

Csapó Tamás Gábor
V. évf.

Beszédszintetizátor prozódiai változatosságának növelése

Konzulensek:
Dr. Németh Géza, Dr. Fék Márk

Budapesti Műszaki és Gazdaságtudományi Egyetem
Távközlési és Médiainformatikai Tanszék

TDK előadás
2007. november 16.



BME TMIT

Tartalom

- A kutatás célja
- Elméleti bevezetés
- Kísérleti munka
- Meghallgatásos vizsgálatok
- További kutatási irányok

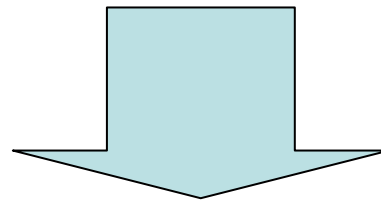
A kutatás célja

Emberi beszédben:

- ugyanaz a mondat többször kiejtve kicsit máshogy hangzik
- változatosság a dallamban, hangsúlyozásban, ritmusban

Szövegfelolvasókban eddig:

- a legtöbb rendszerben egy adott mondat mindig ugyanúgy hangzik
- többnyire determinisztikus rendszer
- nincs jól észlelhető változatosság



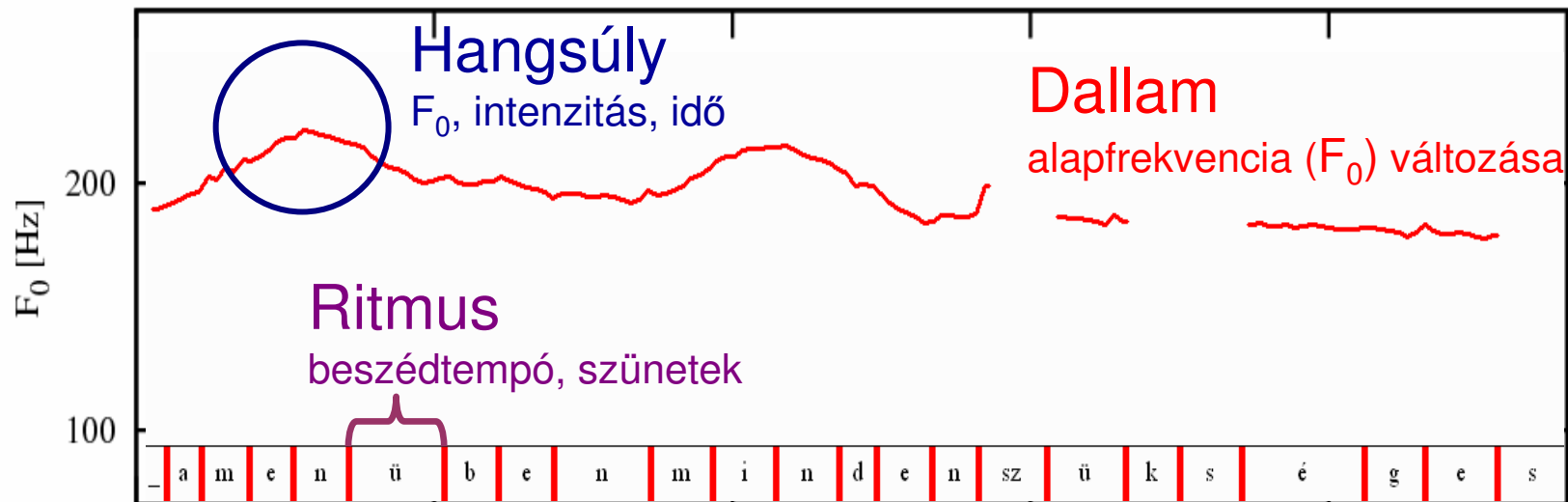
Cél: változatos beszéd megvalósítása a szövegfelolvasókban

