

ALKALMAZÁSOK, ESZKÖZÖK

ULTRAHANGOS NYELVKONTÚRKÖVETÉS AUTOMATIKUSAN: A MÉLY NEURONHÁLÓKON ALAPULÓ AUTOTRACE ELJÁRÁS VIZSGÁLATA

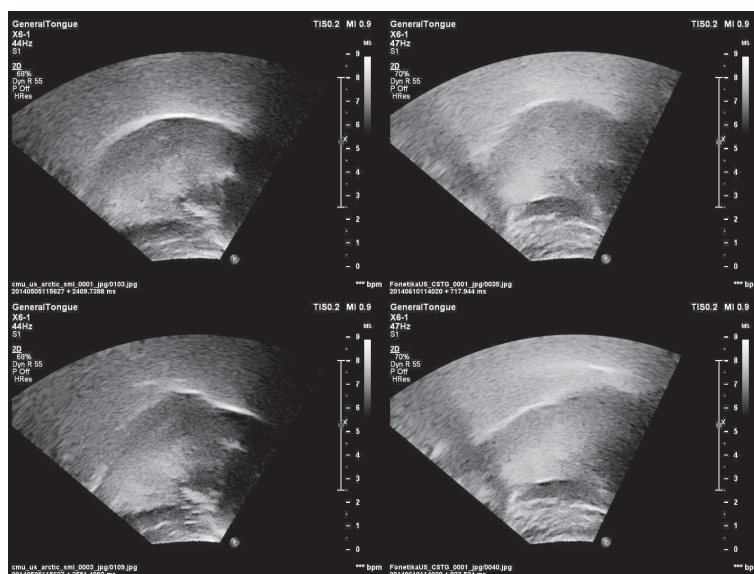
Csapó Tamás Gábor – Csopor Dávid

Bevezetés

Az artikuláció és az akusztikum kapcsolata régóta foglalkoztatja a beszéd-kutatókat. Beszéd közben a nyelv mozgását különböző technológiák segítségével lehet rögzíteni, pl. röntgen (Öhman–Stevens 1963; Bolla 1995), ultrahang (Stone et al. 1983; Stone 2005), EMA – elektromágneses artikulográf (Schönle et al. 1987; Mády 2008), MRI – mágnesrezonancia-képalkotás (Baer et al. 1991; Woo et al. 2012). Az ultrahangos technológia előnye, hogy egyszerűen használható, elérhető árú, valamint nagy felbontású (akár 800×600 pixeles) és nagy sebességű (akár 100 képkocka/s-os) felvétel készíthető vele. Az ultrahang-, MRI- és röntgentechnológiák hátránya viszont ebben a témakörben, hogy a rögzített képsorozatból ki kell nyerni a nyelv körvonalát ahhoz, hogy az adatokon további vizsgálatokat lehessen végezni.

Az 1. ábra néhány példát mutat a nyelvről rögzített ultrahangfelvétellel. A felvételeken bal oldalon látható a nyelvgyök, jobb oldalon a nyelvhegy; a kettő között a nyelv felső felülete. A felvételek során az ultrahangtranszducert az áll alá helyezik; így az ultrahangjelben a legnagyobb változást a nyelv izomzatának felső határa okozza, ami az ultrahangos képeken ideális esetben jól kivehető fehér sávot eredményez. Mivel a hullámok nagy része nem jut tovább a nyelv felső határán, így a távolabbi szövetpontokról, a szájpadról kevesebb az információnk. Az 1. ábrán látható, hogy a képek minősége széles skálán mozog, mivel az ultrahang-technológia nem mindig nyújt teljesen tökéletes nyelvkontúrt. A bal felső képen jól kivehető a nyelv kontúrja; ezzel szemben a bal alsó képen a kontúr nem folytonos, hanem szakadás vagy ugrás látható. A jobb felső képen a nyelvkontúr kevésbé erőteljesen látszik; a jobb alsó képen pedig több kontúr is látható egymás felett, amit valószínűleg a szájpadról közelsége okoz. A kép minősége függ a beszélőtől, általában fiatalabbnál és nőknél jobb, de ez függ a száj hidratációjától is (Stone 2005).

Ennek oka egyrészt az lehet, hogy a nőknek kisebb az állkapcsa, ezért a szájüreg nagyobb részére terjed ki a felvétel; másrészt idősebbeknél nagyobb mértékű az áll elzsírosodása, ami nehezíti a nyelvkontúr követését.



