UNIVERSITY
OF WEST BOHEMIA

Technology Agency
of the Czech Republic

On the Impact of Labialization Contexts on Unit Selection Speech Synthesis

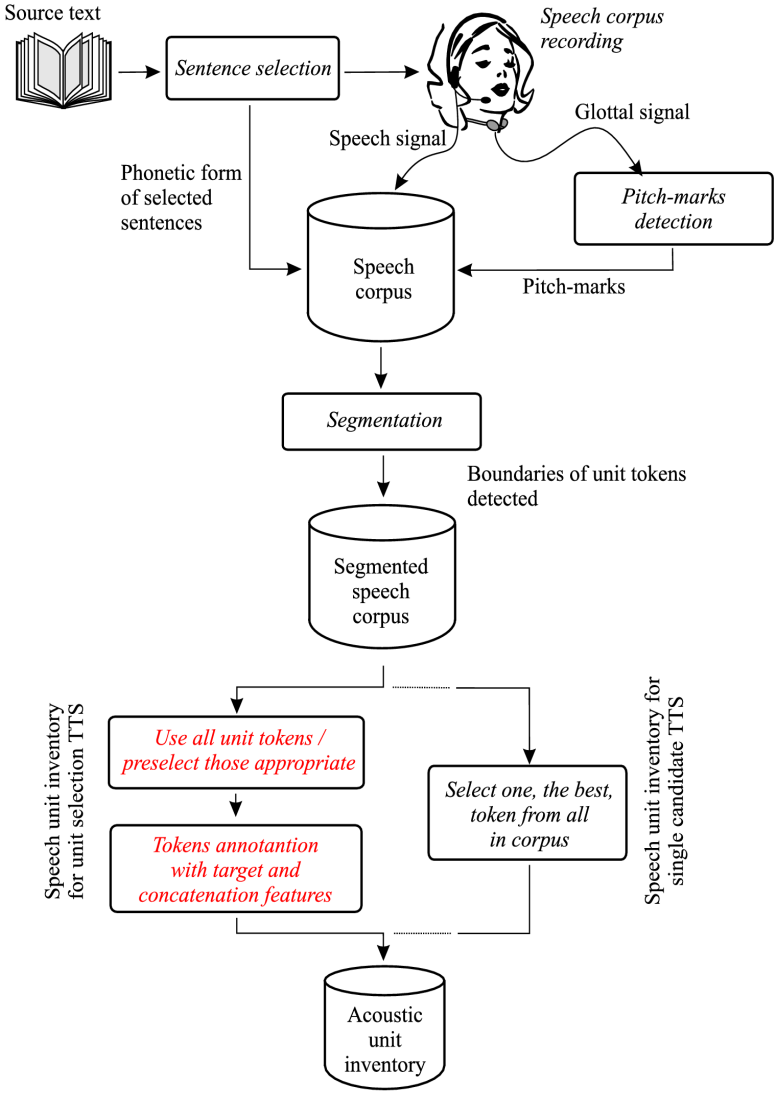
Daniel Tihelka, Zdeněk Hanzlíček, Jindřich Matoušek
Department of Cybernetics, University of West Bohemia in Pilsen, Czech Republic

