt megvalósítani. A tanítás eredményeként kapott függvények paraméterei a vágáshoz legközelebb elhelyezkedő tanuló példák segítségével megadhatók.

Teszteredmények és diszkusszió

A különféle gépi tanulási módszerek találati arányait az egyes jellemzőkészletben a következő táblázat mutatja:

1. sz. táblázat

	ANN	C4.5	GMM	OC1	SVM
FBLE(24)	97,16	90,87	93,10	92,09	95,94
5b(120)	95,94	89,25	93,10	89,86	96,75
Deriv(48)	93,51	90,67	87,02	91,08	94,31
Grav(32)	96,15	89,25	92,90	91,28	96,55
átlag	95,69	90,01	91,53	91,08	95,89

Az összehasonlított algoritmusok mindegyike viszonylag jól osztályozta a magánhangzókat, ami azt jelenti, hogy a kiválasztott tulajdonságok segítségével elkülöníthetők a magánhangzók. Egyes algoritmusok más-más jellemzőkészlet felhasználásával érték el a legjobb eredményt. Az öt szomszéd (5b) készlet általában kissé rosszabbul viselkedett a kiinduló (FBLE) jellemzőkészletnél, ami azt bizonyítja, hogy a szinuszos összetevők (formánsok) viszonylag kevésbé változnak a kitartott magánhangzók esetén, és az osztályozó algoritmusok túl sok jellemző felhasználásával esetlegesen romlanak, kivéve a SVM-et. A táblázat igazolja, hogy a neuronháló (ANN) mellett érdemes döntenünk, mert az egyik legjobb eredményt nyújtja, és a sebessége is elfogadható.

Konklúzió

Az elkészített rendszer jelenleg magánhangzókat ismer fel, és az olvasástanítás mellett eredményesen a részképességeikben sérült iskoláskorúak tanítására és siketek oktatására lehet felhasználni. Felmérések szerint az iskoláskorúak 5-7 százaléka küzd valamilyen beszédhibával. Ez azt jelenti, hogy minden iskolában nagy valószínűséggel előfordul néhány ilyen tanuló. A súlyosabb beszédhibával rendelkező tanulóval csak logopédus tud eredményesen foglalkozni, de kisebb hibán egy megfelelően képzett tanító is segíthet. E programrendszer a tanítók munkáját segítheti, kihasználva azt, hogy a gyerekek többségét a számítógép teljesen magával ragadja.

A beszédhibás gyerekek kiejtésének javításánál is nehezebb feladat a hallássérültek beszédjavítása. A siketek oktatása nagy szaktudást, türelmet és sok közös foglalkozást követel. A visszacsatolást általában az oktatónak kell megteremtenie, így eredményeket csak kis csoportokban lehet elérni. Ezen okok miatt sok oktatóra van szükség, illetve minden lehetséges segédeszközt használatba kell venni az eredményes oktatásért. Egy ilyen programrendszer ezért több, mint csupán egy hasznos segédeszköz, hiszen siketek egy jelentős részénél önálló terápiát is lehetővé tesz.

Mindezen érvek figyelembe vételével a szerzők érdemesnek tartják, hogy a jövőben folytassák az ilyen irányú kutatásaikat, és őszintén remélik, hogy az elkészülő rendszereik hozzájárulnak a sikeres és intenzív számítógéppel segített fonológiai tudatosság-fejlesztéshez.

Didaktikai megjegyzések a gépi tanuláshoz

Az írott nyelv elsajátításával a gyermek egy új nyelvet tanul meg. Írás és olvasás nélkül ma már elképzelhetetlen a tanulás. Hogy az írás és olvasás elsajátítása mennyi idő alatt, milyen szinten valósul meg, vagyis mikorra válik rutinná döntően befolyásolni fogja az összes későbbi ismeretszerzés minőségét és mennyiségét.

A vizuális nyelv (írás-olvasás) elsajátítása nem akkor kezdődik, amikor a gyermek az írás és olvasás órákon irányítottan elkezd foglalkozni a vizuális nyelv tanulásával, hanem jóval előtte.