1. ábra

Ultrahangfelvételek a nyelvről (A képeken bal oldalon látható a nyelvgyök, jobb oldalon a nyelvhegy; a kettő között a nyelv felső felülete.)

Az ultrahangos nyelvkontúrkövetés hagyományosan manuális vagy félautomatikus módon történt, azonban az elmúlt időszakban automatikus megoldások is megjelentek erre a célra.

Az **EdgeTrak** volt az első nyelvkontúrkövető program, amelyet széles körben kezdtek el használni (Li et al. 2005). A módszer a klasszikus „kígyókövető” algoritmusokon alapul (Akgul et al. 1999), és jelenleg az automatikus ultrahangos nyelvkontúr *de facto* szoftverének számít. A **Palatoglossatron** alapvető célja, hogy a szájpadróláshoz képest legyen lehetőség megállapítani a nyelv helyzetét (Mielke et al. 2005). A programban alkalmazott **GloSsatron** félautomatikus nyelvkontúrkövetés a képben lévő gradiensek (hirtelen fényerőváltozás) mérésén alapul (Baker et al. 2005). Az **Ultra-CATS** program a manuális kontúrkövetés sebességének növelésére készült, de rendelkezik félautomatikus nyelvkontúrkövető funkcióval is (Bressmann et al. 2005). A **TongueTrack** módszer a fentiekkel szemben az egymás után következő ultrahangképek térbeli és időbeli kapcsolatát próbálja meg kihasználni, hiszen a nyelv mozgása korlátozott sebességű (Tang et al. 2012).

Az **AutoTrace** szoftver az automatikus nyelvkontúrkövetésre mély neuronhálókat, ezen belül translational Deep Belief Network-öt (tDBN) alkalmaz (Fasel–Berry 2010). Mivel ez géptanulás-alapú eljárás, ezért először jelentős méretű tanítóadatra, azaz manuálisan berajzolt nyelvkontúrra van szükség. Berry és munkatársai (2012) javasolnak egy módszert, amellyel nagy mennyiségű képből kiválasztható az az optimális részhalmoz, amelyet a gépi tanuláshoz érdemes felhasználni. Sung és munkatársai (2013) azt találták, hogy a nyelvkontúrkövetés pontossága növelhető az „iteratív újratanítás” eljárással is. Az AutoTrace a <https://github.com/jjberry/Autotrace> honlapról szabadon letölthető (Hahn-Powell–Archangeli 2014). A legújabb kutatási eredmények szerint a manuális nyelvkontúrkövetés átlagos négyzetes hibája 7 pixel, és a fenti automatikus algoritmusok közül az EdgeTrak, a TongueTrack és az AutoTrace eljárások nyújtják a legpontosabb nyelvkontúrkövetést (Csapó–Lulich 2014a, 2014b).

A tanulmányban bemutatjuk legújabb eredményeinket, melyekhez az AutoTrace eljárás automatikus nyelvkontúrkövetését vizsgáltuk (Csapor 2014). A manuális és az automatikus kontúrok hibájának mérésére hagyományosan alkalmazott átlagos négyzetes hiba mellett új típusú hibamértékeket vezetünk be, melyek jobban mérik az automatikusból hiányzó vagy esetlegesen hozzáadott nyelvkontúrszakaszokat. Emellett meghatározzuk, hogy a tanítóadat mennyiségének függvényében milyen mértékben tudja az automatikus nyelvkontúrkövetés a manuálist közelíteni.

Anyag, módszer

A továbbiakban bemutatjuk a kutatás során rögzített új ultrahangos beszédadatbázist; valamint a manuális és automatikus nyelvkontúrkövetés körülményeit, továbbá ezek összehasonlításának módszerét.

Felvételi körülmények

Egy amerikai angol (jelölés: EN1) és egy magyar anyanyelvű (jelölés: HU1) férfi beszélőtől rögzítettünk felvételeket az Indiana University Speech Production Laboratory csendesített szobájában (Csapó–Lulich 2014b, 2014c). EN1 beszélő a CMU-ARCTIC adatbázis első 135 mondatát (Kominek–Black 2003), míg HU1 beszélő a PPBA adatbázis első 210 mondatát olvasta fel (Olaszy 2013). A beszédfelvételek Shure kondenzátormikrofonnal készültek, melyet 48 kHz-en egy National Instruments jelfeldolgozó kártyával digitalizáltunk. A nyelv mozgását Philips EpiQ-7G ultrahangos rendszerrel és xMatrix 6-1 MHz transzdúccerrel rögzítettük. A felvételek során az ultrahangfej elmozdulásának elkerülése végett egy speciálisan erre a feladatra kialakított sisakot alkalmaztunk (típusa: Ultrasound Stabilisation Headset, Articulate Instruments Ltd), mellyel az ultrahangtranszdúcer szorosan a beszélő álla alá rögzíthető (Wrench 2007). Az ultrahanggépet a háttérzaj csökkentésének céljából a süketszobán kívülre helyeztük, és csak az ultrahangfejet és -kábelt vezettük a beszélőhöz. Az arc alsó részéről (így az ajkak mozgásáról) videofel-

