

## Mutual Intelligibility between closely related languages

› Sebastian Kürschner  
› Charlotte Gooskens  
LOT course 19 - 23 January 2009

## Introduction

### Mutual Intelligibility between closely related languages

#### Intelligibility:

- › The degree to which a speaker of one variety understands the speech of another closely related variety
- › Can be expressed in a single number

#### Closely related languages:

- › Language varieties (dialects and languages)

### Mutual Intelligibility between closely related languages

#### Assumptions:

- › First confrontation (inherent intelligibility)
- › Spoken language only

### Mutual Intelligibility between closely related languages

#### Similarities to:

- › defective speech
- › speech in noise
- › foreign accents
- › talking machines

### Semiconmunication

- › Haugen (1966)
- › ≈ nonconvergent/asymmetric/bilingual discourse, receptive bilingualism
- › Speakers of different but related languages each speak their own language and still comprehend one another's languages
- › Mutual intelligibility is sometimes imperfect and asymmetric

## Semiconmunication

27-1-2009 | 7

### Prerequisites:

- › Language community
- › Interaction
- › Symbolic integration



rijksuniversiteit  
groningen

## Semiconmunication

27-1-2009 | 8

### Observed semiconmunication (Zeevaert 2004):

- › Danish - Norwegian - Swedish (Haugen 1966, Maurud 1976....)
- › Czech - Slovakian (Budovičá 1987)
- › Czech - Polish (Hansen 1987)
- › Spanish - Portuguese (Coseriu 1988, Jensen 1989, Zeevaert 2002)
- › Italian - Spanish (Hansen 1987)
- › German - Dutch (Haz 2002)
- › Frisian - Dutch (Feitsma 1986)
- › Croatian - Serbian (Haugen 1990)
- › Hindi - Urdu (Haugen 1990)
- › Icelandic - Faeroese (Braunmuller & Zeevaert 2001)
- › Macedonian - Bulgarian (Haugen 1990)
- › Russian - Bulgarian (Braunmuller & Zeevaert 2001)
- › Chinese dialects (Cheng 1997, Tang & Van Heuven 2007)
- › Arabian dialects (Haugen 1990)



rijksuniversiteit  
groningen

## Factors explaining intelligibility

27-1-2009 | 9

### Linguistic

- › sounds
- › prosody
- › lexicon
- › morphology
- › syntax

### Extra-linguistic

- › attitude
- › contact
- › linguistic experience
- › orthography



rijksuniversiteit  
groningen

## Factors explaining intelligibility

27-1-2009 | 10

- › A model of intelligibility: the relative importance of the factors
- › Intelligibility measurements can be used to find out how the linguistic factors should be weighed



rijksuniversiteit  
groningen

## Central questions

27-1-2009 | 11

1. How can the mutual intelligibility between closely related languages be measured?
2. How can the relevant (extra-)linguistic factors be measured?
3. To what extent are the (extra-)linguistic factors predictors of intelligibility?



rijksuniversiteit  
groningen

## Day-to-day program

27-1-2009 | 12

### Monday: intelligibility testing

1. Introduction
2. Overview of methods for measuring intelligibility
3. Practical: participation in an online web-based intelligibility experiment



rijksuniversiteit  
groningen

## Day-to-day program

27-1-2009 | 13

### Tuesday: phonetic distances and intelligibility I

1. Practical: extraction of results from the web-based intelligibility experiment
2. Levenshtein distance
3. Relating Levenshtein distances to intelligibility scores



rijksuniversiteit  
groningen

## Day-to-day program

27-1-2009 | 14

### Wednesday: phonetic distances and intelligibility II

1. Practical: exercises on Levenshtein distance
2. Entropies
3. Vowel distance measurements



rijksuniversiteit  
groningen

## Day-to-day program

27-1-2009 | 15

### Thursday: relating linguistic distances to intelligibility scores

1. Practical: linguistic factors determining intelligibility
2. Linguistic factors of Danish-Swedish intelligibility
3. Lexical distances
4. Morpho-syntactic distances



