

Miért fontos és mire jó a beszédtechnológia? – magyar sikerek első kézből

NÉMETH GÉZA

BME Távközlési és Médiainformatikai Tanszék

nemeth@tmit.bme.hu

Kulcsszavak: PDA, magyar nyelvű gépi szövegfelolvasás, TTS, gépi beszéd felismerés, ASR, beszédtechnológia, alkalmazások

A cikkben bemutatjuk, hogy a beszédtechnológia nélkülözhetetlen összetevője a jelen és a közeljövő egyik nagy üzleti lehetőségének; a gépi személyi asszisztensek területének. Ismertetjük, hogy milyen technológiai területeken van szükség jelentős hazai kutatásokra ahhoz, hogy belátható időn belül ilyen megoldások magyar nyelven is működjenek. Egyben áttekintjük a magyar nyelven létező korszerű rendszereket és javaslatot teszünk olyan alkalmazásokra, melyek a ma létező magyar beszédtechnológiai megoldásokra alapozva létrehozhatók.

1. Bevezetés

A beszédtechnológia a természetes beszédlánc bármely elemének gépi megvalósításával foglalkozó tudományág. Természeténél fogva interdiszciplináris terület (nyelvészet, fonetika, akusztika, jelfeldolgozás, gépi tanulás, kognitív tudományok stb.). Az 1. ábrán láthatjuk a szerző munkahelyének ehhez kapcsolódó kompetencia területeit.

Az Ethnologue, a világ egyik nagy és elismert nyelvstatisztikai portálja szerint a világon ma 7102 élő nyelv található. A magyar nyelv a Wikipédia adatai szerint a 74. legtöbb anyanyelvi beszélővel rendelkezik ezek közül. Az infokommunikációs technológia fejlődésével, az eszközök méretének csökkenésével különösen fontosá vált az, hogy ne csak képernyő, egér és billentyűzet segítségével tudjunk információs rendszereket kezelni. A legkézenfekvőbb a természetes kommunikáció alap eszköze a beszéd.

Jelen cikkben a Microsoft egyik vezető kutatójának friss előadása (Sarikaya 2015) alapján tekintjük át, hogy a

legnagyobb szoftvercégek (Microsoft, Apple, Google, Amazon stb.) miért költenek hatalmas összegeket beszéd- (is) kommunikáló, ún. személyi asszisztens szoftverek (Cortana, Siri, Google Now, Echo...) kifejlesztésére és elterjesztésére. Röviden megvizsgáljuk, hogy ezek létrehozásához milyen nyelv- és beszédtechnológiai elemekre van szükség. Ezek összessége egyelőre csak a legnagyobb világnyelveken érhető el (ott és korlátozott tématerületeken). Azonban sok, gyakorlatban is felhasználható megoldás magyar nyelven is elkészült. Ezek ismertetésére és néhány gyakorlati alkalmazási példa bemutatására is sor kerül. A cikket az előttünk álló kihívások összegzése zárja.

2. Helyzetkép

A 2010-es években jelentek meg az első személyi digitális asszisztensek (Personal Digital Assistant) a legnagyobb szoftvercégek ajánlataként (pl. Apple Siri). Milyen jellemzőkkel bír egy ilyen asszisztens (Sarikaya 2015 alapján)?



1. ábra
Beszédtechnológiához kapcsolódó kompetencia területek

- Természetes nyelvi utasításokkal (szöveggel vagy beszéddel) vezérelhető.
- Több (jellemzően külön-külön korábban már létező) szolgáltatás integrált vezérlőjeként/kezelőjeként jelenik meg.
- Strukturált információt szolgáltat multimodális (szöveg, grafika, videó, beszéd...) felhasználói felületen.

Számos komponens-technológia (gépi tanulás, felhasználóra történő automatikus adaptáció, gépi beszédfelismerés és szintézis, természetes nyelv gépi megértése, dialógus-vezérlés, információ-kivonatolás, fontossági rangsorolás stb.) célszerű és magas szintű integrálása szükséges egy ilyen rendszer megvalósításához.

