

Csapó Tamás Gábor  
IV. évf.

# Szintetizált beszéd természetesebbé tétele

Konzulensek:

Dr. Németh Géza, Dr. Fék Márk  
BME TMIT

TDK előadás

2006. november 17.



M Ű E G Y E T E M 1 7 8 2

BME TMIT

# Tartalom

---

- A fejlesztés célja
- Elméleti háttér
- Kísérleti munka
- Meghallgatásos vizsgálatok
- Továbbfejlesztési lehetőségek



# Cél

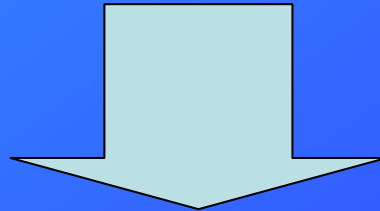
---

## Emberi beszédben:

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

## Szövegfelolvasókban eddig:

- egy adott mondat mindig ugyanúgy hangzik
- determinisztikus rendszer
- nincs semmiféle változatosság



Cél: változatos beszédstílus megvalósítása a szövegfelolvasókban

# Elméleti háttér /1

---

- Beszédstílus (prozódia) összetevői:

## Szubjektív

- Dallam
- Hangsúly
- Ritmus

## Objektív

- alapfrekvencia ( $F_0$ ) változtatás
- $F_0$ , intenzitás emelés, időtartam nyújtás
- beszédtempó változtatása, szünetek

prozódia nélkül



prozódiával



- Mivel a dallam a meghatározó, a továbbiakban ezzel foglalkozunk

# Elméleti háttér /2

---

- Beszédszintetizátorok generációi:

- Formánsszintézis



- Elemösszefűzés (Profivox)



- Korpusz alapú elemkiválasztásos



# Elméleti háttér /3

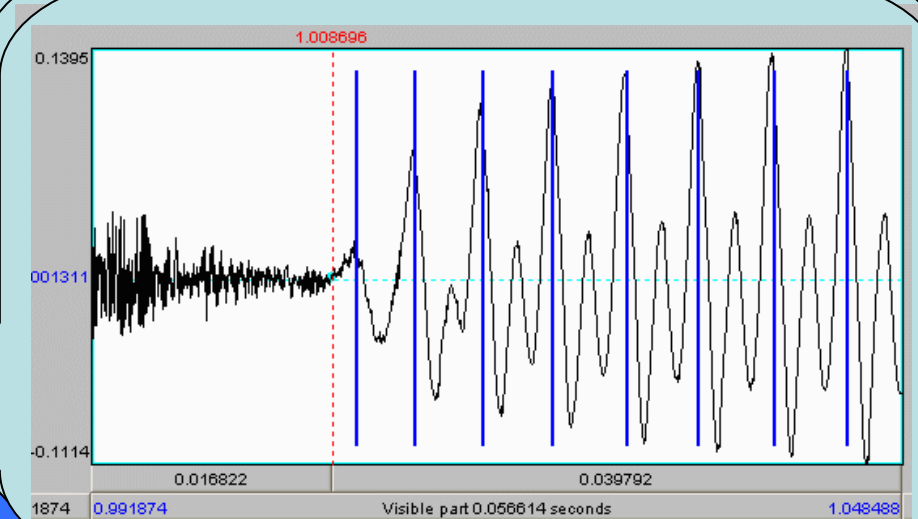
---

- Prozódiai modellek
  - Leíró jellegű (ToBi, GToBi, IViE)
  - Szabály alapú
  - Adatvezérelt (CART-tal, neurális hálóval)
- Profivox: szabály alapú, szuperpozíciós



# Beszéddallam-adatbázis

- Tartalma:
  - Hullámforma fájl
  - Szöveges átírás
  - Fonetikus átírás
  - Hang-, szünethatárok
  - Zöngeperiódus-határok



