

## Uitspraakevaluatie & training met behulp van spraaktechnologie

### Pronunciation assessment & training by means of speech technology

Helmer Strik – and many others  
Centre for Language and Speech Technology (CLST)  
Radboud University Nijmegen, the Netherlands

### Context

'Deviant' pronunciation (e.g., pathology, non-natives) & speech technology (applications) :

- Assessment
  - Diagnosis, monitoring
- Training (therapy, learning)
  - Speaking & listening; reading aloud
- AAC (Augmentative & Alternative Communication)
  - Improve communication

### Our research

#### ➤ Past:

- Fluency assessment - Temporal measures
- CAPT: Computer Assisted Pronunciation Training
- Pronunciation error detection
- Recognition of dysarthric speech

#### ➤ Current, future:

- OSTT: Ontwikkelcentrum voor Spraak- en Taaltechnologie ten behoeve van Spraak- en Taalpathologie en Revalidatietechnologie
- Training & error detection, not only pronunciation, but also other (e.g. morpho-syntactic) aspects

### Our research

#### ➤ Past:

- Fluency assessment - Temporal measures
- **CAPT: Computer Assisted Pronunciation Training**
- **Pronunciation error detection**
- Recognition of dysarthric speech

#### ➤ Current, future:

- OSTT: Ontwikkelcentrum voor Spraak- en Taaltechnologie ten behoeve van Spraak- en Taalpathologie en Revalidatietechnologie
- Training & error detection, not only pronunciation, but also other (e.g. morpho-syntactic) aspects

### CAPT: Computer Assisted Pronunciation Training

Pronunciation errors – detected automatically by means of Automatic Speech Recognition (ASR) → feedback

- Question: ASR-based CAPT: Is it effective?
- Goal: To study the effectiveness and possible advantages of ASR-based CAPT
- Target users :
  - Adult learners of Dutch with different L1's
- Pedagogical goal :
  - Improving segmental quality in pronunciation

### Dutch CAPT: feedback

Content: focus on problematic phonemes

#### Criteria

1. Common across speakers of various L1's
2. Perceptually salient
3. Frequent
4. Persistent
5. Robust for automatic detection (ASR)

#### Result:

11 'targeted phonemes': 9 vowels and 2 consonants

## 11 'targeted phonemes'

IPA symbol	example
/ɔ:/	toch, Scheveningen
/h/	hand, Helmer
/a/	pat
/a:/	naam
/i/	pit
/ɛ:/	put
/y/	vuur
/u/	voer
/ø:/	deur
/ɛi/	fijn
/œy/	huis

Leuven, 28-04-2007

7

Radboud University Nijmegen



## Video (from Nieuwe Buren)



Leuven, 28-04-2007

8

Radboud University Nijmegen



**Sounds** Listen to the whole dialogue first. Then click on the record-button and record your role: you are the female speaker. Remember that if you are the first speaker, you must start speaking immediately.

Play the whole dialogue

Ik heb iets voor je gekocht. Hier.

Dank je wel.

Het staat fantastisch. Het staat goed bij je broek.

Vind je echt?

Ja, natuurlijk.

You had problems with the red sound(s). Listen again to the example and try again.

Ik heb iets voor je gekocht hier.

## Video: dialogue

Leuven, 28-04-2007

10

Radboud University Nijmegen



**Sounds** Click on the buttons to listen to the answers. Choose one answer and record it by clicking on the record-button.

Waarom trekt Tom zijn oude overhemd uit?

omdat karin zegt dat de mouw vies is en dat er een knoop af is.

omdat het binnen erg warm is.

omdat hij zijn nieuwe overhemd wil aantrekken.

