



NYELV- ÉS
BESZÉDTECHNOLÓGIAI
PLATFORM
GÉPEKKEL - EMBERI NYELVEN

A beszédtechnológia jelene és jövője

Németh Géza
BME

Távközlési és Médiainformatikai Tanszék
Beszédtechnológiai Laboratórium
nemeth@tmit.bme.hu

Nyelv- és Beszédtechnológiai Nemzeti Technológiai Platform
Szakmai Fórum
2010. március 18.



BME TMIT 1

www.hlt-platform.hu





- Miért fontos?
- Hazai helyzetkép nemzetközi kontextusban
- Kihívások





Miért fontos?

Mit mond?

Nyelvi tartalom felismerése,
szöveglejegyzés

Ki beszél? Beszélő felismerés

Ideges? Bánatos? Érzelem
felismerés

Milyen nyelven beszél?
Nyelvfelismerés

Meg van fázva? Egészségi állapot
felismerés





Várható forradalmi változások az ICT (Information & Communication Technologies) területen (2005-15)

Hálózatok (networking)

Számítástechnika (computing)

Mobilitás (mobility)

Felhasználói felületek (human interface)

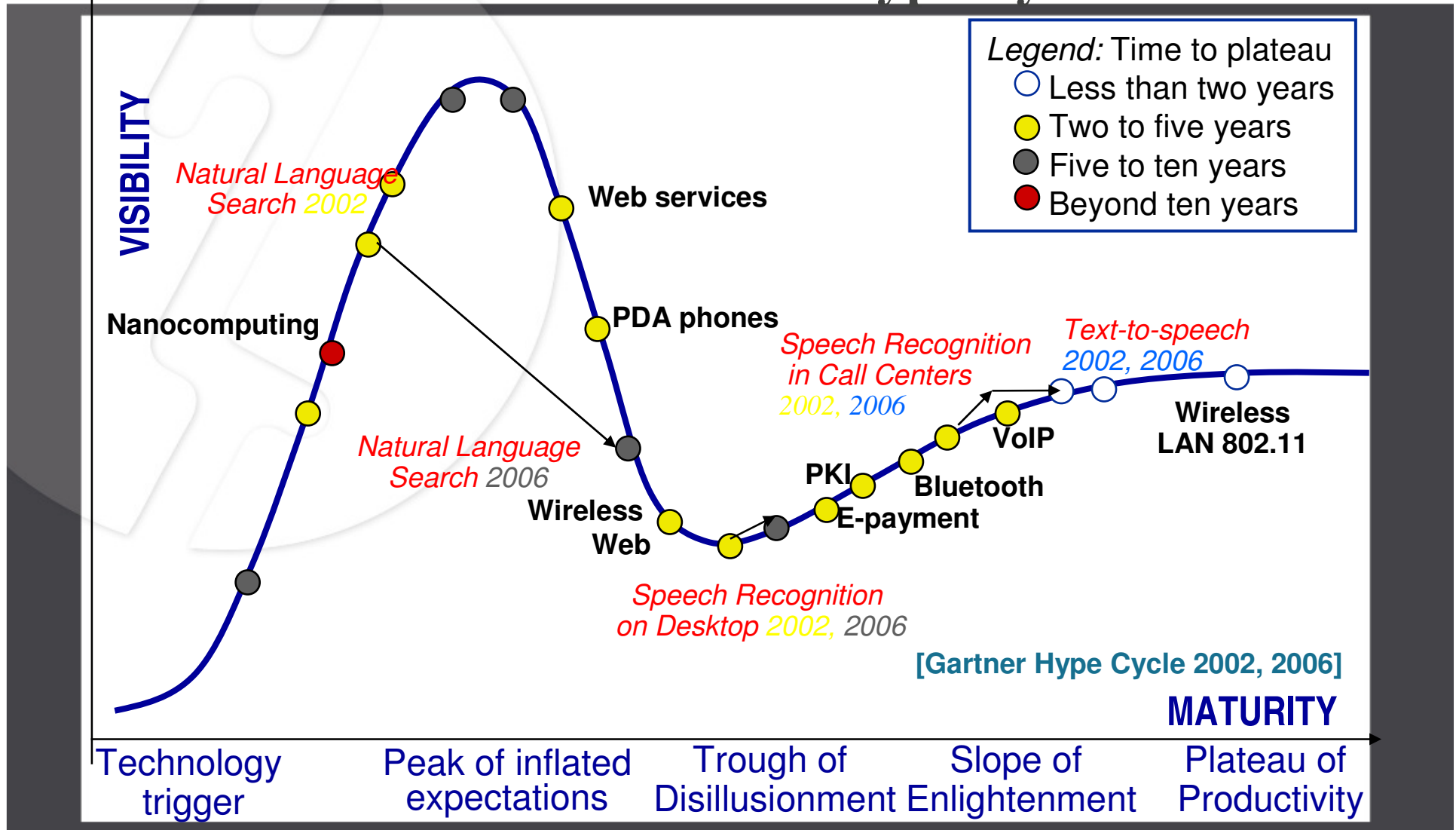
Alkalmazások (applications)

