



Neumann János
Számítógép-tudományi Társaság
Orvos-biológiai Szakosztály



00000000000001111111++/:`~`
oooooooooooooooooooooooooooooooooooo

Orvosi Informatika 2018.

A XXXI. Neumann Kollokvium konferencia-kiadványa

Szegedi Tudományegyetem, Szeged
2018. november 30 - december 1.

Szerkesztők:
Bari Ferenc, Ráosi Ferenc
Szegedi Tudományegyetem

Szerkesztők: Bari Ferenc, Rárosi Ferenc

Borítóterv: Eckert László

Kiadta a Neumann János Számítógép-tudományi
Társaság

Szeged, 2018.

ISBN 978-615-5036-14-9

Orvosi Informatika 2018.

A XXXI. Neumann Kollokvium

Szeged, 2018. november 30 - december 1.

Nyomdai kivitelezés: Innovariant Nyomdaipari Kft.

6750 Algyő, Ipartelep 4.

Felelős vezető: Drágán György

www.innovariant.hu

© Neumann János Számítógép-tudományi Társaság. Minden
jog fenntartva

© John von Neumann Computer Society. All Rights Reserved

Bevezető

„Számítástechnikai és kibernetikai módszerek alkalmazása az orvostudományban és a biológiában” címmel 1970-ben indította útjára a Neumann-kollokvium rendezvénysorozatot Kalmár László akadémikus a szegedi József Attila Tudománytem Kibernetikai Laboratóriumából. A Neumann János Számítógép-tudományi Társaság Orvos-biológiai Szakosztálya néhány éves szünet után 2012 óta ismét rendszeresen, évente rendezi meg a Kollokviumot.

A XXXI., idén kétnapos konferencia újfent lehetőséget teremt arra, hogy a különböző tudományos műhelyek képviselői bemutatkozzanak, közzé tegyék legújabb eredményeiket. A bejelentett előadások mindegyike érdekes területre fókuszál. Tükrözi mindazokat a kihívásokat, amelyekkel az egészségügyi informatika művelői nap, mint nap szembesülnek. Mérünk, adatokat gyűjtünk és tárolunk (ha lehet, előre megtervezett számban és formában), majd szofisztikált módszerekkel törekszünk a lényegi információ megtalálására és bemutatására. Népegészségügyi mutatókra, tendenciákra, ok-okozati összefüggésekre igyekszünk rámutatni. Praktikus megoldásokat keresünk informatikai problémákra. Tesszük mindezt annak érdekében, hogy minél többet megtudjunk az életjelenségekről és az egészségügy makro és mikro folyamatairól, és hogy mindezt a tudást a közjó szolgálatába tudjuk állítani.

A Kollokvium, a szó eredetileg párbeszédet, beszélgetést jelent. A családi légkör, a kötetlenség bizonyára ebben az évben is jellemzője lesz a szegedi rendezvénynek. Így nem lesz akadálya a beszélgetéseknek, a párbeszédnek. Kívánom, hogy ebben az esztendőben se legyenek fel nem tett és megválaszolatlan kérdések.

A szervezők nevében köszöntöm a Kollokvium résztvevőit, eredményes tudományos munkát és termékeny, tartalmas beszélgetéseket kívánok mindannyiunknak.

Szeged, 2018. november

Bari Ferenc

Tudományos bizottság

Elnök: Surján György, ÁEEK, Budapest

Tagok: Bertalan Lóránt, Semmelweis Egyetem

Kósa István, Pannon Egyetem

Nagy István, Országos Kardiológiai Intézet

Nyári Tibor, Szegedi Tudományegyetem

Szanyiné Forczek Erzsébet, Szegedi Tudományegyetem

Tolnai József, Szegedi Tudományegyetem

Vassányi István, Pannon Egyetem

Tartalomjegyzék (rövid közlemények)

