

# Online Experiments for Language Scientists

Lecture 4: Self-paced reading

Kenny Smith

[kenny.smith@ed.ac.uk](mailto:kenny.smith@ed.ac.uk)

# Matters arising from the lab

- “Will I have to write a completely new experiment from scratch for the final assignment?”
- “How do I make the prompt appear **above** the buttons on an html-button-response trial?”

# Enochson & Culbertson (2015)

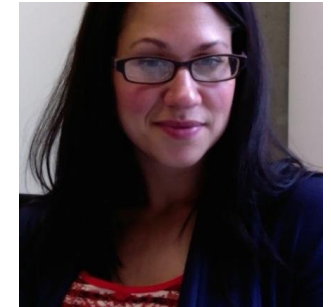
Enochson, K., & Culbertson, J. (2015).  
Collecting Psycholinguistic Response Time  
Data Using Amazon Mechanical Turk. *PLoS  
ONE, 10*, e0116946.

Three self-paced reading experiments on  
MTurk

- Do small but meaningful RT differences  
seen in lab studies replicate online, despite  
reduced experimental control and increased  
variability in e.g. participant hardware?



**Kelly Enochson**  
(formerly George  
Mason University)



**Jenny Culbertson**  
(Edinburgh)

# Self-paced reading

Demo with this week's lab code!

# Sample size, study duration etc

- Self-reported native speakers of English
- N=34 (Exp 1), 82 (Exp 2), 60 (Exp 3)
- 96-120 items per experiment (mainly fillers)
- 20 minutes, **\$1**

# Test items and predicted effects (Experiment 1)

Filler-gap (in all sentences)

- ***Which antique** was the maid polishing \_\_\_\_ in the study?*

Full DP vs pronoun

- *Which antique was **the maid** polishing in the study?*
- *Which antique was **she** polishing in the **upstairs** study?*

