VOX NORWEGIAN AGENCY FOR LIFELONG LEARNING

How valid is the CEFR as a construct for language tests?

Cecilie Hamnes Carlsen

5th ALTE-conference, Paris, April 9th – 10th 2014

"[…] the CEFR levels are neither based on empirical evidence taken from L2-learner performance, nor on any theory in the fields of linguistics or verbal communication (Alderson, 2007, Hulstijn 2007)"

Hulstijn, Alderson, Schoonen 2010:15

	GENERAL LINGUISTIC RANGE
C2	*Can exploit a comprehensive and reliable mastery of a very wide range of language to formulate thoughts precisely, give emphasis, differentiate and eliminate ambiguity . . . No signs of having to restrict what he/she wants to say.*
C1	*Can select an appropriate formulation from a broad range of language to express him/herself clearly, without having to restrict what he/she wants to say.*
B2	*Can express him/herself clearly and without much sign of having to restrict what he/she wants to say.*
	Has a sufficient range of language to be able to give clear descriptions, express viewpoints and develop arguments without much conspicuous searching for words, using some complex sentence forms to do so.
B1	*Has a sufficient range of language to describe unpredictable situations, explain the main points in an idea or problem with reasonable precision and express thoughts on abstract or cultural topics such as music and films.*
	Has enough language to get by, with sufficient vocabulary to express him/herself with some hesitation and circumlocutions on topics such as family, hobbies and interests, work, travel, and current events, but lexical limitations cause repetition and even difficulty with formulation at times.
A2	*Has a repertoire of basic language which enables him/her to deal with everyday situations with predictable content, though he/she will generally have to compromise the message and search for words.*
	Can produce brief everyday expressions in order to satisfy simple needs of a concrete type: personal details, daily routines, wants and needs, requests for information. *Can use basic sentence patterns and communicate with memorised phrases, groups of a few words and formulae about themselves and other people, what they do, places, possessions etc.* *Has a limited repertoire of short memorised phrases covering predictable survival situations; frequent breakdowns and misunderstandings occur in non-routine situations.*
A1	*Has a very basic range of simple expressions about personal details and needs of a concrete type.*

"What we need to know if we want to develop good [proficiency] scales is [...] how somebody acquires language, that is, what the developmental stages in language acquisition are."

(John de Jong 1988:74)

Validating the CEFR against L2-data

Major challenges

1. General and non-language specific

2. A multitude of decriptor scales

3. Lack of suitable L2-data

An example of a validation study

Norwegian

The illustrative scale of coherence and cohesion

1. CEFR is general and non-language specific

Endres i topp-/bunntekst

2. Challenge: A multitude of descriptor scales

> 50 scales

      Functional scales

      Linguistic scales – a challenge

Just *one* linguistic scale needs a lot of research to be validated

	GRAMMATICAL ACCURACY
C2	Maintains consistent grammatical control of complex language, even while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions).
C1	Consistently maintains a high degree of grammatical accuracy; errors are rare and difficult to spot.
B2	Good grammatical control; occasional 'slips' or non-systematic errors and minor flaws in sentence structure may still occur, but they are rare and can often be corrected in retrospect.
	Shows a relatively high degree of grammatical control. Does not make mistakes which lead to misunderstanding.
B1	Communicates with reasonable accuracy in familiar contexts; generally good control though with noticeable mother tongue influence. Errors occur, but it is clear what he/she is trying to express.
	Uses reasonably accurately a repertoire of frequently used 'routines' and patterns associated with more predictable situations.
A2	Uses some simple structures correctly, but still systematically makes basic mistakes – for example tends to mix up tenses and forget to mark agreement; nevertheless, it is usually clear what he/she is trying to say.
A1	Shows only limited control of a few simple grammatical structures and sentence patterns in a learnt repertoire.

