How expertise changes perceptual processes

Evidence from Eye Movement studies in reading

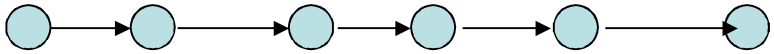
Raymond Bertram

Department of Speech-Language Pathology and Psychology
University of Turku, Finland

Keynote at ETRA 2020 for EMIP, 2.6.2020

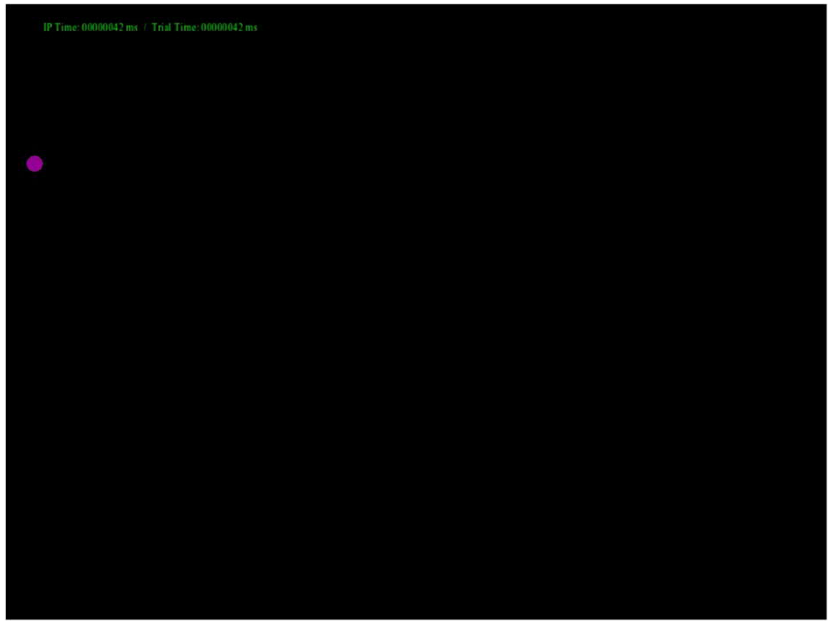
* How does expertise change perceptual processes?
* Comparison: **novice – expert => extreme differences**

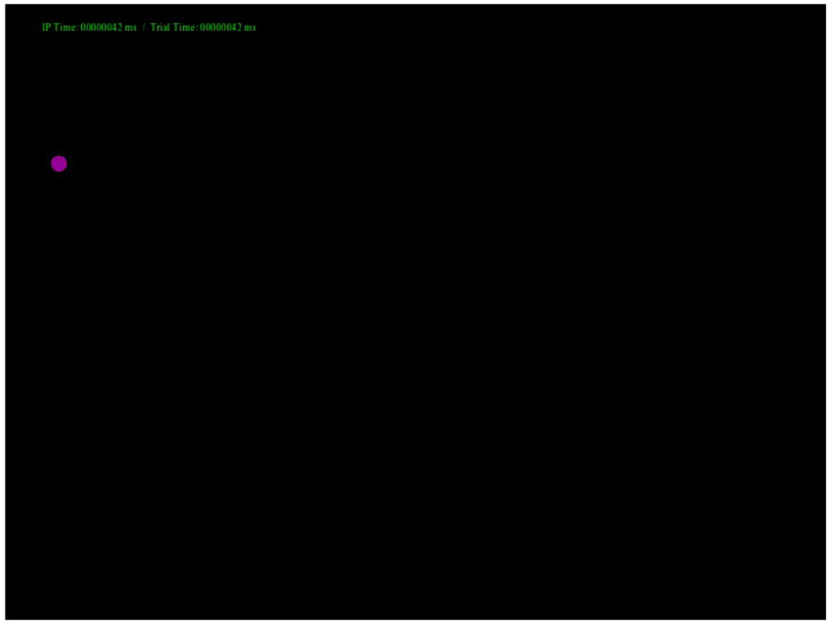
Novel reader (1st grade fall) vs expert reader (adult student)



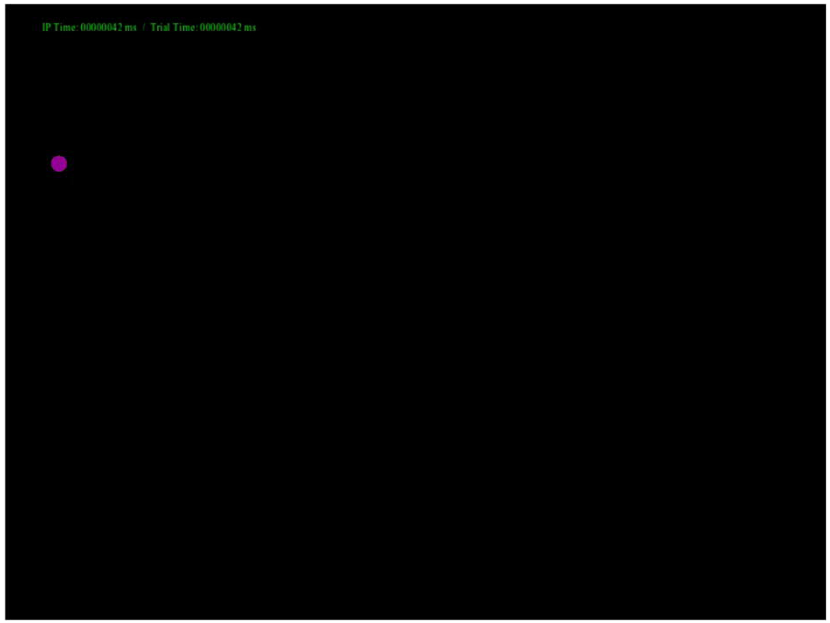
‘Äiti arveli, että juusto oli homeessa’
‘Mom thought, that the cheese was moldy’

Pupil, 1st grade, fall vs. Skilled adult





Pupil, 1st grade, fall vs. Skilled adult

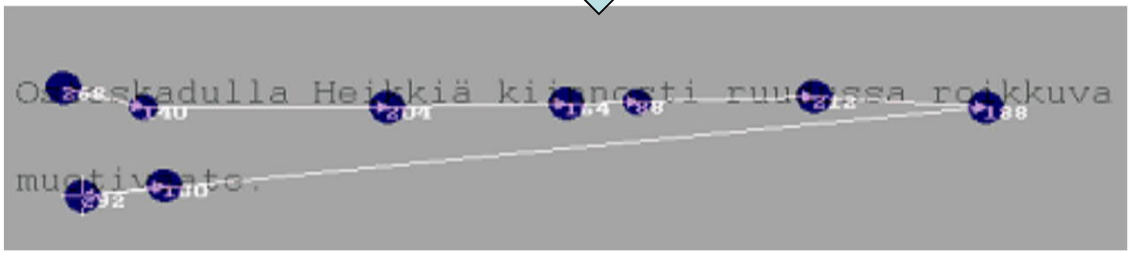


21 fixations, mean fixation duration 375 ms, 6 regressions, no skips, 7.86 seconds altogether, 1.3 s per word