Max. 3 times

You still have some problems with the red sound(s). Try the next exercise: you will have more opportunities to practise with sounds.

omdat karin zegt dat de mouw vies is en dat er een knoop af is

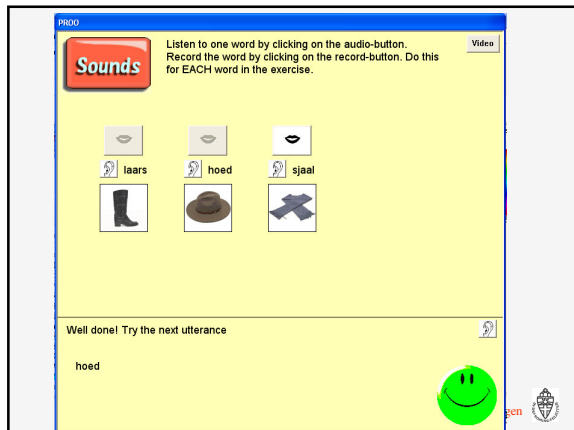
**Sounds** Listen to one word by clicking on its audio-button. Record the word by clicking on the record-button. Do this for EACH word in the exercise.

bos stuk

boos stug

Well done! Try the next utterance.

stuk



## Experiment: participants & training

Regular teacher-fronted lessons: 4-6 hrs per week

- a) Experimental group (EXP): n=15 (10 F, 5 M)  
Dutch CAPT
- b) Control group 1 (NiBu): n=10 (4 F, 6 M)  
reduced version of Nieuwe Buren
- c) Control group 2 (noXT): n=5 (3 F, 2 M)  
no extra training

Extra training: 4 weeks x 1 session 30' – 60'  
1 class – 1 type of training

Leuven, 28-04-2007

14

Radboud University Nijmegen



## Experiment: testing

3 analyses:

- 1. Participants' evaluations: questionnaires on system's usability, accessibility, usefulness etc.
- 2. Global segmental quality: 6 experts rated stimuli on 10-point scale (pretest/posttest, phonetically balanced sentences)
- 3. In-depth analysis of segmental errors: expert annotations

Leuven, 28-04-2007

15

Radboud University Nijmegen



## Results: participants' evaluations

Positive reactions

Enjoyed working with the system

Believed in the usefulness of the system

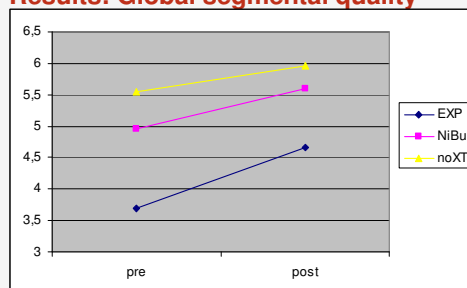
Leuven, 28-04-2007

16

Radboud University Nijmegen



## Results: Global segmental quality



All 3 groups improve (mean improvement)  
EXP improved most

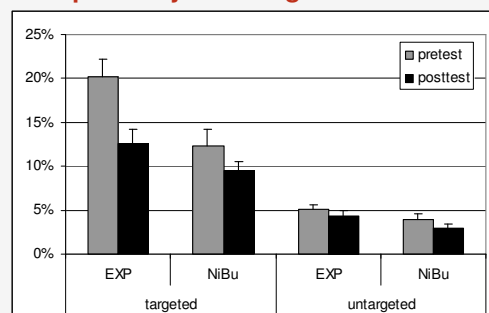
Leuven, 28-04-2007

17

Radboud University Nijmegen



## In-depth analysis of segmental errors



Leuven, 28-04-2007

18

Radboud University Nijmegen



## Conclusions

- Goal: To study the effectiveness and possible advantages of ASR-based CAPT
- Question: ASR-based CAPT: Is it effective?  
Answer: Yes! It is effective in improving the pronunciation of targeted phonemes.
- Advantages :  
ASR-based CAPT can provide automatic, instantaneous, individual feedback on pronunciation in a private environment.