Francia kutatók bizonyították, hogy az újszülött csecsemők sokkal intenzívebben szopják a cumit akkor, amikor a magnóról anyanyelvükön játszanak le szöveget, mint abban az esetben, ha idegen nyelvű szöveget kell hallgatniuk. Ebből a megfigyelésből arra a következtetésre jutottak, hogy az anyanyelv fonémáinak a tudatosodása már a magzati korban elkezdődik.

A beszédtanulás folyamatában (kb. 2 éves kortól) a gyermekekben ki kell hogy alakuljon az a rutin, amelynek a segítségével képesek lesznek érzékelni a hasonlóságokat és a különbségeket a szavak hangjai között. Ez a rutinná váló tudás lesz az alapja annak, hogy a gyermek megtanul beszélni, valamint hogy képes lesz mások beszédét megérteni. Az írás és olvasás tanítását – mint azt a későbbiekben látni fogjuk – megkönnyíti az a tény, hogy a beszédtanulás menetében a gyermekek nagyon sokat játszanak, manipulálnak a nyelvvel. Pl. úgy hoznak létre halandzsaszavakat, hogy egy már az általuk tudott tartalmas szó egy fonémáját elhagyják vagy kicserélik. Valójában ekkor már azt a grammatika elméleti tényt gyakorolják ösztönösen, hogy a fonémák a morféma építő elemei, és hogy a fonémák megkülönböztető jegyként funkcionálnak.

Amikor a szülőktől vagy az óvodában rímelő szövegeket hallgatnak vagy tanulnak (mondókák, rigmusok, versek), már az olvasástanulás tudatos előkészítése folyik, hiszen a rímek felismerése annak felismerését jelenti, hogy két szó egymástól csak 1-2 fonémában különbözik.

A hétköznapi tevékenységekben minden olyan játék és feladat, amely a rész-egész viszony, valamint a hasonlóság, különbség felismerését, alkalmazását teszi szükségessé az írás és olvasás elsajátítását, az intézményesített oktatás befogadását készíti elő.

Az iskolába lépés előtt szinte minden gyermek ismer már egy-két betűt, le tudja írni nevét, nemcsak legalább 10-ig számol, hanem el tud olvasni, le is tud írni néhány számot és szót.

Ma már egyre több gyerek az iskolába lépés előtt megismerkedik a számítógéppel, szakszerűen be és ki tudja kapcsolni, képes néhány játékprogramot megkeresni vagy betölteni, tud azokkal órákon át játszani, vagyis valamilyen szinten képes manipulálni a számítógéppel, igényli a számítógépet, így a számítógépes rendszerek alkalmazása az oktatásban ma még szinte beláthatatlan motivációs erővel bír.

Az olvasáselsajátítás feltétele, hogy hatéves korra a fonológiai tudat kialakuljon a gyermekben.

Adams (1990) a fonológiai tudatnak a következő öt szintjét állapítja meg:

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

Ezek a szintek – vagy legalábbis többségük – 6 éves korra rutinszerűen működnek a legtöbb gyermek elméjében.

Az iskolai olvasástanítás fejlesztő menetében Frith (1985) három szakaszt különböztet meg:

1. „logografikus” szakasz: jellemzője, hogy ebben a szakaszban a gyerekek kevés olyan szót tanulnak meg, amit a grafémakép alapján közvetlenül fel tudnak idézni. E tanítási szakasz a hívóképek és szóképek tanulásának a szakasza. Általában olyan szavak tanítása szerepel, amelyeknek grafémaképe megengedi a gyermek számára, a kiemelkedő grafikai támpontokra való támaszkodást (a homogén gátlás elkerülésének elve).
2. „alfabetikus” szakasz jellemzője, hogy a gyermekek a szavakat kezdik a graféma-fonéma megfeleltetések alapján olvasni.
3. „ortografikus” szakasz: ekkor tanulják meg, hogy helyesírásilag milyen az adott szó, vagyis pl. a magyar nyelvben hogyan kell érvényesíteni a szóelemzés és a hagyomány helyesírásának alapelvét.
4. a rutinszerű olvasás szakasza: amikor a gyermek automatikusan képessé válik az olvasás menetében az előző 3 szakasz közül a lehető legjobbat alkalmazni. Ez általában 8-10 éves korra tehető, de bizonyos esetekben a rutin kialakulása 2-3 évet késhet, diszlexia esetén pedig a diszlexia fajtájától függően egyes szint vagy szintek soha nem alakulnak ki.