vétel készült, valamint a hangszalagok rezgéséről EGG-felvételt rögzítettünk; ezeket azonban a jelen kutatásban nem használtuk fel.

Az ultrahangadatok eredetileg DICOM formátumban készültek 800×600 pixel felbontásban és 40–50 kép/s közötti sebességgel. A DICOM fájlokat az Image-J programmal (National Institutes of Health, USA, v1.46a, <http://imagej.nih.gov/ij>) JPG képekké alakítottuk.

Az elkészült párhuzamos beszéd – nyelvultrahang – videó – EGG adatbázis kutatási célra szabadon elérhető és felhasználható (Csapó–Lulich 2014c).

Manuális nyelvkontúrkövetés

A manuális nyelvkontúrkövetést a fenti felvételek egy kisebb részén (EN1 beszélő: 8 mondat, HU1 beszélő: 9 mondat) egy hallgató végezte egy speciálisan erre a feladatra készült weboldal segítségével. A weboldal az Indiana University, Speech Production Laboratoryban készült, és a lokális hálózaton elérhető. A hallgató a feladat elvégzése előtt általános ismertetést kapott a képek orientációjáról (pl. nyelvhegy a jobb oldalon; a nyelv hátulja a bal oldalon; a jobb oldali sötét folt az állkapocs csontja), illetve utasításokkal láttuk el, hogy az 1. ábrához hasonló helyzetekben hogyan végezze a nyelvkontúrok berajzolását. Minden esetben a nyelv felső részére utaló fehér sáv alján történt a kontúr berajzolása. A feladat elvégzése összesen 5-6 órányi időt vett igénybe. A nyelvkontúrt képenként kb. 200 ponttal reprezentáltuk. A manuális nyelvkontúrkövetés eredménye SQL adatbázisba került, melyből CSV formátumba exportálva lehet az adatokat feldolgozni.

Az EN1 beszélőtől összesen 1140 ultrahangos képet, míg a HU1 beszélőtől összesen 1457 képet használtunk fel a kísérletekben.

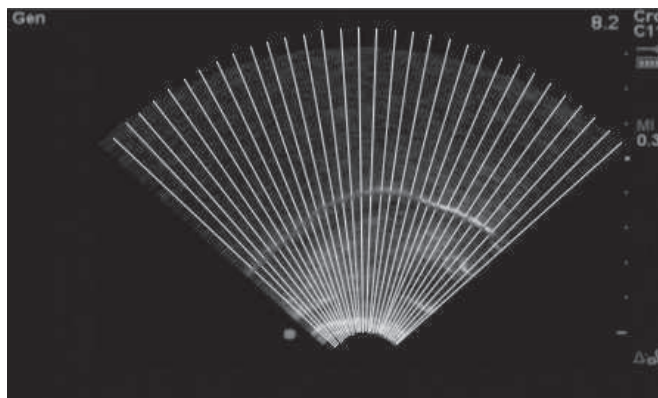
Automatikus nyelvkontúrkövetés

Az automatikus nyelvkontúrkövetést az AutoTrace programmal végeztük. A 800×600 pixel méretű képeket először átméreteztük 550×413 -re, majd eltoltuk vízszintesen 140 pixellel, függőlegesen 30 pixellel, végül 720×480 pixel méretű képet kapva (az origó a bal felső sarokban volt). Az átméretezés és eltolás célja, hogy az AutoTrace-ben definiált polárkoordináta-rendszer illeszkedjen a képeken az ultrahang által rögzített területhez (2. ábra). A képekből kiválasztottuk azt a részt, amely várhatóan tartalmazza a nyelv teljes szakaszát. A manuális nyelvkontúr pontsorozatát interpoláltuk az AutoTrace 32 radiális egyenesére. Az interpoláció során az adott poláregyeneshez tartozó szomszédos pontokat súlyoztuk aszerint, hogy a polárkoordinátára konvertált pontok szöge mennyire van közel az adott egyeneshez. Ennek célja a kvantálási hiba csökkentése volt (vö. Csopor 2014).

