

Miért fontos és mire jó a beszédtechnológia?

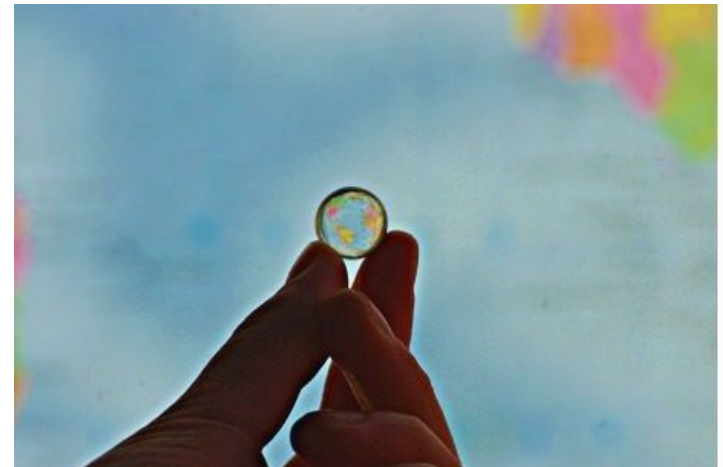
magyar sikerek első kézből

BME Távközlési és Médiainformatikai Tanszék Beszédkommunikáció és Intelligens Interakciók Laboratórium

