Diagnosing reading in L2 – predictors and vocabulary profiles

Ari Huhta, J.Charles Alderson, Lea Nieminen & Riikka Ullakonoja
University of Jyväskylä & Lancaster University

ACTFL CEFR Alignment Conference 2011
August 4–6, 2011, Provo, Utah
Outline of talk

- CEFR-related research at University of Jyväskylä & Lancaster University
- DIALUKI research project & related projects
- Findings from DIALUKI Study 1
  - I Predictors of reading in English as FL (overall)
  - II Vocabulary profiles (frequency) of different CEFR levels in reading in English
Some CEFR related research at University of Jyväskylä & Lancaster University

- Often in co-operation with international partners
- DIALANG 1996–2004 (www.dialang.org)
- The Dutch Construct Project (www.ling.lancs.ac.uk/cefgrid)
- CEFLING & TOPLING 2007–09; 2010–2013
  - www.jyu.fi/cefling ; www.jyu.fi/topling
  - Development of writing across CEFR levels in Finnish as L2 and English & Swedish as FL / L2
- DIALUKI 2010–2013
- SLATE network (Second Language Acquisition and Testing in Europe); www.slate.eu.org

Preliminary results – please do not quote
DIALUKI – Diagnosing reading and writing in a second or foreign language

- a 4-year research project based in Finland

Related projects 2010–12:

- Reading in the language of instruction in PISA
- Reading in L2 in Pearson Test of English Academic

Research Questions include:

- What features of task demands and texts best predict item and task difficulty?
- Which aspects of the constructs underlying SFL reading tests can expert judges agree upon?
- What model of reading processes and text variables will be most helpful for test developers, response coders and teachers, to predict difficulty and to use pedagogically?

More information: J.C. Alderson; also AAAL 2012
Attempts at understanding and operationalising constructs that may be relevant to understanding reading and writing in L2

Paves way for more and better diagnostic tests in the future by increasing our understanding of diagnosing L2 reading and writing
DIALUKI
Diagnosing reading and writing in a second or foreign language /
Toisen tai vieraan kielen lukemisen ja kirjoittamisen diagnosointi /
Диагностика навыков чтения и письма на родном и на втором или
иностранных языках

Choose language / Valitse kiel / Выберите язык:
DIALUKI - Diagnosing reading and writing in a second or foreign language

The Research Project is funded by the Academy of Finland, the University of Jyväskylä and the UK Economic and Social Research Council (ESRC).

Aims

Studies

People

Related projects

Links

Seminar and network for L2 reading researchers

Presentations and publications
DIALUKI – Diagnosing reading and writing in SFL

- Research project 2010–2013: work in progress
- Can different L1 and L2 linguistic & psycholinguistic measures predict difficulties in SFL R/W?
- How does SFL proficiency in R/W develop in psycholinguistic and linguistic terms?
- Which features or combinations of features characterise different CEFR proficiency levels?
- Cooperation between language testers, other applied linguists and psychologists (L1 reading)
DIALUKI informants

- Finnish-speaking learners of English as FL
  - primary school 4th grade (age 10)
  - lower secondary school, 8th grade (age 14)
  - gymnasium (academically oriented upper secondary school), 2nd year students (age 17)

- Russian-speaking learners of Finnish as SL
  - primary school (3-6th grade)
  - lower secondary school (7-9th grade)

- From 111 schools around Finland

Preliminary results – please do not quote
Three major studies

**Study 1** (in 2010): A cross-sectional study with 3 x 200 + 250 students.
- Exploring the value of a range of L1 & L2 measures in predicting L2 reading & writing, in order to select the best predictors for further studies.

**Study 2** (in 2011/12): Several training / experimental studies, each a few weeks in length
- The effects of training on SFL reading and writing
- Using computerized learning games in L1 and L2.
  - **Graphogame** for diagnosing & treating dyslexia

**Study 3** (in 2011 – 2012/13): Longitudinal, 2–3 years.
- The development of literacy skills, and the relationship of this development to the diagnostic measures.
DIALUKI (Study 1) outline

INDEPENDENT VARIABLES

- Cognitive features
- Motivation
- L2 vocabulary size
- Personal and family background
- Self-assessment of L1 & L2 reading & writing