A kutatás lépései

- TDK, 2006 ősz
 - koncepció ellenőrzése
 - 200 rövid mondat vizsgálata
- OTDK, 2007 tavasz
 - 5200 teljes mondat vizsgálata egyben
- TDK, 2007 ősz
 - 5200 teljes mondat vizsgálata
 - mondaton belüli egységek alapján

Elméleti háttér (1)

- Beszéd (prozódia) 3 összetevője:



legfontosabb paraméter: alapfrekvencia

⇒ ezzel foglalkoztunk

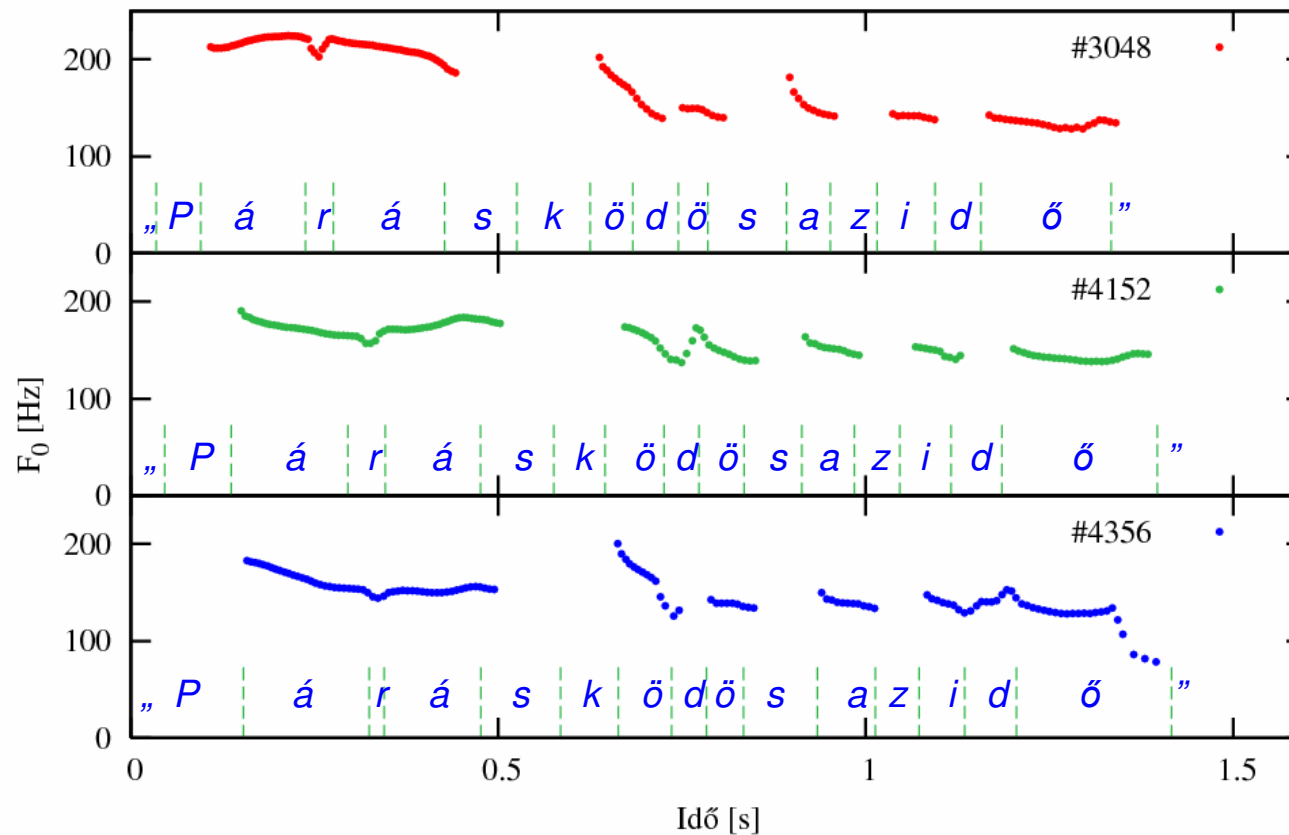
prozódia nélkül



prozódiaval (csak F_0)