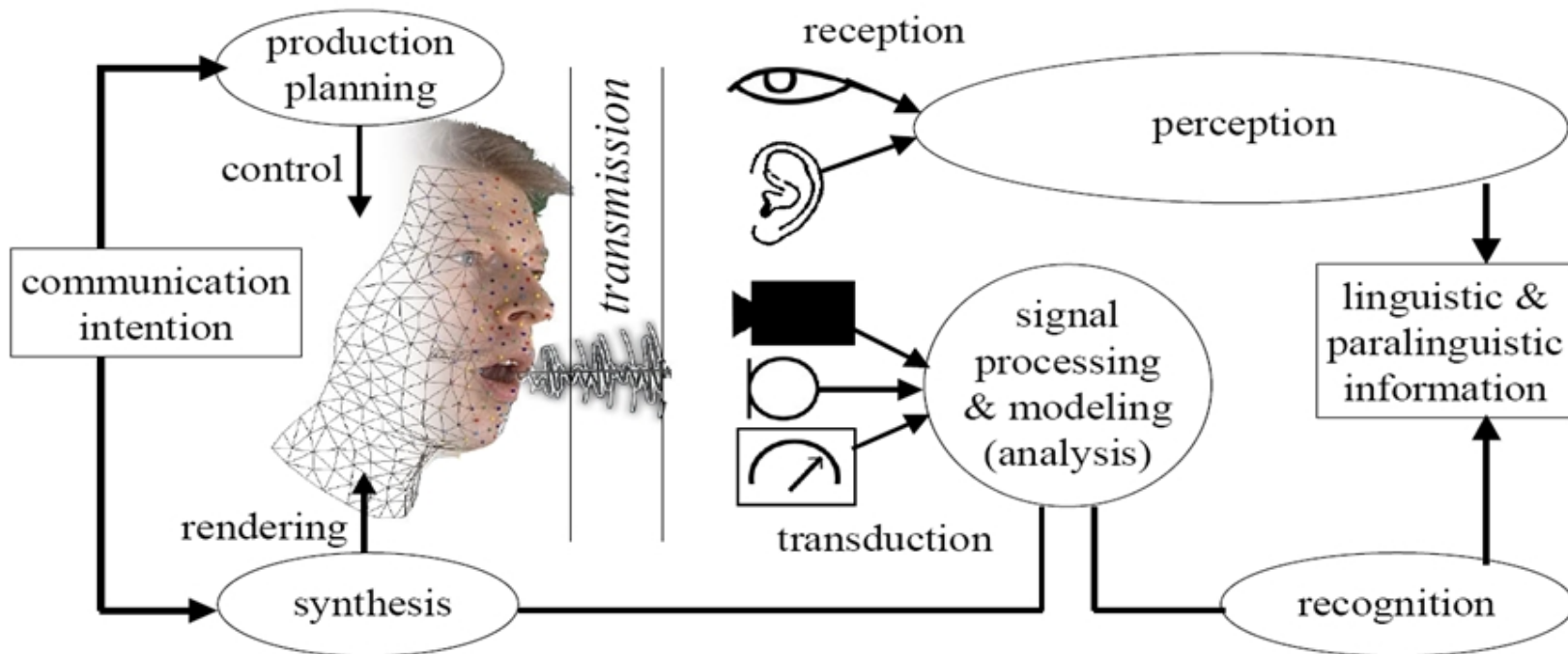


Áttekintés

- Mi is a beszédtechnológia?
- Miért fontos a magyaroknak?
- Történeti áttekintés
- Eredmények
- Erőforrások
- Kutatási kihívások
- Alkalmazási kihívások



Mi is a beszédtechnológia?



**A természetes beszédlánc bármely elemének gépi
megvalósítása
(interdiszciplináris tudomány)**



Miért fontos a magyaroknak?

- Különleges a nyelvünk
(ragozó, szabad szórend)
- Extra befektetés – Közepes piac
(73. a világon [Ethnologue])
- Érdekli a multikat
(Google, Nuance, ...)

de

- Testreszabott, kiváló megoldások
drágák <> „just sufficient effort”