DEPENDENT VARIABLES

- L1 reading
- L1 writing
- L2 reading
- L2 writing

Preliminary results – please do not quote
Study 1 – dependent variable(s)

Reading in a foreign language (English)

- DIALANG reading test of English (2 levels), 30 items; linked with the CEFR levels
- Pearson Test of English (PTE) General, 24 operational reading items (A2 – B2); linked with the CEFR
  - for the first prediction study, 20 + 20 best items combined into a measure of reading in English
  - 4th grade: Pearson Test of English for Young Learners

Preliminary results – please do not quote
I Predictors of reading in English as FL

DIALUKI Study 1

4th graders (10–year–olds)
8th graders (14–year–olds)
Gymnasium students (17–year–olds)

about 200 learners per group
Finnish as L1
Dependent variable: reading in English as FL

- Self-assessment / reading in L1 (Finnish)
  - Effectiveness of working memory (backward digit span score in L1)
- Self-assessment / reading in FL (English)
  - Effectiveness of working memory (backward digit span score in FL)
- Spelling in L1 (= correcting real L1 words)
- Reading in L1 (Finnish) (= PISA test score)
- Size of L1 Finnish vocabulary
  - Size of FL English vocabulary
- Reading in FL (English) (= DIALANG and Pearson PTE General test score)
  - Spelling in L1 (accuracy of spelling non-words in L1)
  - Speed of lexical access (speed of naming in L1)
  - Speed of lexical access (speed of naming in FL)
  - Speed of lexical access (speed of reading a list of words in L1)
  - Speed of lexical access (speed of reading a list of words in FL)
- Listening in L2 (dictation)
- Segmentation in FL
- Segmentation in L1
- Motivation, self-regulation, anxiety
- Frequency of reading in English
- Attitude to reading in the free time

Preliminary results – please do not quote
Independent / predictor variables

Cognitive / psycholinguistic measures

- Size of L1 Finnish vocabulary
- Size of FL English vocabulary
- Listening in L2 (dictation)
- Segmentation in FL
- Segmentation in L1
- Motivation, self-regulation, anxiety
- Frequency of reading in English
- Attitude to reading in the free time

- Self-assessment / reading in L1 (Finnish)
- Spelling in L1 (= correcting real L1 words)
- Reading in L1 (Finnish) (= PISA test score)
- Effectiveness of working memory (= backward digit span score in L1)
- Effectiveness of working memory (= backward digit span score in FL)
- Rapid recognition of words in L1
- Rapid recognition of words in FL
- Speed of lexical access (= speed of reading a list of words in L1)
- Speed of lexical access (= speed of reading a list of words in FL)
- Speed of lexical access (= speed of naming in FL)

Preliminary results – please do not quote
## Backward Digit Span Memory Test in L1 and L2

<table>
<thead>
<tr>
<th></th>
<th>2 - 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5 7 4</td>
</tr>
<tr>
<td>3</td>
<td>7 - 2 - 9 - 6</td>
</tr>
<tr>
<td>4</td>
<td>4 - 1 - 3 - 5 - 7</td>
</tr>
<tr>
<td>5</td>
<td>1 - 6 - 5 - 2 - 9 - 8</td>
</tr>
<tr>
<td>6</td>
<td>8 - 5 - 9 - 2 - 3 - 4 - 2</td>
</tr>
<tr>
<td>7</td>
<td>6 - 9 - 1 - 6 - 3 - 2 - 5 - 8</td>
</tr>
</tbody>
</table>
Rapid Automatic Naming (colours, letters and numbers)

- In L1 and L2
- Say these as fast and as accurately as you can

![Image of a grid with shapes and letters or numbers].

Preliminary results - please do not quote
# Reading as many words as possible in one minute (L2 here but also in L1)