4 fixations, mean fixation duration 235 ms, 1 regression, 2 skips, 0.94 seconds altogether, 150 ms per word

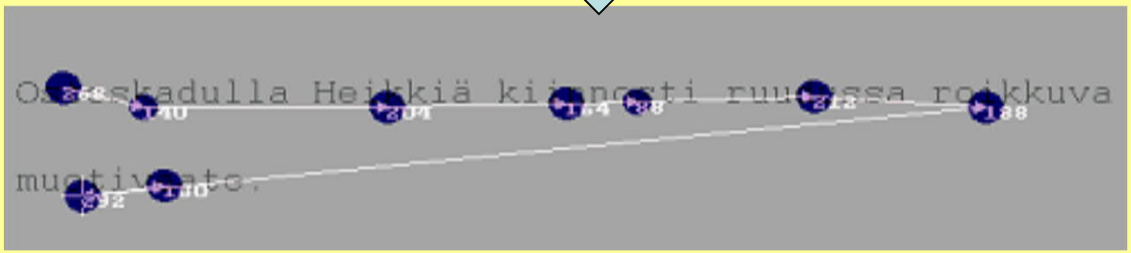
Beginning reader

Expert reader



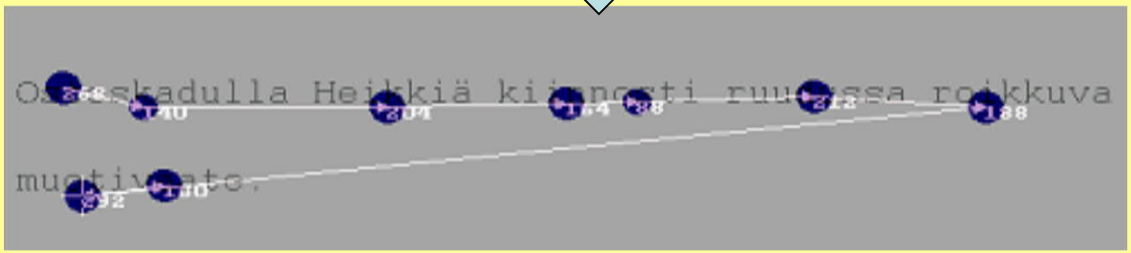
Beginning reader

Expert reader



Beginning reader

Expert reader



```
File  Edit  View  Search  Terminal  Help
 1 #include <iostream>
 2 #include <string>
 3
 4 int g_variable = 10;
 5
 6 auto main(int argc, char* argv[]) -> decltype(0){
 7     int l_variable = 20;
 8
 9     const char* cl_variable = "Const Character String";
10     static const char* scl_variable = " Static Const Character String";
11
12     return 0;
13 }
```

Survey of the presentation

1. **Use of visual cues**
2. **Development of perceptual span => horizontal & vertical**
3. **Speed of information uptake** = foveal & parafoveal
    - The Disappearing text paradigm
    - The Boundary paradigm
4. **Push the limits**
5. **Some remaining issues**
    - Text reading vs code reading
    - Data quality
6. **Summary with suggestions for future EMIP research**

All these issues discussed in context of expertise development

1. Use of visual cues in reading

<u>Visual cues are used to facilitate reading</u>

Text Level: Headings, subheadings, whitelines, indentation, bolding, underlining

$\Rightarrow$ Inform you about organization of the text

Sentence level: Spaces, specific character combinations

$\Rightarrow$ Inform you about word boundaries

1. Use of visual cues in reading

<u>Visual cues are used to facilitate reading</u>

Text Level: Headings, subheadings, whitelines, indentation, bolding, underlining

$\Rightarrow$ Inform you about organization of the text

Sentence level: Spaces, specific character combinations

$\Rightarrow$ Inform you about word boundaries

Readingwithoutspacingis30%slowerthanreadingwithspaces

1. Use of visual cues in reading

<u>Visual cues are used to facilitate reading</u>

Text Level: Headings, subheadings, whitelines, indentation, bolding, underlining

$\Rightarrow$ Inform you about organization of the text

Sentence level: Spaces, specific character combinations

$\Rightarrow$ Inform you about word boundaries

Readingwithoutspacingis30%slowerthanreadingwithspaces

Reading without spacing is 30% slower than reading with spaces
(Rayner & Pollatsek, 1996)

1. Use of visual cues in reading

<u>Visual cues are used to facilitate reading</u>

Text Level: Headings, subheadings, whitelines, indentation, bolding, underlining

⇒ Inform you about organization of the text

Sentence level: Spaces, specific character combinations

⇒ Inform you about word boundaries

Readingwithoutspacingis30%slowerthanreadingwithspaces

Reading without spacing is 30% slower than reading with spaces *(Rayner & Pollatsek, 1996)*

Not in Japanese though (Sainio, Bingushi, Hyönä, Bertram, 2007)

1. Use of visual cues in reading

Visual cues are used to facilitate reading

Word level: To denote sublexical structure

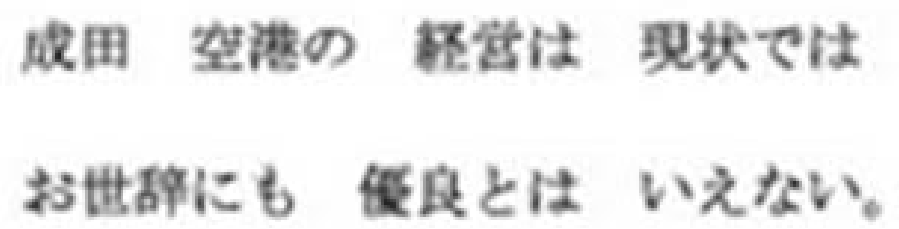
⇒ Inform you about the structure of a word

E.g, the space in English compounds: garden party
E.g., the hyphen in Finnish compounds: vaihto-ohjelma ’exchange program’

1. Use of visual cues in reading

<u>Visual cues are used to facilitate reading:</u>

Assumption is that low-proficient readers need more visual cues



Japanese Children
L2 Readers: used to space but not character sequence info to find word boundaries

1. Use of visual cues in reading

<u>Visual cues are used to facilitate reading:</u>

Assumption is that low-proficient readers need more visual cues

成田　空港の　経営は　現状では

お世辞にも　優良とは　いえない。

Japanese Children
L2 Readers: used to space but not character sequence info to find word boundaries

Varis kirjoittaa kirjeen

Hei serk-ku!