Az EKG parametrizálásának egy biztató kísérlete IV	13
Szövegesen rögzített echokardiográfia leletek numerikus értékeinek strukturálása	20
A szív működés paramétereinek távoli monitorozása és tárolása.....	26
Több szenzort alkalmazva növelhető az indirekt vérnyomásmérés pontossága	32
Látens változók hatása dichotom kimenetű vizsgálatok kiértékelésére.....	37
Stroke-on átesett dysarthriás betegek beszédének gépi elemzése – kezdeti eredmények	43
Betegségek beazonosítása a NEAK forgalmi adataiból	50
Betegutak elemzése az akut stroke ellátásban	54
Új ellátóhelyek létrehozásának hatása az ellátórendszerre	59
Ischaemiás stroke magyarországi túlélési adatainak vizsgálata	64
Az öregedés hatása az agykérgi terjedő depolari-záció spektrális és multifraktál-mintázatára	68
A magyar közsférabeli weboldalak használhatóságának, akadálymentesítésének és biztonságának vizsgálata	74
Magyarországi egészségüggyel foglalkozó weblapok akadálymentességi tesztelése	80
Személyre szabható inzulin terápia az intenzív ápolásban	86
Az emberi hibátényező vizsgálata CT felvételek manuális kiértékelésekor	93
Magyarország gyógyszerfogyasztásának előrejelzése a demográfiai viszonyok változásának alapján	98
Metasztatikus kolorektális daganatos betegek terápiás mintázatának meghatározása	102

Öngyilkosság általi halálozások szezonális változása 1995 és 2014 között Magyarországon.....	108
A magzati halálozások alakulása Magyarországon.....	114
A légáramlás paramétereinek vizsgálata végeelem analízis segítségével, számítógéppel generált 3 dimenziós gégemodellekben	118
Development of a Wingsuit-style gamified application.....	122
Full-stack felhő alapú mobil backend offline képes WebDAO-val	128
Felhőben tárolt egészségügyi adatok védelme ABAC modellel	134
Ismételt mérések figyelembe vétele a statisztikai elemzés során	140
A mesterséges intelligencia egészségügyi alkalmazásai	144
Természetesebb irányítást biztosító eszköz számítógépes műtéttervezéshez.....	150
Beszéd felismerés objektív vizsgálása pupillometriás vizsgálatok segítségével	153
Egészségügyi informatikai rendszerek biztonsági kérdései	158
Implantálható hallókészülékkel rendelkező pácienseket nyilvántartó rendszer újabb fejlesztései.....	164
Névmutató.....	169

Stroke-on átesett dysarthriás betegek beszédének gépi elemzése – kezdeti eredmények

Tóth László¹, Kovács György², Ivaskó Livia^{3,4}, Tóth Alinka^{3,5},
Jakab Katalin⁵, Vécsei László^{5,6}

¹Számítógépes Algoritmusok és Mesterséges Intelligencia Tanszék, SZTE

²MTA-SZTE Mesterséges Intelligencia Kutatócsoport

³Fejlődéses és Neuropragmatikai Kutatócsoport, Általános Nyelvészeti Tanszék, BTK, SZTE

⁴MTA-SZTE-DE Elméleti Nyelvészeti Kutatócsoport

⁵Neurorehabilitációs Osztály, Neurológia Klinika, ÁOK, SZTE

⁶MTA Idegtudományi Kutatócsoport

Összefoglaló: A stroke okozta kortikális dysarthria jelentősen megnehezítheti a stroke-on átesett páciensek érthető artikulációját. Munkánk hosszú távú célja olyan műszaki megoldások kialakítása, amelyek segítik ezen betegek beszédkommunikációját. Jelen cikkben egy kézen hozzáférhető, de eredetileg más célra készített beszédfelismerő rendszer dysarthriás beszédre való alkalmazhatóságát vizsgáltuk. Mivel alacsony felismerési pontosságot kaptunk, így második lépésben objektív beszédelemzési lépéseket végeztük annak megértése céljából, hogy a dysarthriás beszéd milyen akusztikai sajátosságokban tér el a normál beszédétől.

Bevezető

A kortikális sérülésből eredő dysarthriás beszédfolyamatok elsősorban úgy jellemezhetők, mint a beszéd primer motoros tervezésének és kivitelezésének nem megfelelő működéséből eredő specifikus mintázatok [4]. A stroke eredetű dysarthria a motoros funkciók érintettségéből eredő beszédzavar, mely (a stroke kiterjedésétől függően) nem érinti a nyelvi tervezési folyamatokat, a legtöbb esetben nem kíséri afázia. A klinikai differenciáldiagnosztika során az ilyen területek érintettségéből eredő hangzókülönbségek auditív úton is elkülöníthetők a hangképzőszerveket érintő más atípusos formáktól [1]. A stroke okozta kortikális dysarthriák általában jelentős kommunikációs hátrányt jelentenek azon személyek számára, akik megváltozott artikulációs folyamataik miatt nehezen érthető, nem könnyen feldolgozható formában tudják verbálisan megfogalmazott gondolataikat közvetíteni. A humán jelfeldolgozás számára is nehézséget jelent ezekben az esetekben az újonnan keletkező hangzók, a normától

eltérő formánsok azonosítása. Ezen betegek életvitelét nagyban segítené bármilyen, a kommunikációjukat támogató technológia.