rijksuniversiteit  
groningen

## Day-to-day program

27-1-2009 | 16

### Friday: the role of extra-linguistic factors for intelligibility

1. Overview of extra-linguistic factors
2. Practical: Exercises on extra-linguistic factors
3. Excluding extra-linguistic factors in Danish-Swedish intelligibility
4. Conclusions



rijksuniversiteit  
groningen

## Project I

27-1-2009 | 17

### *Linguistic determinants of mutual intelligibility in Scandinavia*

- VIDL, financed by NWO
- 1 January 2006 – 1 January 2011
- Project members:
  - Charlotte Gooskens
  - Sebastian Kürschner
  - Renée van Bezooijen
  - Anja Schüppert



rijksuniversiteit  
groningen

## Project II

27-1-2009 | 18

### *Mutual intelligibility of language varieties in the Low Countries: linguistic and attitudinal determinants*

- VNC, financed by NWO and FWO
- 1 January 2007 – 1 January 2011
- Project members from Leuven, Nijmegen and Groningen



rijksuniversiteit  
groningen


 rijksuniversiteit  
 groningen

27-1-2009 | 19

## Measuring intelligibility

## Measuring intelligibility

- › Opinion testing
- › Functional testing
- › Observations


 rijksuniversiteit  
 groningen

## Opinion testing

27-1-2009 | 21

How well does the listener **think** he understands the other language variety (opinion scores)?

**Advantages:**

- efficient
- the same words can be tested in each variety

**Disadvantages:**

- listeners may not be able to judge intelligibility


 rijksuniversiteit  
 groningen

## Opinion testing

27-1-2009 | 22

Questionnaire **without** speech samples

- Judgments on a scale

Example Haugen (1953) :

- 300 persons in Scandinavia chosen randomly from the phone book
- ‘When you met a X for the first time, how well could you understand him?’

0 = ‘not at all’  
5 = ‘understood everything’


 rijksuniversiteit  
 groningen

## Opinion testing

27-1-2009 | 23

Questionnaire **with** speech samples

- Judgments on a scale

Example Tang & Van Heuven (2007):

- Recordings of 15 Chinese dialects
- The fable “The North Wind and the Sun”.
- 24 listeners from each of the places where the dialects were spoken
- ‘How well do you believe monolingual listeners of your own dialect would understand the speaker?’

0 = ‘They will not understand a word of the speaker’  
10 = ‘They will understand the other speaker perfectly’


 rijksuniversiteit  
 groningen

## Functional testing

27-1-2009 | 24

How well does the listener **actually** understand the other language variety?

**Advantages:**

- actually measures intelligibility

**Disadvantages:**

- priming effects must be avoided
- heavy memory load
- time consuming


 rijksuniversiteit  
 groningen

## Functional testing

27-1-2009 | 25

### Text intelligibility:

- tests language as a whole
- resembles a natural situation

### Word intelligibility:

- gives researcher the opportunity to investigate the role of specific linguistic factors
- artificial situation



rijksuniversiteit  
groningen

## Functional testing: texts

27-1-2009 | 26

### Questions about content of text

- % correct answers

Example Delsing/Lundin Åkesson 2003:

- Speech samples: recordings of Danish, Norwegian and Swedish news items
- Danish, Norwegian and Swedish listeners
- 5 open questions, for example:  
'What kind of animal is this text about?'  
'For how long did he do this job?'



rijksuniversiteit  
groningen

## Functional testing: texts

27-1-2009 | 27

### Translation of whole text

- % correctly translated words

Example Beijering/Gooskens 2008:

- The Fable 'The North wind and the Sun'
- Presented in small parts (max. 8 words)
- 18 Nordic varieties
- Listeners from Copenhagen
- % correct translations per dialect



rijksuniversiteit  
groningen

## Functional testing: texts

27-1-2009 | 28

### Translation of whole text

- Proportion of content covered by translation

Examples from ethnolinguistic research:

- Hickerson, Turner & Hickerson (1952): Iroquois
- Pierce (1952): Algonquian
- Biggs (1957): Yuman



rijksuniversiteit  
groningen

## Functional testing: texts

27-1-2009 | 29

### Recorded Text Testing (RTT)

- Questions about story or retelling

Example Kluge (2007):

- Recording of an autobiographical text (3 min)
- Divided into 10-12 segments with short pauses between
- Identification of core elements by L1 speakers
- Listeners retell each segment
- % correctly retold core elements



rijksuniversiteit  
groningen

## Functional testing: texts

27-1-2009 | 30

Example RTT, Kluge (2007), Western Sentani (Indonesia)

As I pulled, the back of the canoe nudged a Sago leaf which startled the crocodile. Immediately, it jumped out of the water to throw itself behind me onto the canoe.

4 core elements = 4 points

- (he/I) pulled (in the net); canoe touches/nudges Sago leaf/tree/trunk; crocodile startled; crocodile jumps/throws itself onto/beside canoe

RTT testing response:

While pulling in the net, the canoe hit a Sago leaf. The man was startled. The crocodile jumped (to get into) onto the boat.

Score: 3.5



rijksuniversiteit  
groningen

Functional testing: texts

27-1-2009 | 31

**Cloze test**

- text with certain words removed
- the listener is asked to replace the missing words
- % correctly inserted words

Example (Van Bezooijen & Van den Berg 1999):

- Semi-spontaneous speech samples of various Dutch dialects
- Task: translate nouns into Dutch
- E.g. '..... liggen d'r in de .....

 rijksuniversiteit  
groningen

Functional testing: texts

27-1-2009 | 32

**Cloze test**

Example:

No, I \_\_\_\_\_ . 

 rijksuniversiteit  
groningen

Functional testing: texts

27-1-2009 | 33

**Cloze test**

Example:

No, I \_\_\_\_\_ .

No, I hated school. 

 rijksuniversiteit  
groningen

Functional testing: texts

27-1-2009 | 34

**Cloze test**

Example:

I \_\_\_\_\_ didnae like \_\_\_\_\_ at a \_\_\_\_\_ and being \_\_\_\_\_ what to dae. 

 rijksuniversiteit  
groningen

Functional testing: texts

27-1-2009 | 35

**Cloze test**

Example:

I just didnae \_\_\_\_\_ at a \_\_\_\_\_ and being \_\_\_\_\_ what to dae.

I just didnae like sitting at a table and being telt what to dae. 

 rijksuniversiteit  
groningen

Functional testing: texts

27-1-2009 | 36

**Cloze test**

I was pretty much a, an outdoor person, I never s, I was a bit like Jimmy's school, you was glad to see, hear a bell at the end of the day so you could get back hame I stayed on a farm, as well, so farming's been my life since I was born and you just got hame and you, you threwed the schoolbooks in the door, went and got changed and went outside and if I wasnae helping my dad on the farm I was playing with a football in the garden, so that was pretty much my younger day.

 rijksuniversiteit  
groningen

## Functional testing: texts

27-1-2009 | 37

**SPIN-test (Speech Perception in Noise, Kalikow, Stevens and Elliott (1977).**

- Write down last word of simple sentences
- e.g.  
Low-predictability: *We were talking about the **road***  
High-predictability: *The car drove down the **road***

Example Tang/Van Heuven (2008):

- Chinese dialects
- High-predictability sentences



rijksuniversiteit  
groningen

## Functional testing: texts

27-1-2009 | 38

**Semantically Unpredictable Sentences (SUS, Benoît, Grice & Hazan 1996)**

- SUS-generator generates sentences automatically
- Five basic syntactic structures
- Frequent mini-syllabic words
- Ex. *How does the cloud watch the low text?*
- % correctly translated (content) words or sentences



rijksuniversiteit  
groningen

## Functional testing: words

27-1-2009 | 39

**% correctly translated words**

Example Kürschner/Van Bezooijen/Gooskens (2008):

- 384 isolated words
- 7 Germanic languages
- Listeners are native speakers of the 7 languages
- Translations into mother tongue
- Via internet



rijksuniversiteit  
groningen

## Functional testing: words

27-1-2009 | 40

**Reaction time**

Example Impe (in progress):