VOX — NORWEGIAN AGENCY FOR LIFELONG LEARNING

	GRAMMATICAL ACCURACY
C2	**Morphology** trol of complex language, even w... nonitoring others' reactions). **Syntax**
C1	Consistently maintains ... matical a... difficult to **Finitness** **Word order**
B2	**Verb phrase** ... **Tenses** ... 'slips' or non-systemati... structure may still occur, but they are rare and can often be c... **Compulsory subject?** **Aspect** Shows a relatively high a... trol. Does not make mistakes which lead to misunderstanding.
B1	**Noun Phrase** ...nab... ...iliar contexts; ge... **Gender** influence. Errors occur, but it is clear what he/she is trying to express. **Verb second?** Uses reasonably ac... re...gy...quently... an... predictable situati... **Number** **Definitness** **Subordinate clauses**
A2	**Articles** ...s correctly, but still syste... **Pronouns** ...mistakes – for example tends to mix up tenses and fo... ...nevertheless, it is usually clear what he/she is trying to say. **Adjectives**
A1	**Adverbs** ...l of a few s... **Prepositions** ...sentence patterns in a learnt repertoire.

3. Lack of suitable L2-data

To validate the CEFR we need large data sets of L2-texts

- from different L1-groups
- at different CEFR-levels
- tagged for grammatical traits
- tagged for errors

3. Lack of suitable L2-data

To validate the CEFR we need large data sets of L2-texts

Computer learner corpora

An example of a validation study

- Norweigan Profile

- April 2011- September 2012

- 8 project members
  - 5 language test developers (Norsk språktest)
  - 3 SLA-researchers (University of Bergen)

- Reference Level Descriptions (RLD) for Norwegian

- Financial support: Ministry of Children, Equality and Social inclusion

- Carlsen, C. (ed.) Norsk profil 2103
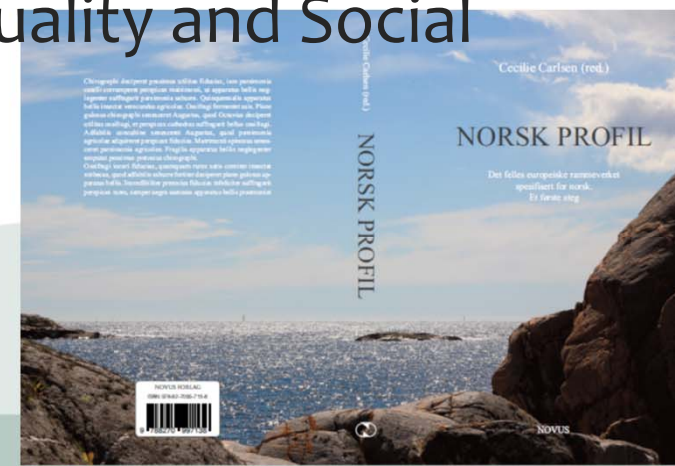
Discourse connectives

Fraser (1996) defines discourse connectives as elements "which signal a relation between the discourse segment which hosts them, and the prior discourse segment" (p. 190)

"Linking words" such as: *and, for example, in addition, but, despite, however, because, since* etc.

In my study: 36 connectives

	COHERENCE AND COHESION
C2	*Can create coherent and cohesive text making full and appropriate use of a variety of organisational patterns and a wide range of cohesive devices.*
C1	*Can produce clear, smoothly flowing, well-structured speech, showing controlled use of organisational patterns, connectors and cohesive devices.*
B2	*Can use a variety of linking words efficiently to mark clearly the relationships between ideas.*
	Can use a limited number of cohesive devices to link his/her utterances into clear, coherent discourse, though there may be some "jumpiness" in a long contribution.
B1	*Can link a series of shorter, discrete simple elements into a connected, linear sequence of points.*
A2	*Can use the most frequently occurring connectors to link simple sentences in order to tell a story or describe something as a simple list of points.*
	Can link groups of words with simple connectors like "and", "but" and "because".
A1	*Can link words or groups of words with the very basic linear connectors like "and" or "then".*