<table>
<thead>
<tr>
<th></th>
<th>Word</th>
<th>Response</th>
<th></th>
<th>Word</th>
<th>Response</th>
<th></th>
<th>Word</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>be</td>
<td></td>
<td>36</td>
<td>house</td>
<td></td>
<td>71</td>
<td>because</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>as</td>
<td></td>
<td>37</td>
<td>under</td>
<td></td>
<td>72</td>
<td>develop</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>on</td>
<td></td>
<td>38</td>
<td>thing</td>
<td></td>
<td>73</td>
<td>between</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>at</td>
<td></td>
<td>39</td>
<td>write</td>
<td></td>
<td>74</td>
<td>general</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>do</td>
<td></td>
<td>40</td>
<td>tell</td>
<td></td>
<td>75</td>
<td>another</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>or</td>
<td></td>
<td>41</td>
<td>hand</td>
<td></td>
<td>76</td>
<td>present</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>if</td>
<td></td>
<td>42</td>
<td>both</td>
<td></td>
<td>77</td>
<td>without</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>we</td>
<td></td>
<td>43</td>
<td>move</td>
<td></td>
<td>78</td>
<td>problem</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>all</td>
<td></td>
<td>44</td>
<td>much</td>
<td></td>
<td>79</td>
<td>however</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>for</td>
<td></td>
<td>45</td>
<td>back</td>
<td></td>
<td>80</td>
<td>country</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>you</td>
<td></td>
<td>46</td>
<td>little</td>
<td></td>
<td>81</td>
<td>several</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>say</td>
<td></td>
<td>47</td>
<td>point</td>
<td></td>
<td>82</td>
<td>student</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>who</td>
<td></td>
<td>48</td>
<td>world</td>
<td></td>
<td>83</td>
<td>inform</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>can</td>
<td></td>
<td>49</td>
<td>place</td>
<td></td>
<td>84</td>
<td>control</td>
<td></td>
</tr>
</tbody>
</table>
Stepwise multiple regression analysis, **cognitive** variables with EFL reading

<table>
<thead>
<tr>
<th></th>
<th>Adjusted R Square</th>
<th>% variance</th>
<th>First variable</th>
<th>Second variable</th>
<th>Third variable</th>
<th>Fourth variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>4th Grade</td>
<td>.192</td>
<td>19%</td>
<td>Rapidly presented words in English (.375)</td>
<td>Rapid naming of colours etc in English (-.367)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8th Grade</td>
<td>.294</td>
<td>29%</td>
<td>Rapid naming of colours etc in English (.509)</td>
<td>Backward digit span in English (.336)</td>
<td>Rapidly presented words in English (.289)</td>
<td></td>
</tr>
<tr>
<td>Gym</td>
<td>.317</td>
<td>31%</td>
<td>Rapid naming of colours etc in English (-.476)</td>
<td>Backwards digit span in English (.377)</td>
<td>Reading word list in Finnish (.046)</td>
<td>Rapidly presented words in English (.245)</td>
</tr>
</tbody>
</table>
Indep 2
Motivation & (Background)

Preliminary results – please do not quote
Motivation and background variables

- Iwaniec Motivation measure of Instrumentality, Intrinsic Interest, Motivational Intensity, Parental Encouragement, Anxiety, Self-regulation and English Self-concept

- (Some) background variables based on questionnaires from PISA and previous research projects
## Stepwise multiple regression analysis, motivation factors with EFL reading

<table>
<thead>
<tr>
<th></th>
<th>Adjusted R Square</th>
<th>% variance explained</th>
<th>First variable</th>
<th>Second variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>4th grade (10-year-olds)</td>
<td>.226</td>
<td>23%</td>
<td>English Self-concept (.445)</td>
<td>Parental encouragement (-.176)</td>
</tr>
<tr>
<td>8th grade (14-year-olds)</td>
<td>.382</td>
<td>38%</td>
<td>English Self-concept (.605)</td>
<td>Self-regulation (.044)</td>
</tr>
<tr>
<td>Gymnasium (17-18-year-olds)</td>
<td>.382</td>
<td>38%</td>
<td>English Self-concept (.609)</td>
<td>Self-regulation (.169)</td>
</tr>
</tbody>
</table>
Independent / predictor variables 3 – Linguistic measures
Linguistic constructs and their measures

Receptive vocabulary
○ FL (Nation / Schmitt et al. Vocabulary Levels test of English)
○ L1 (Finnish vocabulary test from DIALANG)

L1 reading
○ PISA reading tests
○ ALLU test (4th grade)

Spelling (& phonology)
○ correcting spelling errors in real L1 words

Syntax: segmentation task in L1 and FL

Self-assessed reading and writing in L1 and FL
## Stepwise multiple regression analysis, linguistic variables with EFL reading