Tekintsünk egy potenciális felhasználói példát. Üzleti ügyben Budapestről Párizsba készülünk utazni. Már fél évvel korábban egy webes felületen lefoglaltuk a repülőjegyet és elektronikus levélben megkaptuk a visszajelzést róla. A személyi asszisztens

- az elektronikus levél elemzése alapján
- meghatározta az utazási szándékot,
- és azt rögzítette a naptárban,
- majd az utazás napján lekérdezi a felhasználó aktuális pozícióját,
- ellenőrzi a járatinformációkat,
- kiszámítja az optimális útvonalat és
- figyelmeztet arra (írásban és szóban), hogy mikor kell elindulni ahhoz, hogy időben kiérjünk a repülőtérré.

Ezek az elemi funkciók külön-külön vagy kisebb csoportban már ma is léteznek, de integrált működésük a hasznosságot és a felhasználói élményt (UX, user experience) ugrásszerűen javíthatja. Így együtt elsőre talán futurisztikusnak tűnnek és talán ijesztő, „nagy testvér” (Big Brother) hatást is kiválthatnak, különösen hazánkban és a közép-kelet-európai régióban. Mégis mi lehet az oka annak, hogy nagy világcégek sora hatalmas összegeket fektet ilyen jellegű fejlesztésekbe? Sarikaya az alábbi fő indokokat sorolja fel.

A jelenlegi piaci tendenciákat figyelembe véve 2017-re gyakorlatilag minden 12 évnél idősebb embernek lesz telefonja (döntően okostelefonja), tehát nem lesz jelentős növekedési lehetőség az okostelefonon keresztül az internethez kapcsolódó üzleti területeken a készülékeladások tekintetében. Viszont a tőzsde növekedést vár a cégektől. A kérdés az, hogy mi lesz a következő „húzó” üzleti lehetőség.

Kézenfekvő a szoftveráruházakra gondolni, hiszen évről-évre egyre több alkalmazást fejlesztenek az okostelefonokra. Az Apple és a Google alkalmazás boltjaiban egyaránt több mint másfél millió megoldás található. A letöltések száma is folyamatosan nő. Megjegyzendő, hogy 2003-ban az egyik hazai mobilszolgáltató meghatározó marketingese azzal utasította el egy okostelefonos szoftverbolt fejlesztését, hogy „Ki az a hülye, aki szoftvert tölt a telefonjára?”. Viszont az eszközönként egy év alatt letöltött és üzembe helyezett alkalmazások száma (kb. 33), az évente legalább egyszer elindítottak (kb. 25) és az egy hónap alatt átlagosan használtak (kb.

12) számértéke nagyjából állandó. Az alkalmazások mintegy 80%-át tehát „döglöttnek” (zombi) tekinthetjük. A fejlődés egyik szűk keresztmetszete tehát a számunkra érdekes alkalmazások megtalálásának korlátja. A másik szűk keresztmetszet az alkalmazások megtanulására fordítható idő (kognitív sávzsélesség).

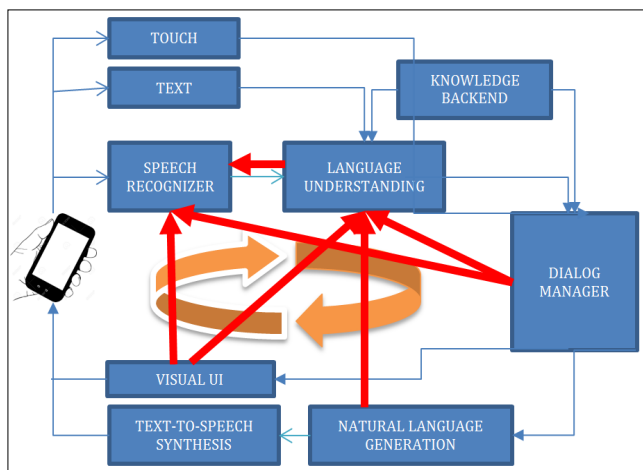
A harmadik szűk keresztmetszet pedig az alkalmazások használatára fordítható idő. Friss USA-beli statisztikák (2014) szerint a mintegy 1000 perces napi ébrenlétből a TV nézésnél (168 perc) több időt töltenek a felhasználók az okostelefonjukkal (180 perc). A dolgok internete (Internet of Things, IoT) keretében számos újabb eszköz kapcsolódik az internetre, de nem világos, hogyan léphetünk kapcsolatba velük és mennyi időbe kerül ez nekünk. Egyre több az „okostelefonozással” átlapolódó tevékenység. A munka és a személyes használat közötti határ is elmosódik. Fontos az időbeosztás és gazdálkodás optimalizálása. A negyedik szűk keresztmetszet pedig az okostelefonok érintőképernyőinek nyomogatásával és az azon megjelenő billentyűzeten történő adatbevitellel és a kapcsolódó kijelzéssel történő korlátozott sebességű információcsere.