## Kísérleteinkben

Időjárás előrejelzés  
hangkorpusz

Harangok története  
hangkorpusz

# Mondatok vizsgálata /1

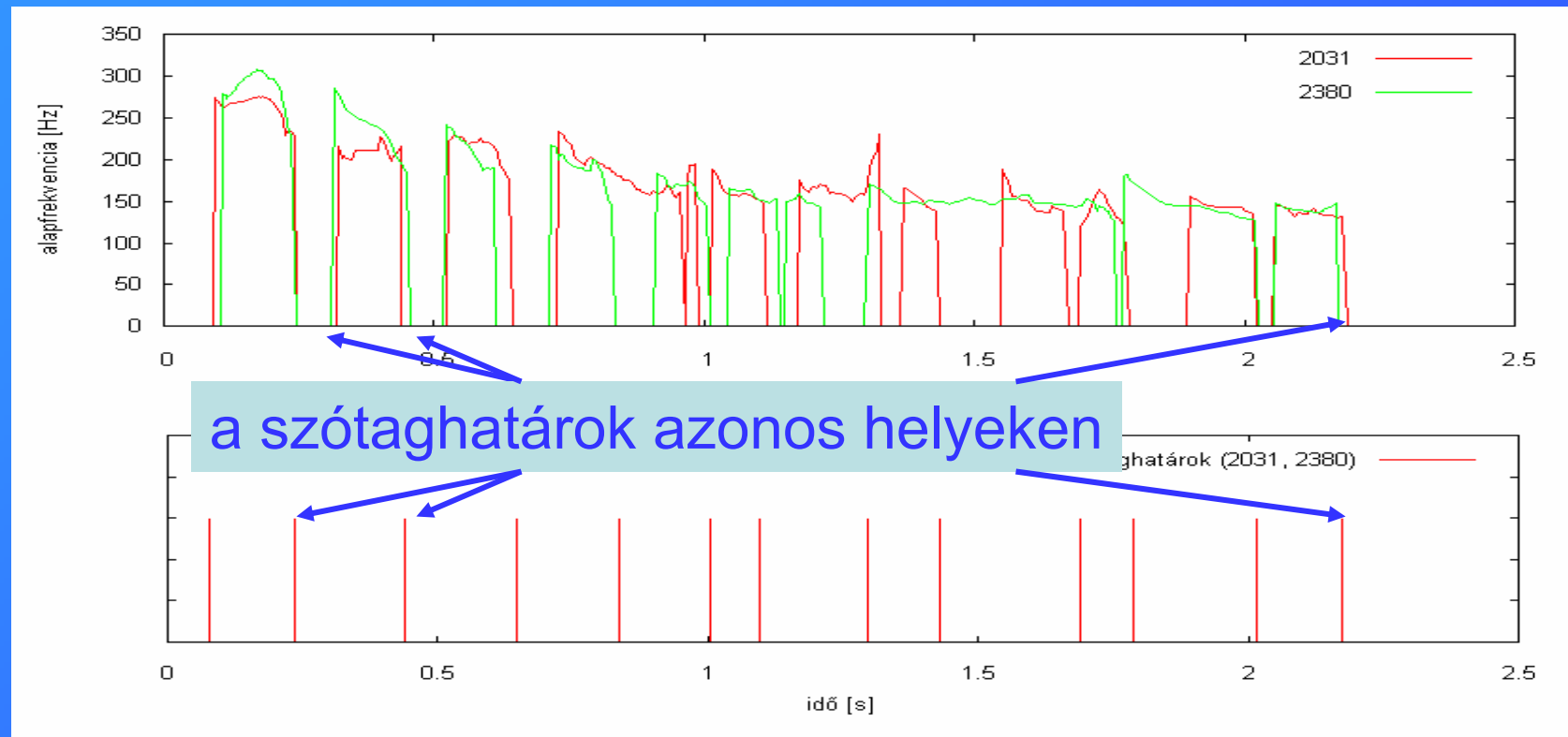
- Hasonlóságok keresése mondatok között
- Hasonlóság alapja:
  - időtartam
  - szótagszám
    - teljes mondat
    - egyes szavak
- Azért szótagszám, hogy a mondatok dallama hasonló legyen

mondat azonosító	szótag-szám	szó-szám	szavak szótagszáma
2203	10	4	6 1 1 2
2563	10	4	3 4 2 1
3056	10	3	4 3 3
2019	11	5	1 1 3 3 3
2150	11	4	4 1 5 1
2565	11	4	3 2 2 4
3024	11	4	4 1 3 3
2031	12	4	3 4 2 3
2380	12	5	3 4 1 2 2
3031	12	7	3 1 1 1 3 1 2