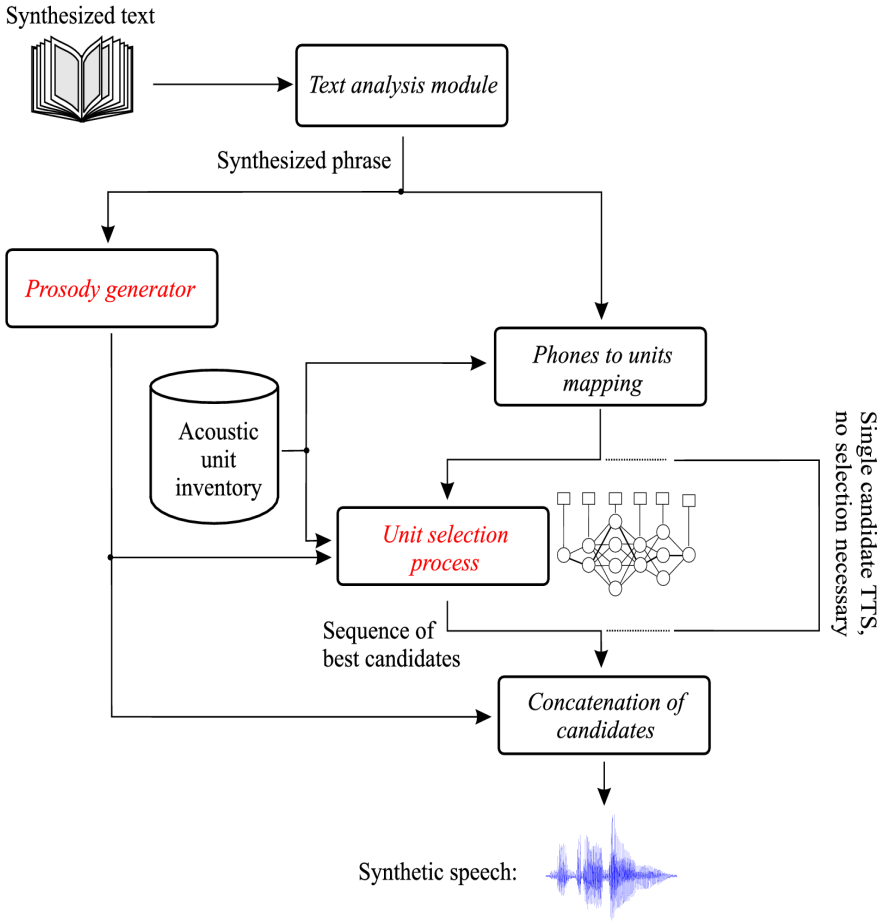
Pavel Machač, Radek Skarnitzl
Institute of Phonetics, Charles University in Prague, Czech Republic

Outline

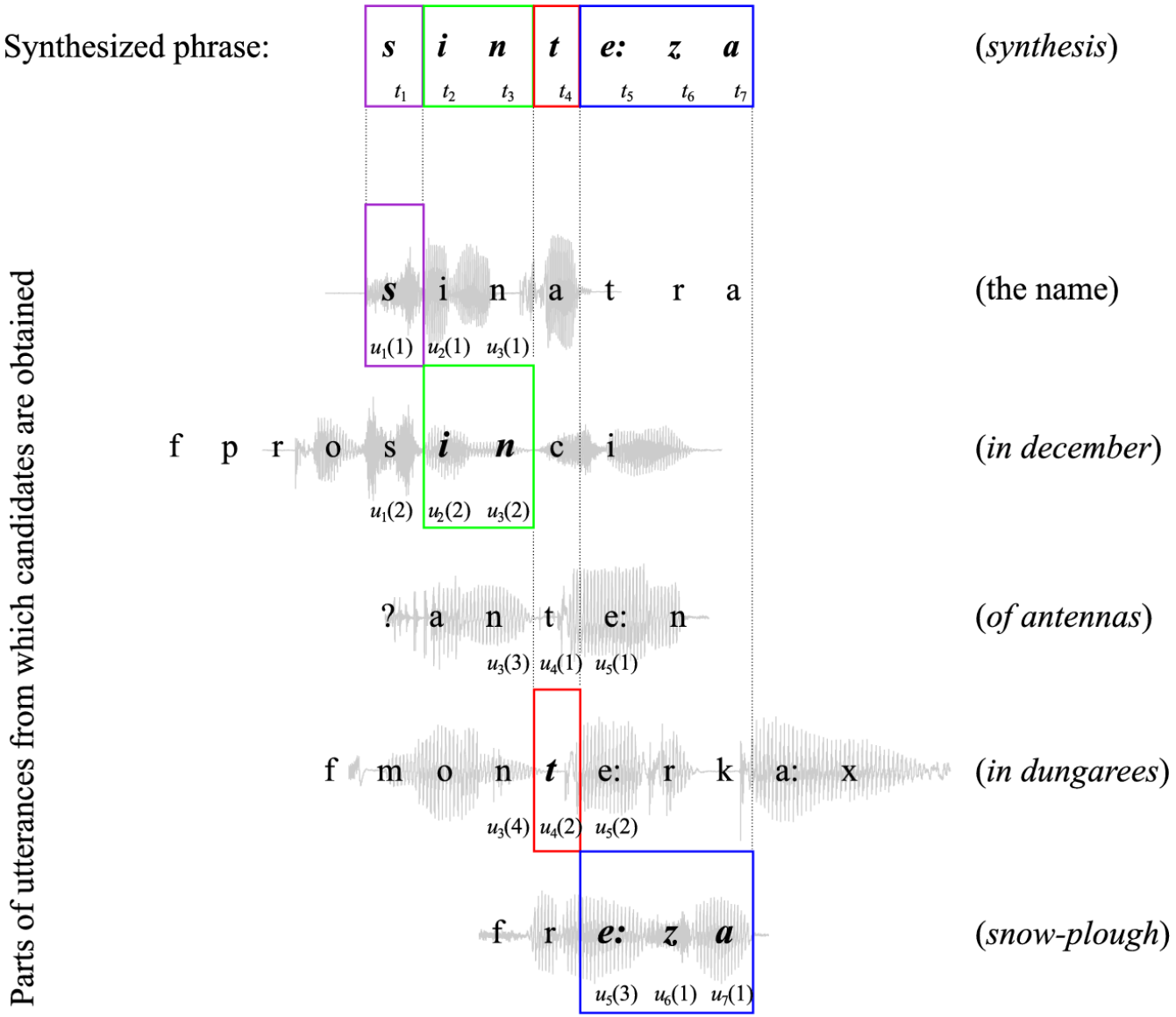
- Unit selection speech synthesis
- Coarticulatory labialization
- Experiments and results
- Conclusion
- Speech samples