A fentiek figyelembe vételével már meg is indult az a folyamat, melynek során egyre jobban integrálják a beszédtechnológiai elemeket az Apple, a Microsoft és a Google operációs rendszereibe. A mély neurális hálózatok (Deep Neural Networks, DNN) beszédtechnológiai alkalmazása jelentősen csökkentette a hibarányt. Ezért egyre nagyobb arányban használják például szövegbevitelre és keresésre a beszédet. A nagyobb arányú használat pedig a minőséget is javítja, hiszen a felhő alapú megoldások lehetővé teszik, hogy a felhasználó által kijavított felismerési hibát visszacsatoljuk a rendszerbe.

Esély van tehát arra, hogy olyan rendszereket (lásd a 2. ábrát; a vastag nyilak az elemek között a hagyományos architektúrán túllépő, korszerű megoldást jelzik) hozzunk létre, melyek

- megértik a feladatot, szét tudják osztani azt a hagyományos alkalmazások között és meg is tudják keresni az adott feladathoz jól illeszkedő(ke)t,

2. ábra
Személyi asszisztent támogató dialógus rendszer egy lehetséges blokkdiagramja (forrás: Sarikaya, 2015)



- lehetővé teszik a rutinfeladatok delegálását, automatizálását (számlafizetés, eseményekre figyelmeztetés...),
- mindig kéznél levő asszisztensként „bármikor bármire” választ tudnak adni (taxi rendelés, útvonal tervezés, üzenetek fogadása és küldése, fájl keresés...).

Tehát életünk hatékonyabb menedzselését és termelékenységünk növelését teszik lehetővé. Ez potenciálisan dollárban mérve 100 milliárdos nagyságrendű üzlet. Aki megnyeri ezt a területet, az az alkalmazások felett is meghatározó ellenőrzést szerez. Ezért fektetnek be a nagy cégek a személyi asszisztensekbe és a kapcsolódó területekbe hatalmas összegeket.

2.1. Mi a helyzet Magyarországon?

A szűken vett hazai beszéd- (Németh, Olasz, 2010) és nyelvtéchnológiai (Prószéky, 2015) területről viszonylag jó áttekintést kaphatunk összefoglaló kötetekből. A nyelvtéchnológia magyarországi gyakorlati alkalmazási területei elsősorban számítógépes és a számítógéppel segített fordítás, a keresés támogatása és a kivonatolás. A személyi asszisztensek fejlesztéséhez viszont elengedhetetlen a kommunikációs kontextust megérteni képes dialógus-vezérlő (dialog manager), a nyelvi tartalmat értelmezni képes nyelvmegértő (language understanding) és a természetes nyelvű szöveget generáló (language generation) alrendszerek kutatás-fejlesztése. A tudástár (knowledge backend) tématerület specifikus kialakítása és a többi komponenshez csatolt metaadatok létrehozása is szükséges egy hatékony megoldáshoz.