Leuven, 28-04-2007

19



## Video: pronouncing words

Leuven, 28-04-2007

20



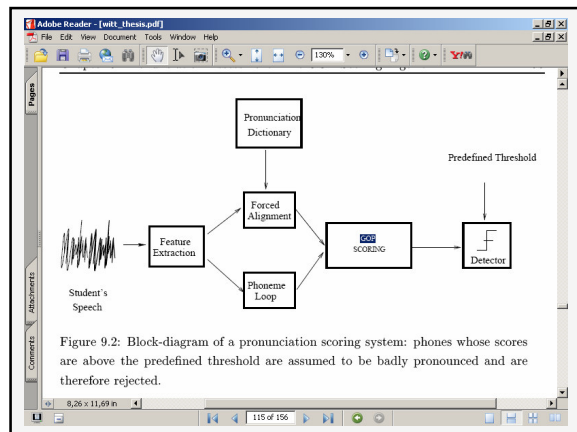
## Error detection

Detection of pronunciation errors

- Goodness Of Pronunciation (GOP)
  - Silke Witt & Steve Young
  - Acoustic-phonetic features (APF)
    - Khiet Truong et al.
- Goal: improve error detection

Leuven, 28-04-2007

21



## Goodness Of Pronunciation (GOP): Accuracy

15 participants  
2174 target phones

	Accept	Reject	Total
Correct	CA: 59.5%	CR: 26.5%	C: 86.0%
False	FA: 9.2%	FR: 4.8%	F: 14.0%

Leuven, 28-04-2007

23



## Acoustic-phonetic features (APF)

Selection of segmental pronunciation errors:

/A/ mispronounced as /a:/ (man - maan)

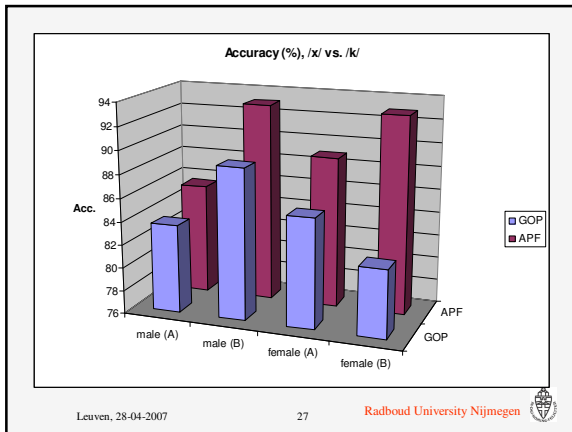
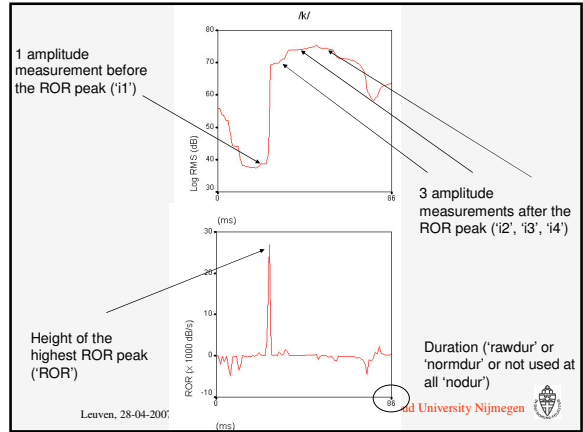
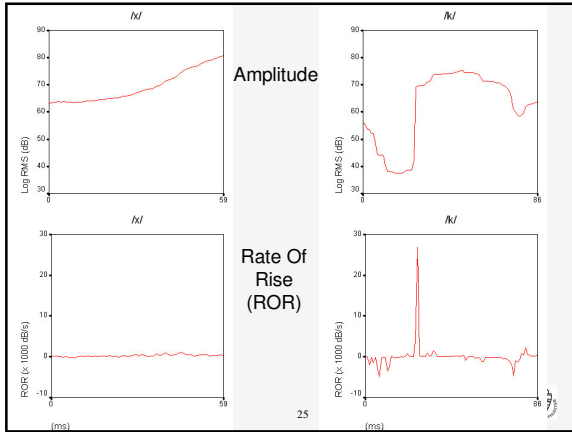
/Y/ mispronounced as /u/ or /y/ (tut - toet or tuut)

/x/ mispronounced as /k/ or /g/ (gat - kat or /g/at)