<table>
<thead>
<tr>
<th>Grade</th>
<th>Dependent variable</th>
<th>Adjusted R Squared</th>
<th>% variance</th>
<th>1st IV</th>
<th>2nd IV</th>
<th>3rd IV</th>
<th>4th IV</th>
<th>5th IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>4th</td>
<td>Pearson Young Learners Test in English</td>
<td>.494</td>
<td>49%</td>
<td>Size of English vocab (.664)</td>
<td>L1 Finnish Reading (ALLU) (.403)</td>
<td>English dictation (.602)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8th</td>
<td>Pearson General + DIALANG Medium</td>
<td>.650</td>
<td>65%</td>
<td>Size of English vocab (.740)</td>
<td>L1 Finnish Reading (PISA) (.403)</td>
<td>English dictation (.670)</td>
<td>Size of Finnish vocab (.282)</td>
<td>Self assessment of English reading (.272)</td>
</tr>
<tr>
<td>Gym</td>
<td>Pearson General + DIALANG Advanced</td>
<td>.692</td>
<td>69%</td>
<td>English dictation (.795)</td>
<td>Size of English vocab (.747)</td>
<td>L1 Finnish Reading (PISA) (.418)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Preliminary results – please do not quote

Copyright by DIALUKI
Stepwise multiple regression analysis, linguistic, cognitive and motivation variables

<table>
<thead>
<tr>
<th>Grade</th>
<th>Dependent variable</th>
<th>Adjusted R Squared</th>
<th>% variance</th>
<th>First IV</th>
<th>Second IV</th>
<th>Third IV</th>
<th>Fourth IV</th>
<th>Fifth IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>4th Grade</td>
<td>Pearson Young Learners Test in English</td>
<td>.526</td>
<td>53%</td>
<td>Size of English Vocab (.664)</td>
<td>Writing in L1 Finnish (.419)</td>
<td>L2 segmentation accuracy (-.584)</td>
<td>L1 Finnish Reading (ALLU) (.403)</td>
<td></td>
</tr>
<tr>
<td>8th Grade</td>
<td>Pearson General + DIALANG Medium</td>
<td>.671</td>
<td>67%</td>
<td>Size of English Vocab (.740)</td>
<td>Writing in English (.696)</td>
<td>L2 segmentation accuracy (-.641)</td>
<td>Size of Finnish Vocab (.282)</td>
<td></td>
</tr>
<tr>
<td>Gymnasi um</td>
<td>Pearson General + DIALANG Advanced</td>
<td>.708</td>
<td>71%</td>
<td>English dictation (.795)</td>
<td>Size of English Vocab (.747)</td>
<td>L1 Finnish Reading (PISA) (.418)</td>
<td>L2 segmentation accuracy (-.677)</td>
<td>Writing in English (.680)</td>
</tr>
</tbody>
</table>

Preliminary results – please do not quote
II Vocabulary profile at different CEFR reading levels

English as a foreign language

DIALUKI Study 1 gymnasium students (17–year-olds, Finnish as L1)
Vocabulary measures (English as FL)

- Nation / Schmitt et al. Vocabulary Levels Test of English
- 2000, 3000 & 5000 level tests (Version 1)
- Academic Vocabulary test (Version 2)
- for the 4th graders, a 500 and 1000 level test was created following the same design

Preliminary results – please do not quote
The Vocabulary Levels Test

- Measure the quality (depth) of vocabulary knowledge
  - “to recognize the word forms which match the definitions given” (Schmitt 2010)

- Original basis for the levels:
  - 2,000 word families were sufficient to engage in daily conversation
  - 3,000 families were thought to enable initial access to authentic reading, and
  - 5,000 families independent reading of that material (Schmitt 2010)

- Current view: knowledge of 8000 – 9000 word families needed for reading academic texts (Schmitt, Jiang & Grabe 2011)
Original Vocabulary Levels Test – a 2000 level example

1 birth
2 dust
3 operation
4 row
5 sport
6 victory

_____ game
_____ winning
_____ being born

– each level test consists of 30 such sets of words sampled from the particular frequency band (definitions easier than the target words)

Preliminary results – please do not quote
**Descriptive results – all 4 vocabulary levels tests**
(only the gymnasium students / 17–year-olds)

