

Sinteza hrvatskog govora uporabom sustava Festival

Renato Šoić

Uvod

- ▶ sinteza govora
- ▶ sustavi za sintezu govora
- ▶ korištene tehnologije
- ▶ modeliranje glasa u Festivalu
- ▶ rezultat i moguća poboljšanja



Sinteza govora

- ▶ postupak umjetnog generiranja ljudskoga govora
 - ▶ razvoj započeo još u 18. stoljeću
 - ▶ od sredine 20. stoljeća usmjeren na računala
- ▶ primjena
 - ▶ telekomunikacijski servisi, pomoć osobama s poteškoćama, učenje stranih jezika, igre, razni uređaji, komunikacija čovjeka i računala



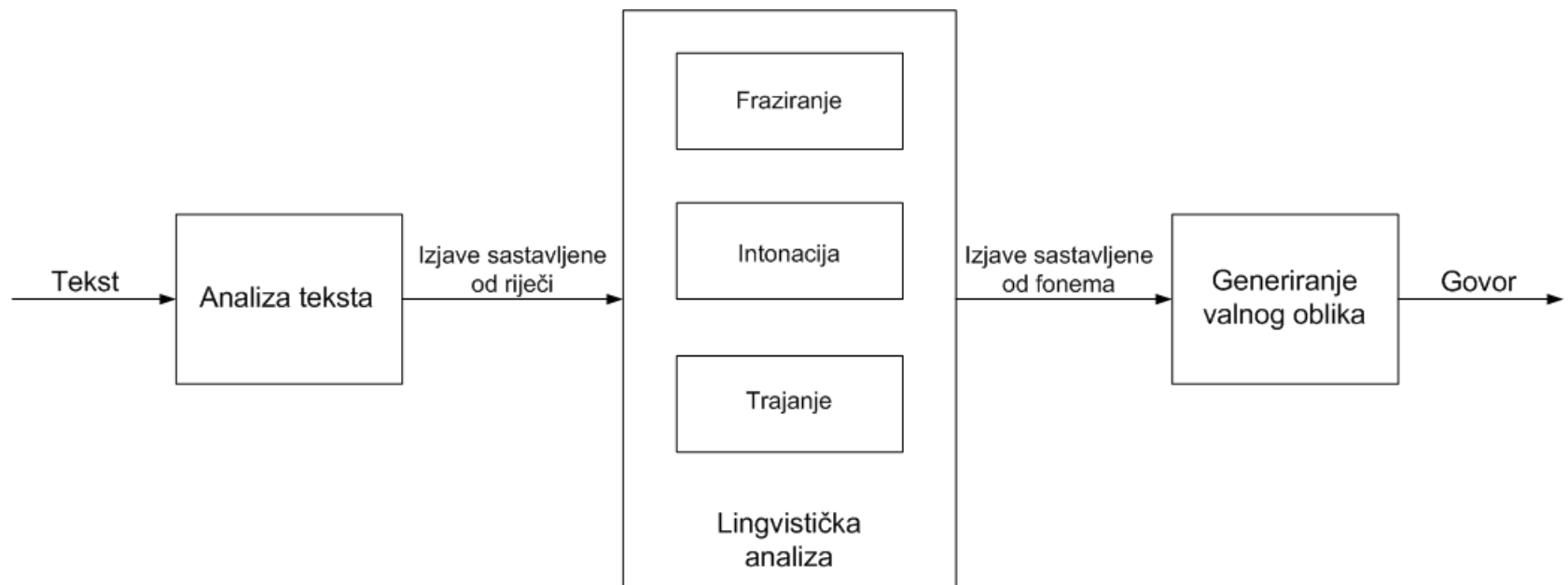
Metode sinteze govora

- ▶ sinteza govora ulančavanjem
 - ▶ govor se reproducira ulančavanjem snimljenih segmenata prirodnog govora
 - ▶ diphone, unit selection, domain-specific
- ▶ tvorba metodom formantata
 - ▶ propuštanje pobudnih signala kroz linearni filter podešen prema rezonantnim frekvencijama ljudskog glasa
- ▶ artikulacijski model tvorbe
 - ▶ temelji se na modelu ljudskog govornog trakta



Sustavi za sintezu govora

- ▶ tri osnovne komponente:
 - ▶ analiza teksta
 - ▶ lingvistička analiza
 - ▶ generiranje valnog oblika



Analiza teksta

- ▶ **podjela ulaznog teksta na manje dijelove(izjave)**
 - ▶ izjave su uglavnom rečenice
 - ▶ omogućavaju da se govor sintetizira po dijelovima
- ▶ **normalizacija teksta**
 - ▶ segmentacija riječi
 - ▶ pronalaženje tokena i zamjena riječima
 - ▶ prevođenje skraćenica u neskraćeni oblik
- ▶ **primjer**
 - ▶ Prema izjavi g. Perića, iz banke je nestalo 12 milijuna kn.