[Gartner symp. 2002. nov.]





Miért fontos? Hype Cycle 2002-2006





- **Internetes számítógép**
 - tényleges használó a lakosság 51.6%-a (KSH, 2007)
 - ténylegesen interneten vásárló a lakosság 6.9%-a
- **Mobiltelefon**
 - SMS-írásra képes kb. 60%
 - **hangkapcsolatra képes közel 100%**
- **Vezetékes telefon**
 - **Hangkapcsolatra képes közel 100%**

Mindenkit elérő automatizált interaktív információ-szolgáltatást csak a beszédtechnológia eredményeinek felhasználásával lehet nyújtani.





- Ahogy egy cég vizuális arculatát szakértők csapata alakítja ki, egy beszédinformációs rendszereket alkalmazó cégnél szakértőknek kéne az **akusztikus arculattal** is foglalkozni.
- Furcsa, hogy olyan cégek, amelyek elsődlegesen beszédkapcsolatból élnek, a saját beszédminőségük, -beleértve annak nyelvi színvonalát- fontosságát lassan ismerik fel





Az áttörés megtörtént – a beszédtechnológia használható!

- Használható: kisebb költséggel, többet tud a gép, mint az ember?!
- **Igen!**
 - Ki tudja 200 óra **hangzó híryananyagban** megtalálni a megadott 1000 kulcsszó ~50%-át 24 óra alatt (pl. www.mindroom.hu)?
 - 2 db (4 magos) CPU képes rá! (RTF<1 /mag)
 - 2 ember, > 2 hét – csak lehallgatás (RTF>1 /fő)
 - 12 ember, > 2 hét – teljes lejegyzés (RTF>6 /fő)

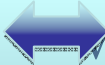
Áramdíj < 2 eFt ↔ Munkadíj > 200 eFt / 1.2 MFt





Az áttörés megtörtént – a beszédtechnológia használható!

- Használható: többet tud a gép, mint az ember?!
- **Igen!**
 - Ki képes egy nagyváros összes lakosának ~250.000 fő nevét és lakcímét első hallásra 10 esetből min. 9-szer hibátlanul leírni?
 - 1 db CPU képes rá (+ beszédtechnológia)!
 - 0 sec válaszidő!!

Áramdíj, válaszidő: ~ 0  visszakérdezések, hibák





Az áttörés megtörtént – a beszédtechnológia használható!

- Használható: kisebb költséggel, többet tud a gép, mint az ember?!
- *Igen, igen!*
 - A **beszédszintetizátor** bármit kimond
 - Nem fárad el, nem reked be
 - Ha kérik, gyorsabban beszél, ha kérik, lassabban
 - Nem sértődik meg, ha a felmenőit emlegetik
 - pl. www.metnet.hu
 - ...