<table>
<thead>
<tr>
<th></th>
<th>2000 level</th>
<th>3000 level</th>
<th>5000 level</th>
<th>Acad. W.List</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean (%)</strong></td>
<td>92.2</td>
<td>79.7</td>
<td>60.1</td>
<td>69.5</td>
</tr>
<tr>
<td><strong>Median (%)</strong></td>
<td>96.7</td>
<td>86.7</td>
<td>63.3</td>
<td>73.3</td>
</tr>
<tr>
<td><strong>Std. Dev.</strong></td>
<td>10.8</td>
<td>18.7</td>
<td>21.6</td>
<td>18.6</td>
</tr>
</tbody>
</table>

n = 192 (for AWL 189)
## Vocabulary profiles at different CEFR reading levels (gymnasium)

<table>
<thead>
<tr>
<th>DIALANG READING TEST RESULT</th>
<th>n</th>
<th>2000 level (% correct)</th>
<th>3000 level (% correct)</th>
<th>5000 level (% correct)</th>
<th>Acad. W.List (% correct)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>17</td>
<td>76.9</td>
<td>58.8</td>
<td>35.5</td>
<td>44.0</td>
</tr>
<tr>
<td>A2</td>
<td>42</td>
<td>88.1</td>
<td>68.3</td>
<td>47.3</td>
<td>58.6</td>
</tr>
<tr>
<td>B1</td>
<td>48</td>
<td>93.2</td>
<td>79.1</td>
<td>58.2</td>
<td>71.0</td>
</tr>
<tr>
<td>B2</td>
<td>43</td>
<td>98.5</td>
<td>91.2</td>
<td>71.0</td>
<td>78.0</td>
</tr>
<tr>
<td>C1</td>
<td>22</td>
<td>99.2</td>
<td>96.4</td>
<td>82.6</td>
<td>85.0</td>
</tr>
<tr>
<td>C2</td>
<td>4</td>
<td>100</td>
<td>98.3</td>
<td>93.3</td>
<td>90.0</td>
</tr>
</tbody>
</table>

Note: only the DIALANG reading tests are used in this part of the study because relating the PTE General items used in the study to the CEFR levels has not yet been done.
The same in graphical format: 2000 level test: mean test result at different CEFR reading levels
3000 level test: mean test result at different CEFR reading levels

Mean of English vocabulary 3000 level part of the test - percentage correct

CEFR level based on the weighted total DIALANG score (English reading)

Preliminary results - please do not quote
5000 level test: mean test result at different CEFR reading levels

Mean of English vocabulary 5000 level part of the test - percentage correct

CEFR level based on the weighted total DIALANG score (English reading)

Preliminary results – please do not quote
Academic Word List: mean test result at different CEFR reading levels

Mean of English vocabulary Academic Word List part of the test - percentage correct

CEFR level based on the weighted total DIALANG score (English reading)
Which differences are statistically significant?

- based on 1-way Anova and Kruskal-Wallis 1-way (non-parametric) Anova analyses & Post Hoc comparisons of group means
- 2000 & 3000 level test results were skewed and group variances were unequal
- 5000 & AWL test results were normally distributed & group variances were equal
## Statistically significant differences: 2000 level test (gymnasium)

<table>
<thead>
<tr>
<th>DIALANG READING TEST RESULT</th>
<th>n</th>
<th>2000 level (% correct)</th>
<th>0.05 level differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>17</td>
<td>76.9</td>
<td>A1 vs. A2: probably no difference</td>
</tr>
<tr>
<td>A2</td>
<td>42</td>
<td>88.1</td>
<td>A2 vs. B1: difference</td>
</tr>
<tr>
<td>B1</td>
<td>48</td>
<td>93.2</td>
<td>B1 vs. B2: uncertain result</td>
</tr>
<tr>
<td>B2</td>
<td>43</td>
<td>98.5</td>
<td>B2 vs. C1/C2: no difference</td>
</tr>
<tr>
<td>C1 – C2</td>
<td>26</td>
<td>99.4</td>
<td></td>
</tr>
</tbody>
</table>

Note: C1 and C2 combined in these analyses
## Statistically significant differences: 3000 level test (gymnasium)