- 10 Dutch Standard varieties in Belgium and The Netherlands
- 20 existing words and 20 non-words in each variety
- Listeners from the same places as the speakers
- Lexical decision task + multiple choice task
- Reaction times measured with E-prime
- Reaction times of correctly recognized words



rijksuniversiteit  
groningen

## Functional testing: words

27-1-2009 | 41

**Semantic categorization**

- Listeners categorize test words

Example Tang/Van Heuven (2008):

- Chinese dialects
- Listeners indicate to which of ten pre-given semantic categories a spoken word belongs
- for example, 'apple' should be categorized as a member of the category 'fruit'



rijksuniversiteit  
groningen

## observations

27-1-2009 | 42

**Observations of real language situations**

- Number of misunderstandings, repairs, reformulations, pauses, turn taking etc.
- Arranged or real conversations

Example Börestam Uhlmann (1994):

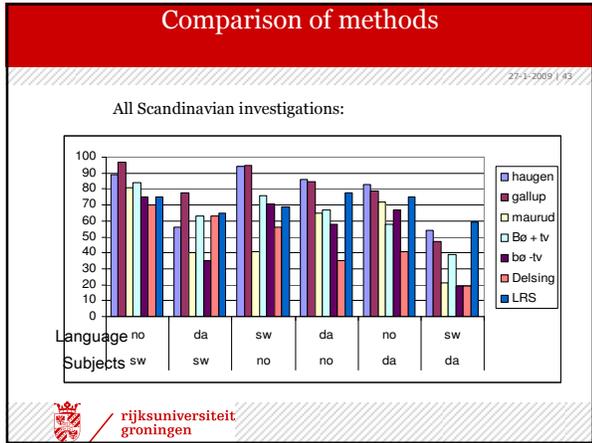
- Pairs of Scandinavians were asked to arrange a low-budget party

Example Zeevaert (2005):

- Observations of real Nordic meetings



rijksuniversiteit  
groningen



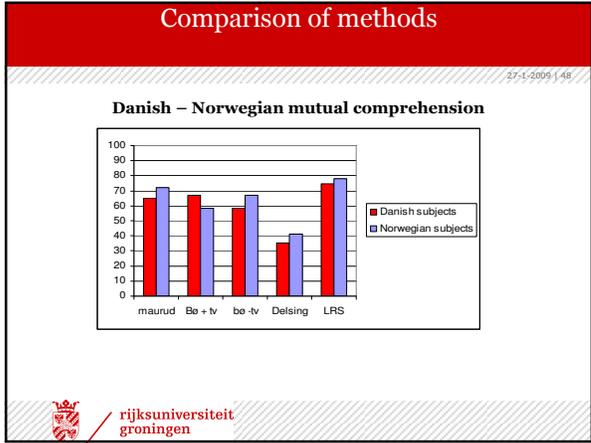
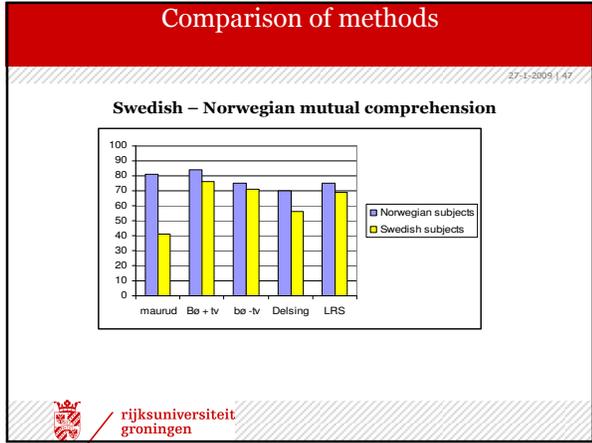
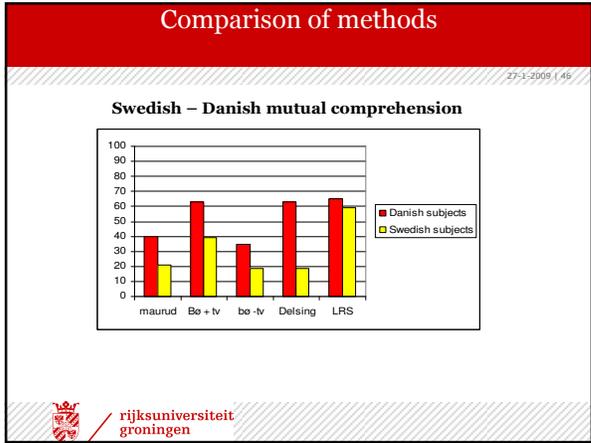
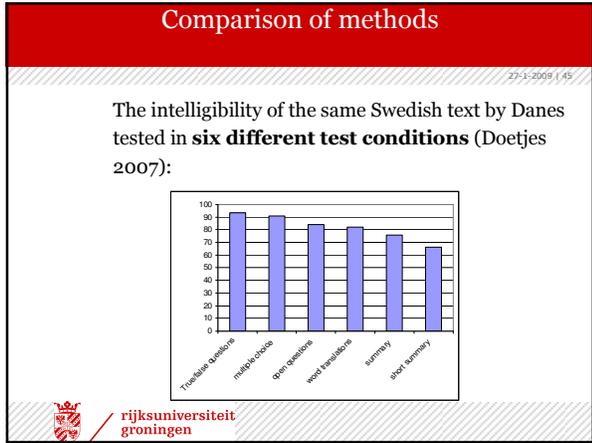
### Comparison of methods