A beszédtechnológia nemzetközi területén jellemző a közös adatbázisokon alapuló rendszerfejlesztés eredményeinek kompetitív összehasonlítása. Az 1. táblázatban láthatjuk, hogy az egyre összetettebb feladatokat (elsősorban angol, majd kínai és hindi nyelvekre) az emberi szövegfelolvasási teljesítményhez képest milyen szinten tudták megoldani a gépi megoldások fejlesztői. Figye-

lemreméltó, hogy a legjobb ember sem éri el az 5-öst és a legjobb gépi rendszer sem haladja meg a 4,2-es értéket.

Elmondhatjuk, hogy a TTS rendszerek területén a magyar nyelvű megoldások minősége követi a nemzetközi trendeket. A magyar látássérült emberek már 15 éve használják a Jaws for Windows szoftver magyar változatában a ProfiVox rendszer diádós elemösszfűzésen alapuló változatát (Olasz és tsai, 2000). Ez a rendszer szól a Stephen Hawking életét bemutató „A mindenség elmélete” című film magyar változatában is.

A BME TMIT kutatóinak legújabb fejlesztése, a korpusz-alapú technológiát alkalmazó többnyelvű vasúti hangos utastájékoztató rendszer (Németh és tsai, 2015 és Zainkó és tsai, 2015) magyar nyelvű változatának 50 anyanyelvű beszélővel történő értékelésekor a legrosszabbnak ítélt mondat átlagpontszáma 4,28, a legjobb pedig 4,76 volt. A rendszer tehát az adott szűk tématerületen képes az emberi bemondással összemérhető hangot előállítani. Az angol változat 54 fő nem anyanyelvű, de angol nyelvtudással rendelkező személyek által minősített legjobb értéke 4,1, az átlag pedig 3,6 volt. Tekintve, hogy sem az adatbázist felolvasó személy, sem az értékelők nem voltak anyanyelvűek, ez a megoldás is versenyképesnek tekinthető. A rendszer ma már több mint 60 magyarországi vasútállomáson és megállóhelyen (pl. Bp-Keleti, Kelenföld, Debrecen, Székesfehérvár, Szeged, Keszthely, Celldömölk) meghallgatható. Kézenfekvően kiterjeszhető lehet más járművekre (pl. BKV, VOLÁN) és fedélzeti információs helyzetekre is.

A rejtett Markov-modellen (Hidden Markov Model, HMM) alapuló legújabb általános célú felolvasó rendszerünk (Nagy és tsai, 2015) pedig 43 tesztelő értékelése szerint 3,9-4,3 közötti eredményeket ért el.

A nagy szótáras, magyar nyelvű gép beszédfelismerés területén is jelentős eredmények születtek ebben az évtizedben. Kísérleti üzemben van az MTVA-nál egy valószínű idejű adásfeliratozó rendszer időjárásjelentés és híradó témakörre optimalizálva (Varga és tsai, 2015).

Több magyarországi bank és biztosítótársaság is használja ügyfélszolgálati adatbányászatra és minőségellenőrzésre az SPSS szoftverrel kombinálva a BME TMIT munkatársai által kifejlesztett automatikus ügyfélirányítási és leiratozási technológiát (Sárosi és tsai, 2012). Ígéretes kísérletek folynak a diktálási területeken is (pl. bíróságok, orvosi leletezés, fordítás támogatás).

Bizonyos szűk területeken a multinacionális vállalatok is létrehoztak magyar nyelvű megoldásokat. Például a Waze és a Google Térkép navigációs rendszereiben jó minőségben lehet magyarul bediktálni a keresett címet (akár Németországban is). A Google Translate pedig képes magyar nyelven is felolvasni a szöveget. Az Apple

1. táblázat

A gépi szövegfelolvasás (text-to-speech, TTS) fejlődése (1-legrosszabb, 5-legjobb, forrás: <http://festvox.org/blizzard>)