Mi-tä si-nul-le kuu-luu?
Mi-nul-le kuu-luu tä-nään hy-vää.

Ei-len mi-nul-le kuu-lui huo-no-a.
Jäin luk-ko-jen taak-se.
Ei se mi-nun syy-ni ol-lut.
Jo-ku o-li jät-tä-nyt hä-kin vää-rään paik-kaan.

Finnish 1st and 2nd graders

1. Use of visual cues in reading

<u>Visual cues are used to facilitate reading:</u>

Assumption is that low-proficient readers need more visual cues

There is evidence that supports this notion

vaihto-ohjelma **ulko-ovi**

Adults facilitated by hyphen when processing long but not short compounds:

2nd graders: faciltitated by hyphen when processing long & short compounds

vaihto-ohjelma **ulko-ovi**

Bertram & Hyönä, 2013; Häikiö, Bertram, Hyönä, 2011

Probably need of extra visual cues not only on word level

1. Use of visual cues in code reading

Use of visual cues in reading code (code comprehension, addition, debugging)

Code conventions: Colors, Indentation, Capitalization

* **Visual cues more important for novice/beginning coders?**

```
if  x == 1:
      print (”x is 1”)
      print (”x is odd”)
```

```
if  x == 1:
print (”x is 1”)
print (”x is odd”)
```

1. Use of visual cues in code reading

Use of visual cues in reading code (code comprehension, addition, debugging)

Code conventions: Colors, Indentation, Capitalization

- **In general, to what extent are they used by novices vs expert**
  ⇒ EM guidance

* **Visual cues more difficult in L2 when not compatible with L1**

cross-language interference? (Strong L1 Javascript – Weak L2 C++)

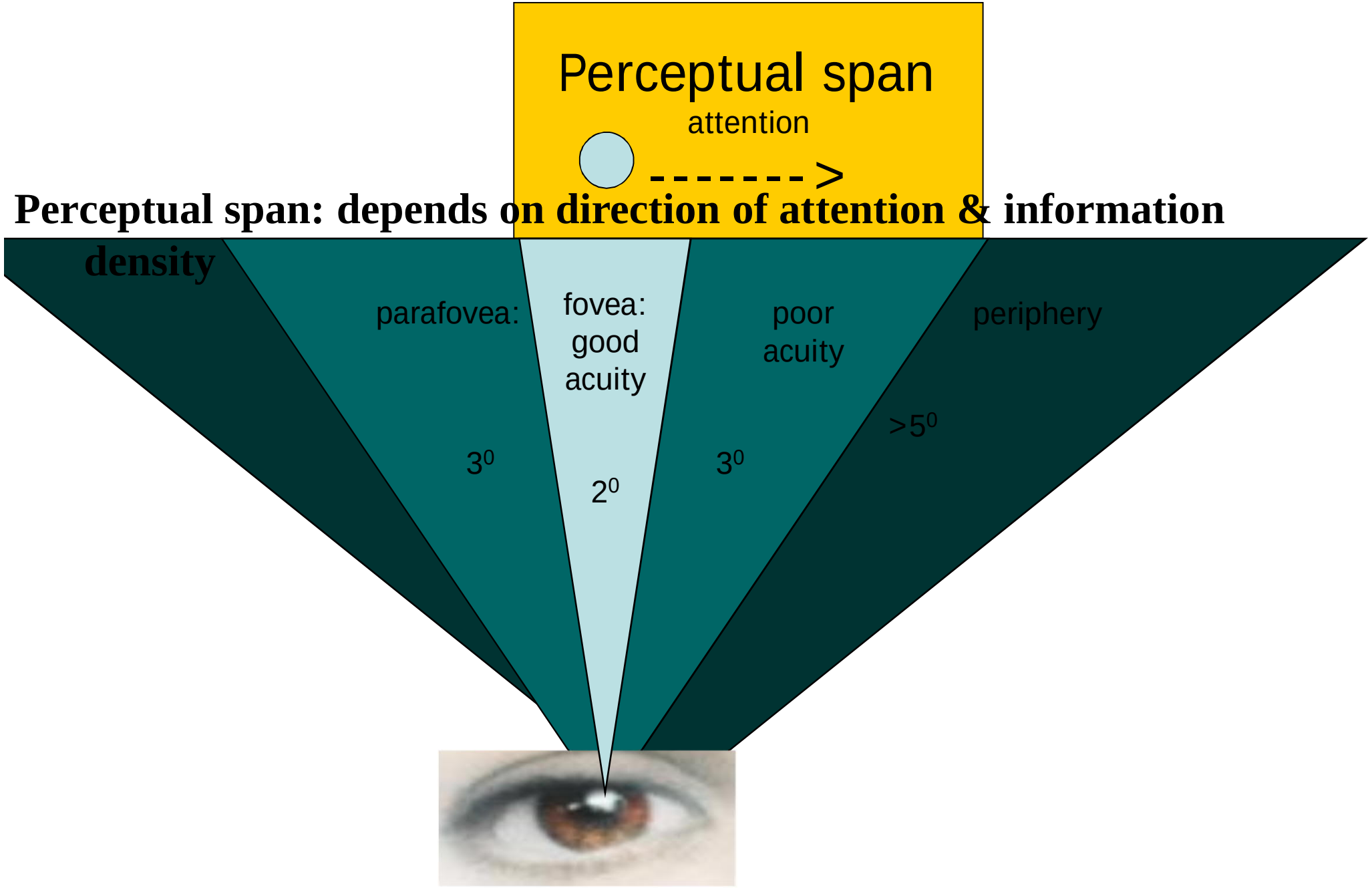
=> Few EM exp. in EMIP on use visual cues – expertise development

=> Measures affected by vc: durations, saccadic amplitude, scan paths

2. Development of the perceptual span

- Perceptual span: Area from which we extract useful information during reading
- Size of perceptual span measured by moving window paradigm