The Illustrative scale of Coherence and Cohesion, CEFR p. 125., my emphasis

COHERENCE AND COHESION

C2	*Can create coherent and cohesive text making **full and appropriate use** of a variety of organisational patterns and a **wide range of cohesive devices**.*
C1	*Can produce clear, smoothly flowing, well-structured speech, showing **controlled use** of organisational patterns, **connectors and cohesive devices**.*
B2	*Can use a **variety of linking words efficiently** to mark clearly the relationships between ideas.*
	*Can use a **limited number of cohesive devices** to link his/her utterances into clear, coherent discourse, though there may be some "jumpiness" in a long contribution.*
B1	***Can link** a series of shorter, discrete simple elements into a connected, linear sequence of points.*
A2	*Can use **the most frequently occurring connectors** to link simple sentences in order to tell a story or describe something as a simple list of points.*
	*Can link groups of words with **simple connectors** like "and", "but" and "because".*
A1	*Can link words or groups of words with **the very basic linear connectors** like "and" or "then".*

The Illustrative scale of Coherence and Cohesion, CEFR p. 125., my emphasis

CEFR-predictions about connectives

P1: Texts at higher levels contain a *broader range* of different cohesive devices than texts at lower levels

P2: Texts at higher levels contain more *low-frequency* connectives than texts at lower levels

P3: Texts at higher levels show a *greater degree of control* of the cohesive devices used than texts at lower levels

CEFR-based hypotheses about connectives

H1: Texts at higher levels contain a *broader range* of different cohesive devices than texts at lower levels

H2: Texts at higher levels contain more *low-frequency* connectives than texts at lower levels

H3: Texts at higher levels show a *greater degree of control* of the cohesive devices used than texts at lower levels

Computer learner corpus of Norwegian

- Collaborative effort (Langauge testers, SLA-researchers, computer spesialists)

- Texts from two standardized tests of Norwegian for adult immigrants

- 1700 texts (ca. 700 000 words)

