



Københavns Universitet



Inter-coder reliability and the effects of coder background and linguistic structure.

Thøgersen, Jacob; Heegård, Jan

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Thøgersen, J., & Heegård, J. (2013). Inter-coder reliability and the effects of coder background and linguistic structure.. Poster session presented at ExAPP 2013, København, Denmark.

Inter-coder reliability and the effects of coder background and linguistic structure

Jacob Thøgersen, University of Copenhagen
Jan Heegård, Copenhagen Business School

1. Research question

- Phonetic coding and transcriptions are not objective.
- Acoustic structure plays a significant role for the reliability of a transcription: some structures are easier to differentiate than others.
- Coders' disciplinary backgrounds affect what they hear.
- We propose
 - Also *semantics*, i.e. phonological contrast, must be considered.
 - Listener (dis)agreement correlates with level of difficulty in perceptability.
 - Reduction phenomena should be view as graded: listener disagreement reveals intermediate stages in word forms' pronunciation.

2. Variables

+Semantic		-Semantic
+Easy	-Easy	
/-də/	/-ədə/	/-ŋə/

Semantic contrast

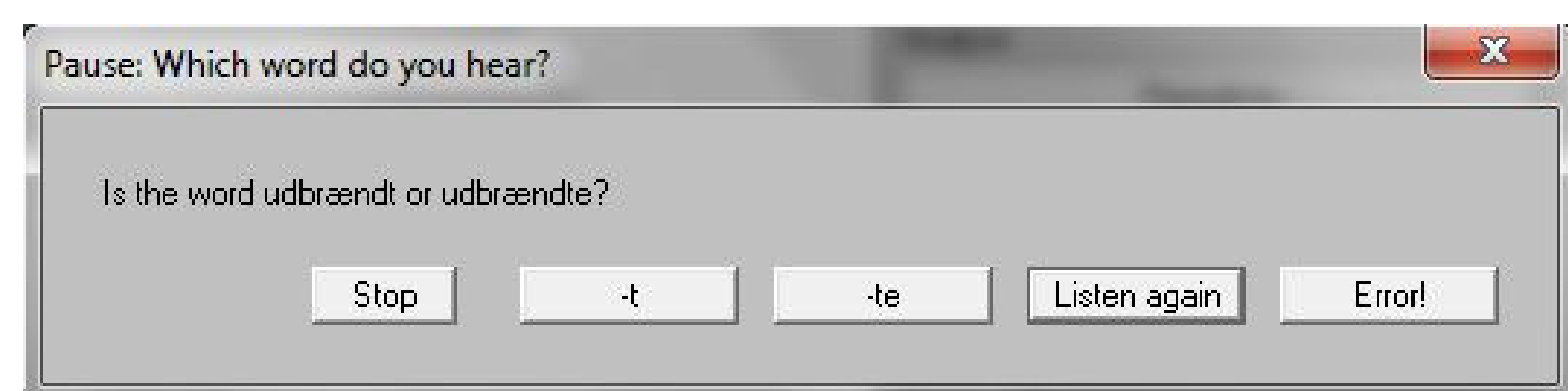
	Long form	Short form
/-də/	brugte 'used' (pret.)/(plur.)/(defin.)	brugt 'used' (partc.)/(sing.)/(indefin.)
/-ədə/	blandede 'mixed' (pret.)/(plur.)/(defin.)	blandet 'mixed' (partc.)/(sing.)/(indefin.)
/-ŋə/	mange 'many' (plur.)/(defin.)	

'Easiness'