- Név szerinti kapcsolás név-visszamondással (100-3000 név, izolált szavas)
 - egy biztosítónál kísérleti rendszer
- Gyógyszervonal (5000 gyógyszernév, izolált szavas)
- Pozíció alapú hangos kereső (ATM, étterem, stb... címe)
- Beszéd alapú híradó tartalom kereső (www.mindroom.hu)
- **Lehetséges lenne:**
 - földhivatali ingatlan nyilvántartás helyrajzi szám,
 - cégnyilvántartás lekérése adószám,
 - gépkocsi nyilvántartás rendszám, tulajdonos neve, címe
 - adótartozás adószám

... alapján történő automatikus elérése telefonon





- ***Gyógyszervonal információs rendszer***
(www.gyogyszervonal.hu, tel: 886 94 90)
- Gyógyszerek betegtájékoztatójának gépi felolvasása
- 5500 adatrekord feldolgozása alapján 2005-7 között a GVOP Program projekt keretében kifejlesztve
- Üzemelteti az Országos Gyógyszerészeti Intézet
- Nagyszótárú telefonos, személyfüggetlen, magyar nyelvű beszéd felismerő és gyógyszer-tájékoztatóra adaptált szöveg felolvasó + Web és Wap
- Világszerte egyedülálló megoldás (EU szerint minta projekt)
- Hasonló témákban ilyen jellegű információs rendszerek létrehozására partnereket keresünk





- Jelentősen eltérő kategóriák
 - létező szolgáltatás (személyes ismerős) automatizálása
 - részben (ismeretlen emberi kezelő)
 - teljesen (csak gépi hang)
 - új szolgáltatás létrehozása (testreszabott információk, pl. tömeges egységes üzenetkezelés, egyéni út- és hójelentés, Google voice search, MS project Natal...)
 - multik nyelvi „sorozatfejlesztése” <-> használhatóság
 - pl. Google „nyomulás” – magyar média tájékozatlansága
- A minőségi beszédtechnológia
NEM VÁSÁROLHATÓ KÉSZTERMÉK,
a jó integráció a siker kulcsa.





Kulcsterületek

- Infrastruktúrák
 - Adatbázisok
 - Eszközrendszerek
- Zajos környezet (pl. gépjármű)
- Távoli mikrofon (pl. környezeti intelligencia)
- Természetes (spontán) beszéd feldolgozása (felismerés és szintézis)
- Nyelvcsaládokra alkalmazható módszerek
- Skálázható technológiák (kis fogyasztású, szenzor jellegű – mobiltelefon – telepített számítógép)