Célkitűzés

Munkánk célja olyan műszaki eljárások vizsgálata, amelyek segítik a stroke eredetű, (elsősorban) cortikális dysarthria jeleit mutató betegek beszédkommunikációját. A beszédtechnológia ehhez kétféle megközelítést kínál. Az egyik a gépi beszédfelismerés alkalmazása [14, 15, 20], a másik pedig a dysarthriás beszédjel konverziója, érthetőségének feljavítása [5, 16]. A gépi felismeréssel leírt beszédet írott vagy hangzó formában (beszédszintézissel [19]) lehetne eljuttatni a kommunikációs partnerhez. Sajnos azonban a jelenlegi beszédfelismerők nagyon érzékenyek az átlagostól eltérő ejtésmódokra, így dysarthriás beszéd esetén pontosságuk drasztikusan leromlik [6, 14, 15, 20]. A Scientific American cikke szerint az USA-ban népszerű Siri nevű rendszer sem boldogul a dysarthriás beszélőkkel [11]. Pozitív beszámolókat inkább csak a beszédfelismerők rehabilitációs terápiában való alkalmazásáról találhatunk [3], valamint az aktuális kiejtés érthetőségének objektív kiértékelésében használják még őket sikeresen [7].

Munkánknak ebben az első fázisában egy alapvetően más célra készített beszédfelismerő rendszer dysarthriás beszédre való alkalmazhatóságát vizsgáltuk meg – magyar nyelven, hiszen betegeink is magyarok. Az egyetemünkön fejlesztett felismerő rendszer híradók feliratozására készült, így várható volt, hogy a jelentősen eltérő akusztikai és artikulációs viszonyok miatt elég rossz eredményeket fogunk kapni. Második lépésben ezért különféle beszédlemzési méréseket végeztük annak megértése céljából, hogy a dysarthriás beszéd milyen akusztikai sajátosságokban tér el a normál beszédétől. Mivel az emberi beszédkeltés rendkívül összetett folyamat, a beszédjel a stroke helyétől és kiterjedésétől függően eltérő módokon torzulhat. A motoros funkciók érintettségéből eredő beszédzavarok közül artikulációs szervek vezérlésének zavara a hangképzési folyamatot befolyásolja, például hibás formáns szerkezetű magánhangzókat eredményezhet. Ha az artikulációs szervek összehangolása sérül, akkor időben elkent hangzókat kapunk. A hangok adott ideig és hangmagasságon való kitartásának nehézsége a beszéd szupraszegmentális szintjének, a prozódiaéknak a torzulásaként jelentkezik. Végül, a hangszalagok vezérlésének zavara a hangminőség romlását okozza, ez az ún. diszfónia [8] gyakran van jelen a dysarthriával egyidejűleg [2].

Módszer

A dysarthriás hangfelvételeket az SZTE Neurológiai Klinikájának Neurorehabilitációs osztályán gyűjtöttük. A betegeket először spontán beszédre (képleírás) készítettük, a WAB-teszt magyar változatának [13] alkalmazásával. Másrészt, a Meixner-féle [9] olvasólapok segítségével az izoláltan vagy hangkapcsolatban kiejtett hangok differenciálását vizsgáltuk. Harmadrészt, egy fonológiailag kiegyenlített szöveg (A szél és a Nap című mese) felolvastatása révén olvasott hanganyagot is gyűjtöttünk.

A vizsgálatban az SZTE Informatikai Intézeténél készült, magyar nyelvű hírműsorok feliratozására optimalizált felismerő rendszer pontosságát értékeltük ki a dysarthriás felvételeken, annak is az olvasott részén, ugyanis ezt a szövegrészt éreztük beszédstílusban a híradókhoz leginkább illeszkedőnek.