<table>
<thead>
<tr>
<th>DIALANG READING TEST RESULT</th>
<th>n</th>
<th>3000 level (% correct)</th>
<th>0.05 level differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>17</td>
<td>58.8</td>
<td>A1 vs. A2: no difference</td>
</tr>
<tr>
<td>A2</td>
<td>42</td>
<td>68.3</td>
<td>A2 vs. B1: difference</td>
</tr>
<tr>
<td>B1</td>
<td>48</td>
<td>79.1</td>
<td>B1 vs. B2: difference</td>
</tr>
<tr>
<td>B2</td>
<td>43</td>
<td>91.2</td>
<td>B2 vs. C1 / C2: uncertain results</td>
</tr>
<tr>
<td>C1 – C2</td>
<td>26</td>
<td>96.7</td>
<td></td>
</tr>
</tbody>
</table>
## Statistically significant differences: 5000 level test (gymnasium)

<table>
<thead>
<tr>
<th>DIALANG READING TEST RESULT</th>
<th>n</th>
<th>5000 level (% correct)</th>
<th>0.05 level differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>17</td>
<td>35.5</td>
<td>A1 vs. A2: possibly differed</td>
</tr>
<tr>
<td>A2</td>
<td>42</td>
<td>47.3</td>
<td>A2 vs. B1: possibly differed</td>
</tr>
<tr>
<td>B1</td>
<td>48</td>
<td>58.2</td>
<td>B1 vs. B2: possibly differed</td>
</tr>
<tr>
<td>B2</td>
<td>43</td>
<td>71.0</td>
<td>B2 vs. C1/C2: difference</td>
</tr>
<tr>
<td>C1 – C2</td>
<td>26</td>
<td>84.2</td>
<td></td>
</tr>
</tbody>
</table>

Note: Post Hoc results varied depending on the method used: according to the most conservative Scheffé approach, only B2 vs C1/C2 differed; Duncan’s approach indicated all differed significantly.
# Statistically significant differences: Academic Vocabulary (gymnasium)

<table>
<thead>
<tr>
<th>DIALANG READING TEST RESULT</th>
<th>n</th>
<th>Acad. W.List (%correct)</th>
<th>0.05 level differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>17</td>
<td>44.0</td>
<td>A1 vs. A2: differed</td>
</tr>
<tr>
<td>A2</td>
<td>42</td>
<td>58.6</td>
<td>A2 vs. B1: differed</td>
</tr>
<tr>
<td>B1</td>
<td>48</td>
<td>71.0</td>
<td>B1 vs. B2: no difference</td>
</tr>
<tr>
<td>B2</td>
<td>43</td>
<td>78.0</td>
<td>B2 vs. C1/C2: possibly differed</td>
</tr>
<tr>
<td>C1 – C2</td>
<td>26</td>
<td>85.8</td>
<td></td>
</tr>
</tbody>
</table>

Note: Scheffé indicated B2 vs. C1/C2 did not differ whereas Duncan did; both indicated B1 & B2 did not differ

Preliminary results – please do not quote
Summary 1: predictors of reading in English:

- Cognitive / psycholinguistic variables predicted about 20–30% of variance
  - we don’t yet have coded all tasks (e.g., phonological processing)
- Tasks in L2 (English) usually better predictors than similar / same tasks in L2
- Motivation predicted about 20–40% of variance
  - English Self-concept best predictor; also Self-Regulation
Summary 2: predictors of reading in English

- Linguistic variables predicted about 50–70% of variables
- Better predictors than cognitive or motivation variables
- L2 vocabulary consistently good predictor
- L2 dictation even better for one age group
- Reading in L1 also a significant predictor
- But the better predictors are more integrated / holistic. So in what sense are they “diagnostic”, rather than simply overlapping?
Conclusions about CEFR levels in reading (so far)

- Which linguistic and other features / combinations of features characterise CEFR levels (in reading, writing, speaking, listening, ...)?
- Can the levels be distinguished in terms of structures, vocabulary or other features?
- This study: vocabulary level profiles at CEFR reading levels A1 to C1 / C2
- 2000, 3000, 5000 levels & Academic Vocabulary: consistent, linear relationship between vocabulary and reading comprehension
  - e.g. B2 level: 99 – 91 – 71 – 78%