Leuven, 28-04-2007

24





### Error detection

Goodness Of Pronunciation (GOP):

- One general method for all sounds
- Error specific knowledge is not used

Acoustic-phonetic features (APF)

- Error specific knowledge is used
- Works well
- How to generalize? (artic. + other features)

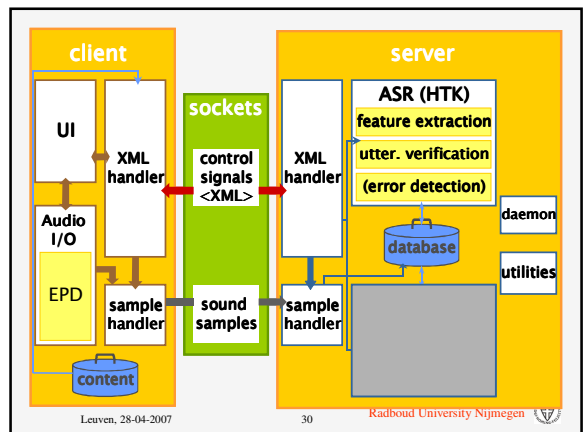
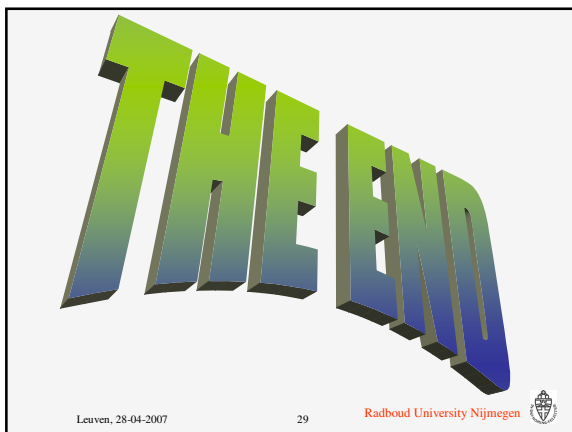
Combination?

Other approaches, e.g. post. prob's (ANN)?

Leuven, 28-04-2007

28

Radboud University Nijmegen



## Dutch CAPT

Gender-specific, Dutch & English version.

- 4 units, each containing:
  - 1 video (from Nieuwe Buren) with real-life + amusing situations
  - + ca. 30 exercises based on video: dialogues, question-answer, minimal pairs, word repetition

Sequential, constrained navigation: min. one attempt needed to proceed to next exercise, maximum 3

Leuven, 28-04-2007

31

Radboud University Nijmegen



## Results: reliability global ratings

Cronbach's  $\alpha$ :

Intrarater: 0.94 – 1.00

Interrater: 0.83 - 0.96

Leuven, 28-04-2007

32

Radboud University Nijmegen



L1	Training group			Total
	EXP	NiBu	NoXt	
Arabic	6			6
Bengali		1		1
Catalan		2		2
English	1	1	1	3
German		1		1
Greek		2		2
Hebrew	1			1
Italian		1	1	2
Lithuanian			1	1
Polish	2	1	2	5
Russian	1			1
Spanish	1			1
Swedish	1			1
Turkish	2			2
Ukrainian		1		1
<b>Total</b>	<b>15</b>	<b>10</b>	<b>5</b>	<b>30</b>

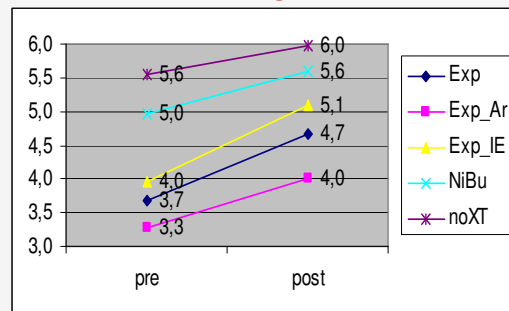
Leuven, 28-04-2007

33

Radboud University Nijmegen



## Results: Global ratings



Leuven, 28-04-2007

34

Radboud University Nijmegen



## Possible improvements

- Increase sample size (more participants)
- Increase training intensity (more training)
- Match training groups: L1's, proficiency, etc.
- Give feedback on more phonemes
- More targeted systems for fixed L1-L2 pairs.
- Give feedback on suprasegmentals
- Improve error detection?

