

Szintetizált beszéd prozódiai változatosságának növelése spontán beszéd alapján

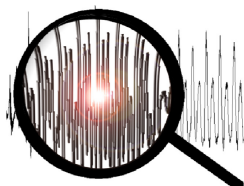
Csapó Tamás Gábor,
Zainkó Csaba, Németh Géza

{csapot,zainko,nemeth}@tmit.bme.hu

BME Távközlési- és Médiainformatikai Tanszék



Beszédkutatás, 2009. október 16.



Tartalom

Cél: spontán jellegű gépi beszéd,
amely nem zavarja a megértést

- Spontán beszéd és a szintetizátorok
- Felhasznált beszédadatbázis
- Alkalmazott beszéd szintetizátor
- Spontán jellegű szintetizált beszéd

Háttér: spontán beszéd

Különbségek a felolvasott beszédhez képest

- F0 menet és változatossága
- Szünetek (csendes, kitöltött)
- Prozódia és szintaktika kapcsolata lazább
- Tagolás
- Hezitálások, megakadásjelenségek
- Szerkesztettség hiánya
- Redundancia, Hiány
- Hang redukció
- Egyebek

Csapó, Zainkó, Németh: Szintetizált beszéd prozódiai változatosságának növelése spontán beszéd alapján

Háttér: prozódiai változatosság

- 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

Csapó, Zainkó, Németh: Szintetizált beszéd prozódiai változatosságának növelése spontán beszéd alapján

Háttér: felhasznált adatbázis

- EKR, rádió felvételei
- Spontán jellegű beszédrészletek
- Folyamatos beszéd, dialógusok, interjú
- 2650 „mondat”, sok beszélő
- Automatikus feldolgozás
 - annotáció alapján
- Kitöltött szünetek - statisztika
 - Levegővétel – 95%
 - Hümmögések – a mondatok 20%-ában
 - Nevetés – 2%

Csapó, Zainkó, Németh: Szintetizált beszéd prozódiai változatosságának növelése spontán beszéd alapján

Háttér: korpusz alapú TTS

- Korpusz alapú általános elemkiválasztásos beszéd szintetizátor
 - Női bemondó,
 - Felolvasott beszéd adatbázis
 - 9 300 mondat, 144 000 szó



Csapó, Zainkó, Németh: Szintetizált beszéd prozódiai változatosságának növelése spontán beszéd alapján

Módszer: Spontán jellegű beszéd előállítása

- Mondatok újratagolása
- Normalizált F0 menet
 - „Hasonló” intonációs egységek keresése az F0-adatbázisból
 - Több lehetséges jelölt
 - Egy kiválasztása „véletlenszerűen”
- Szintetizálás a korpuszos TTS-sel
- Szünetek kitöltése

Csapó, Zainkó, Németh: Szintetizált beszéd prozódiai változatosságának növelése spontán beszéd alapján

Eredmény: szintetizált beszédváltozatok

- Előállított beszédrészlet
 - Eredeti általános szintetizálás 
 - Spontán jellegű gépi beszéd 

Csapó, Zainkó, Németh: Szintetizált beszéd prozódiai változatosságának növelése spontán beszéd alapján

Összefoglalás

□ Általános szintetizátor

- Korlátozottan spontán jellegű beszédet tudunk létrehozni
- Nagy spontán adatbázist igényel

□ Alkalmazási lehetőségek

- Természetesebb ember-gép kapcsolat

Köszönöm a figyelmet!

Csapó Tamás Gábor, Zainkó Csaba, Németh Géza
{csapot,zainko,nemeth}@tmit.bme.hu

Csapó, Zainkó, Németh: Szintetizált beszéd prozódiai
változatosságának növelése spontán beszéd alapján