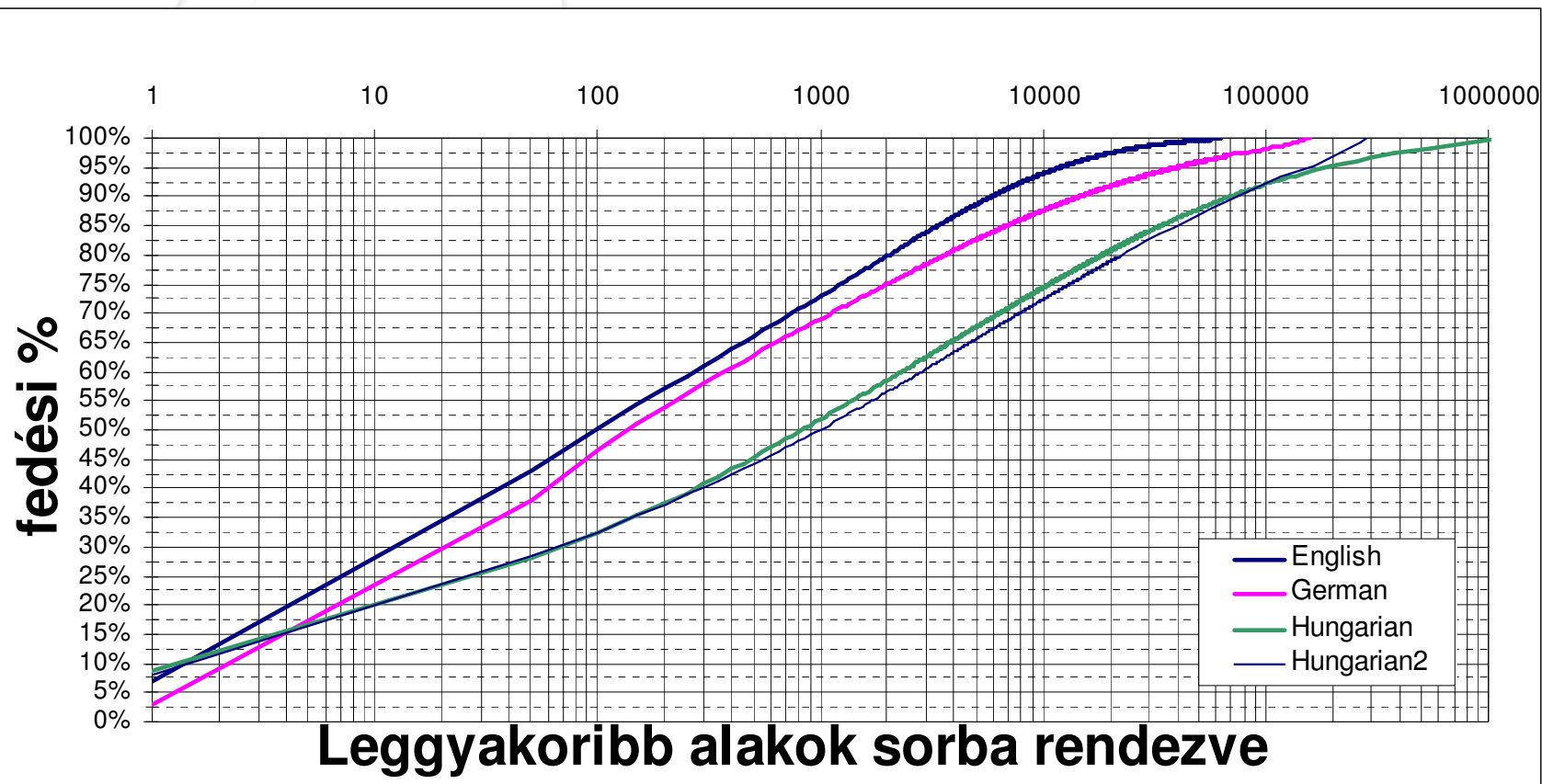
Szakpolitikai és politikai tényezők

- közbeszerzéseknél előírhatná a kormány bizonyos nyelv- és beszédtechnológiai erőforrások használatát, illetve ezek folyamatos karbantartását
- a kormányzat sokat tehetne (ha ismerné a technológiai lehetőségeket) szabályozással, illetve a kormányzati projektek (pl. portál, esélyegyenlőség) szolgáltatásainak és folyamatainak a fejlesztésével
- Nemzetközi kutatási projektekhez, hazai interdiszciplináris együttműködésekhez való kapcsolódás támogatása
- az ingyenes (de rosszabb minőségű) megoldások eltéríthetik a kormányzati figyelmet
- tipikus (ál)hír: az adott problémához tartozó erőforrást már „létrehozta” valamelyik nagy multi (Google, MS stb.)
- közbeszerzési szempontrendszer normatív kialakítása





Angol, német és magyar korpuszok fedési tulajdonságai





Mindenkit elérő *felhasználóbarát* automatizált interaktív információ-szolgáltatást már ma is lehet hazánkban nyújtani a beszédtechnológia eredményeinek felhasználásával !!!

Ma a kritikus kérdés a köztudatba és a hétköznapi életbe bevinni az itthon létrejött eredményeket ill. biztosítani a K+F infrastruktúra folyamatosságát!





**Érdeklődéssel várom javaslataikat,
kérdéseiket, megjegyzéseiket itt
(hozzászólás időben)**

valamint a

nemeth@tmit.bme.hu

emil (drótposta,...) címen

Köszönjük az  **támogatását.**

Nemzeti Kutatási és Technológiai Hivatal