2. Development of the perceptual span



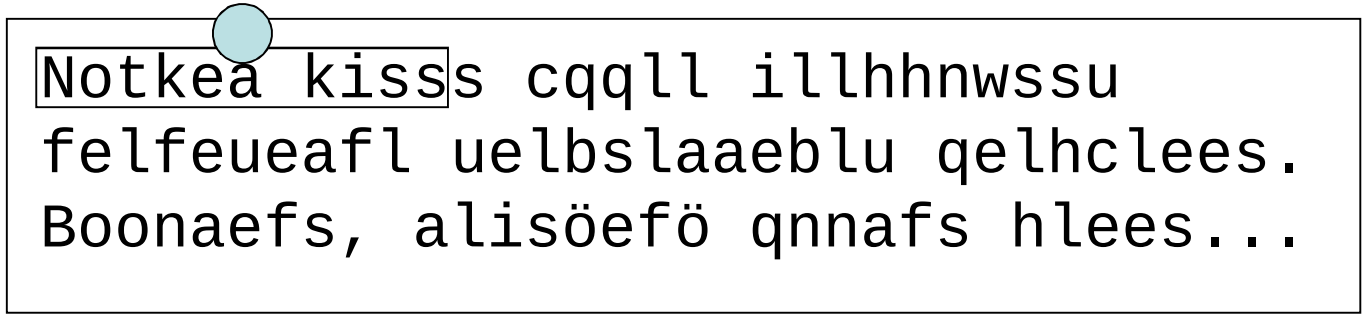
Perceptual span: depends on direction of attention & information

2. Development of the perceptual span

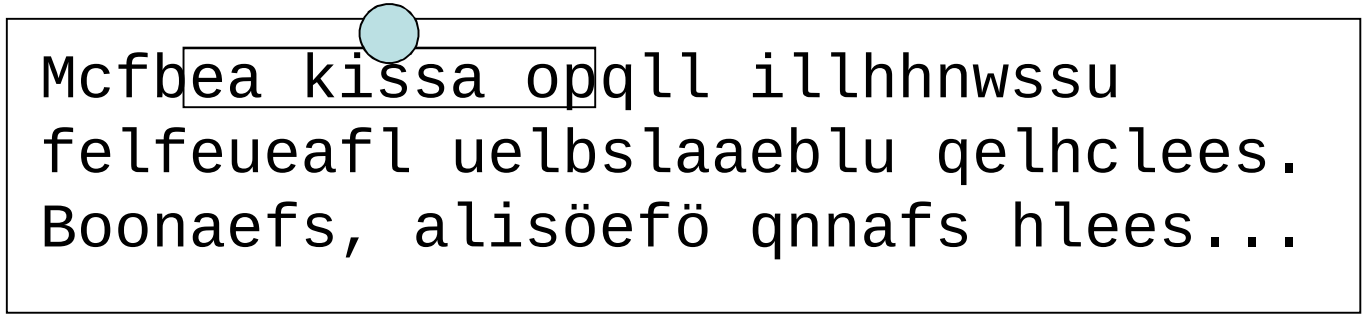
Moving window (Häikiö, Bertram, & Hyönä, 2009)

```
Notkea kissa oppii liikkumaan
taitavasti vaikeissakin paikoissa.
Suuresta, sileästä puusta kissa...
```

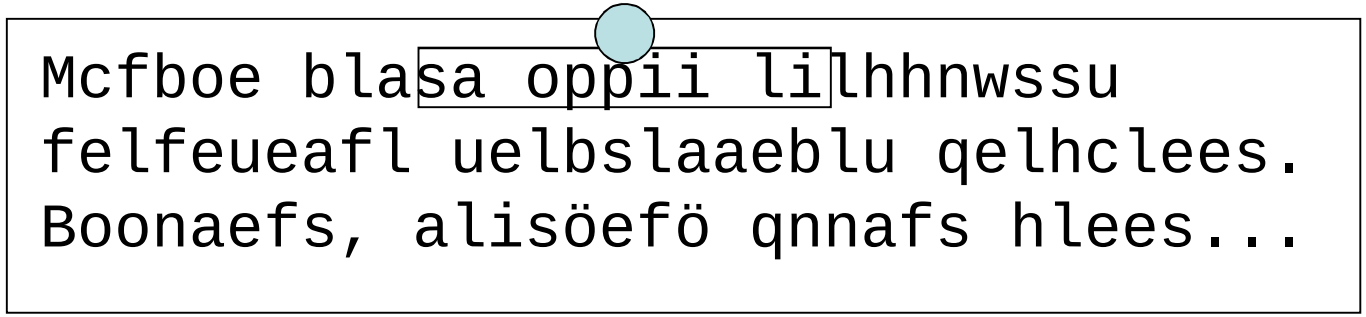
[The flexible cat learns to move skillfully even in hard places. The cat [gets down] from a large smooth tree...]



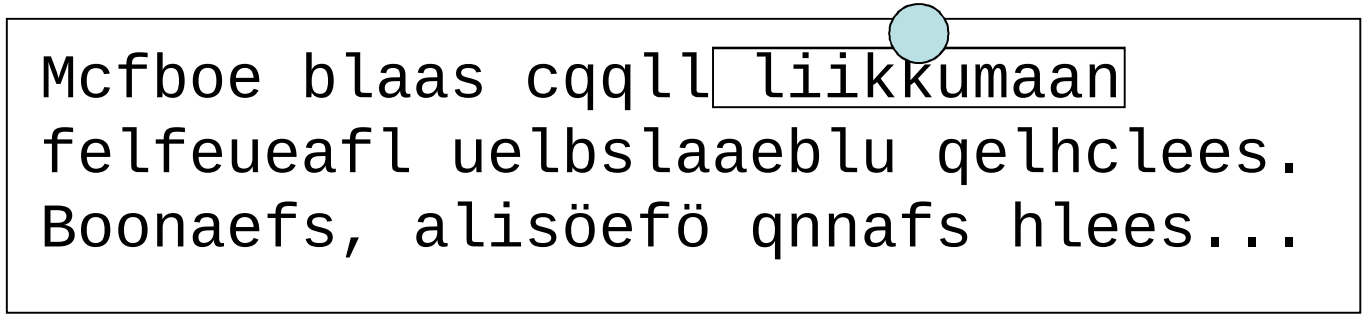
Window size 11 characters



Window size 11 characters



Window size 11 characters



Window size 11 characters

2. Development of the perceptual span

Table 3

Type of information obtained to the right of fixation in the current study and Rayner (1986)