Év	Legjobb ember	Legjobb TTS	Legrosszabb TTS	Megjegyzés
2005	4,76	3,19	1,98	
2006	4,66	3,74	1,34	nagyobb adatbázis (5000 mondat)
2007	4,7	3,9	1,3	nagyobb adatbázis (8 óra)
2008	4,8	4,1	2,0	UK English (15 óra) + Mandarin (6.5 óra)
2009	4,9	4,2	1,9	
2010	4,8	4,2	1,6	zaj, kisebb adatbázisok
2013	4,8	3,9	1,2	300 órányi angol hangoskönyv címkézés nélkül

és a Samsung okostelefonokban elérhető a Nuance által forgalmazott magyar nyelvű TTS technológia. Azonban jelzi a korlátokat, hogy például ez a rendszer elfogadható kiejtéssel csak kijelentő mondatokat tud felolvasni, a kérdő mondatok helyes intonációjával már nem tud megbirkózni.

Általánosságban elmondható, hogy a nagyvállalatok a magyar EU-tagságra tekintettel előbb-utóbb létrehozák technológiáik magyar nyelvű változatát. Azonban ezek sokszor az „éppen megfelelő” minőségi cél szerint készülnek, ezért használhatóságuk gyakran korlátozott.

3. Kutatási kihívások

Rövidtávon (5 éven belül) még a legnagyobb nyelvekre sem várható egy tetszőleges természetes nyelvű kérdést többnyire helyesen megválaszoló gépi asszisztens létrehozása. Ebben az időtávban (vagy akár középtávon, azaz 10 éven belül) az sem várható, hogy a nagyvállalatok sokoldalú magyar nyelvű megoldást hozzanak létre. Ezért magyar finanszírozással magyar kutatóknak kellene jól definiált, fontos tématerületeken szükséges rendszerelemeket és egy működő platformot létrehozni. Például ilyen terület lehet az üzenetkezelés, amiben korábban ipari szolgáltatásként működő megoldásokkal szereztünk tapasztalatokat (Németh és tsai, 2000 és 2007). Hasonlóan fontos lehet a nagy költségű kórházi kezelést és rehabilitációt igénylő egészségügyi alkalmazások (rák, stroke diagnózis és rehabilitáció, csontvelő transzplantáltak támogatása stb., pl. Csala és tsai, 2012) köre is.

A spontán beszédkommunikáció kezelése még a nagy nyelveken sem megoldott. Ehhez feltétlenül szükséges reprezentatív méretű, annotált adatbázisok létrehozása a megcélzott alkalmazási területeken. Ezek segítségével lehetne kiindulásként használható működő, felhő alapú demonstrációs alkalmazásokat létrehozni. Ezek elterjesztése után sok valós, érvényes adat gyűjthető be és a felhasználók egyben tömeges annotálási/javítási (crowdsourcing) feladatokat is elvégezhetnek. Sokat ígérő megközelítés lehet a statisztikai módszerekkel lefedhető gyakori feladatok és a ritkán előforduló, de fontos események szabály-alapú kezelésének kombinációja is.

4. Alkalmazási kihívások

Az alábbiakban néhány olyan alkalmazási területet sorolunk fel, melyekben a ma rendelkezésre álló beszédtechnológiai elemek felhasználásával is jelentős új alkalmazások hozhatók létre.

4.1. Információs akadálymentesítés

A 15-69 éves magyar lakosok 70%-a internet felhasználó (nrc.hu, 2015). Tőlük elvárható a webes információs szolgáltatások alkalmazása (pl. magyarorszag.hu). Viszont mi legyen a többi 30%-nyi magyar állampolgárral. Számukra telefonos hangkapcsolaton keresztül jelentős arányban megoldható a szolgáltatások automatizálása

(pl. 112, MÁV, BKV, Volán tudakozó) DTMF menükben történő bolyongás és emberi kezelők túlzott költségű bevonása nélkül. Ebben a körbe tartozik a fogyatékos emberek (látás- és beszédsérültek) támogatása is (pl. Tóth és Németh, 2006).