Az AutoTrace-ben az automatikus nyelvkontúrkövetéshez egy betanító lépésre van szükség. Tanítóadatként először EN1 beszélő első mondatát használtuk fel (összesen 169 kép), majd az eredményként kapott betanított neurális háló modellel a 8. mondaton végeztünk tesztelést. A tanítóadat méretét fokozatosan, mondatonként növeltük; majd a tesztelést ugyanazon a mondaton végeztük. Ezután HU1 beszélő ultrahangos képeivel ismételtük meg a fenti

kísérletet. Minden tanítási és tesztelési lépés után megvizsgáltuk az automatikus nyelvkontúrkövetés pontosságát, melyet a következő fejezetben ismertett hibamértékekkel számítottunk.

Az AutoTrace mélyneuronháló-alapú gépi tanulást végez, melyre a translational Deep Belief Network architektúrát alkalmazza. A neurális hálózatban alapértelmezésben két rejtett réteget alkalmaznak (Fasel–Berry 2010).



2. ábra

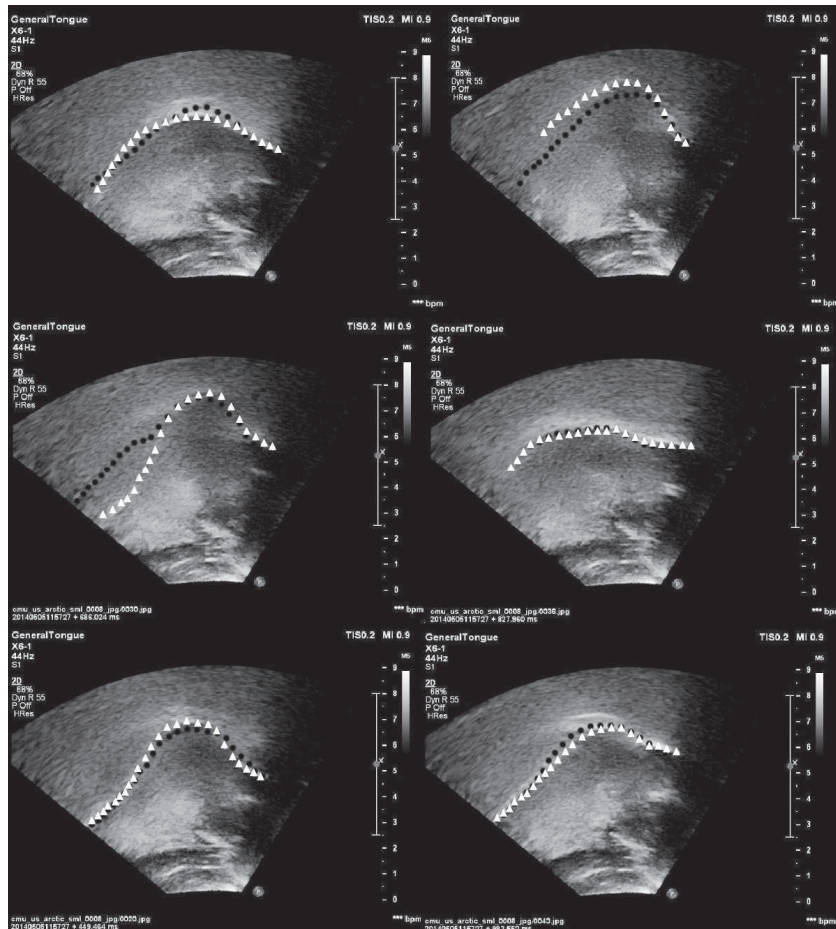
Az AutoTrace-ben definiált polárkoordináta-rendszer (A nyelvkontúr és a 32 radiális egyenes metszéspontjai az AutoTrace-ben eltárolt értékek. Forrás: [https://github.com/jjberry/Autotrace.](https://github.com/jjberry/Autotrace))