Lingvistička analiza

- ▶ pronalaženje izgovora riječi i dodjeljivanje prozodijske strukture izgovoru
- ▶ svakoj riječi dodjeljuje se fonetska transkripcija
- ▶ postupak prozodijskog fraziranja
- ▶ izgrađuju se modeli trajanja i intonacije



Generiranje valnog oblika

- ▶ ulazni podaci – izjave sastavljene od fonema
- ▶ postupak stvaranja valnog oblika ovisi o korištenoj metodi sinteze govora
- ▶ **Statistical Parametric Synthesis**
 - ▶ temelji se na korištenju skrivenih Markovljevih modela
 - ▶ dobri rezultati s malim bazama podataka



Korištene tehnologije

▶ Festival i pomoćni alati

- ▶ najpoznatiji open source sustav za sintezu govora
- ▶ razvoj glasa neovisan o jeziku
- ▶ podržava više metoda sinteze govora
- ▶ konzolno sučelje, naredbe i parametri glasova u jeziku Scheme
- ▶ FestVox – alati za razvoj novih glasova

▶ Audacity

- ▶ postupak snimanja govora



Modeliranje glasa u Festivalu

▶ nekoliko osnovnih koraka:

- ▶ stvaranje tekstualnog i govornog korpusa
- ▶ definiranje skupa fonema
- ▶ pravila obrade tokena (brojevi, kratice...)
- ▶ leksikon i LTS pravila
- ▶ fraziranje, intonacija, trajanje
- ▶ sinteza valnog oblika

▶ specifičnosti hrvatskog jezika

- ▶ složen postupak normalizacije teksta i model intonacije
- ▶ nema potrebe za leksikonom i LTS pravilima



Stvaranje korpusa

- ▶ tekst prikupljen iz različitih izvora
 - ▶ mora zadovoljavati neke uvjete (format, strane riječi, simboli)
 - ▶ 4500 rečenica, 110 000 riječi
- ▶ analiza tekstualnog korpusa
 - ▶ učestalosti pojavljivanja riječi i znakova
 - ▶ pronalaženje prikladnih rečenica, odabir prema fonetskoj zastupljenosti
- ▶ govorni korpus
 - ▶ 270 odabranih prikladnih rečenica
 - ▶ 17 minuta snimljenog materijala



Koraci u izgradnji glasa

▶ Definiranje skupa fonema

- ▶ postupak definiranja traženih značajki za sve foneme u jeziku
 - ▶ vokal ili konsonant
 - ▶ klasifikacija vokala po mjestu tvorbe
 - ▶ tip konsonanta
 - ▶ mjesto artikulacije
 - ▶ naglašenost

▶ označavanje podataka

- ▶ pridruživanje fonema dobivenih iz transkripata s pripadnim zvučnim zapisima



Koraci u izgradnji glasa (nastavak)

- ▶ **izgradnja strukture izjava**
 - ▶ za svaku snimljenu rečenicu stvara se struktura kao da je rečenica pravilno sintetizirana
- ▶ **ekstrakcija parametara fundamentalne frekvencije**
 - ▶ analiza svih pohranjenih rečenica, određivanje granica vrijednosti prema glasu govornika
- ▶ **stvaranje CLUSTERGEN glasa**
 - ▶ stvaranje datoteke s imenima stanja, izgradnja parametarskog modela, izgradnja modela trajanja



Rezultat i moguća poboljšanja

- ▶ razumljiv govor, ali prilično neprirodan
 - ▶ “isprekidanost” govora
 - ▶ pogrešno naglašavanje, brzina izgovora
- ▶ moguća poboljšanja
 - ▶ stvoriti opsežniju, kvalitetniju bazu snimljenog govora
 - ▶ ručna provjera dobivenih granica fonema
 - ▶ normalizacija teksta
 - ▶ unaprijediti modele intonacije i trajanja



Hvala na pažnji!