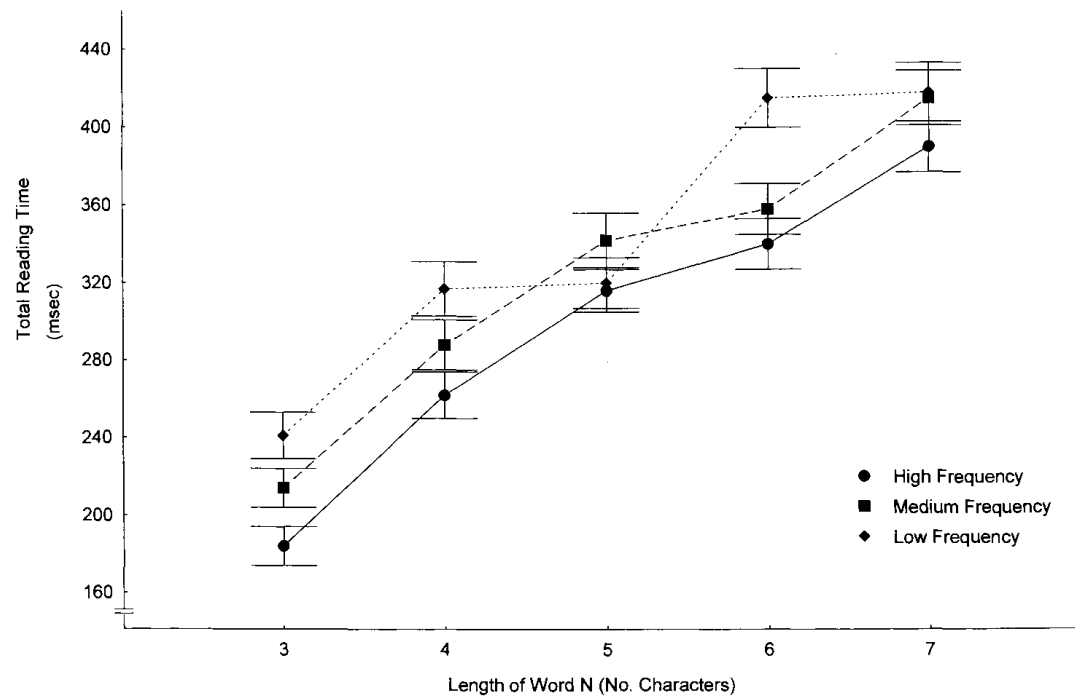
Agreement attraction

***Which antique** was the maid polishing in the study?*

***Which antiques** was the maid polishing in the study?*

# Residual reading times

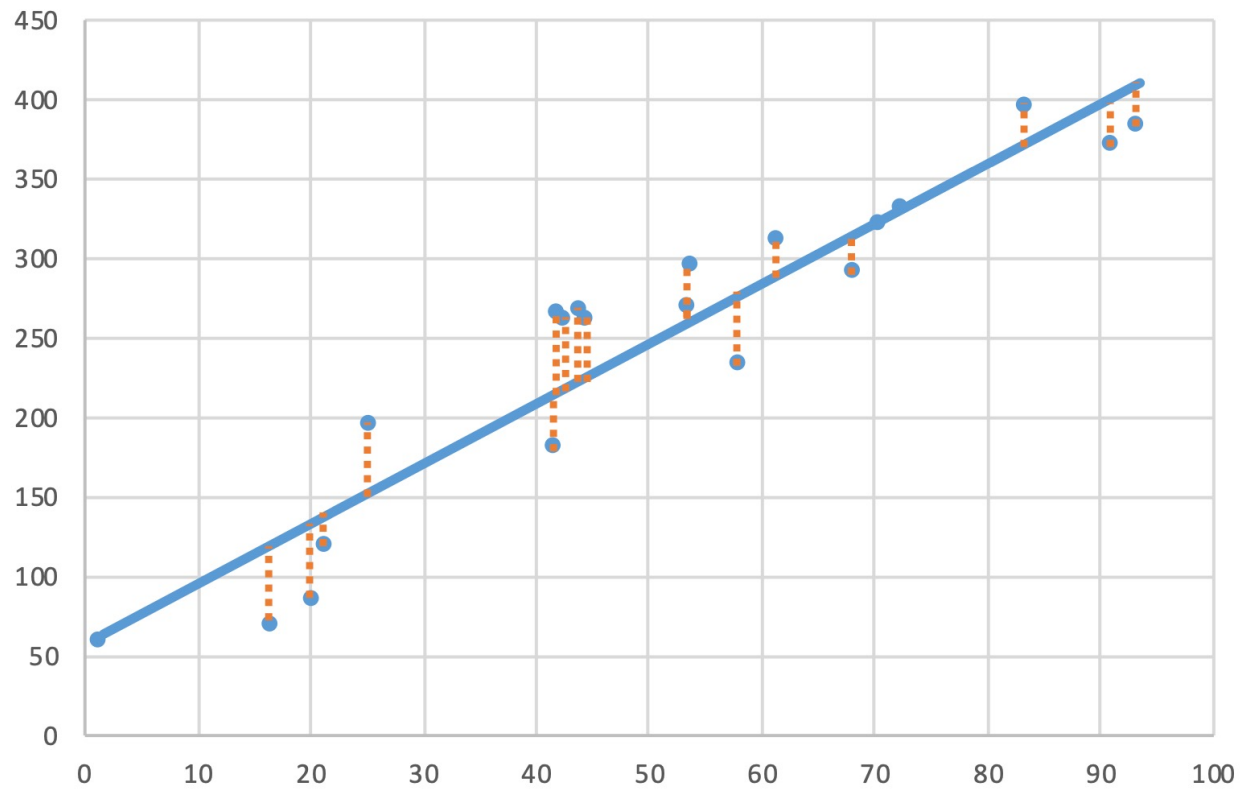
You would expect e.g. word length and frequency to influence reading time



From Underwood, G., Binns, A., & Walker, S. (2000). Attentional Demands on the Processing of Neighbouring Words. In I. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a Perceptual Process* (pp 247-268). Amsterdam: North-Holland.