Az összehasonlítás módszere

Az automatikus nyelvkontúrkövetés tipikus hibái a 3. ábrán láthatóak: fehér háromszögekkel jelöltük a manuálisan berajzolt nyelvkontúrt; míg fekete körökkel az AutoTrace-ből származó automatikus kontúrt. Referenciának a manuális kontúrt használjuk. Három fő hibatípust állapítottunk meg: a) az automatikus kontúr bizonyos távolságnál messzebb van a referenciától (pl. bal középső kép bal oldala a 3. ábrán); b) a referenciában egy adott szakaszon nincs berajzolt pont, de az automatikus kontúrban van (pl. jobb felső kép bal oldala a 3. ábrán); c) a referenciában egy adott szakaszon van berajzolt pont, de az automatikus kontúrban nincs (pl. jobb középső kép bal oldala a 3. ábrán). A nyelvkontúrkövetés hibáinak mérésére négy hibamértéket alkalmazunk és vizsgálunk: (i) átlagos négyzetes hiba (Root Mean Squared Error, RMSE), (ii) két kontúr közötti terület (Area), (iii) hozzáadások aránya (Insertion), (iv) törlések aránya (Deletion).

A szakirodalomban általában az RMSE hibamértéket vagy ehhez hasonlókat (például átlagos abszolút hiba) alkalmaznak (Li et al. 2005; Sung et al. 2013; Hahn-Powell–Archangeli 2014). Az RMSE számítása során páronként vesz-

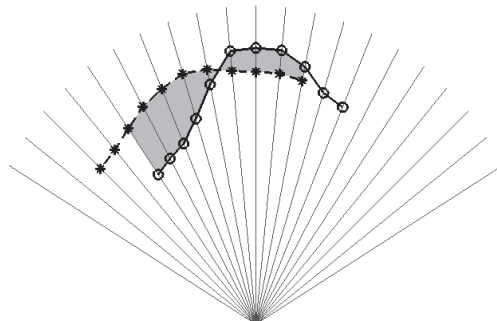
szük a manuális és automatikus kontúrban lévő pontokat; koordinátaikat egymásból kivonjuk; ezen távolságokat négyzetre emeljük, majd összegezzük; végül az összegből gyököt vonunk. Azokat a pontokat nem vesszük figyelembe az RMSE számításánál, ahol nincs automatikusan és manuálisan is meghatározott pont. Az RMSE tehát jól méri az a) típusú hibát, de ugyanakkor nem jellemzi a b) és c) típusú hibákat.



3. ábra

Néhány példa manuálisan berajzolt (fehér háromszögek) és automatikusan előállított (fekete körök) nyelvkontúrra

Az Area hibamérték számítását a 4. ábra mutatja (a szürkével jelölt terület). A második paraméterünk bevezetésének motivációja az volt, hogy az RMSE-vel szemben ez várhatóan pontosabban méri a manuális és automatikus nyelvkontúr közötti hibát (Csopor 2014). A pontokat a számítás során egyenes szakaszokkal kötöttük össze, így meghatározva a két kontúr közötti területet. Az RMSE-hez hasonlóan itt is csak akkor lehet területet számolni, ha adott poláregyenesen van automatikusan meghatározott és manuálisan be rajzolt pont is, azaz az Area szintén az a) típusú hibát méri.



4. ábra

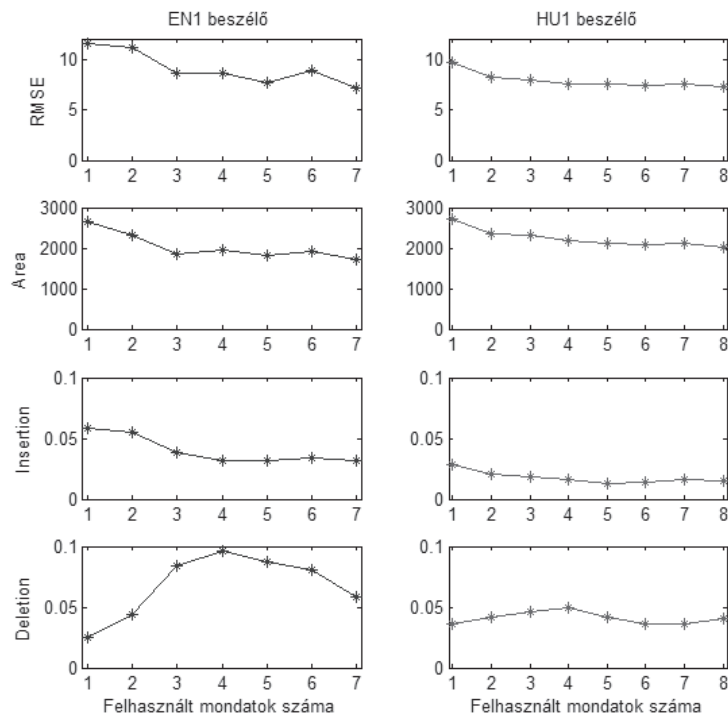
Az Area hibamérték számítása két nyelvkontúr között
(Manuális: folytonos vonal, körök; automatikus: szaggatott vonal, csillagok.)