	Unreduced	Reduced
/-də/	ɖə	ɖ
/-ədə/	ɔ̃ə/ɔ̃	ɔ̃
/-ŋə/	ŋə	ŋ

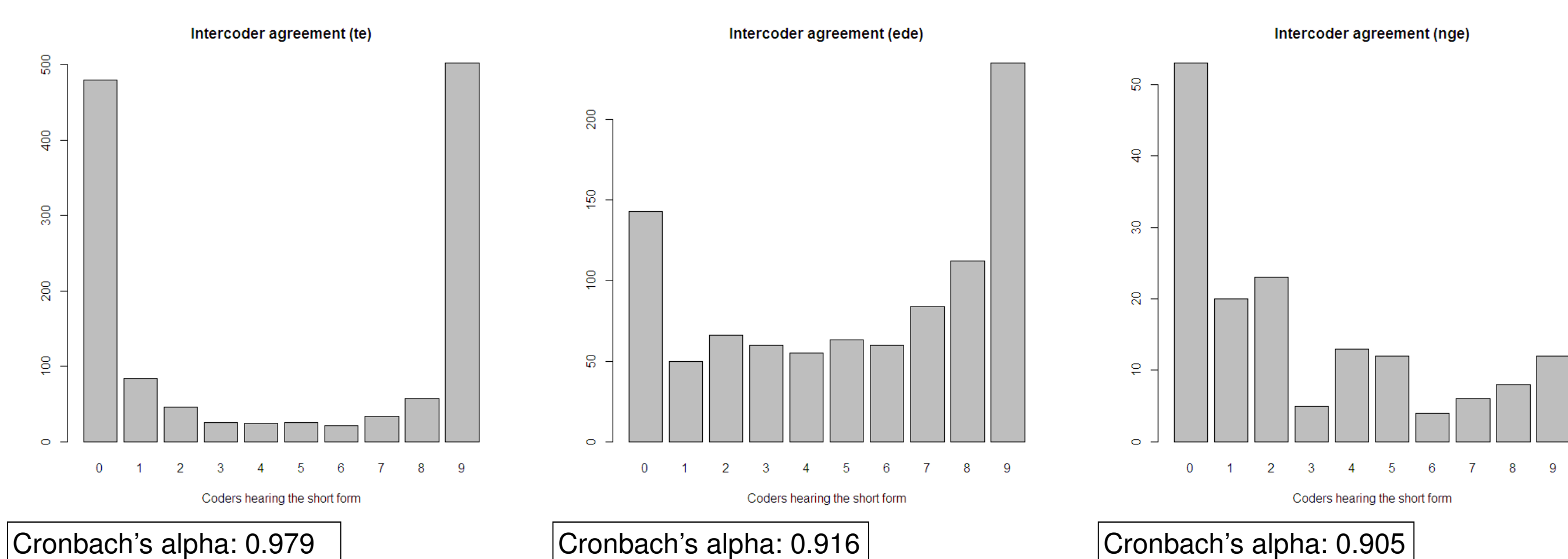
3. Method

- 9 listeners (4 trained linguists, 4 students of other subjects, 1 experimenter with different background).
- 2303 tokens (1289 for /-də/, 872 for /-ədə/, 149 for /-ŋə/).
- Listened to tokens in isolation (through Praat script).
- Task for /-də/ and /-ədə/: Which word is heard (semantic contrast)?
- Task for /-ŋə/: Is the schwa assimilated?
- Perceived short forms = The number of listeners who heard the token in its short form.



4. Results, intercoder agreement

- /-də/ (+Semantic, +Easy): Highest agreement
- /-ədə/ (+Semantic, -Easy): Medium agreement
- /-ŋə/ (-Semantic): Least agreement