Preliminary results – please do not quote
Next steps

- More robust measurement of reading in English (DIALANG + PTE General)
- More detailed analysis of reading levels’ in terms of types / ’subskills’ of reading
  - PISA: Access & retrieve; Interpret & integrate; Reflect & evaluate (cf. the PISA and Pearson PTE Academic studies)
- More analyses in relation to the CEFR levels
  - other features of language and cognition besides L2 vocabulary (e.g. segmentation task)
- In Study 2 & 3: more comprehensive measurement of structures
- Extending the analyses to the 8th graders’ data (already done for L2 vocabulary) and to the Finnish / Russian population

Preliminary results – please do not quote
Thank you for your attention!

ari.huht@jyu.fi

c.alderson@lancaster.ac.uk

www.jyu.fi/dialuki
# Overall L2 vocabulary results: 8th graders vs. gymnasium students

<table>
<thead>
<tr>
<th></th>
<th>1000 level</th>
<th>2000 level</th>
<th>3000 level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>8th only</td>
<td>8th gym</td>
<td>8th gym</td>
</tr>
<tr>
<td>Mean (%)</td>
<td>95.3</td>
<td>71.6</td>
<td>92.2</td>
</tr>
<tr>
<td>Median (%)</td>
<td>100</td>
<td>76.7</td>
<td>96.7</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>7.1</td>
<td>23.3</td>
<td>10.8</td>
</tr>
<tr>
<td>Minimum</td>
<td>50</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>Maximum</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Preliminary results – please do not quote

Copyright by DIALUKI
Vocabulary results at different CEFR reading levels: 8th grade vs. gymnasium

<table>
<thead>
<tr>
<th>DIALANG READING TEST</th>
<th>n</th>
<th>1000 level</th>
<th>2000 level (% correct)</th>
<th>3000 level (% correct)</th>
<th>5000 level (% correct)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>8 only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>76</td>
<td>17</td>
<td>92.3 60.9 76.9</td>
<td>34.2 58.8 35.5</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>87</td>
<td>42</td>
<td>97.0 75.0 88.1</td>
<td>48.8 68.3 47.3</td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>22</td>
<td>48</td>
<td>99.4 93.3 93.2</td>
<td>76.7 79.1 58.2</td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>5</td>
<td>43</td>
<td>100 96.0 98.5</td>
<td>92.0 91.2 71.0</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>–</td>
<td>22</td>
<td>99.2</td>
<td></td>
<td>96.4 82.6</td>
</tr>
<tr>
<td>C2</td>
<td>–</td>
<td>4</td>
<td>100 98.3 93.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Preliminary results - please do not quote
Which of the differences are statistically significant?

- based on 1-way Anova and Kruskal-Wallis 1-way (non-parametric) Anova analyses & Post Hoc comparisons of group means
- The results of all the three vocabulary tests (1000, 2000 & 3000 level) were skewed and group variances were unequal
### Statistically significant differences: 1000 level test (8th grade)

<table>
<thead>
<tr>
<th>DIALANG READING TEST RESULT</th>
<th>n</th>
<th>1000 level (% correct)</th>
<th>Only possible difference between A1 and A2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>76</td>
<td>92.3</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>87</td>
<td>97.0</td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>22</td>
<td>99.4</td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>5</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
## Statistically significant differences: 2000 level test (8th grade)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>76</td>
<td>60.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>87</td>
<td>75.0</td>
<td>A1 vs. A2: no difference</td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>22</td>
<td>93.3</td>
<td></td>
<td>B1 vs. B2: no difference</td>
</tr>
<tr>
<td>B2</td>
<td>5</td>
<td>96.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Preliminary results – please do not quote
### Statistically significant differences: 3000 level test (8th graders)

<table>
<thead>
<tr>
<th>DIALANG READING TEST RESULT</th>
<th>n</th>
<th>3000 level (% correct)</th>
<th>Note: also the difference between A2 and B1 is uncertain as different post hoc measures gave conflicting results</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>76</td>
<td>34.2</td>
<td>A1 vs. A2: no difference</td>
</tr>
<tr>
<td>A2</td>
<td>87</td>
<td>48.8</td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>22</td>
<td>76.7</td>
<td>B1 vs. B2: no difference</td>
</tr>
<tr>
<td>B2</td>
<td>5</td>
<td>92.0</td>
<td></td>
</tr>
</tbody>
</table>

Copyright by DIALUKI