A b) és c) típusú hibák méréséhez a beszédfelismerésben használt Word Error Rate-hez hasonlóan kialakítottuk a Tracking Error Rate (TER) változóhármast, mely tartalmazza az Insertiont (beillesztést), a Deletion (törlést) és a Substitution (helyettesítést) (Csopor 2014). A Substitution az RMSE-nek megfelelő, ezért azzal a továbbiakban nem foglalkozunk. A TER-Insertion számítása során megszámloljuk azokat a pontokat, melyek az automatikus kontúrban megjelennek, de a referenciában nem; majd ezt elosztjuk az összes referenciában lévő pontok számával. A TER-Deletion számítása épp fordított: megszámloljuk azokat a pontokat, melyek az automatikus kontúrban nem jelennek meg, de a referenciában igen; majd ezt elosztjuk az összes referenciában lévő pontok számával. Mivel a TER-Insertion és a TER-Deletion is arányszámok, ezért értéküket százalékban adjuk meg.

Eredmények

A mérési eredmények az 5. ábrán láthatóak, amelyen a négyféle hiba változását ábrázoltuk a tanítóadat mennyiségének függvényében, külön a két beszélőre. Megfigyelhető, hogy az RMSE és az Area mindkét beszélő harmadik modelljénél (azaz amikor összesen 3 mondatot, kb. 600 képet használunk tanításra) beáll egy bizonyos értékre, és egyre alacsonyabb mértékben csökken

további tanítóadat hozzáadásával. Mind EN1, mind HU1 beszélő esetén három mondatot felhasználva az RMSE 8 pixelre csökken, amely közel meg-
egyezik a manuális nyelvkontúrkövetés átlagosan 7 pixel hibájával (Csapó-
Lulich 2014b). Az Area hibamérték az RMSE-hez hasonló tendenciát mutat:
a két hibamérték között a korreláció magas (EN1: 0,97; HU1: 0,99).



5. ábra

A négyféle hibamérték változása a tanítóadat mennyiségének függvényében a két beszélőre

Az Insertion és Deletion értékei nem változnak egyértelműen a tanítóadat növekedésével. Az Insertion először csökken az ötödik mondat hozzáadásáig, majd kis mértékben növekszik. A Deletion hiba a negyedik mondat hozzáadásáig növekszik mindkét beszélő esetén (EN1-nél egészen 9,5%-ig); további tanítóadat hozzáadásával pedig csökkenni kezd. Ezen két hibamérték értékei valószínűleg nem közvetlenül függnek a tanítóadat mennyiségétől. Az Insertion és Deletion típusú hibák csökkentéséhez a tanítóadat előzetes változtatása vagy más neurális háló architektúra lehet szükséges.

Következtetések

Az eredmények szerint az AutoTrace programban két rejtett rétegből álló translational Deep Belief Network típusú mély neuronháló architektúrával, kb. 600 ultrahangos képet tanítóadatként használva már közelíthető a manuális nyelvkontúrkövetés átlagos négyzetes hibaként mért pontossága. Az átlagos négyzetes hiba a tanítóadat mennyiségét növelve folyamatosan csökken. Az automatikusan meghatározott nyelvkontúrszakasz hossza is változik a tanítóadat függvényében: eredményeink szerint a tanítóadat felét felhasználva volt a legrövidebb a nyelvkontúr. Előzetes kísérleteink alapján az Insertion és Deletion típusú hiba a neurális hálózat architektúrájának változtatásával (pl. rejtett rétegek számának növelése) tovább csökkenthető (Csopor 2014).