4.2. Beszédtechnológia az oktatásban

A beszédtechnológia kiválóan alkalmas tanulást segítő interaktív „játékok” létrehozására óvodásoknak és iskolásoknak. Segítségével interaktív multimodális oktatási anyag hozható létre akár napi hírek kivonatolásával. Ez kiválóan alkalmas lehet kisebbségi helyzetű magyar gyermekek motiválására és afáziás, autista, ill. más hasonló nehézségekkel küzdő emberek támogatására.

4.3. Beszédtechnológia az egészség iparban

A beszédtechnológia segítségével megoldható a műtétek bizonyos fokú automatizálása (utasítások, jegyzetelés), ami egyben egységes, elektronikusan elérhető adatbázisba is kerülhet. Szintén költséghatékony megoldás lehet a leletek diktálása, a beszédhang alapján gégeproblémák, depresszió stb. korai diagnózisa és rehabilitációja. Táv(fel)ügyeleti alkalmazások (pl. gyógyszerbevitelre figyelmeztetés, ablak zárása stb.), demencia, Alzheimer és más hasonló betegségekben szenvedők részben automatizálható felügyelete.

4.4. Beszédtechnológia a tartalom- és a kreatív iparban

Interdiszciplináris integráció segítségével a beszédtechnológia közreműködhet az egészségügy és a szociális rendszer más szereplői közötti feladatmegosztásban. A digitális közoktatás, intelligens otthon, intelligens város programokban újszerű szolgáltatásokat tesz lehetővé. Multi-modális tartalomelemzés, közvéleménykutatás hatékonyan támogathatja az üzleti szféra és a kormányzat kommunikációs céljait. Bankok, biztosítók és kiskereskedelmi láncok információs szolgáltatásai is hatékonyabbá tehetők a beszédtechnológia segítségével (pl. áruk vagy szolgáltatások vonalkódja vagy QR-kódja alapján okostelefonos információs szolgáltatások).

A gépjárművekben alapvető probléma, hogy az információs és szórakoztató szolgáltatások (car infotainment) kezelése és keresése elvonja a figyelmet a vezetéstől. A beszédfelismerés és a gépi szövegfelolvasás hatékonyabbá teheti ezeket a feladatokat és csökkentheti a balesetveszélyt. A beszédvezérelt otthon (okostelefon, okosTV, okos mosógép...) kritikus jelentőségű lehet beteg emberek számára, de jól megtervezve mindenki számára hasznos kényelmi szolgáltatás lehet.

4.5. Beszédtechnológia a gyártásban

A raktár-logisztika automatizálása területén egy felügyeleti rendszer automatikusan utasíthatja a targoncák kezelőit, hogy mit hova vigyenek, így folyamatosan a környezetükre figyelhetnek, ami csökkenti a hibázás lehetőségét és a balesetveszélyt egyaránt. Gyártás közbeni információ, figyelmeztetés automatizáltan megoldható. Beszélő gépkönyvek gyorsíthatják a szervizelési feladatok megoldását.

5. Összefoglalás

A cikkben bemutatottuk, hogy a beszédtechnológia nélkülözhetetlen összetevője a jelen és a közeljövő egyik nagy üzleti lehetőségének a gépi személyi asszisztensek területének. Ismertettük, hogy milyen technológiai területeken van szükség jelentős hazai kutatásokra ahhoz, hogy belátható időn belül ilyen megoldások magyar nyelven is működjének. Egyben áttekintettük a magyar nyelven létező korszerű rendszereket és javaslatot tettünk olyan alkalmazásokra, melyek a ma létező magyar beszédtechnológiai megoldásokra alapozva létrehozhatók. Készen állunk az érdeklődő partnerekkel az együttműködésre.