A második lépésben néhány klasszikus akusztikus paraméter vizsgálatát végeztük el. A jitter, shimmer és harmonicitás-zaj viszony (HNR) standard mérőszámok a diszfónia vizsgálatában [2, 8, 18]. A jitter az alapprofrendencia, a shimmer az amplitúdó ingadozását számszerűsíti, míg a HNR a hangszalagok rezgésének szabálytalanságát méri. Az elemzést az openSMILE szoftverrel, Vicsi és társaihoz hasonlóan [18] a folyamatos, felolvasott szövegen végeztük, mivel a „kontroll” híradós felvételek is ilyen jellegűek voltak.

Az artikuláció pontosságának vizsgálatára elvégeztük a beszélők magánhangzóinak formánselemzését (a Praat szoftverrel), ezúttal a kitarított ejtésű felvételeken. Végezetül, a folyamatos beszéd tulajdonságainak elemzése céljából elvégeztük a beszédsebesség és az artikulációs sebesség becslését.

Eredmények

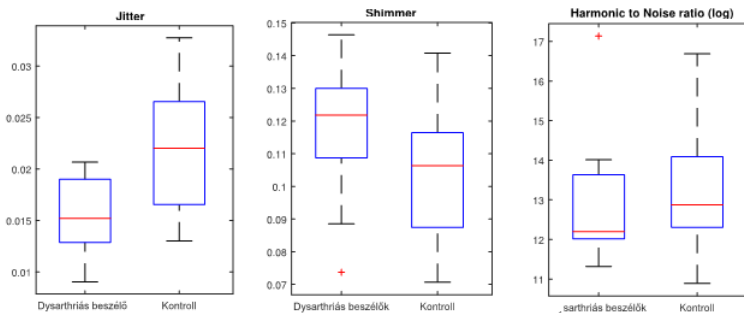
A felismerőrendszer kiértékelését leszűkítettük a legenyhébb dysarthriát mutató nyolc beteg felvételeire. Ezek minimális dysarthriás jegyeket mutattak, főleg a hangkeltés és a prozódia vonatkozásában, de artikulációjuk egészen jól érthető volt. Ennek ellenére a beszédfelismerő gyakorlatilag használhatatlan, 50% alatti szószintű pontosságot adott (hírműsorokon a pontosság 85% fölötti). Egy példa a helyes és a gép által adott átiratra:

Helyes átirat: MINDKETTŐ AZT ÁLLÍTVÁN MAGÁRÓL HOGY ERŐSEBB MINT A MÁSIK VÉGÜL MEGEGYEZTEK HOGY KIPRÓBÁLJÁK EREJÜKET EGY VÁNDORON

Felismeret kimenet: ÍGY KEDDTŐL LITVÁN MOND LE ARRÓL HOGY ERŐSEBB MINT A MÁÁS VÉGÜL MEGEGYEZTEK HOGY KIPRÓBÁLJÁK EL ŐKET EGY VÁNDOROL

A kimeneteket vizsgálva úgy láttuk, hogy nem a magánhangzók torzulása és tévesztése, inkább a szótagok elvesztése-beszúrása jelentkezt fő hibaként, amit a hangerő és a beszédsebesség ingadozása okozhat.

A felismerő rossz teljesítményének okait kutatva elsőként a felvételek jitter, shimmer, és HNR értékeit vizsgáltuk. Kontroll anyagként a híradófelvételekből vett véletlenszerű minta szolgált. Az 1. ábra mutatja a kapott jitter