Az automatikus nyelvkontúrkövetés a beszédkutatás alapkérdéseinek megválaszolása mellett hasznos lehet a nyelvoktatásban, a beszédrehabilitációban, illetve a beszédtechnológiában, audiovizuális beszéd-szintézisben is (Hueber et al. 2011). A további kutatás során tervezzük a fenti adatbázis képeiből ki-nyerhető nyelvkontúr alakját összehasonlítani a beszédjelből mérhető formánsértékekkel. A fenti kutatásban a kétdimenziós ultrahang-technológiát alkalmaztuk, melynek során a nyelv középvonaláról rendelkezünk információval. A legújabb, háromdimenziós technológiával azonban a nyelv teljes felületéről információt kaphatunk beszéd közben. Előzetes kísérleteink szerint a nyelv mozgása közel sem szimmetrikus, így a háromdimenziós vizsgálattal olyan kérdésekre is válasz kapható, hogy a nyelv mozgása és alakja milyen mértékben járul hozzá az akusztikai kimenet formálásához.

Az elkészült párhuzamos beszéd – nyelvultrahang – videó – EGG adatbázis kutatási célra szabadon elérhető és felhasználható (Csapó–Lulich 2014c).

Irodalom

- Akgul, Yusuf Sinan – Kambhamettu, Chandra – Stone, Maureen 1999. Automatic extraction and tracking of the tongue contours. *IEEE Transactions on Medical Imaging* 18/10. 1035–1045.
- Baer, Thomas – Gore, John C. – Gracco, L. C. – Nye, Patrick W. 1991. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels. *The Journal of the Acoustical Society of America* 90/2. 799–828.
- Baker, Adam – Mielke, Jeff – Archangeli, Diana 2005. *Tracing the tongue with GLoSsatron*. Előadás. Ultrafest III. Tucson, AZ, USA. <http://apil.arizona.edu/presentations/GLoSsatronPresentation.ppt>. (A letöltés ideje: 2015. február 1.)
- Berry, Jeff – Fasel, Ian – Fadiga, Luciano – Archangeli, Diana 2012. Training deep nets with imbalanced and unlabeled data. In: *Proceedings of Interspeech*. Portland, OR, USA. 1756–1759.
- Bolla, Kálmán 1995. *Magyar fonetikai atlasz. A szegmentális hangszerkezet elemei*. Nemzeti Tankönyvkiadó, Budapest.
- Bressmann, Tim – Heng, Chiang-Le – Irish, Jonathan C. 2005. Applications of 2D and 3D ultrasound imaging in speech-language pathology. *Journal of Speech-Language Pathology and Audiology* 29/4. 158–168.

- Csapó, Tamás Gábor – Lulich, Steven M. 2014a. *Comparison of tongue contour extraction methods from ultrasound images for use in text-to-speech synthesis*. Előadás. Inaugural Conference of the Hungarian Cultural Association. Bloomington, IN, USA. <http://speechlab.tmit.bme.hu/csapo/downloads/CsapoLulich-HungStudies2014-presentation.pdf>. (A letöltés ideje: 2015. február 1.)
- Csapó, Tamás Gábor – Lulich, Steven M. 2014b. Tongue contour tracings from 2D ultrasound image sequences: quantification of measurement error using manual and automatic tracing methods. (Előkészületben.)
- Csapó, Tamás Gábor – Lulich, Steven M. 2014c. IU_ULTRASOUND data set. http://speechlab.tmit.bme.hu/csapo/IU_ULTRASOUND/. (A letöltés ideje: 2015. január 7.)
- Csopor Dávid 2014. Mély neuronhálók alkalmazása ultrahangos nyelvkontúr követésre. TDK-dolgozat. BME VIK. <https://tdk.bme.hu/VIK/DownloadPaper/Mely-neuronhalok-alkalmazasa-ultrahangos>. (A letöltés ideje: 2014. november 5.)
- Fasel, Ian – Berry, Jeff 2010. Deep Belief Networks for Real-Time Extraction of Tongue Contours from Ultrasound During Speech. In: *Proceedings of ICPR*. Istanbul, Turkey, 1493–1496.
- Hahn-Powell, Gustave V. – Archangeli, Diana 2014. AutoTrace: An automatic system for tracing tongue contours. *The Journal of the Acoustical Society of America* 136/4. 2104.
- Hueber, Thomas – Benaroya, Elie-Laurent – Denby, Bruce – Chollet, Gérard 2011. Statistical mapping between articulatory and acoustic data for an ultrasound-based silent speech interface. In: *Proceedings of Interspeech*. Florence, Italy. 593–596.
- Kominek, John – Black, Alan W. 2003. *CMU ARCTIC databases for speech synthesis*. Carnegie Mellon University.
- Li, Min – Kambhamettu, Chandra – Stone, Maureen 2005. Automatic contour tracking in ultrasound images. *Clinical Linguistics and Phonetics* 19/6-7. 545–554.
- Mády Katalin 2008. Magyar magánhangzók vizsgálata elektromágneses artikulográfiával normál és gyors beszédben. *Beszédkutatás* 2008. 52–66.
- Mielke, Jeff – Baker, Adam – Archangeli, Diana – Racy, Sumayya 2005. Palatron: A technique for aligning ultrasound images of the tongue and palate. *Coyote Papers* 14. 97–108.
- Olaszy Gábor 2013. Precíziós, párhuzamos magyar beszédatbázis fejlesztése és szolgáltatásai. *Beszédkutatás* 2013. 261–270.
- Öhman, Sven – Stevens, Kenneth N. 1963. Cineradiographic studies of speech: Procedures and objectives. *The Journal of the Acoustical Society of America* 35. 1889.
- Schönle, Paul W. – Gräbe, Klaus – Wenig, Peter – Höhne, Jörg – Schrader, Jörg – Conrad, Bastian 1987. Electromagnetic articulography: Use of alternating magnetic fields for tracking movements of multiple points inside and outside the vocal tract. *Brain and Language* 31/1. 26–35.
- Stone, Maureen – Sonies, Barbara C. – Shawker, Thomas H. – Weiss, George – Nadel, Lawrence 1983. Analysis of real-time ultrasound images of tongue configuration using a grid-digitizing system. *Journal of Phonetics* 11. 207–218.
- Stone, Maureen 2005. A guide to analysing tongue motion from ultrasound images. *Clinical Linguistics and Phonetics* 19/6-7. 455–501.