Történelem

közlekedés és beszédtechnológia



• 1791



• 2015



Hawking gépi hangja angolul és magyarul

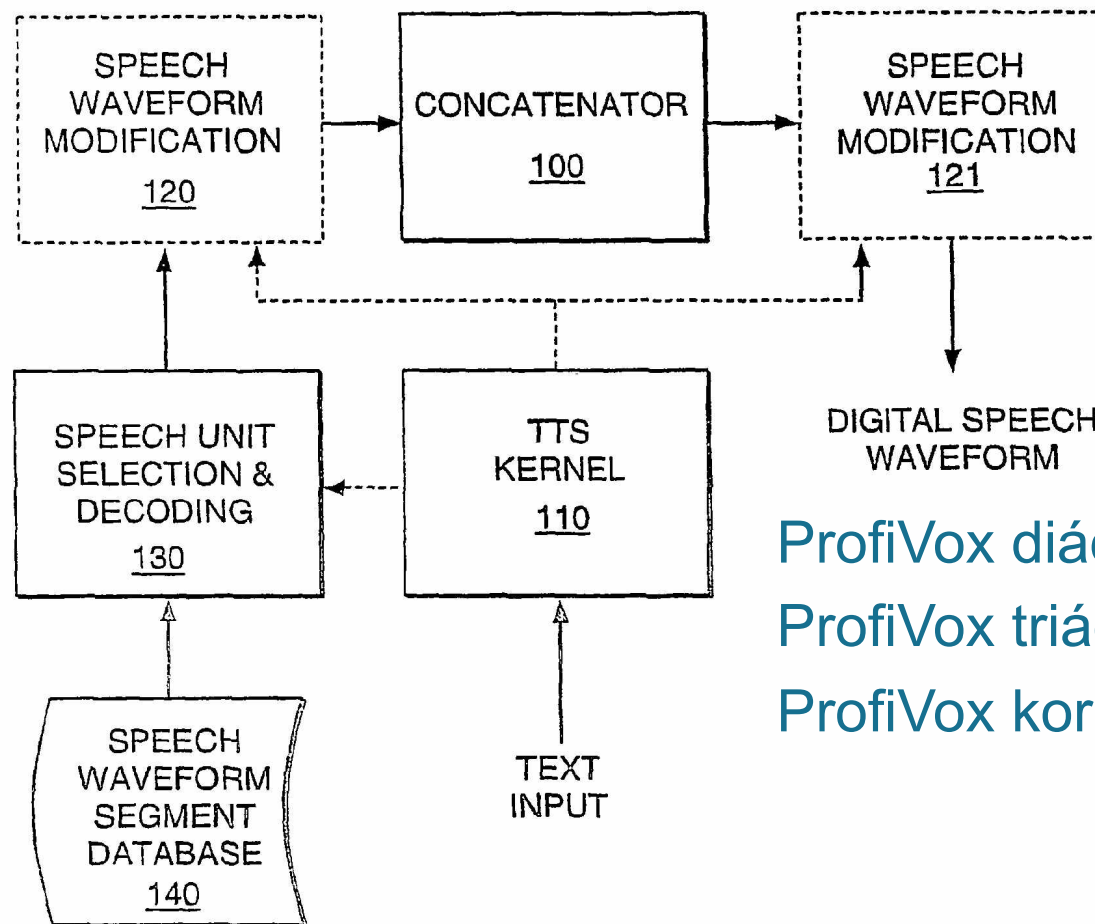
Dectalk 1982



ProfiVox 2000 - 2014



Hullámforma összefűzés

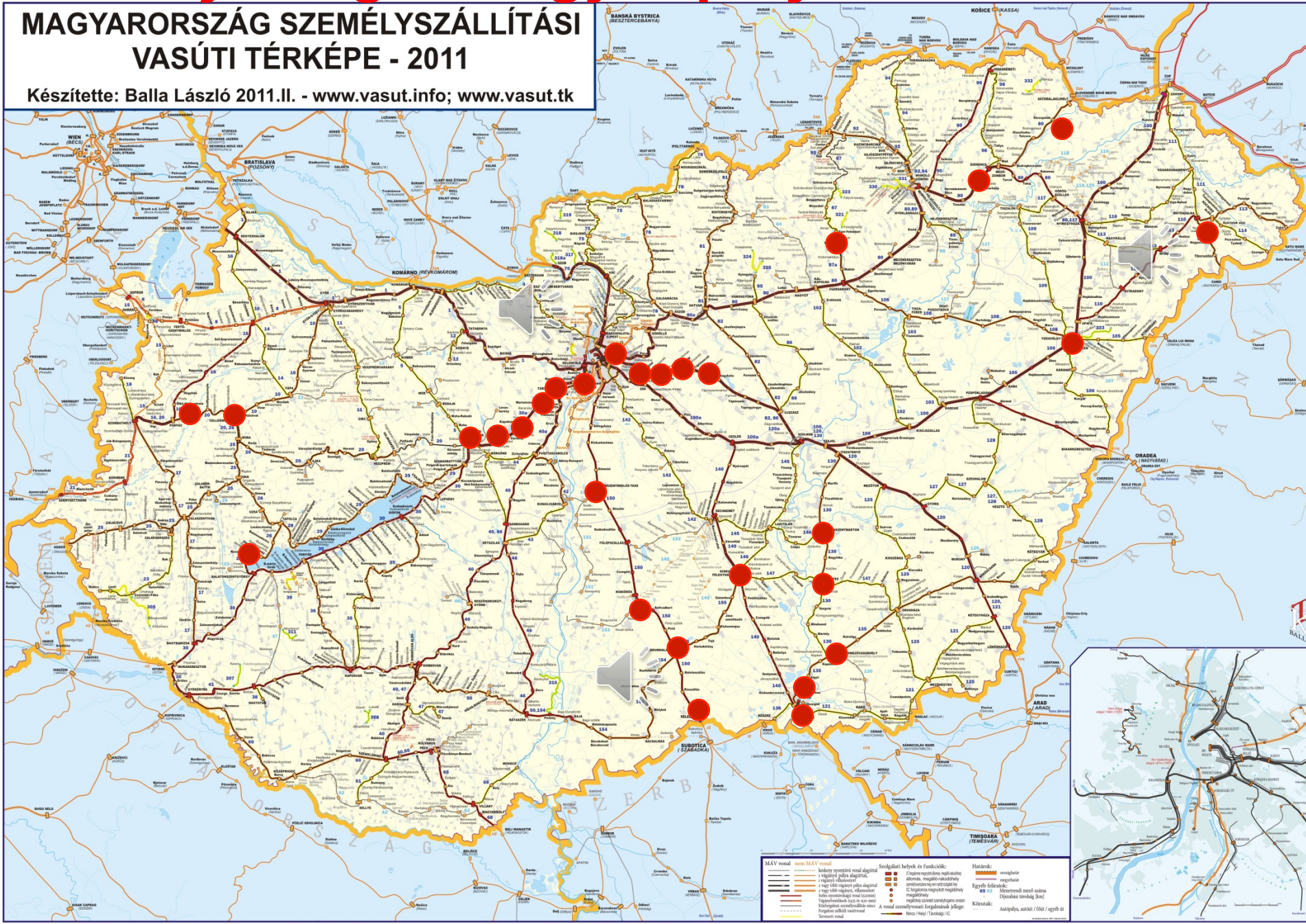


BLOCK 120 AND 121 ARE OPTIONAL IN CORPUS-BASED SYTHESIS

ProfiVox diád 1995-
ProfiVox triád 2000-
ProfiVox korpusz 2002-



Készítette: Balla László 2011.II. - www.vasut.info; www.vasut.tk

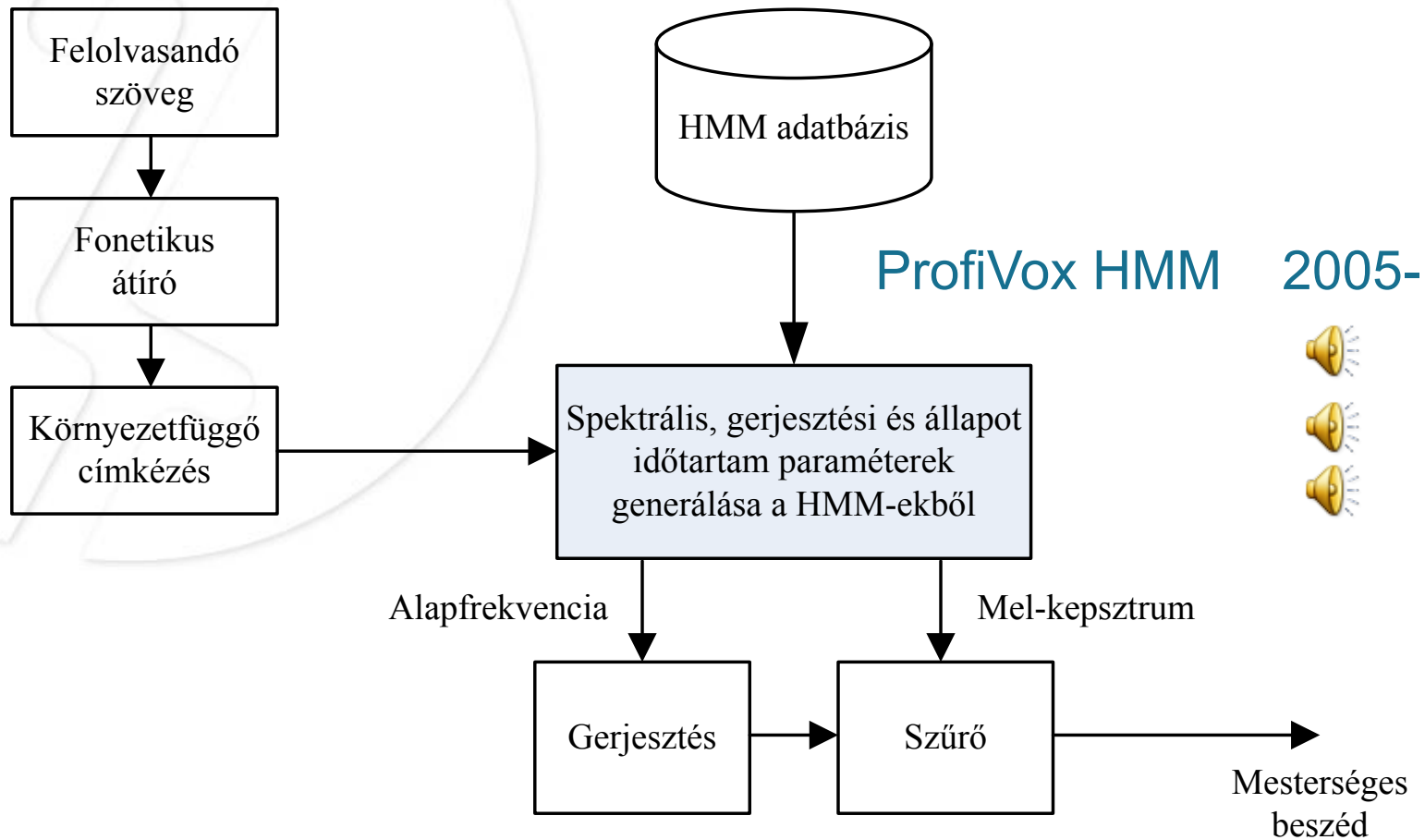


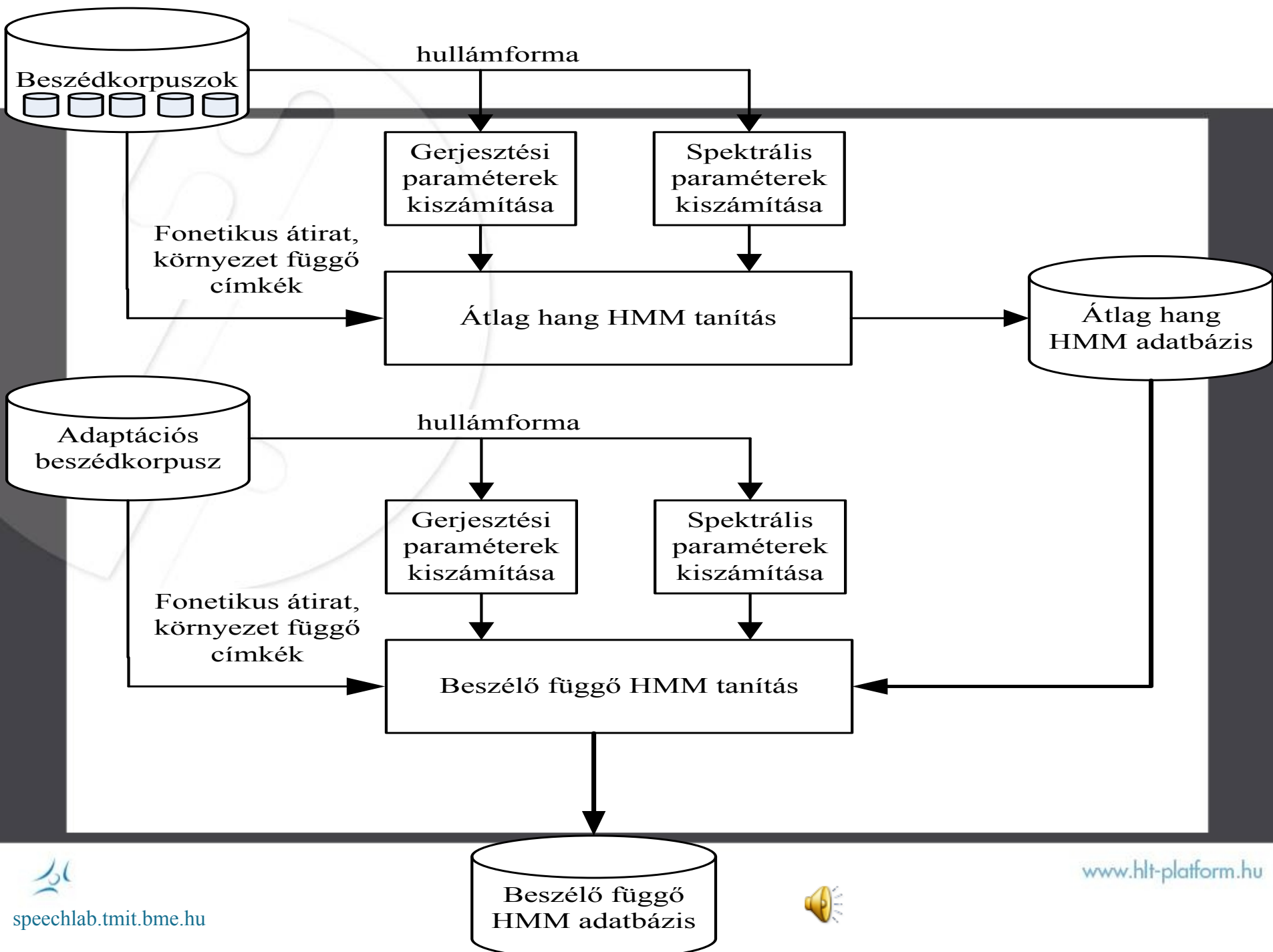
A fejlődés útja

A szabály-alapú modellek
(artikulációs csatorna, prozódia)
mellett és helyett
Természetes elemek
egyre nagyobb halmaza
statisztikai modellépítés
minimális jelfeldolgozás