Az intézményesített olvasástanítás menetét döntően az határozza meg, hogy az olvasástanítási metodikák közül egy tanító melyiket választja.

Az olvasás tanításában háromféle módszert különböztetünk meg, melyekhez különböző betűtanítási programok társulnak.

1. Szintetikus módszer, mely három szakaszra bontható:
 - előkészítő szakasz, melyben a hangok tanítása kezdődik meg. Fejleszti a tiszta fonémahallást és fonémaejtést.
 - betűtanítás szakasza, általában egyszerre történik a kis és nagy nyomtatott betűk tanítása és az írott kisbetűk tanítása. Majd értelmetlen és értelmes szótagok, szavak olvasása következik.
 - készségfejlesztő szakasz: az olvasás gyakorlása, olvasástechnika fejlesztése, szövegértés kialakítása, megalapozása, fejlesztése. Az olvasás tanítása szótagoló módszerrel történik.
2. Globális módszer: szóképtanulásra épít. Egy szókép megtanítása a nyomtatottan leírt szó és az általa jelölt dolog grafikusan ábrázolt képe között asszociációs kapcsolaton alapul. A szóképek tanítása után következik a hang és betűtanítás.
3. Kombinált módszer: Szintetikus és globális elemek alkalmazásán alapul (pl. NYIK.). A tanítás tárgykép és szókép egyeztetésével folyik, majd

belép a betűtanítás, szó, szószerkezetek olvasása. E program nem szótagoltat, azonnal egész szavakat olvastat.

A XX. század második feléig szinte kizárólagos módszer volt hazánkban a hangoztató – elemző – összetevő olvasástanítási módszer. A másik két módszert (globális és kombinált módszer) nem nyelvészek dolgozták ki, hanem pedagógusok azzal a szándékkal, hogy a gyermek és tanító számára egyaránt hosszú, fáradságos és unalmas fonéma-graféma megfeleltetési szakaszt valamivel kiválthassák. Nem tartjuk feladatunknak a már említett két módszer kritikáját, de nekünk meggyőződésünk, hogy anyanyelvünk olvasástanításában és helyesírásának tanításában a legcélravezetőbb és eredményesebb a magyar nyelv sajátosságaira épülő szintetikus módszer. Ezért olyan metodikai eljárást szeretnénk kifejleszteni, amely megőrzi a szintetikus módszer lényegét, de nem válik unalmassá, éppen ellenkezőleg óriási motivációs erővel bír.

Eddig megvalósítottuk az 'Easy Speech-Vowel Teacher' valós idejű magánhangzó felismerő rendszert, amely képes a szintetikus módszerben alkalmazható graféma-fonéma megfeleltetésnek valós időben.

A hívóképeken megjelenő magánhangzók segítségével a gyermekekben könnyen és gyorsan rögzül a megfelelő fonéma-graféma pár. Mivel a rendszer a szín élességével az ejtett hang „tisztaságát” is visszajelzi segíti az artikulációs bázisban a beszédhang helyes ejtésének rögzülését.

A szoftver nemcsak az olvasástanításban nyújt segítséget, hanem segíthet a siket gyermekek beszédtanításában, mivel a fonéma felismerés valós idejű, az ilyen gyermekeknél kiesett vagy erősen sérült audotórikus percepciót képes lehet vizuális síkon helyettesíteni. (Kaposváron az Óvoda Általános Iskola Diákotthon és Gyermekotthonban (Siketek Iskolája) már jelenleg folynak kísérletek a szoftver alkalmazhatóságáról.)