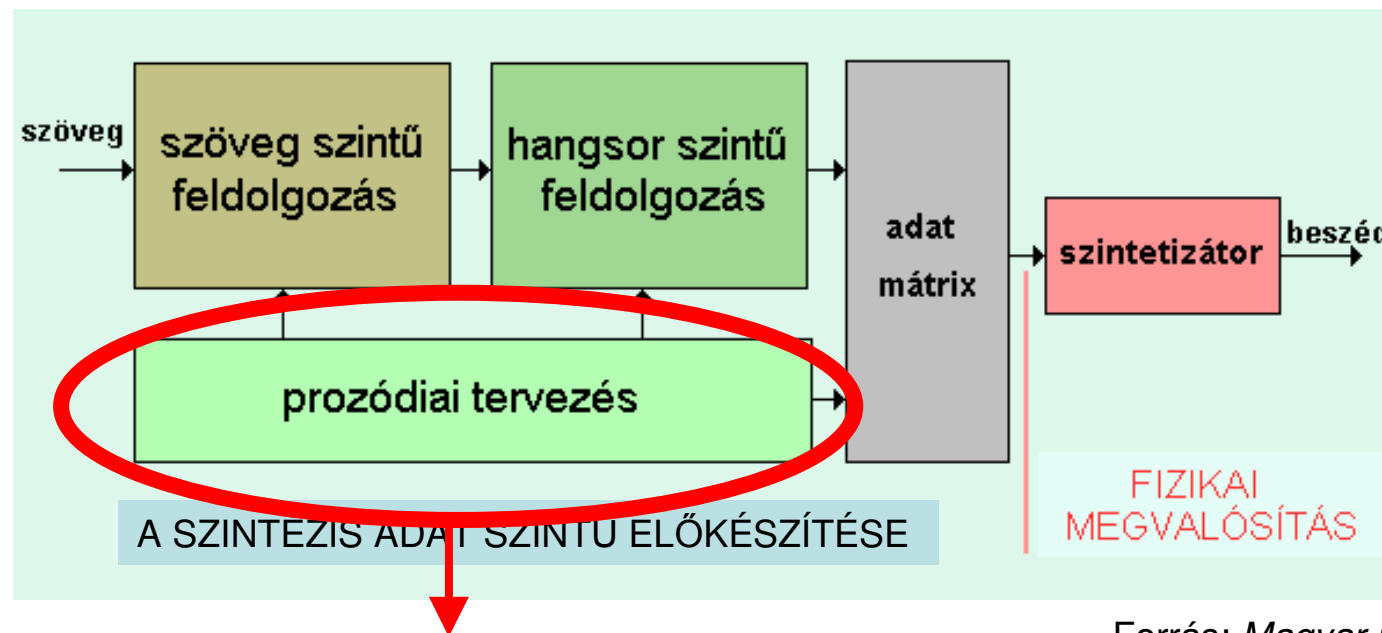
Elméleti háttér (2)

- Változatosság a természetes beszédben:



Elméleti háttér (3)

- Beszédszintézis: szöveg → beszéd
(Text-To-Speech = TTS)



Dallammenet módosítása

Forrás: Magyar nyelvi
beszédtechnológiai
alapismeretek CD

Elméleti háttér (4)

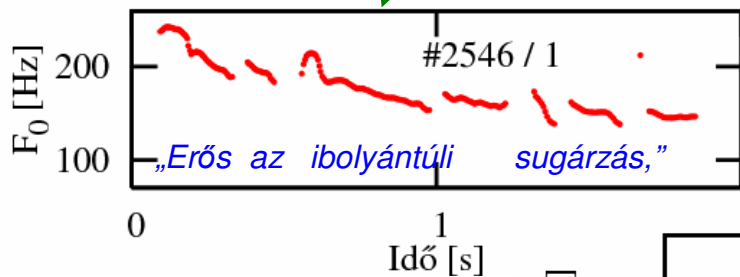
- **Prozódiai modellek**
 - Szabály alapú, szuperpozíciós
 - BME TMIT: Profivox
 - Korpusz alapú
 - beszéd-adatbázisbeli minták alapján
 - különbség: a minták méretében
 - hang, szótag, prozódiai egység, tagmondat
 - szuperpozíciós