- Learners with 10 different L1s

- Automatically tagged for grammatical traits

- Manually tagged for errors

- 1222 of the texts were reassessed using the CEFR-scale A1-C2

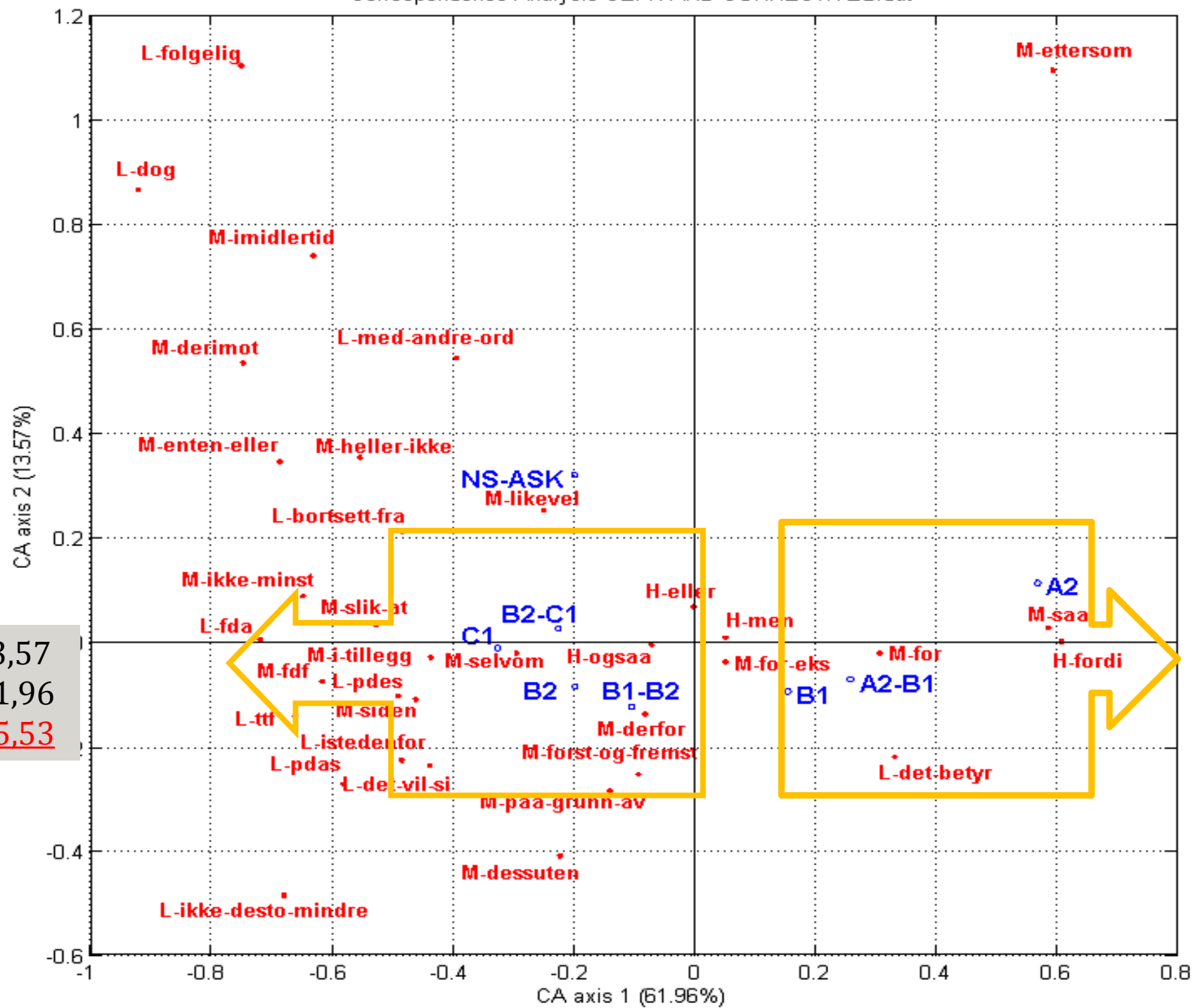(5-10 raters, Carlsen & Kaftandjieva 2009)