27-1-2009 | 44

The intelligibility of the same Swedish text by Danes tested in **six different test conditions** (Doetjes 2007):

1. Open questions
2. True/false questions
3. Multiple choice questions
4. Word translation
5. Summary
6. Short summary

riijksuniversiteit groningen



## Comparison of methods

27-1-2009 | 49

### Word tests versus content tests

Maurud (1976):

- Mutual intelligibility in Scandinavia
- Word tests and content tests based on same texts
- Correlations between the two test types between .60 and .80

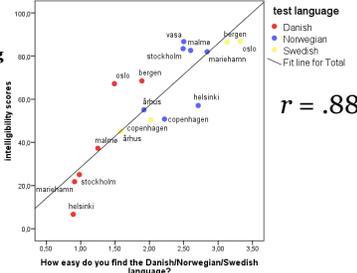

 rijksuniversiteit  
groningen

## Comparison of methods

27-1-2009 | 50

### Opinion versus functional testing

**Delsing & Lundin Åkesson (2003):**




 rijksuniversiteit  
groningen

## Comparison of methods

27-1-2009 | 51

### Opinion versus functional testing

Tang & Van Heuven:

Correlations between opinion tests and functional tests of intelligibility among 15 Chinese dialects:

$r = .80$

→

35% unexplained variance


 rijksuniversiteit  
groningen

## Methodological considerations

27-1-2009 | 52

### Avoiding priming effects

Latin square design:

- each listener hears a proportion of the words in each of the test languages, and yet hears words in each of the languages in equal proportions, and never hears the same word twice

languages	Test version			
	I	II	III	IV
1	A	D	C	B
2	B	A	D	C
3	C	B	A	D
4	D	C	B	A


 rijksuniversiteit  
groningen

## Methodological considerations

27-1-2009 | 53

### Avoiding ceiling effects

- Reaction times
- Make listeners task more difficult:
  - filtering
  - signal compression
  - noise


 rijksuniversiteit  
groningen

## Methodological considerations

27-1-2009 | 54

### Test persons should be matched:

- Age
- Gender
- Social class
- Level of education
- Geography


 rijksuniversiteit  
groningen

## Methodological considerations

27-1-2009 | 35

### **Stimulus material:**

- Same for all varieties
- Comparable speakers (gender, age, voice quality etc.)
- Sentences not too long
- Suitable level of difficulty
- Not translations from one test language
- Control language



rijksuniversiteit  
groningen