Beszédkorpusz (1)

- 5200 magyar nyelvű kijelentő mondat
- időjárás-előrejelzés témakör
- tartalma:
 - hang
 - szöveg
 - címkézés (hangidőtartam, F_0 -menet)
- prozódiai egységekre bontás
 - szóban: szünetek határolják
 - írásban: írásjelek (vessző, pont, ...) határolják

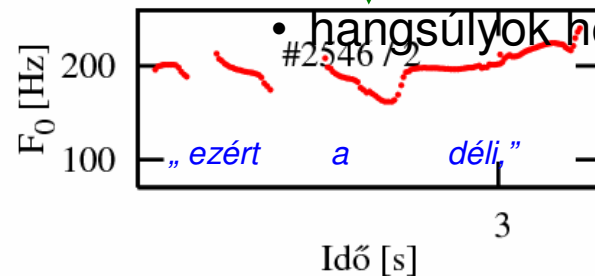
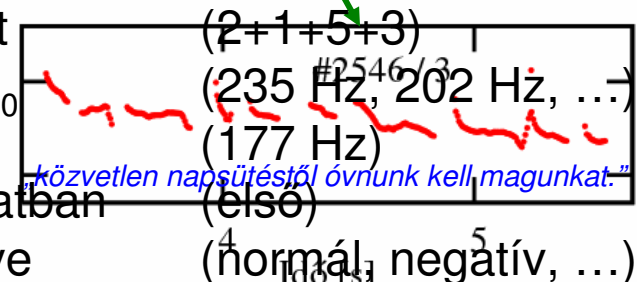
Beszédkorpusz (2)

- Prozódiai egységekre bontás



Eltávolítva:

- szótagszerkezet
- szótagonkénti F_0
- átlagos F_0
- pozíció a mondatban
- hangsúlyok helye



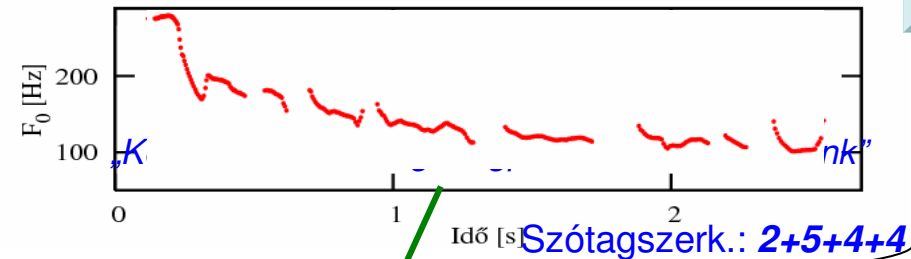
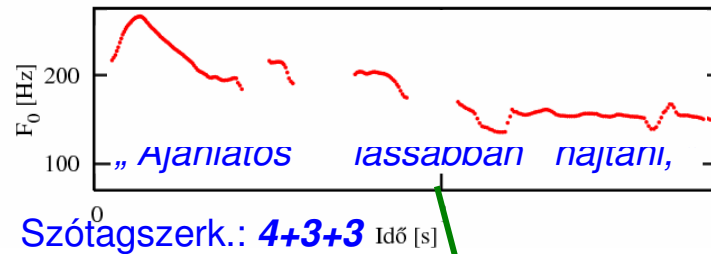
Dallammásolás