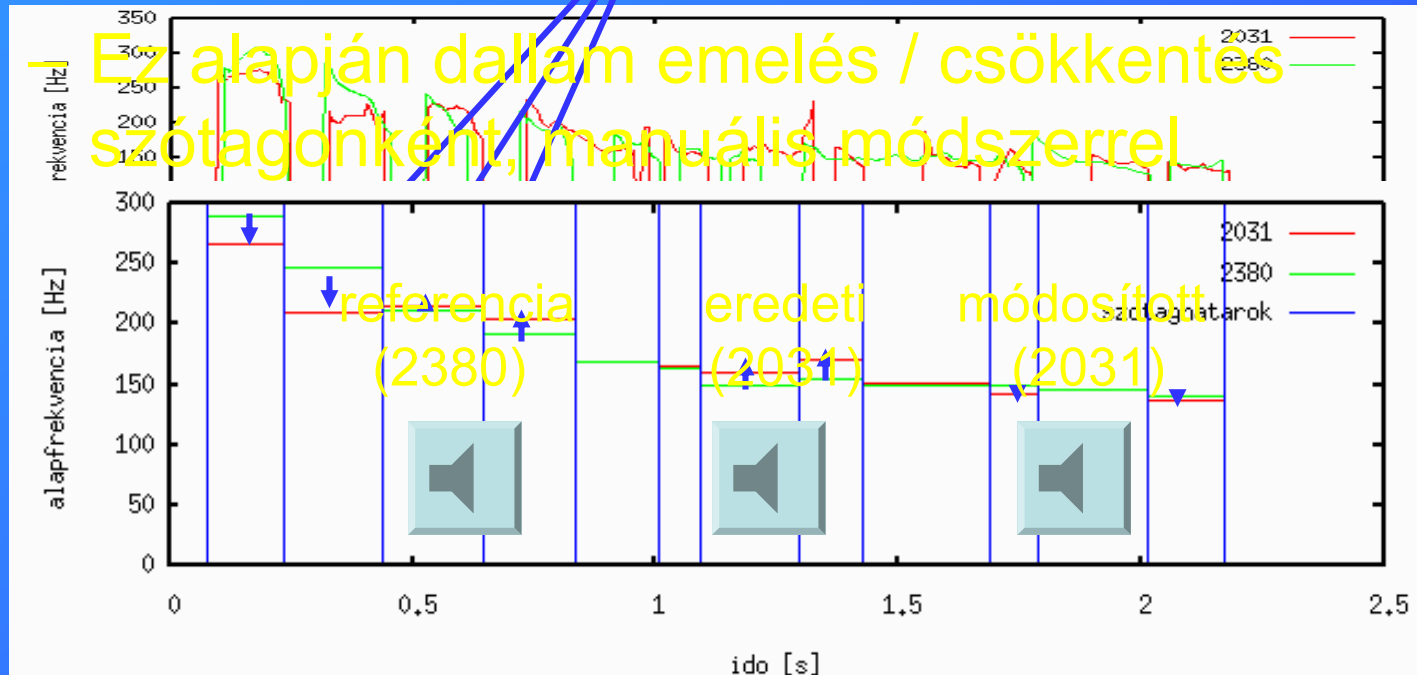
# Mondatok vizsgálata /2

- Dallammenetek összehasonlítása idővetemítéssel



# Dallamcsere /1

- Természetes bemondások között: Praat [[www.praat.org](http://www.praat.org)]
  - Szótagonkénti átlagos alapfrekvencia kiszámítása mindkét mondatban



# Dallamcsere /2

---

- Szintetizált mondatok között: szabály alapú szövegfelolvasóval (Profivox)

– bemenet:

- szöveges átírás
- eredeti hangfájl

– kimenet:

- eredeti mondat dallamával szintetizált változat



– bemenet:

- szöveges átírás
- referencia hangfájl

– kimenet:

- referenciamondat dallamával szintetizált változat



Profivox  
alapbeállításával:



# Vizsgálatok

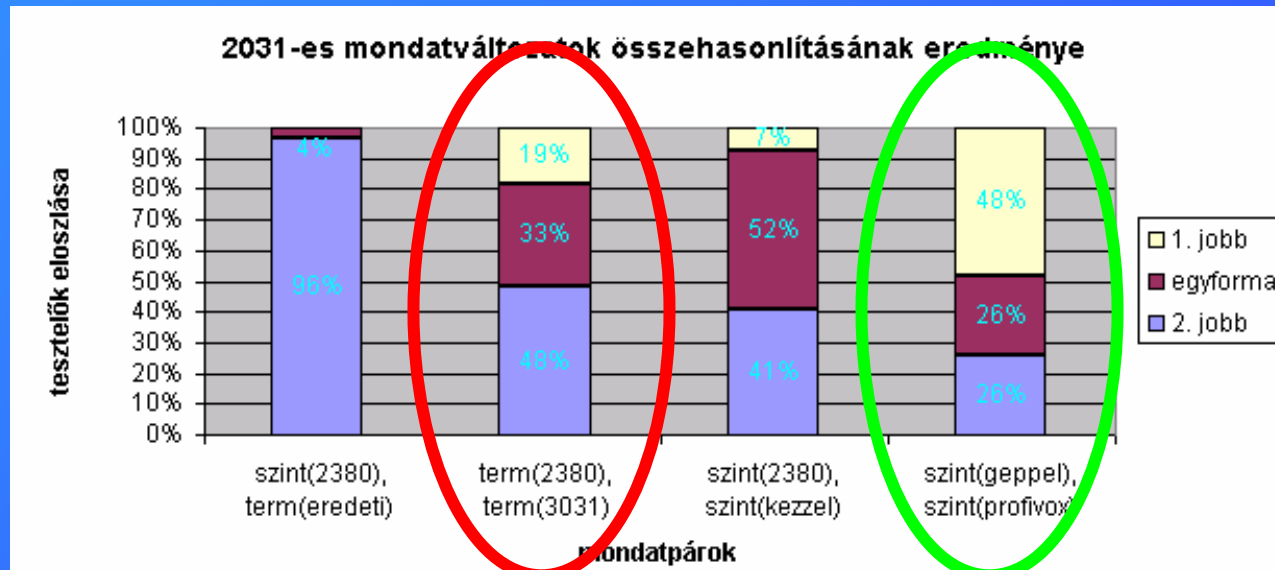
---

- Meghallgatásos vizsgálatok weben
- 28 mondatpár, egy-egy mondat több változata:
  - eredeti, eredeti módosítva, szintetizált, szintetizált módosítva
- 26 tesztelő, el kellett dönteniük, hogy melyik változat „természetesebb” (1. jobb / 2. jobb / egyformák)
- Mit akartunk megtudni?
  - Vajon a megváltoztatott prozódia észlelhető-e?
  - Ha igen, elviselhető-e a változás?



# Kísérletek eredménye

- Az eredeti változat jobb, mint a módosított
  - Különböző referenciával módosított változatok között hallható a különbség, minőségük hasonló
- ➔ sikerült elérni a prozódiai változatosságot



# Elért eredmények összefoglalása

- Szakirodalom áttekintése (összesen 17 tétel, ebből 5 magyar, 12 angol)
- Mondatdallamok összehasonlítása idővetemítéssel
- Mondatok dallamának cseréje
- Meghallgatásos vizsgálat 26 tesztelővel
- Bebizonyosodott, hogy a prozódiai változatosság megvalósítható



# Továbbfejlesztési lehetőségek

- „hasonló” mondatok keresése pontosabban
- nagyobb beszéd-adatbázis:
  - NAVA (Nemzeti Audiovizuális Archívum)
  - NDA (Nemzeti Digitális Adattár)
- prozódiai változatosságért felelős modul beépítése a Profivox-ba



Köszönöm a figyelmet!

Csapó Tamás Gábor  
csapszi@sch.bme.hu