	Second grade	Fourth grade	Sixth grade	Adults
Letter identity (current study)	5	7	9	9
Letter feature (Rayner Experiments 2 and 3)	7	11–12	11–12	11–12
Word length (Rayner Experiment 1)	11	11	14	14-15

SPAN	W=5	MANIPULATION	EXCLUDE
Letter identity (LI):	oppii	illhhnwssu (liikkumaan)	– LI
Letter feature (LF):	oppii	xxxxxxxxxx (liikkumaan)	– LI + LF
Word length (WL):	oppii	xxxxxxxxxx (liikkumaan)	– LI, LF + WL

2. Development of the perceptual span

LI LF WL

Table 3

Type of information obtained to the right of fixation in the current study and Rayner (1986)

	Second grade	Fourth grade	Sixth grade	Adults
Letter identity (current study)	5	7	9	9
Letter feature (Rayner Experiments 2 and 3)	7	11–12	11–12	11–12
Word length (Rayner Experiment 1)	11	11	14	14-15

2. Development of the perceptual span

Table 3

Type of information obtained to the right of fixation in the current study and Rayner (1986)

	Second grade	Fourth grade	Sixth grade	Adults
Letter identity (current study)	5	7	9	9
Letter feature (Rayner Experiments 2 and 3)	7	11–12	11–12	11–12
Word length (Rayner Experiment 1)	11	11	14	14-15

- Growing expertise, growing span; assymptote reached after 6 years of reading experience
- Would that be the same when gaining expertise in reading code?

2. Development of the perceptual span

Window size 11 characters

Notkea kisss cqqll illhhnwssu felfeueafl uelbslaaeblu qelhclees. Boonaefs, alisöefö qnnafs hlees...

* Additional need: vertical perceptual span to capture non-linear reading

INTERNATIONAL JOURNAL OF PSYCHOLOGY, 2010, 45 (1), 40–47

Psychology Press
Taylor & Francis Group

Vertical perceptual span and the processing of visual signals in reading

Fabrice Cauchard and Hélène Eyrolle

University of Toulouse, Toulouse, France

Jean-Marie Cellier

Ecole Pratique des Hautes Etudes, Paris and University of Toulouse, Toulouse, France

Jukka Hyönä

University of Turku, Turku, Finland

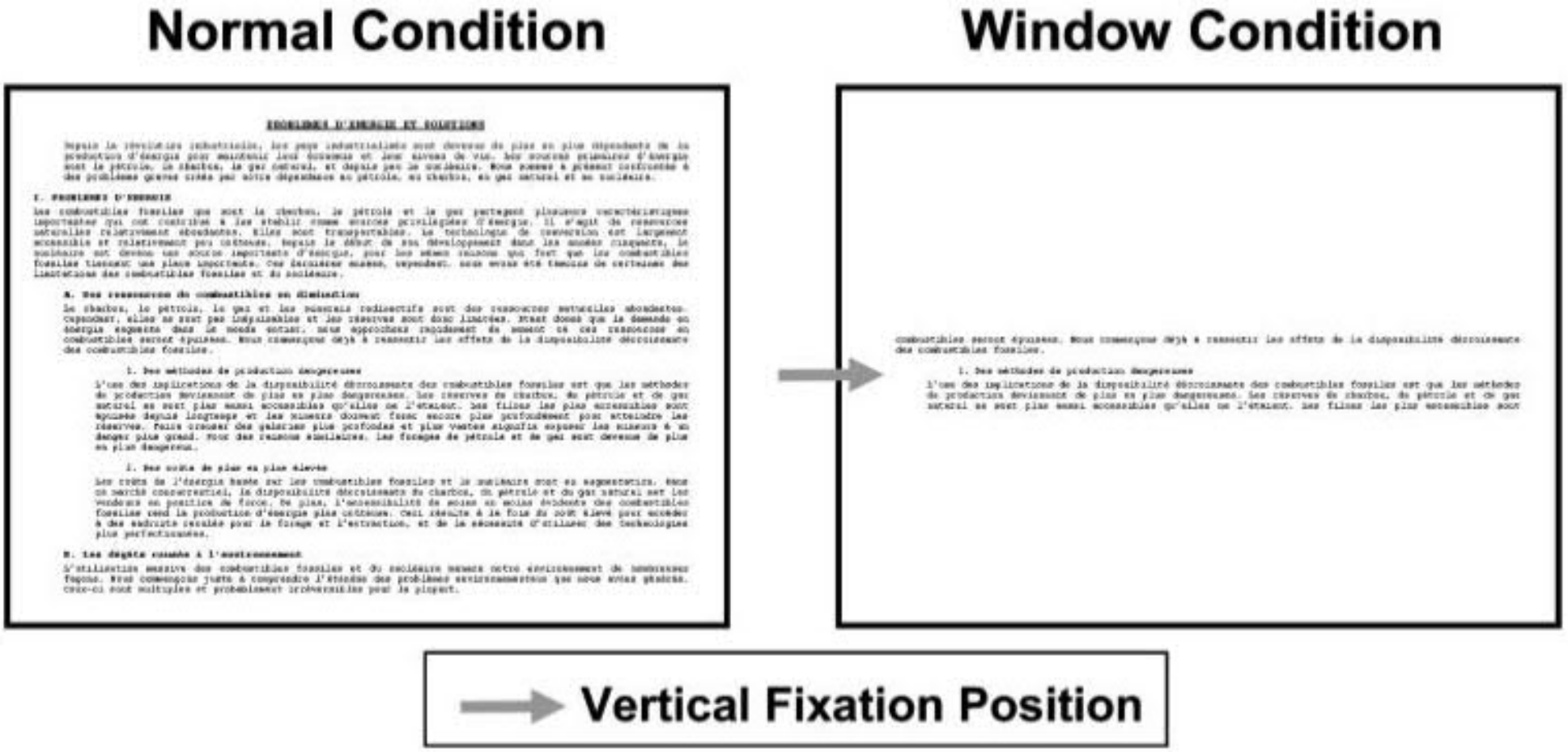


Figure 1. Illustration of the gaze-contingent window technique used in the experiment.