A rendszer fejlesztése során arra is rájöttünk, hogy a program alkalmas a különböző típusú beszédhibák javítására. A szoftver a jelenlegi verzióban még csak a magánhangzók ejtэшhibáinak terápiájára alkalmazható. A számítógépes program alkalmazásával lehetőség nyílik a logopédiában oly nehézkesnek tartott, időigényes és fáradságos ún. „tükrő-módszer” kiváltására. Vagyis nem kell a logopédusnak ölébe venni vagy maga mellé ültetni a gyermeket, és nagyon sokszor elismételnie egy-egy fonémát magyarázva annak jellemzőit: képzés helye, módja, stb., hanem a gyermek kezébe veszi a mikrofont és addig realizálja az adott fonémát, ameddig a megfelelő visszajelzést meg nem kapja a géptől.

Terveink között szerepel egy olyan szoftver kifejlesztése, amely játékos formában a legtöbb ejtэшhiba terápiájában alkalmazható lesz. Így a logopédusok segítséget kapnak a pöszeség, hangtorzítás, hanghelyettesítés, az orrhangzós beszéd stb. terápiás kezeléséhez.

Az eddigiekben már ígéretes kísérleteket végeztünk a szoftver segítségével bizonyos típusú dadogások terápiájában. Egyes típusú dadogások valószínűsíthető oka az, hogy a gyermek a szókezdő vagy az első szótagban található magánhangzót nem ejti egyenletes intonációval, így a hosszan elnyúló „vibráló” intonáció miatt a levegő elfogy, ezért újra ismétli az első szótagot. A jelenség alapoka az, hogy egyes magánhangzók helyes ejtésének módja nem vagy rosszul rögzült az artikulációs bázisban. A számítógéppel segített magánhangzó tanulás során a megfelelő hangképzés rögzül, így a dadogás enyhülhet vagy megszűnhet.

A szoftver alkalmas lehet a diszlexia terápiában, olyan esetekben, ahol a diszlexia okát a vizuális érzékelés elégtelenségében vagy helytelenségében diagnosztizálták. A szoftver alkalmazásával ugyanis a fonéma vagy szó ejtése során azonnali visszacsatolás nyerhető. A homogén gátlás miatt a hasonló formájú betűket gyakran keverik össze, cserélik fel a diszlexiások. Jelenleg már dolgozunk a programnak egy olyan variációján, amely a leggyakrabban előforduló szavakban azonos időben kijelzi és gyakoroltatja ezeket a kritikus hangpárokat.

A valós idejű hangfelismerő rendszer még nagyon sok didaktikai lehetőséget rejt.

Végző célunk egy olyan multimédiális szoftver kifejlesztése, amelynek a segítségével egy gyermek a számára az optimális idő alatt önállóan képes lesz az olvasás elsajátítására.

Irodalomjegyzék

- [1] Adams, M.I. (1990) *Beginning to read: Thinking and learning about print*. MIT Press, Cambridge
- [2] Bácsi János (2001) A megkésett beszédfejlődés nyelvészeti aspektusai. *Pediáter* 10., 75-81.
- [3] Bishop, C.M. (1995) *Neural Networks for Pattern Recognition*. Oxford University Press.
- [4] Brian, C. J. Moore (1995) *Hearing*. Academic Press.
- [5] Frith, U. (1985) Beneath the surface of developmental dyslexia. In Peterson, K.E., Marshall, J.C., Coltheart, M.C. (eds.) *Surface dyslexia*. Routledge, London, 301-330.
- [6] Kocsor András, Toth László, Paczolay Dénes (2001) A Nonlinearized Discriminant Analysis and Its Application to Speech Impediment Therapy. *Text Speech and Dialogue 2001*. Springer. LNAI 2166. 249-257.
- [7] Lengyel Zsolt (1992) *Nyelvelsajátítási formák*. Szeged.
- [8] Pinker, S. (1999) *A nyelvi ösztön. Hogyan hozza létre az elme a nyelvet*. Typotex.
- [9] Richard, O. Duda, Peter E. Hart, David, G. Stork (2001) *Pattern Classification*. John Wiley & Sons Inc.
- [10] Vapnik, V.N. (1998) *Statistical Learning Theory*. John Wiley & Sons Inc.