- Prozódiai egység alapján
– keresés: szótagszerkezet szerint

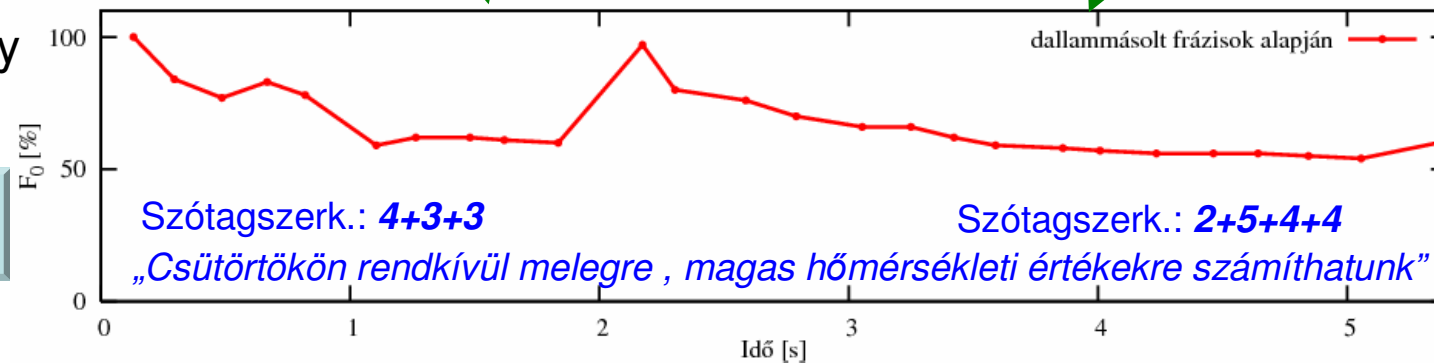
#1091 /1

#2514 /1

Beszéddallam-adatbázis prozódiai egységekkel



szabály alapú



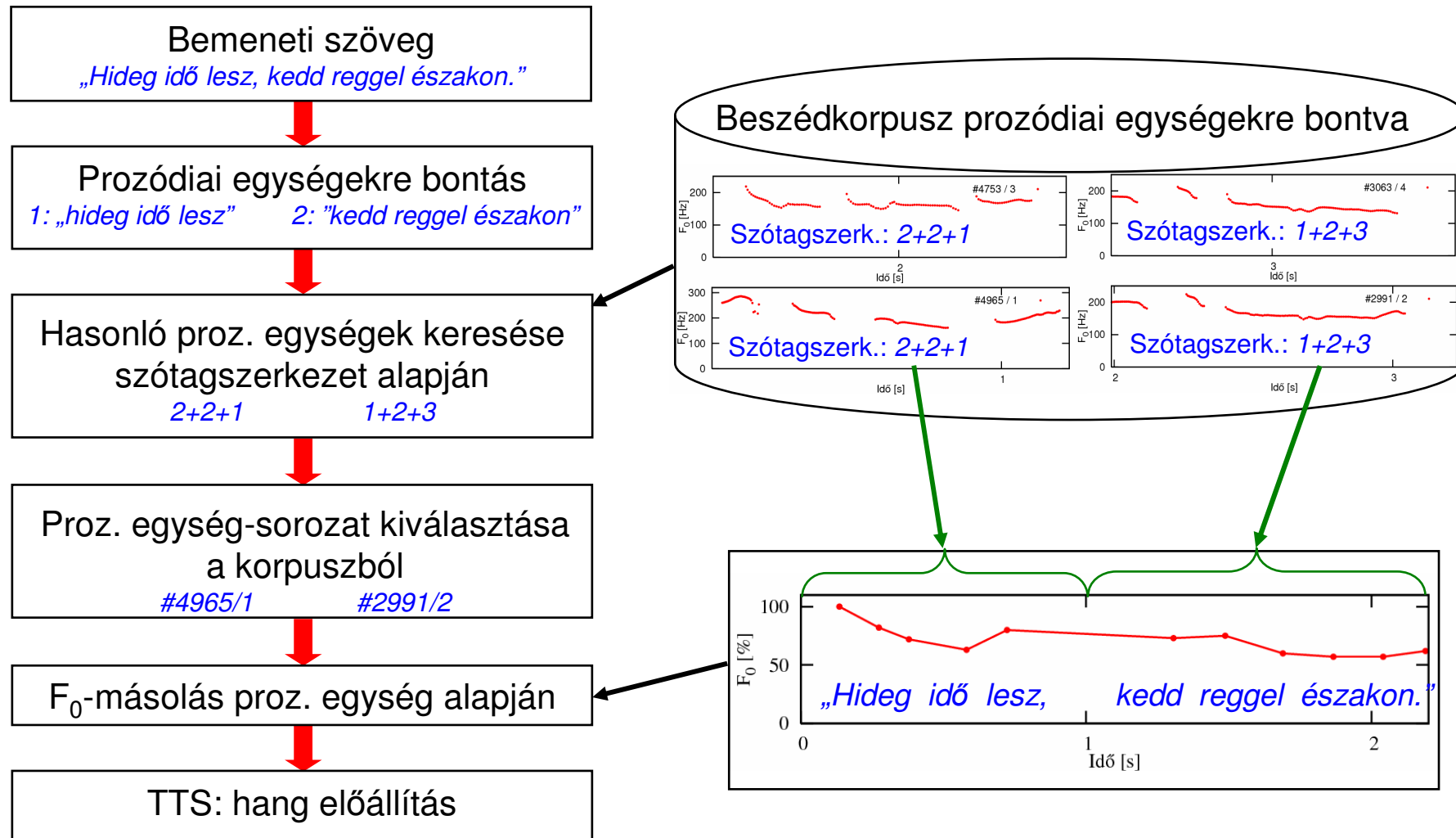
dallammásolt frázisok alapján

dallam-
másolt