Egységes(re törekvő) kiértékelés



Rejtett Markov Modell (HMM)





A fejlődés egy mértéke

Blizzard Challenge (<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



A magyar nyelvű gépi szövegfelolvasási technológia alkalmazott eredményei

MailMondó



**Westel
T-Mobile**

BME TMIT

1999

**Freedom
Scientific**

BME TMIT

2002

Informatika a Látássérültekért



**Westel
T-Mobile**

**BME TMIT
MIT Systems**

2003



Nat. Danish Center for Vis. Imp. BME TMIT 2013
Informatika a Látássérültekért



IT.DOT Kft

BME TMIT

2014



Személyre szabott beszédfelismerés

- **A beszédfelismerő rendszerek statisztikai alapelveken működnek, a beszéd akusztikai és lexikai szerkezetét 1-1 modell írja le:**
 - **Beszédhang-modellek**
 - betanításához több 100 órányi pontosan lejegyzett hanganyag.
 - A modell a beszédstílusra és a környezeti zajokra is érzékeny.
 - **Nyelvi modell**
 - szavak, szókapcsolatok valószínűségi hálózata.
 - szöveges nagy adatbázisokon tanul (parlamenti beszédek, hír szövegek, jogi anyagok, lejegyzett beszélgetések, stb.)
 - A nyelvi modell tematikus, azaz amely szövegkörnyezetre rátanult, arra működik jól a felismerő.

Nincs univerzális, azaz minden területen egyformán jól működő beszédfelismerő modell.



A magyar nyelvű gépi beszédfelismerési technológia alkalmazott eredményei (BME TMIT, Aitia, SpeechTex)

- **Telefonos ügyfélszolgálati beszélgetések**