Connectives

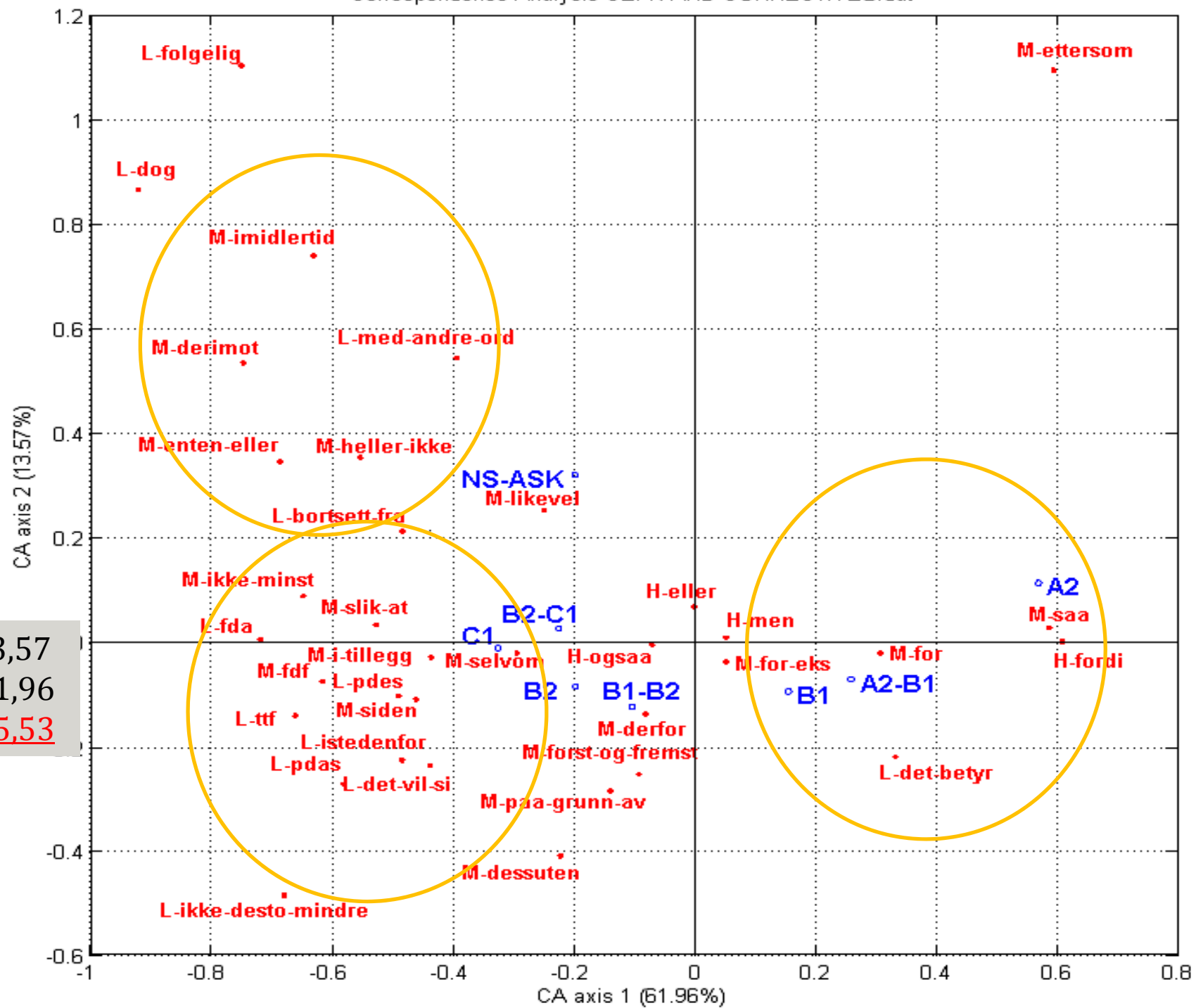High frequent (H-)
Medium frequent (M-)
Low frequent (L-)

FUNKSJON	HØYFREKVENTE	MIDDELSFREKVENTE	LAVFREKVENTE
ADDITIVE	eller også	for eksempel i tillegg enten… eller ikke minst heller ikke først og fremst for det første (fdf) dessuten	det vil si med andre ord for det andre (fda) bortsett fra det betyr
ADVERSATIVE	men	likevel selv om derimot imidlertid	istedenfor til tross for (ttf) dog på den ene siden på den andre siden ikke desto mindre
KAUSATIVE	fordi	derfor slik at så (saa) på grunn av siden ettersom for	følgelig

Correspondence Analysis CEFR-AND-CONNECTIVES.dat

Correspondence Analysis CEFR-AND-CONNECTIVES.dat

13,57
+ 61,96
= 75,53

Results

H1: Texts at higher levels contain a *broader range* of different cohesive devices than texts at lower levels

> Supported by the data

H2: Texts at higher levels contain more *low-frequency* connectives than texts at lower levels

> Supported by the data

H3: Texts at higher levels show a *greater degree of control* of the cohesive devices used than texts at lower levels

> Supported by the data

An interesting finding

The CEFR predicts that it is only at the B2+ level that a "variety of linking words" are used

This study shows that learners even at a B1/B2 level use a range of different connectives, and to an extent which separates them sharply from the lower levels

Revision of the scale warranted?

Paris Dream

Thank you!

Wouldnt' it be lovely if we could...

1) make an overview of prior validation studies?

2) compare findings?

3) revise of the CEFR based on empirical findings?