BME TMIT

Beszédszintetizátor változatossággal

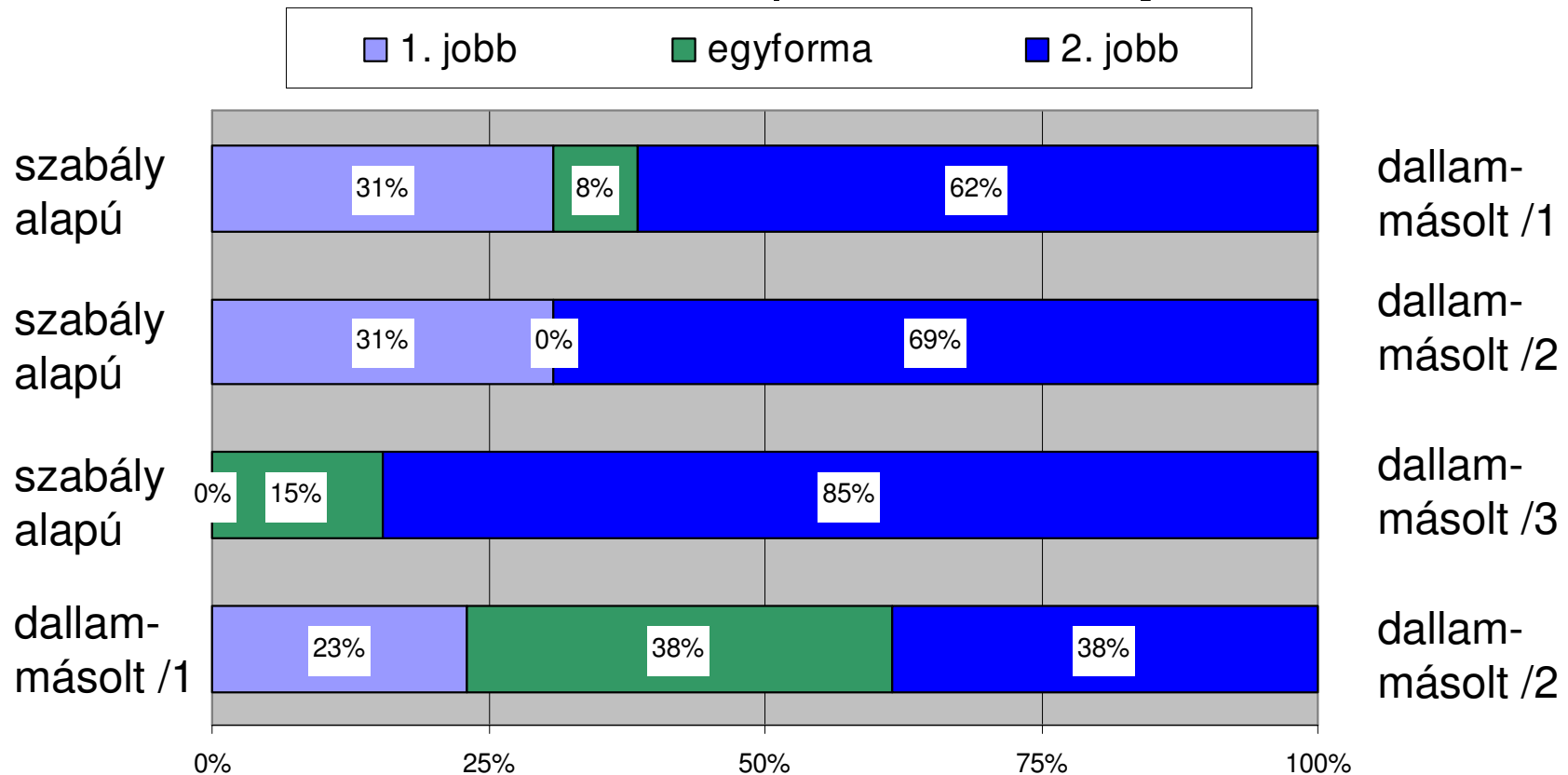


Vizsgálatok

- Meghallgatásos vizsgálatok weben
- 37 mondatpár, egy-egy mondat több változata:
 - szintetizált: szabály alapú dallammenettel (1 db)
 - szintetizált: korpuszbeli mondatokról másolt dallammenettel (1-3 db)