Concatenative TTS system





Unit selection scheme



Unit selection

- each speech unit described by:
  - description / context of its original occurence
    - neighbouring (preceding and following) speech units
    - position in sentence / phrase / word / syllable
    - type of phrase, accent
  - its acoustical properties
    - spectral parameters (MFCC, LSP, formant frequencies) fundamental frequency, energy
- selection criteria:
  - **target cost** – selected unit should originate from similar context
  - **join cost** – neighbouring units should be smoothly concatenated (no abrupt changes in acoustical properties)
  - optimizing both target and join cost through the whole sequence of units – dynamic programming
  - crucial problem – setting the proper weighs for particular costs

Labialization

- **coarticulation**
  - mutual influencing of neighbouring speechsounds during speech production
  - has to be considered during unit selection and concatenation
- **coarticulary labialization**
  - inherent phonetic feature of back vowels
  - lowers formant frequencies F_2 and partly F_3
  - influences phonetic makeup of neighbouring consonants

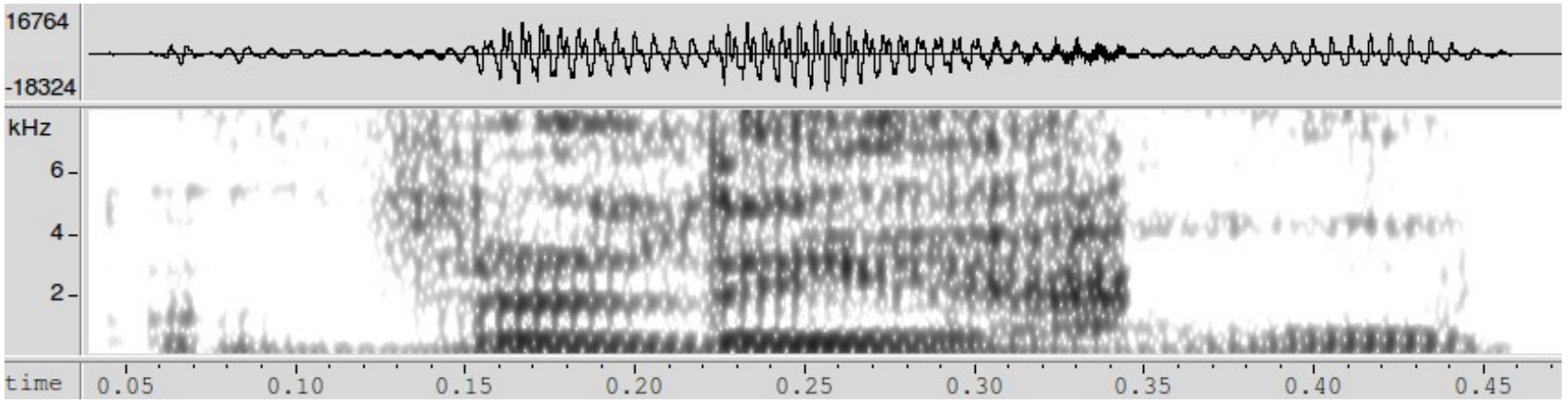
Labialization

- examples for voiced laryngeal fricative ***h*** (high probability of coarticulatory effects)