- Normal condition: => More returns and time spent on (sub)headings
- Code readers: => Low expertise: Linear reading => small windows sufficient

3. Speed of information processing: foveal

- Speed information uptake foveal area: **disappearing text paradigm**
- Adults have no cost when fixated word disappears after 60ms
  (Rayner et al., 2003; Liversedge et al.,2004).

```
0 ms:  Sam wore the horrid coat though his pretty girlfriend complained.
                             *
60ms:  Sam wore the horrid      though his pretty girlfriend complained.
                             *
```

Sentence RT	Normal	Disappearing	No Difference
N	3286	3327	*t*s < 1

- Different with 8-9 year olds reading longer words e.g., sairaala
  (Blythe, Häikiö, Bertram, Liversedge, Hyönä, 2011)

⇒ Can be used as a way to measure expertise, also in code reading

3. Speed of information processing: parafoveal

- Cost when parafoveal word disappears after 60ms => underlines importance of parafoveal word in reading (Rayner, Liversedge, & White., 2005).

```
0 ms:  Sam wore the horrid coat though his pretty girlfriend complained.
                               *

60ms:  Sam wore the horrid coat        his pretty girlfriend complained.
                               *
```

Sentence RT	Normal	Disappearing	Difference
N+1	3449	4045	$p < .05$

- Different with children making less use of parafoveal info?

⇒ Can be used as a way to measure expertise, also in code reading

⇒ Also boundary paradigm used to assess nohana of parafov. proc.

3. Speed of information processing: parafoveal

- Cost when parafoveal word disappears after 60ms => underlines importance of parafoveal word in reading (Rayner, Liversedge, & White., 2005).

```
0 ms:  Sam wore the horrid coat though his pretty girlfriend complained.
                               *

60ms:  Sam wore the horrid coat        his pretty girlfriend complained.
                               *
```

Sentence RT	Normal	Disappearing	Difference
N+1	3449	4045	$p < .05$

- Different with children making less use of parafoveal info?

⇒ Can be used as a way to measure expertise, also in code reading

⇒ Also boundary paradigm used to assess nature of parafov. proc.

3. Speed of information processing: parafoveal

- Research demonstrates that parafoveal processing is critical for normal efficient reading.
- Parafoveal preview benefit
- How important is parafoveal processing for reading code?
- Disappearing code paradigm, or code boundary paradigm
- Also these paradigms could be vertically orientated
- Development in parafoveal processing ⇔ Gaining expertise

4. Push the limits

- Finnish L2 speakers processing short high-frequency words (auto, posti)

  $\Rightarrow$ no insight in level of expertise

- Real challenge => Morphophonological variation

Morphophonological Variation

consonant gradation (p, k, t):

	ka**t**u =>	ka**d**ulla ’on the street’
	il**t**a =>	il**l**alla ’evening’ => ’in the evening’
	tau**k**o =>	tauolla ’in the break’
But	katu =>	katua ’street (ptv)
	katu =>	katuna ’as a street’
keitin =>	keittimellä	’with the cooker’
	(4 morphophon. rules in one inflection)	

ystävyy**s**, ystävyy**d**en, ystävyy**tt**ä, ystävyy**t**iin, ystävyy**ks**ien

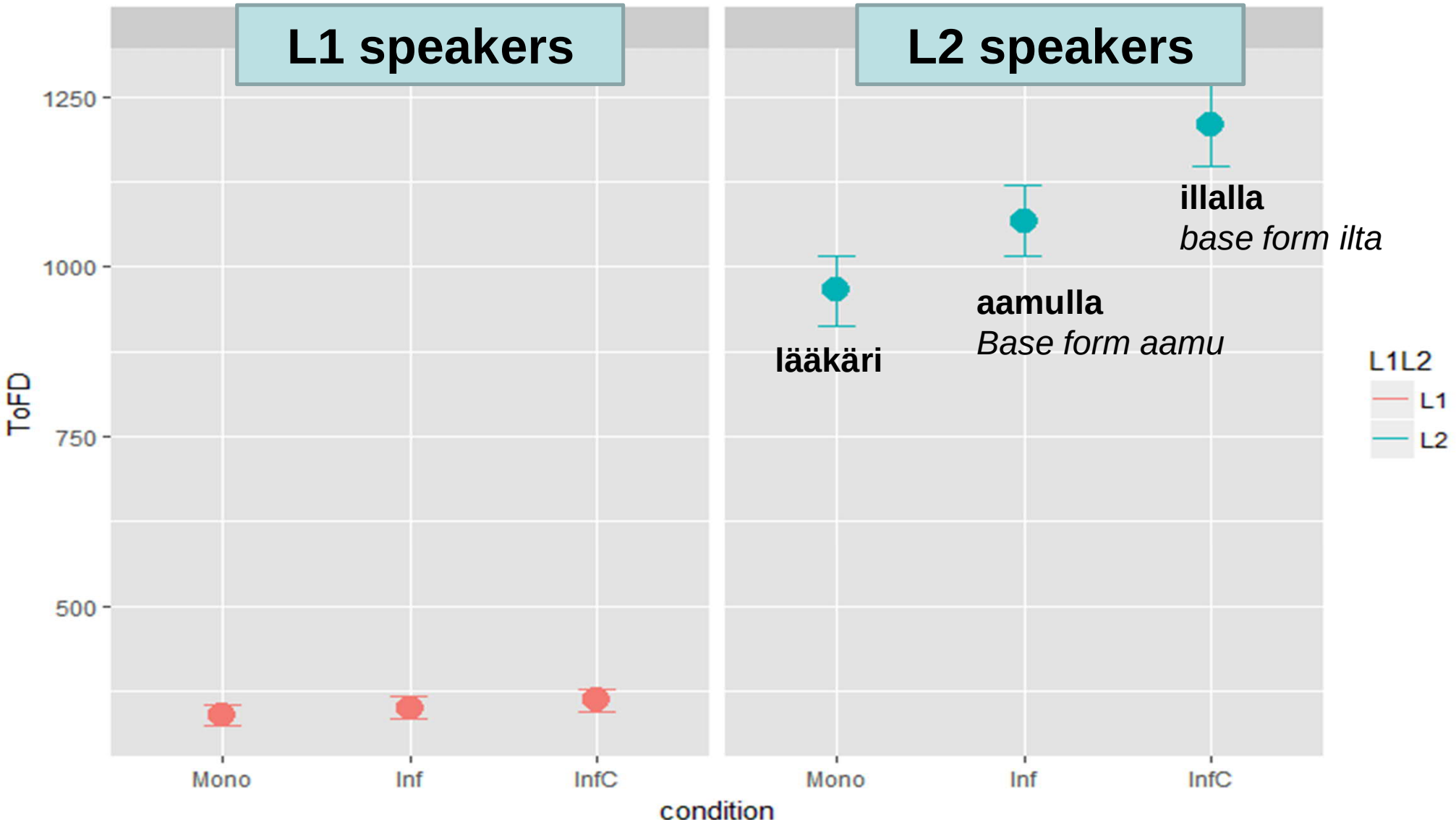
Morphophonological Variation

- L2-speakers disproportionate problems with Consonant Gradation (*Salmela*, *Bertram*, *Lehtonen*, *& Vainio*, *submitted*).