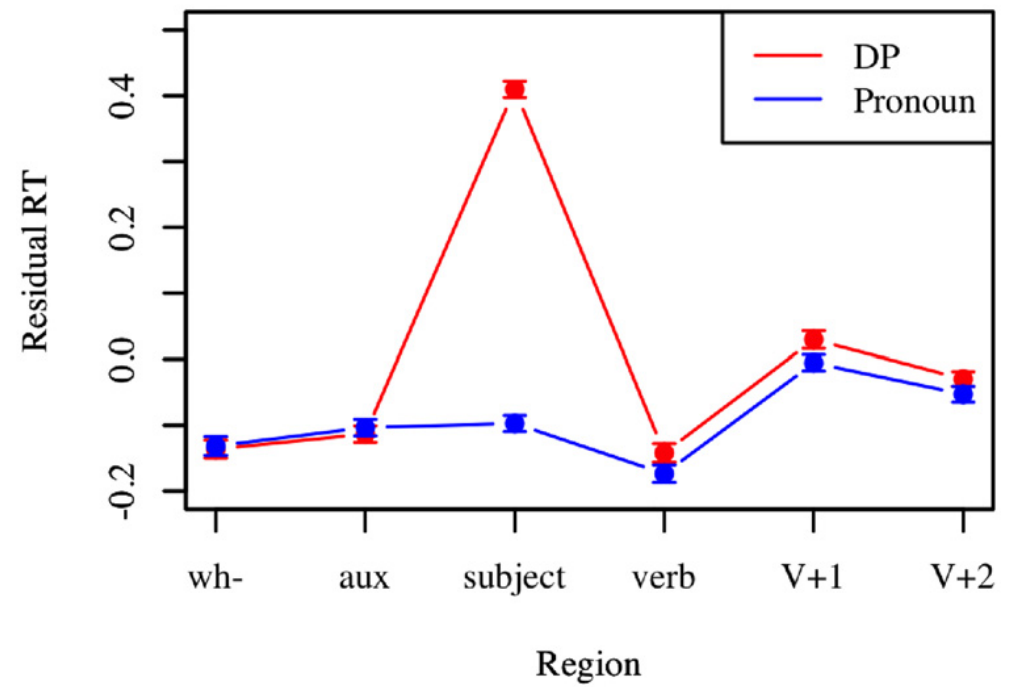
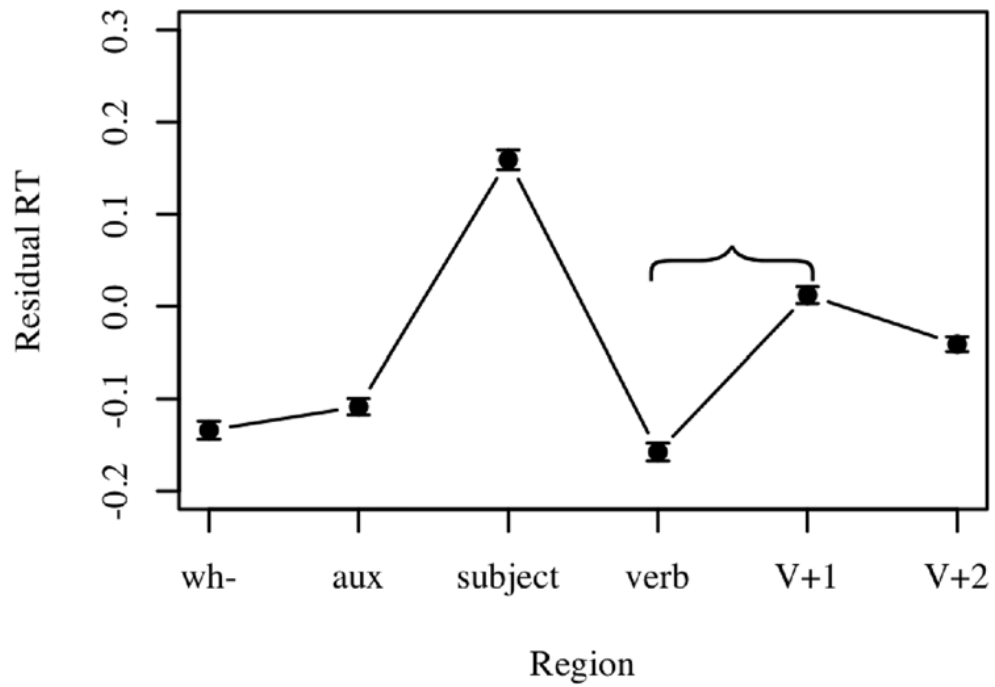
# Residual reading times