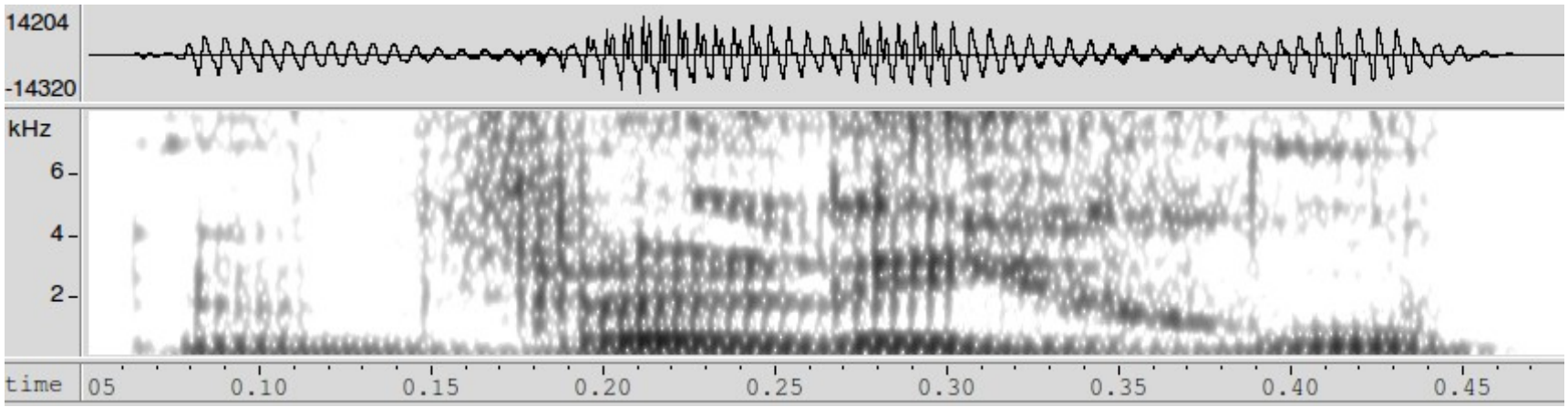
	labial. combination	example	transcription	diphones
1	V(lab0) - C - V(lab0)	v**yhy**nul	v**ihi**nul	#v **vi ih hi in** nu ul l#
2	V(lab+) - C - V(lab+)	v kr**uhu**	fkr**uhu**	#f fk kr **ru uh hu u#**
3	V(lab0) - C - V(lab+)	v l**ihu**	vl**ihu**	#v vl **li ih hu u#**
4	V(lab+) - C - V(lab0)	v kr**uhy**	fkr**uhi**	#f fk kr **ru uh hi i#**

- each diphone could be influenced by the labialization of both left and right neighbouring unit
- labialization context within speech synthesis:
- respected or violated
- partly (one side) or fully (both sides)

Labialization



violated labialization



respected labialization

Experiments

- running TTS-system without considering labialization
  - 5 000 utterances ~ 195 964 diphones

Context	Mismatched phone context	Mismatched labial context
left	23.9 %	4.3 %
right	23.8 %	4.1 %
both	5.6 %	---

Experiments

- 2 preference listening tests (pairwise comparison)
- for non-phoneticians
  - simplified setup - pairs with fully respected (both sides) and fully violated labialization only
  - participants: 19 phonetic laymen
  - 40 queries
- for phoneticians
  - complex setup – all possible combinations of fully and partly (one side) respected and violated labialization
  - participants: 8 students of phonetics
  - 112 queries

Experiments

- test for non-phoneticians
  - preference respected labial context 55.4 %
  - preference for violated labial context 19.5 %
  - no preference 25.1 %

- test for phoneticians
  - preference respected labial context 74.2 %
  - preference for violated labial context 10.6 %
  - no preference 15.2 %

  - other combinations of partly/fully respected/violated labialization – respected labialization always preferred (detailed results in article)

- consistency of rating
  - non-phoneticians 75,2 (+ 17,3) %
  - phoneticians 83.9 (+ 12.5) %

Conclusion

- importance of considering coarticulatory labialization was confirmed
- future work
  - more detailed study for particular phones in specific contexts
  - incorporating labialization feature into TTS system (modify target cost, specify new weighs in unit selection criterion...)
  - other coarticulation-related features of speech (e.g.nasalization)

Speech samples

Respected labialization

Violated labialization

UNIVERSITY
OF WEST BOHEMIA

Technology Agency
of the Czech Republic

Thank you for your attention!