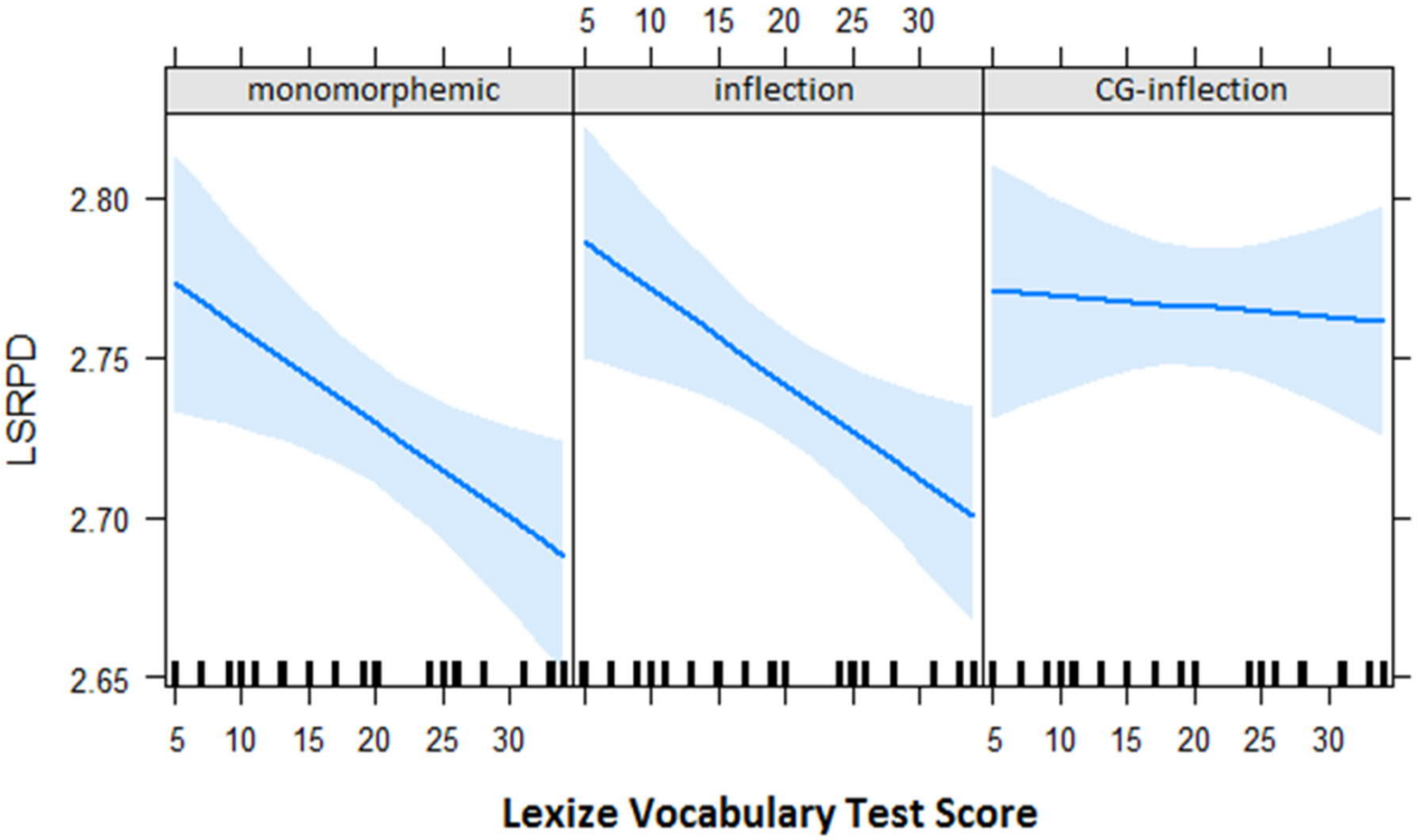
Mono: Koska **lääkäri** oli neuvonut minua lepäämään, päätin jäädä kotiin.
Inf: Koska **aamulla** oli pilvistä, otin mukaani sateenvarjon. - **aamu**
InfC: Koska **illalla** oli huono ilma, en mennyt ulos – **ilta (CG)**