1. ábra. A jitter, shimmer és harmonicitás-zaj viszonyértékek eloszlása

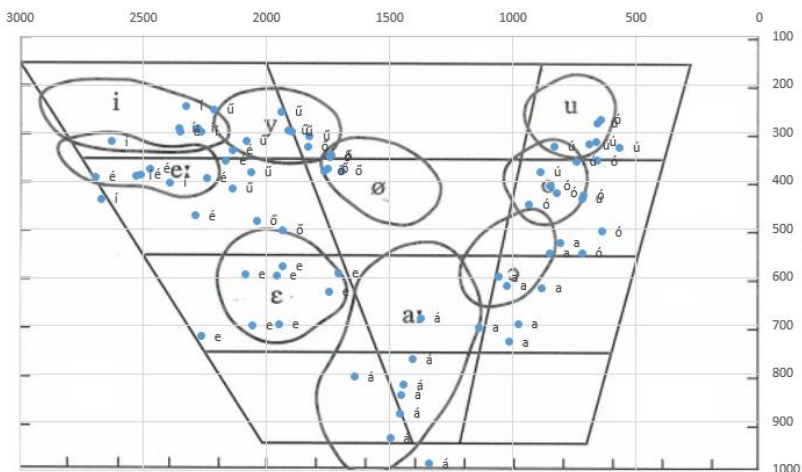
és shimmer értékeket, amelyek az irodalom szerint dysarthria esetén gyakran megnövekednek [2]. Mi a dysarthriás betegeink esetén *kisebb* jitter értékeket kaptunk, mint a kontroll felvételekre, a shimmer értékek azonban az irodalomnak megfelelően tényleg rosszabbnak adódtak. Mivel a shimmer a hangerő ingadozását méri, ez egybeesni látszik előzetes benyomásunkkal, miszerint a dysarthriás pácienseknek nehézséget okoz a hangerő egyenletes tartása. Végezetül, a HNR értékek szintén rosszabbak lettek a dysarthriás beszélőknél (ez esetben a *magas* érték jelenti a jobb beszédminőséget), az eltérés azonban elég kicsi.

Következőként formánselemzést végeztünk a betegek kitarotottan ejtett magánhangzóin. Mivel a híradófelvételek nem tartalmaznak ilyen hangokat, ezért összevetés céljából a mért F1-F2 értékeket a Bolla Kálmán közismert fonetikai atlaszából vett formánstérképre montíroztuk rá (2. ábra). A mért eredmények túlnyomórészt egybeesnek a Bolla-féle formánstérképpel, megerősítve benyomásunkat, hogy a felismerési hibákért – legalábbis ilyen enyhe fokú dysarthria esetén – nem a formánsszerkezet durva torzulása a felelős.

Végezetül összevetettük a dysarthriás és a kontroll anyag beszédsebességét, illetve artikulációs sebességét, a felismerő fonetikai kimenete alapján. A kapott sebességértékeket az 1. tábla összegzi (beszédhang/sec, a zárhangok zár- és zörejrészét külön hangnak tekintve a felismerő technikai sajátosságai miatt). Látható, hogy a híradók beszéd- és artikulációs tempója közel kétszerese a betegeinkre kapott értékeknek. Ez a hatalmas eltérés magyarázhatja a felismerőrendszer nagyszámú ún. beszúrási hibáját.

1. sz. táblázat – Artikulációs és beszédtempó

	Artikulációs tempó	Beszédtempó
Dysarthriás betegek	9,25	8,04
Híradó	16,10	15,69



2. ábra. Kitarottnan ejtett magánhangzók F1-F2 formánstérképe

Következtetések

Legfontosabb következtetésként megállapítottuk, hogy dysarthriás betegek esetén a beszédfelismerő rendszerek csakis a beszélő hangjához való adaptálással érhetnek el általánosan használható hatásfokot. A beszélőadaptáció ma már standard technika a beszédfelismerésben, akár dysarthriás betegek esetére is [6, 12, 14]. Ehhez azonban betegenként hosszabb hangmintákra van szükség, mint ami jelen esetben rendelkezésünkre állt.

A gépi beszédfelismerés kimenetét beszéddé alakíthatjuk beszédszintézissel. Ha rendelkezésünkre állnának stroke előtti hangminták a betegtől, akkor a szintetizátor a beteg saját hangján szólalhatna meg [19], ellenkező esetben

egy „donor” hangra lesz szükség a szintézishez. Egy alternatív technológiai lehetőség, ha a felismerés-szintézis lépések kihagyásával a beteg hangját közvetlenül próbáljuk „feljavítani”, jobb minőségűvé konvertálni [5, 16], ún. hangkonverziós technológiával [10, 17]. Sajnos ebben az esetben is szükség lehet egy „donor” hangmintájára, amelyre átkonvertáljuk a betegünk hangját, illetve ez esetben is hosszabb (1-2 órányi) hangminta szükséges egy-egy betegtől, hiszen itt is egy személyre szabott rendszert kell készítenünk. A jövőben ezen megoldások alkalmazhatóságát szeretnénk vizsgálni.