- 13 tesztelő
 - el kellett dönteniük, hogy melyik változat „természetesebb” (1. jobb / egyformák / 2. jobb)
- Mit akartunk megtudni?
 - az egyes dallamváltozatok ugyanannyira jók-e?

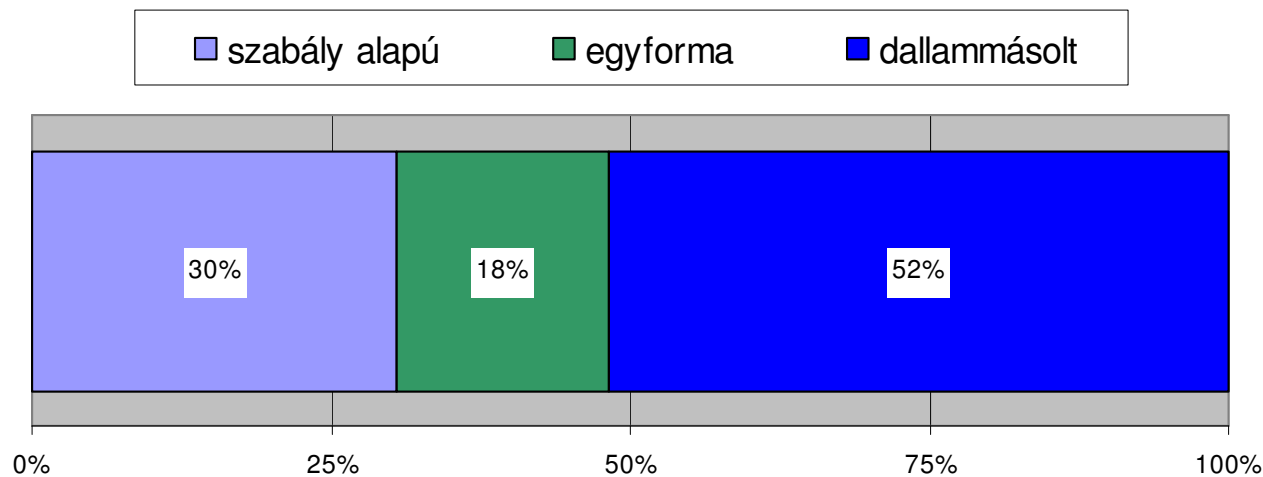
Teszt-eredmények (1)