 - Utólagos leiratozás és elemzés (minőségbiztosítás, folyamat elemzés, értékesítés támogatás, clemvoice: OTP Bank, AEGON Biztosító)
 - Valós idejű leiratozás (IVR automatizálás, operátor támogatás)
- **Rádió és televíziós műsorok, illetve Interneten tárolt hang- és videó tartalom leiratozása**
 - valós idejű teletext feliratozás (MTVA, 2015-től törvényi kötelezettség)
 - médiafigyelés és médiaelemzés
 - keresés (Mindroom)
- **Tematikus diktálás, fordítás támogatás**



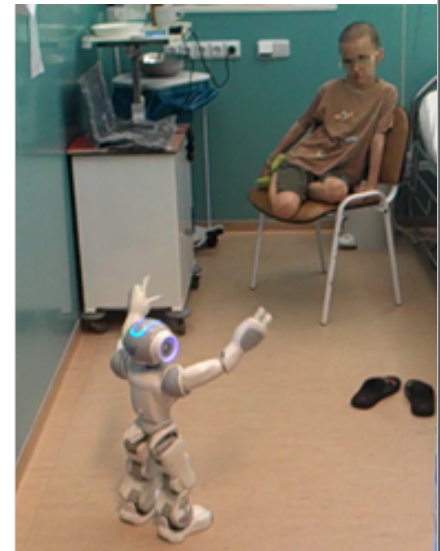
Erőforrások

- Világszínvonalú nyelv- és beszédtechnológiai együttműködő K+F kapacitás (www.hlt-platform.hu)
- Cégek (AITIA, SpeechTex, Morphologic, Nextent, ...)
- Nemzetközi hálózatok **META-NET**