Irodalom

- [1] E. Csala, G. Németh, Cs. Zainkó, „Application of the NAO humanoid robot in the treatment of marrow-transplanted children”, 3rd IEEE Int. Conf. on Cognitive Infocommunications, Kosice, Slovakia 2012, pp.655–658.
- [2] Ethnologue, <https://www.ethnologue.com/world> (hozzáférés, 2015. november 14)
- [3] P. Nagy, Cs. Zainkó, G. Németh, „Synthesis of Speaking Styles with Corpus- and HMM-Based Approaches”, Proc. of 6th IEEE Conf. on Cognitive Infocommunications Győr, Hungary 19-21/10/2015., pp.195–200.
- [4] Németh G., Zainkó C., Fekete L., Olasz G., Endrédi G., Olaszi P., Kiss G., Kis P., „The design, implementation and operation of a Hungarian e-mail reader”, International Journal of Speech Technology 3:(3-4), pp.217–236., 2000.
- [5] Németh G., Kiss G., Tóth B., „Cross Platform Solution of Communication and Voice/ Graphical User Interface for Mobile Devices in Vehicles” In: Abut H., Hansen J.H.L., Takeda K. (szerk.), Advances for In-Vehicle and Mobile Systems: Challenges for International Standards, Springer, New York 2007., pp.237–250.
- [6] Németh G., Olasz G. (szerk.), „A magyar beszéd: Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek”, Akadémiai Kiadó, Budapest 2010., p.708.
- [7] Németh G., Zainkó Cs., Bartalis M., Olasz G., „Többnyelvű vasúti hangos utastájékoztató korpusz alapú TTS módszerrel”, Beszédkutatás 23 (ISSN 1218-8727), pp.233–241, 2015.
- [8] Olasz G., Németh G., Olaszi P., Kiss G., Zainkó Cs., Gordos G., „Profivox – a Hungarian TTS System for Telecommunications Applications”, International Journal of Speech Technology 3:(3-4), pp.201–215., 2000.
- [9] Prószyk Gábor, „Számítógépes nyelvészet”, Átdolgozott, elektr. kiadás, 2015. Morphologic Kft. http://www.morphologic.hu/szamitogepes_nyelveszet_nka (hozzáférés: 2015. november 14.)
- [10] Ruhi Sarikaya, The Technology Powering Personal Digital Assistants, Interspeech 2015 Keynote, Dresden, Germany 2015.
- [11] G. Sárosi, T. Fegyó, P. Mihajlik, B. Tarján, J. Pancza, Z. Hans, „LVCSR-based Speech Analytics of a Hungarian Language Call-Center”, In: IAST 2012 – Workshop on Innovation and Applications in Speech Technology, Dublin, Ireland 09-10/03/2012.
- [12] Tóth B., Németh G., „VoxAid 2006: Telephone Communication for Hearing and/or Vocally Impaired People”, In: Miesenberger et al. (szerk.) ICCHP 2006, Computers Helping People with Special Needs. Springer, Berlin 2006., pp.651–658.
- [13] Á. Varga, B. Tarján, Z. Tobler, G. Szaszák, T. Fegyó, C. Bordás, P. Mihajlik, „Automatic Close Captioning for Live Hungarian Television Broadcast Speech: A Fast and Resource-Efficient Approach”, Proc. of Speech and Computer, SPECOM 2015, Athens, Greece 2015., pp.105–112.
- [14] Wikipedia, https://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers (hozzáférés: 2015. november 14.)
- [15] Cs. Zainkó, M. Bartalis, G. Németh, G. Olasz, „A Polyglot Domain Optimised Text-To-Speech System for Railway Station Announcements”, In: Interspeech 2015 – Proc. of the 16th Annual Conf. of the Int. Speech Communication Association, Dresden, Germany 2015., pp.1236–1240.

A szerzőről



NÉMETH GÉZA a BME-n 1983-ban szerzett villamosmérnöki oklevelet műsorközlő ágazaton. A BEAG ösztöndíjasaként 1983-85 között szakmérnöki tanulmányokat folytatott. 1985-87 között a BEAG fejlesztőmérnöke volt. 1987 óta a BME TMIT (ill. jogelődjei) munkatársa. 1997-ben PhD fokozatot szerzett, 2013-ban habilitált egyetemi docensként. Kutatási területe a beszédinformációs rendszereknek, a multimodális és a mobil ember-gép interfészeknek az alapkutatótól, a technológiafejlesztésen át az információs társadalom alkalmazásaiig terjedő köre.