Köszönetnyilvánítás

A kutatást az EFOP-3.6.1-16-2016-00008 azonosítójú, EU társfinanszírozású projekt támogatja. Tóth Lászlót az Emberi Erőforrások Minisztériuma UNKP-18-4 kódszámú Új Nemzeti Kiválóság Programja támogatta.

Hivatkozások

- [1] Aronson, A. E. (1981) Motor Speech Signs of Neurologic Disease In: Darby, J. K. ed. *Speech Evaluation in Medicine*. Grune and Stratton. New York, 159-180.
- [2] Camillo, L., Ortiz, K.Z. (2007) Vocal Analysis (auditory-perceptual and acoustic) in dysarthrias. *Pro-Fono Revista de Atualizacao Cientifica*, 19(4), 381-6.
- [3] Fager, K.S. (2017) Speech Recognition as a Practice Tool for Dysarthria, *Semin Speech Lang* 38(3), 220-228.
- [4] Horváth Szabolcs , Hirschberg Jenő (2013) Diszartria/diszartrofónia (Dysarthria/dysarthrophonia) In: Hirschberg J. , Hacki T. , Mészáros K. szerk. *Foniátria és társtudományok II. Eötvös Kiadó*. 80-86.
- [5] Kain, A.B., et al., (2007) Improving the intelligibility of dysarthric speech. *Speech Communication* 49, 743-759.
- [6] Kim, M.J., Yoo, J., Kim, H. (2013) Dysarthric Speech Recognition using Dysarthria-Severity-Dependent and Speaker-Adaptive Models, *Interspeech* 2013, pp. 3622-3626.
- [7] Kitzing, P., Mayer, A, Ahlander, VL. (2009) Automatic speech recognition and its use as a tool for assessment or therapy of voice, speech and langugae disorders. *Logopedics Phoniatrics Vocology* 34(2), 91-96.
- [8] Markó, A., Grácsi, T.E., Bajnócziné Szucsák K. (2012) A diszfónia terápiájának hatékonysága a beteg beszédtechnikai képzettségének függvényében, *Alkalmazott nyelvtudomány*, 12(1-2), 83-103.
- [9] Meixner Ildikó (1995) A dyslexia prevenció, redukáció módszere. *Ranschburg Pál Kollégium, BGGYTF Budapest*
- [10] Mohammadi, S.H., Kain, A. (2017) An overview of voice conversion systems. *Speech Communication* 88, 65-82.
- [11] Mullin, E. (2016) Why Siri won't listen to millions of people with disabilities. <https://www.scientificamerican.com/article/why-siri-won-t-listen-to-millions-of-people-with-disabilities>
- [12] Mustafa et al. (2014) Severity-Based Adaptation with Limited Data for ASR to Aid Dysarthric Speakers. *PLoS ONE* 9(5): e97665
- [13] Osman - Sági J. (1991) Az afázia klasszifikációja és diagnosztikája. *Ideggyógyászati Szemle* 8. 339-351, 351-361.
- [14] Raghavendra, P. (2001) An investigation of different degrees of dysarthric speech as

- input to speaker-adaptive and speaker-dependent recognition systems. *Augmentative and Alternative Communication* 17(4),
- [15] Rosen, K., Yampolski, S. (2009) Automatic speech recognition and a review of its functioning with dysarthric speech. *Augmentative and Alternative Communication* 16(1), 48-60.
- [16] Rudzicz, F. (2012) Adjusting dysarthric speech signals to be more intelligible, *Computer Speech and Language*, 27, 1163-1177
- [17] Sun, L., Kang, S., Li, K., Meng, H. (2015) Voice conversion using deep Bidirectional Long Short-Term Memory based Recurrent Neural Networks. *Proceedings of ICASSP 2015*, 4869-4873.
- [18] Vicsi, K., Imre, V., Mészáros, K. (2011) Voice Disorder Detection on the Basis of Continuous Speech, *European Conference of the International Federation for Medical and Biological Engineering*, 86-91
- [19] Yamagishi, J., Veaux, C., King, S., Renals, S. (2012) Speech Synthesis technologies for individuals with vocal disabilities: Voice banking and reconstruction, *Acoust. Sci & Tech*, 33(1), 1-5.
- [20] Young, V., Mihailidis, A. (2010) Difficulties in Automatic Speech Recognition of Dysarthric Speakers and Implications for Speech-Based Applications Used by the Elderly: A Literature Review, *Assitive Technology*, 22(2), 99-112.