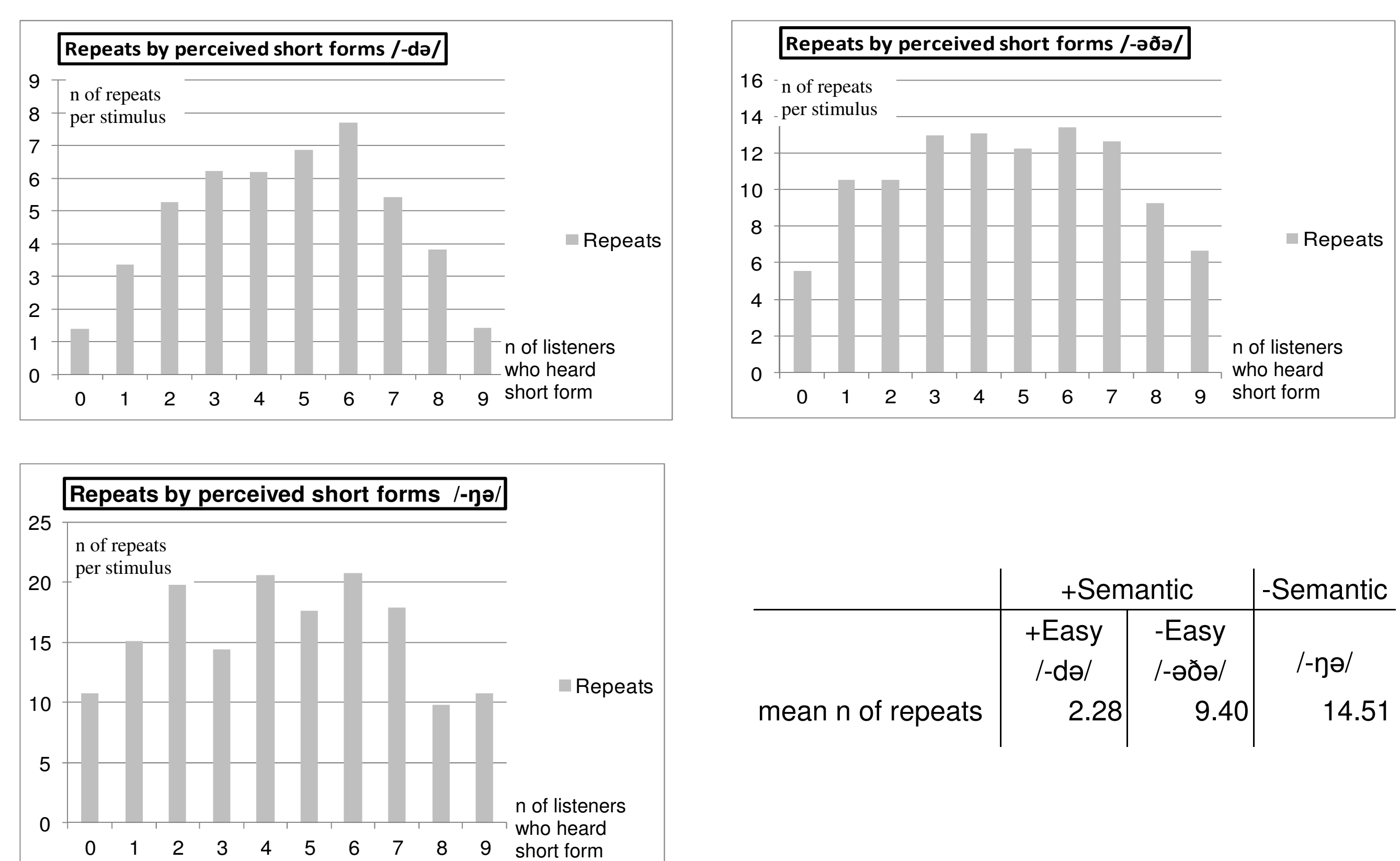


5. Results, listener background

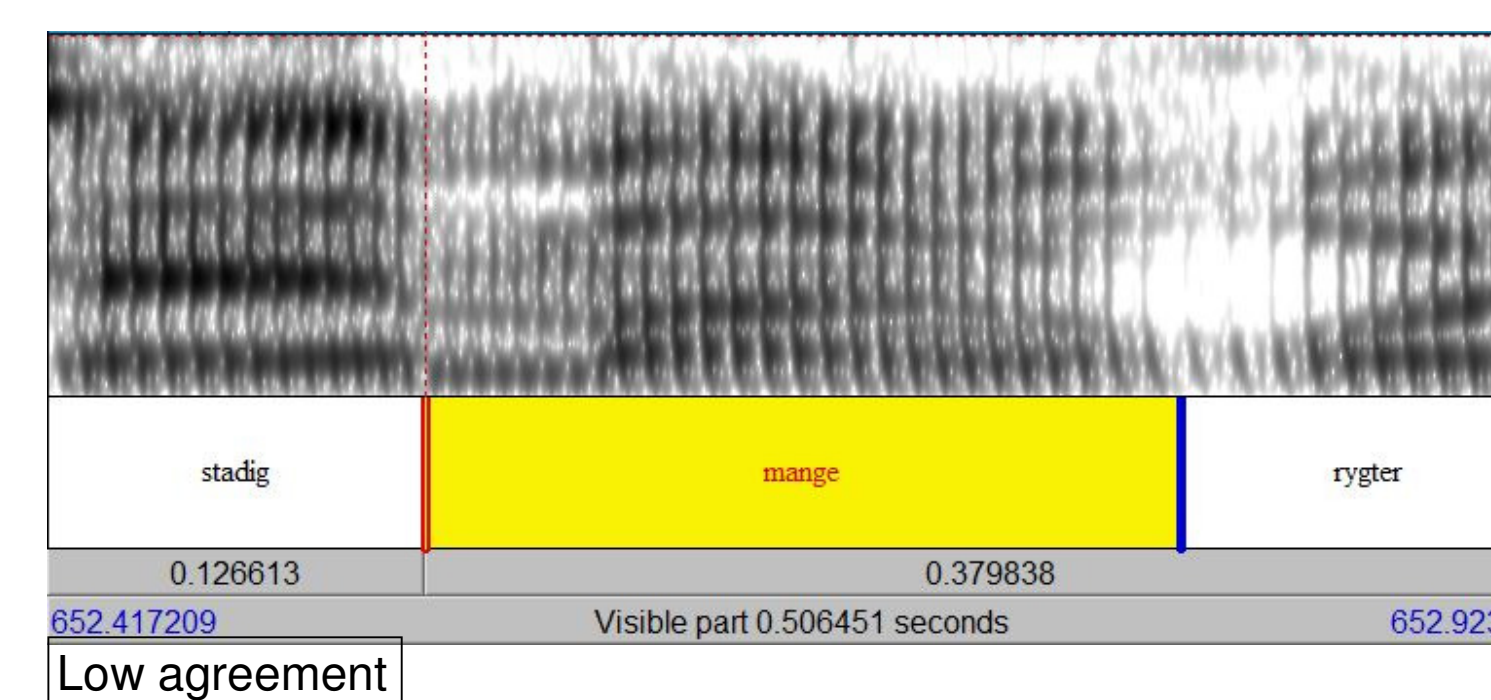
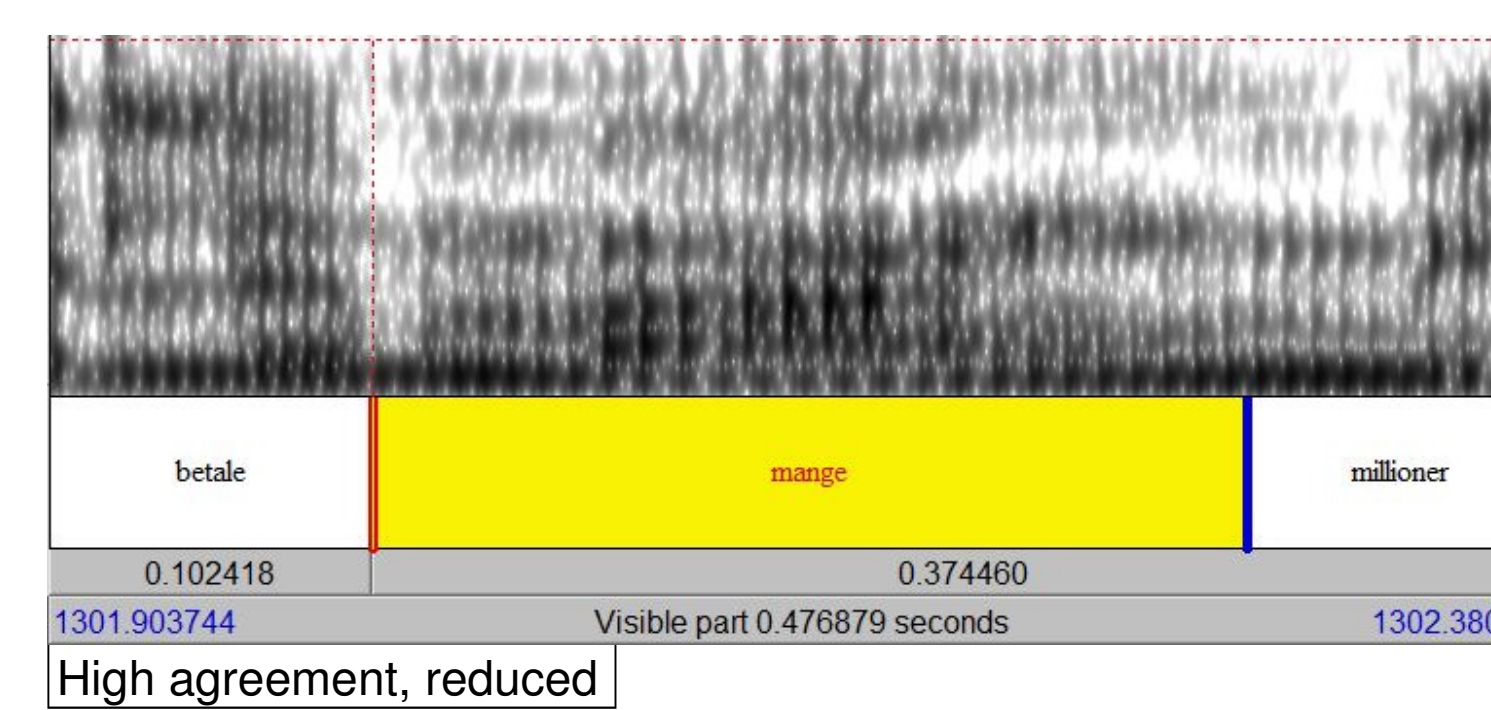
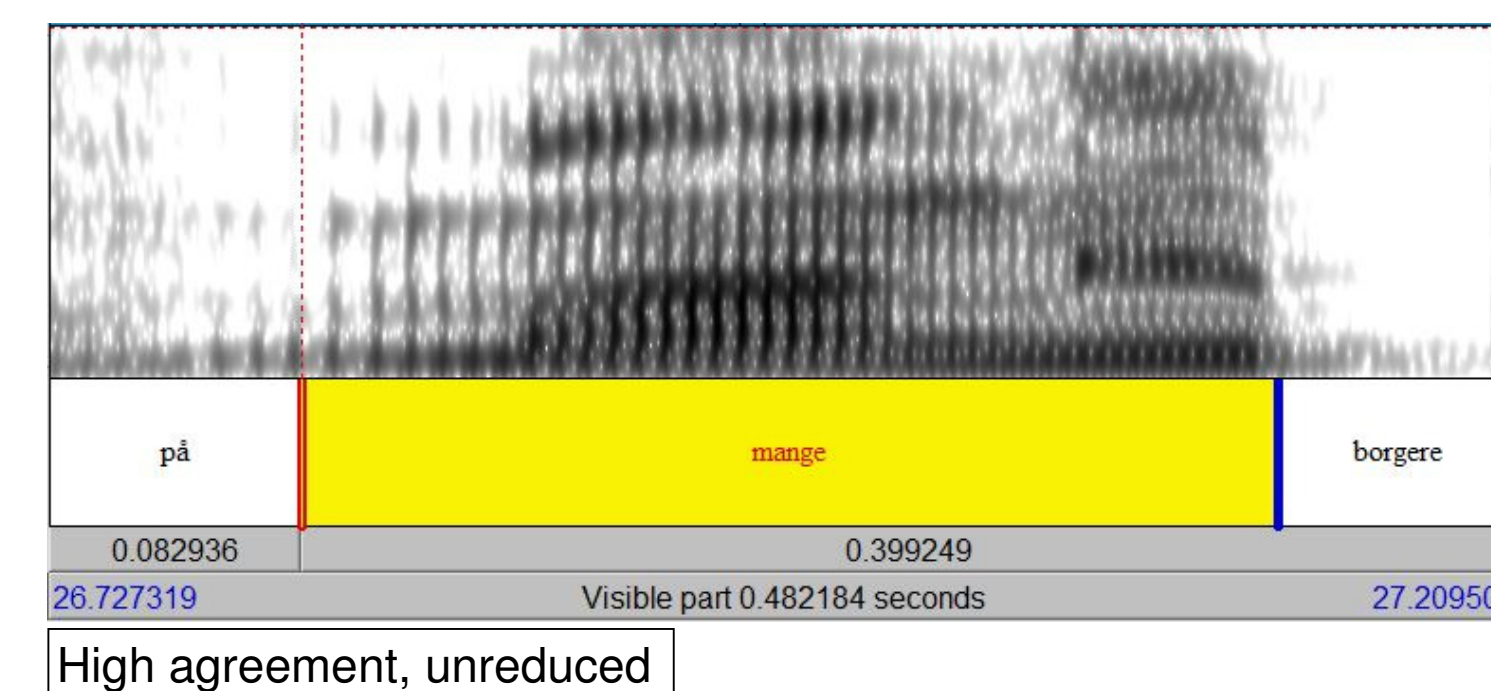
	/-də/		/-ədə/		/-ŋə/	
	Linguists	Non-linguists	Linguists	Non-linguists	Linguists	Non-linguists
Intergroup agreement (Cronbach's alpha)	0.948	0.963	0.867	0.840	0.842	0.784
'Short form' codes	65.4%	52.4%	51.4%	47.4%	31.2%	35.6%
Between group differences	Chi-squared = 123, df = 1, p<0.001		Chi-squared = 16.7, df = 1, p<0.001		Chi-squared = 2.53, df = 1, p=0.112	

6. Results, coder confidence

- Coders could choose to listen again to tokens if they were in doubt.
- The graphs show the relationship between such stimulus repeats and listener (dis)agreement.



7. Results, tokens with high and low agreement



8. Conclusions

- Coding reliability depends on semantic functional load of target. Phonetic contrasts which are distinctive are inherently easier to perceive.
- Some phonetic contrasts are inherently more difficult to perceive.
- Trained linguists have an advantage over lay listeners on some but not all variables. In our study, linguist showed higher agreement only on the more *difficult* variables.
- Intercoder agreement correlates highly with intercoder confidence as measured by repeats of stimulus.
- Intercoder (dis)agreement appears to correlate with segmental ambiguity.
- Intercoder disagreement is not *noise* in our investigation, but evidence against a polar interpretation of reduction phenomena.