Morphophonological Variation

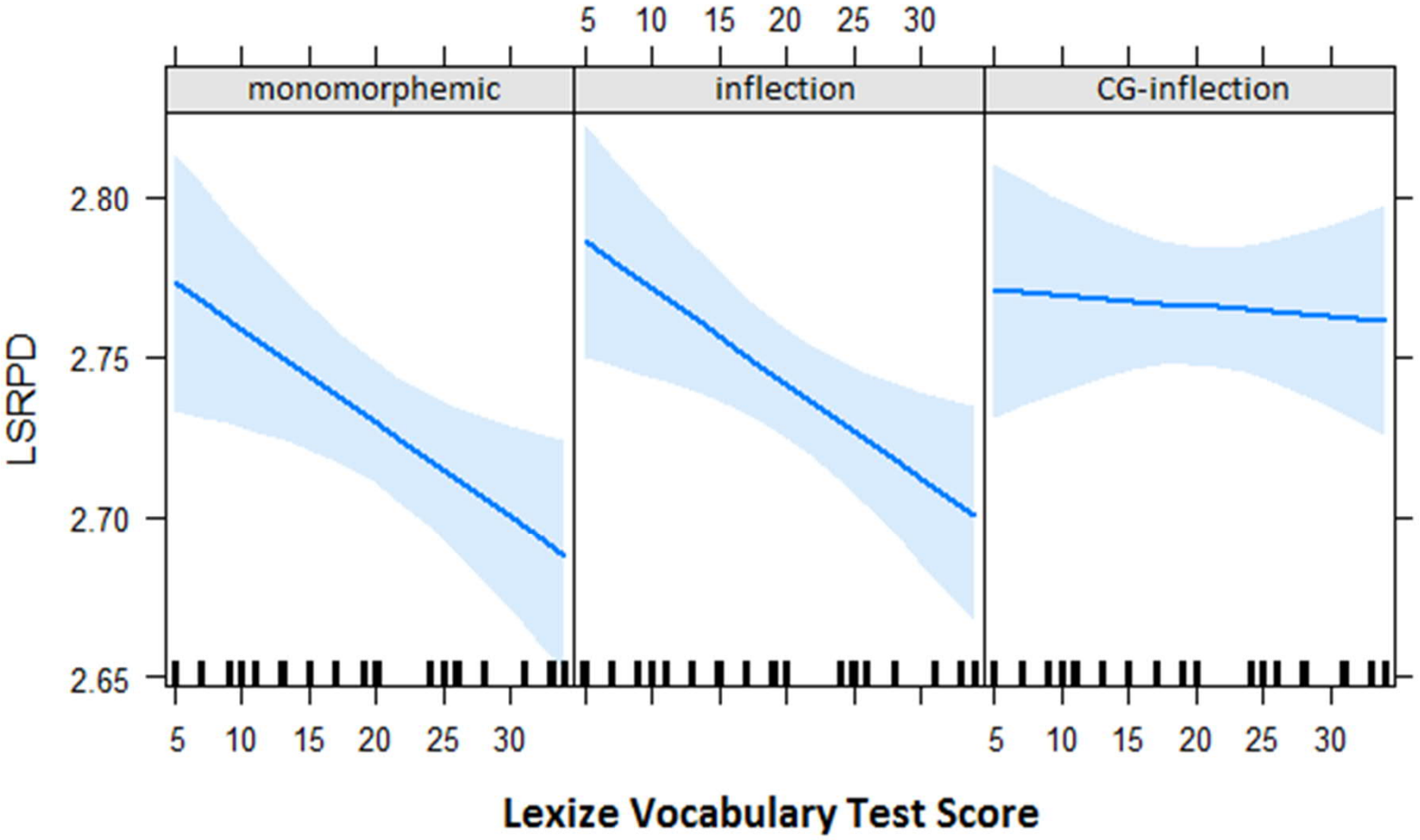
- L2-speakers disproportionate problems with Consonant Gradation (*Salmela, Bertram, Lehtonen, & Vainio, submitted*).



Morphophonological Variation



4. Push the limits



=> Don’t use too easy code snippets when trying to assess EM-patterns related to expertise

5. Text vs Code Reading: Some Remaining issues

a. Comparing text reading with code reading

- What drives EM behavior?
- Text Reading: word length, frequency, predictability directly related to EM measures (durations, saccadic amplitude, skips, regressions).
- These factors have been assessed wrt expertise development as well (e.g., effect of WL decreases with age).
- Code Reading: Factors that may be linked to EM measures: Place in hierarchy, degree of repetition, relevance, line length, visual cues

$\Rightarrow$ systematic investigations

- Interaction of expertise and these factors?

5. Text vs Code Reading: Some Remaining issues

b. Data quality: 5 experts vs 12 novices (also in text reading studies)

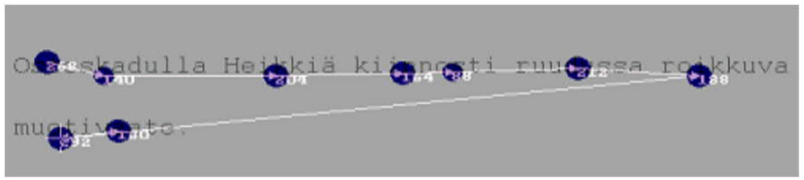
Collaborative research efforts => Next to cross-laboratory and cross-linguistic studies, development of Eye-tracking corpora for reading:

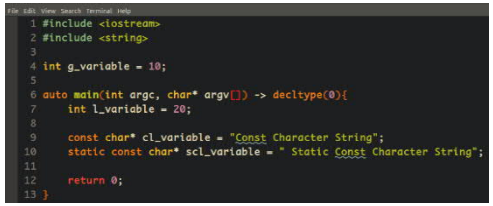
- English (Frank et al., 2013; Luke & Christianson, 2018)
- German (Kliegl et al., 2006)
- Hindi (Husain et al., 2014)
- Russian (Laurinavichyute et al., 2019).
- English-French (Pynte & Kennedy, 2006; Whitford & Titone, 2012)
- English-Dutch: GECO (Cop et al., 2017)

- MECO – (Siegelman, Kuperman et al., in prepartion): Multilingual eye movement corpus https://meco-read.com/

  English, Dutch, Finnish, German, Hebrew, Spanish, Turkish, Russian, Italian, Norvegian, Greek, Korean as L1 & all of these for English as L2

5. Summary & suggestions for future EMIP research





1. Visual cues heavily used on word, sententence and text level

⇒ Segmentation into smaller units, text organisation

⇒ Language proficiency related to usefulness of visual cues

⇒ Code: visual cue exploitment related to expertise?

2. Perceptual span grows with development, levels off at grade 6

⇒ Developmental pattern in code reading

⇒ worth considering to investigate vertical span

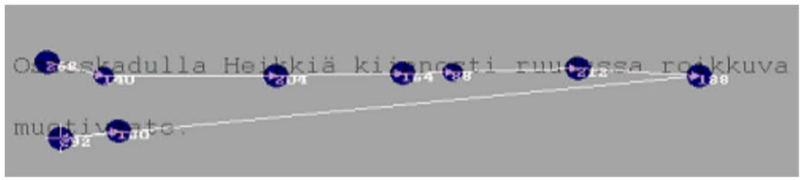
3. Parafoveal processing crucial for fluent reading.

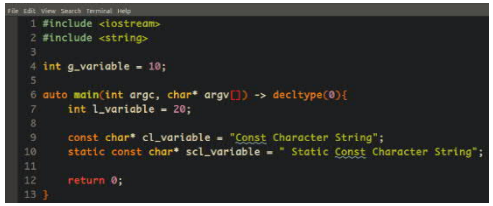
⇒ Information extraction of upcoming word affords fast reading rate

⇒ Code disappearing and code boundary paradigm to assess development of parafoveal processing in code reading

⇒ Possibly also in vertical direction

5. Summary & suggestions for future EMIP research





4. Large-scale corpora or cross-linguistic enterprises

⇒ plenty of subjects to enhance data quality

⇒ background questionnaires and tests that can be linked to expertise development (e.g., **memory span**)

5. Direct comparison between text reading & code reading

⇒ Factors that predict eye movement behavior in reading known, in coding less established.

⇒ Interaction of factors with expertise

⇒ Expertise may play out in certain measures (e.g., scan path) but not in others (e.g., dwell time, saccadic amplitude)

KIITOS

&

HAVE A GOOD CONFERENCE!