- Hiányoznak a nagy ipari K+F központok
- Hiányzik a fókuszált (kormányzati) odafigyelés, a minőségi elvárások

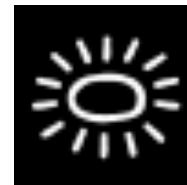
Kutatási kihívások₁

- Pontos referencia beszédfeldolgozási infrastruktúra (platform)
- Spontán interakciók feldolgozása
- Elégséges (?) adat gyűjtése és annotálása
- Finanszírozás nélküli nagy projektek (pl. U-STAR)
- Szabály-adatvezérelt kombináció
- Cognitive Infocommunications
- Cognitive Robotics
- Eto – communications
- Életközeli alkalmazások



Alkalmazási kihívások ₁

- A 15-69 éves magyar lakosok 62%-a internet felhasználó
 - Mi legyen a többiekkel (38%)?
 - Információs akadálymentesítés ???
 - A beszédtechnológia segíthet
(magyarorszag.hu, 112, MÁV, BKV, Volán)
 - Példák: www.gyogyszervonal.hu, www.metnet.hu
- Alkalmazások fogyatékosoknak
 - Képernyő olvasó vakoknak
 - Elektronikus hozzáférés írott anyagokhoz
 - Példák: www.robobrace.org, StrokeAid



Alkalmazási kihívások ₂

- Beszédtechnológia az oktatásban
 - „Játékok” óvodásoknak és iskolásoknak
 - Példa: GOH hallásvizsgálat 3 éveseknek
 - Interaktív multimodális oktatási anyag
 - Kisebbségi helyzetű magyar gyermekek motiválása
 - Afáziás, autista, stb. támogatása



Alkalmazási kihívások ₃

- Beszédtechnológia az egészség iparban
 - Műtétek automatizálása (utasítások, jegyzetelés)
 - Leletek diktálása
 - Hang alapján gége problémák, depresszió, stb. korai diagnózisa és rehabilitációja
 - Táv(fel)ügyeleti alkalmazások (pl. gyógyszer bevétel figyelmeztetés, ablak zárás, stb.)
 - Demencia, Alzheimer, ... felügyelete



Alkalmazási kihívások 4

- Beszédtechnológia a tartalom- és a kreatív iparban
 - Interdiszciplináris integráció
 - Beszédtechnológia – orvosok – szociális munkások képzése (pl. IBM –kormányzat)
 - Digitális közoktatás és intelligens otthon program (pl. Microsoft – kormányzat)
 - Multi-modális tartalomelemzés (közvélemény kutatás)
 - Bankok, kisker., információs szolgáltatások
 - Car infotainment (Audi, Daimler – kormányzat)
 - Beszédvezérelt otthon
 - okostelefon, okosTV
 - okos mosógép,



Alkalmazási kihívások ₅

- Beszédtechnológia a gyártásban
 - Raktár-logisztika automatizálás
 - Gyártásközbeni információ, figyelmeztetés
 - Beszéd utasítások
 - Beszélő gépkönyvek
 - 3DICC

**3D Internet Based
Control and
Communication**



Hozzászólások

**Mélyebb érdeklődőknek: <http://speechlab.tmit.bme.hu/>
<http://magyarbeszed.tmit.bme.hu/>**

Köszönjük



EUROPEAN
COMMISSION

támogatását.



(Teleauto, BelAmi, EtoCom -TÁMOP-4.2.2-08/1/KMR-2008-0007- , BME
Kutatóegyetemi -TÁMOP-4.2.1/B-09/1/KMR-2010-0002- ,
CIP CESAR, AAL PAELIFE projektek)