Leuven, 28-04-2007

35

Radboud University Nijmegen



## Error detection

Pronunciation errors

- 11 'problematic sounds': 9 V + 2 C
- Goal: give feedback on more sounds

Morpho-syntactic errors

- maak / maakt / maken
  - o Ik maak
  - o Hij/zij maakt
  - o Wij maken
- Goal: also give feedback on morpho-syntactic aspects

Leuven, 28-04-2007

36

Radboud University Nijmegen



## Goodness Of Pronunciation (GOP)

GOP has been applied in the exp. system.  
The exp. system was effective.

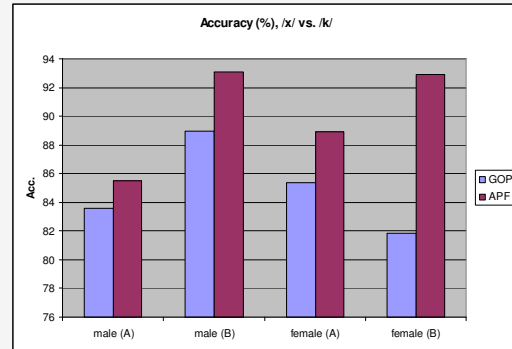
Evaluate GOP

- Correct vs. errors
- Patterns
- Pros & cons
- Improve

Leuven, 28-04-2007

37

Radboud University Nijmegen



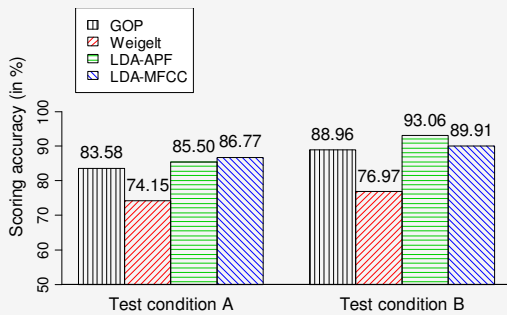
Leuven, 28-04-2007

38

Radboud University Nijmegen



## Results /x/ vs /k/, male speakers



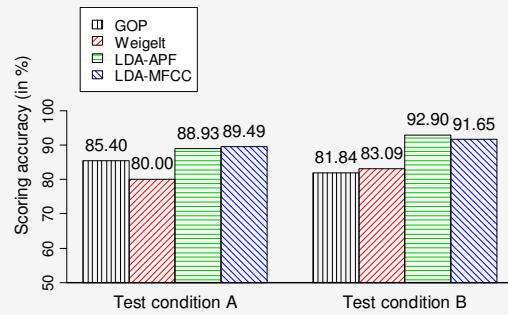
Leuven, 28-04-2007

39

Radboud University Nijmegen



## Results /x/ vs /k/, female speakers



Leuven, 28-04-2007

40

Radboud University Nijmegen



## Results method II (LDA)

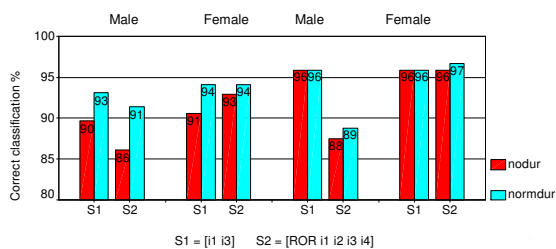
/x/ vs /k/

Training = DL2N1-Nat

Test = DL2N1-Nat

Training = DL2N1-NN

Test = DL2N1-NN



## 11 'targeted phonemes'

/ɔ/, /ɔ̃/, /ɑ/, /y/, /œy/, /a/, /ɛi/, /h/, /u/, /ø/, /i/

Leuven, 28-04-2007

42

Radboud University Nijmegen