- #2546-os mondat változatai
– soronként 1-1 mondatpár eredménye



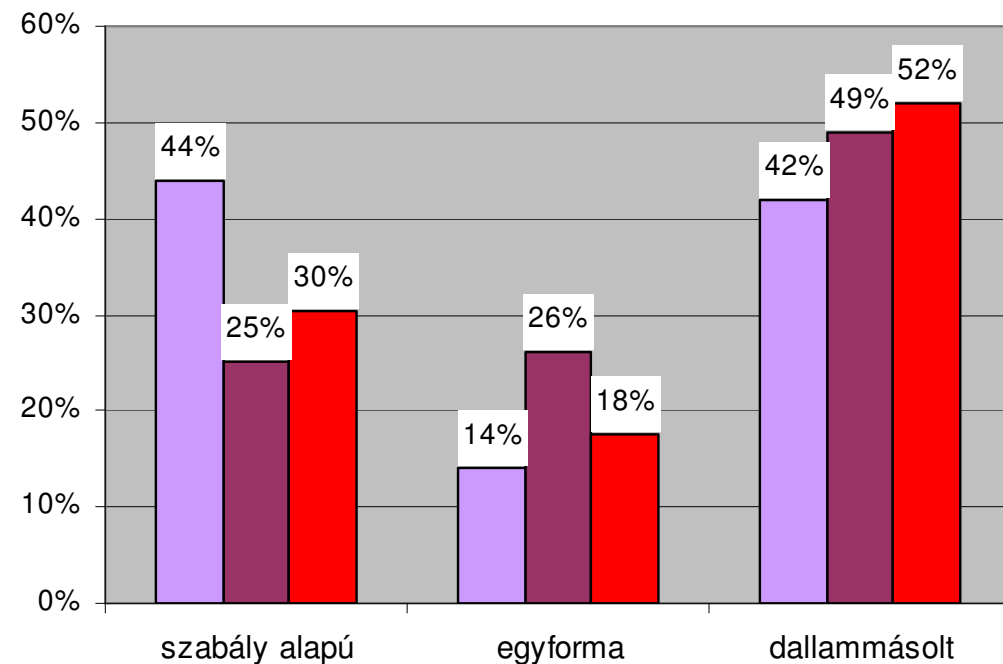
Teszt-eredmények (2)

- Összesített eredmény
 - 37 mondatpár
 - 11 esetben dallammásolt szignifikánsan jobb
 - 3 esetben szabály alapú szignifikánsan jobb
 - F_0 -másolt változat végén szokatlanul magas volt a dallam
 - többi esetben egyik sem jobb szignifikánsan a másiknál



Összehasonlítás a korábbi eredményekkel

- TDK, 2006 ősz
 - dallammásolás manuális módszerrel
 - rövid mondatok
- OTDK, 2007 tavasz
 - hosszabb mondatok
- TDK, 2007
 - dallammásolás automatikus módszerrel
 - prozódiai egységek alapján



Előadás összefoglalása

- Elméleti bevezetés
- Beszédkorpusz prozódiai egységekre bontása
- Dallammásolás prozódiai egység alapján
- Meghallgatásos vizsgálat 13 tesztelővel
- Dallammásolás: a legtöbb esetben jó eredményt nyújt, de nem mindig

További kutatási irányok

- Nagyobb beszéd-adatbázis:
 - rövidebb mondatok
 - más téma
- Prozódia többi komponense
 - időzítés
- Dallammásolás során alkalmazott hasonlósági mérték finomítása
- Prozódiai változatosságért felelős modul beépítése a Profivox-ba

Köszönöm a figyelmet!

Csapó Tamás Gábor

csapszi@sch.bme.hu