- Sung, Jae-Hyun – Berry, Jeff – Cooper, Marissa – Hahn-Powell, Gustave – Archangeli, Diana 2013. Testing AutoTrace: A machine-learning approach to automated tongue contour data extraction. *Ultrafest VI*. Edinburgh, UK. 9–10. http://www.qmu.ac.uk/casl/conf/ultrafest_2013/docs/G_HahnPowel_1_ultrafest.pdf. (A letöltés ideje: 2015. február 1.)
- Tang, Lisa – Bressmann, Tim – Hamarneh, Ghassan 2012. Tongue contour tracking in dynamic ultrasound via higher-order MRFs and efficient fusion moves. *Medical Image Analysis* 16/8. 1503–1520.
- Woo, Jonghye – Murano, Emi Z. – Stone, Maureen – Prince, Jerry L. 2012. Reconstruction of high-resolution tongue volumes from MRI. *IEEE Transactions on Biomedical Engineering* 59/12. 3511–3524.
- Wrench, Alan 2007. Articulate Assistant Advanced: Ultrasound module. *Ultrafest IV*. New York, NY, USA.

Az első szerzőt részben a Fulbright ösztöndíj támogatta a kutatás során. A szerzők köszönetüket fejezik ki Steven M. Lulichnak (Indiana University, Bloomington, IN, USA), hogy a felvételek a Speech Production Laboratory eszközei segítségével elkészülhettek. A szerzők emellett köszönetet mondanak Elizabeth Mazzoconak (Indiana University, Bloomington, IN, USA) a manuális nyelvkontúrkövetésért, valamint Gustave Hahn-Powellnek (Arizona Phonological Imaging Lab, University of Arizona, Tucson, AZ, USA) az AutoTrace eljárás használatában nyújtott segítségével.

VIZUÁLIS PERCEPCIÓ ÉS NYELVI FELDOLGOZÁS

Hunyadi László – Szekrényes István – Sziklai István

Bevezetés

A jelen tanulmányban olyan kísérletek és próbálkozások eredményeiről, tapasztalatairól számolunk be, amelyek bizonyos lényeges elemeiben akár nehezen értelmezhetőknél is tűnhetnek a fonetika és a fonológia szorosabban vett megközelítésében, azonban amelyek ugyanakkor tanulságokkal is szolgálhatnak a beszédfeldolgozás általános mechanizmusainak jobb megértése szempontjából.

A két éve folyó, több résztvevős, inter- és multidiszciplináris kutatás középpontjában az a problematika áll, milyen módon lehetne elősegíteni a születésüktől fogva siket vagy erősen nagyothalló gyermekek magyar nyelvi beszédfejlődését, különös tekintettel a hangzók helyes kiejtésére és a beszéd prozódijára. Az általános tapasztalat szerint ugyanis a halláson keresztül kapott visszajelzés teljes vagy részleges hiánya jelentősen megnehezíti a természe-