Regression line – line of best fit, **minimising residuals**





# Exp 1 results



# Test items and predicted effects (Exps 2-3)

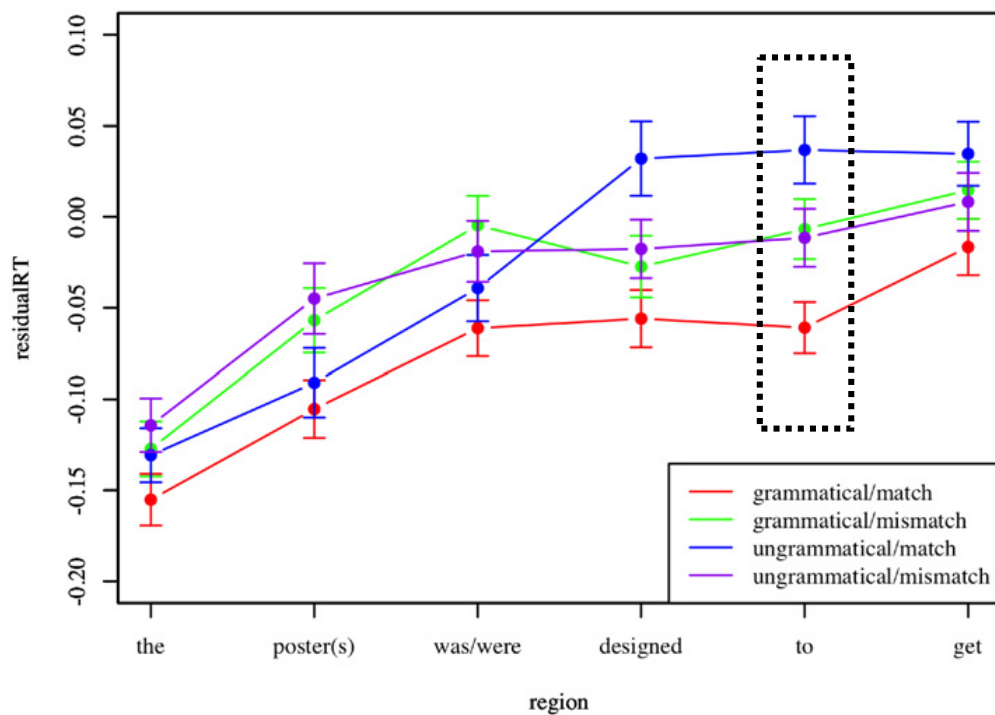
## Experiment 2

- *The slogan on the **poster** was designed to get attention*
- *The slogan on the **posters** was designed to get attention*
- *\*The slogan on the **poster** were designed to get attention*
- *\*The slogan on the **posters** were designed to get attention*

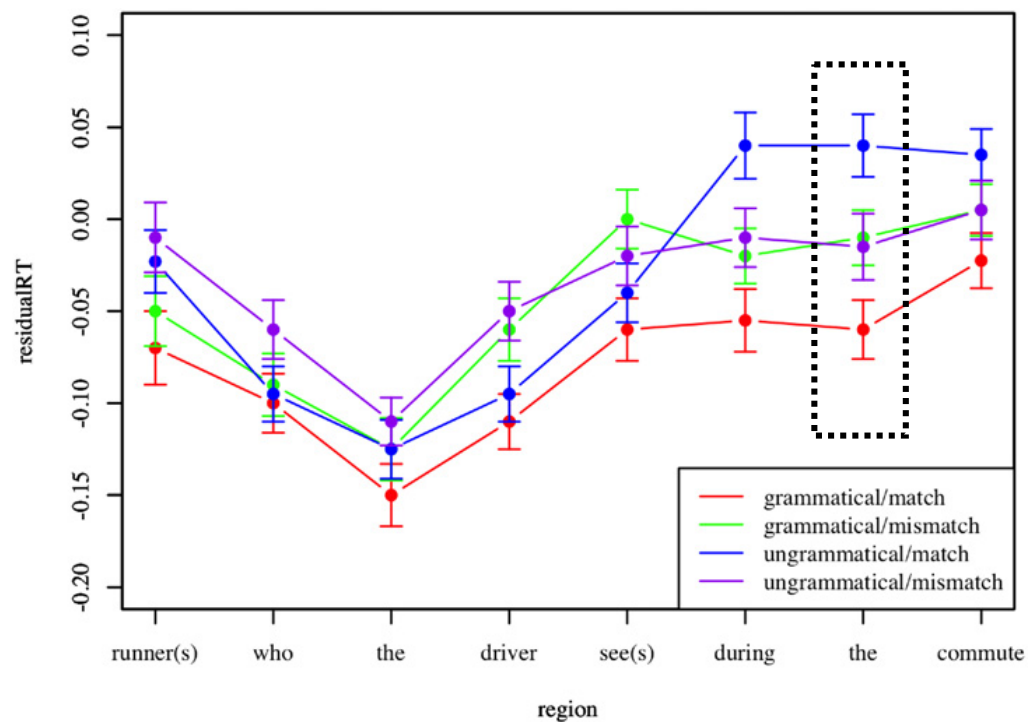
## Experiment 3

- *The runner who the driver sees during the commute...*
- *The runners who the driver sees during the commute...*
- *\***The runner** who the driver see during the commute...*
- *\***The runners** who the driver see during the commute...*

*The slogan on the poster was designed to get attention*  
*The slogan on the posters was designed to get attention*  
*\*The slogan on the poster were designed to get attention*  
*\*The slogan on the posters were designed to get attention*



*The runner who the driver sees during the commute...*  
*The runners who the driver sees during the commute...*  
*\*The runner who the driver see during the commute...*  
*\*The runners who the driver see during the commute...*



# Enochson & Culbertson's conclusions

MTurk is suitable for collecting reading-time data in self-paced reading tasks

- Similar patterns of results to those seen in lab tasks
- (Paper also includes lab replication of Exp 1)

They also make some suggestions re. Masters qualifications and batch sizes that I don't necessarily agree with – see my reading notes!

Time for Q&A/discussion on this week's reading

# Next up

Wednesday, 9am: lab on Gather

- Our second proper experiment: self-paced reading

Next week:

- **No lecture on 18<sup>th</sup> October**
- **Lab on 20<sup>th</sup> October is for catchup – no new material**