Online Experiments for Language Scientists

Lecture 1: Introduction

Kenny Smith

kenny.smith@ed.ac.uk

What is this course about?

Language is something that humans do, so all subfields of linguistics involve collecting data from humans

- Grammaticality judgments, naturalistic recordings, sociolinguistic interviews, preferential looking/listening in infants, reading times, reaction times and/or choices in psycholinguistic experiments, …

This kind of data can be collected in person

Or it can be collected online (i.e. via a web browser)

This course shows you how to build language-relevant experiments that run in browsers, and how to crowdsource experiment participants

Who teaches on it?

Kenny Smith
Prof

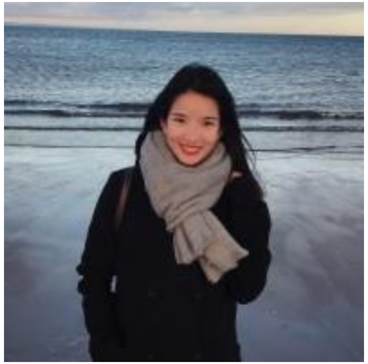
Course organiser
Lectures
(Labs)

Alisdair Tullo
Programming and Apps Manager

Labs

Maisy Hallam
PhD student, Linguistics

Labs



Yajun Liu
PhD student, Linguistics

Labs

How is it delivered?

Lectures (Mondays 9am-9.50am)

Labs (Wednesdays 9am-10.50am)

Lectures

- Each lecture has associated pre-reading, **do the readings before the lectures** so we can discuss any questions/thoughts you have
- Bit of context by me, **plus discussion / Q&A time**

Labs: Appleton Tower M2 (Teaching Studio, bring a laptop)

- Work through the practicals, with support on hand!
- Particularly in the first few weeks, worth at least reading through the materials in advance
- If you don’t complete the practicals in lab time, finish them in your own time - try not to fall too far behind.
- **Attendance will be taken each week**

Additional drop-in labs

A small number, in the run-in to the final assignment

- Details TBC

For **undergraduates**, how is it assessed?

Assessment 1: annotated bibliography, worth 30%, due 7th November

- Brief summary plus evaluation of 4 papers
- Papers can come from course readings or elsewhere

Assessment 2: coding project + report, worth 70%, due 5th December

- A functioning web experiment
- A short report explaining the motivation for that experiment, discussing and evaluating critical implementation decisions

Lots of information re. rationale and expectations available in the assignment brief.

There will be a cut-off date for questions on the assignments!

For **postgraduates**, how is it assessed?

Single assessment: coding project + report, worth 100%, due 5th December

- A functioning web experiment
- A short report explaining the motivation for that experiment, discussing and evaluating critical implementation decisions

Lots of information re. rationale and expectations available in the assignment brief.

There will be a cut-off date for questions on the assignments!

Where can I find all this information?

Course Learn page

- Links to course pages on github
- Assignment submission links

Course pages on github

- Everything else

Any questions on course organization, admin?

Three components of running an online experiment

Building an experiment that will run in a web browser

- We'll be using javascript and jsPsych
- Also useful for running experiments in-person!

Making it openly available online

- PPLS / the Uni provide servers

Connecting with experiment participants

- E.g. through **crowdsourcing websites**

A look at some simple experiments

Javascript and jsPsych

Javascript: a programming language that runs in web browsers

jsPsych: a library that makes it easy to build experiments (https://www.jspsych.org)

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*, 1-12. doi:10.3758/s13428-014-0458-y.

Josh de Leeuw
Vassar College

Plugins and timelines

Plugins: basic building blocks

```
var hello_trial = {
  type: jsPsychHtmlKeyboardResponse,
  stimulus: 'Hello world!'
}
```

Timeline: a sequence of those building blocks

```
jsPsych.run([hello_trial]);
```

A wide range of plugins available

See https://www.jspsych.org/latest/plugins/list-of-plugins/

Building an experiment involves

- Knowing how to use plugins
- Figuring out how to piece them together to make the experiment you want
- Some tiny bits of html and javascript to connect the plugins and make them do what you want
- (Occasionally, and optionally, making your own plugin)

A quick word about coding and realistic expectations!

A quick word about reading experimental papers

pared to pronouns. To test whether we can capture this effect using AMT, a linear mixed-effects model was fit using residual reading time as the dependent variable and subject type as a fixed effect. Results indicate that at the subject region, pronouns are read significantly faster than DPs ($\beta = -0.508 \pm 0.02$, $p < 0.0001$). This is illustrated in Fig. 2.

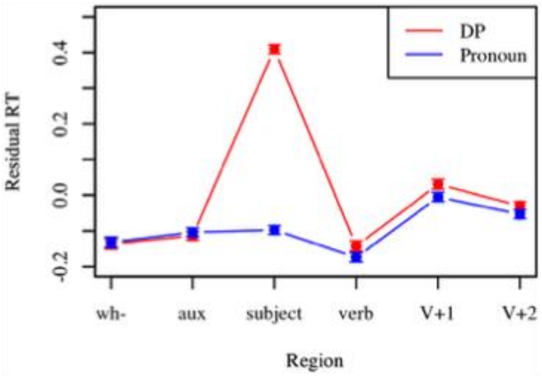


Fig 2. Subject definiteness results. Mean residual reading time is plotted by region for DP and pronoun sentences in Experiment 1. Error bars represent standard error of the mean.

From Enochson & Culbertson (2015) - week 4

A quick word about reading experimental papers

To assess the significance of differences in regularization within and between conditions, a linear mixed effects regression analysis was performed using R (R Core Team, 2013) and *lme4* (Bates, Maechler, Bolker, & Walker, 2013). The dependent variable was the change in entropy of the input-output ratios. Experimental condition was the independent variable. Participant was entered as a random effect (with random intercepts). No obvious deviations from normality or homoscedasticity were apparent in the residual plots.

Within-condition changes were assessed by re-leveling the model to obtain the intercept value for each condition. The intercept equals the condition's mean change in entropy and the regression analysis provides a t-statistic to evaluate whether this mean is significantly different from zero. Three of the four experimental conditions elicited a significant amount of regularization behavior (Fig. 6). Participants regularized an average of 0.17 bits in *marbles6* ($S.E. = 0.03$, $t(1152) = -5.53$, $p < .001$), 0.19 bits in *words1* ($S.E. = 0.03$, $t(1152) = -6.52$, $p < .001$), and 0.36 bits in *words6* ($S.E. = 0.03$, $t(1152) = -11.34$, $p < .001$). In *marbles1*, the mean loss of 0.01 bits was not significantly different from zero, which indicates that participants are probability matching in this condition ($S.E. = 0.03$, $t(1152) = -0.35$, $p = 0.73$). Overall, participants regular-

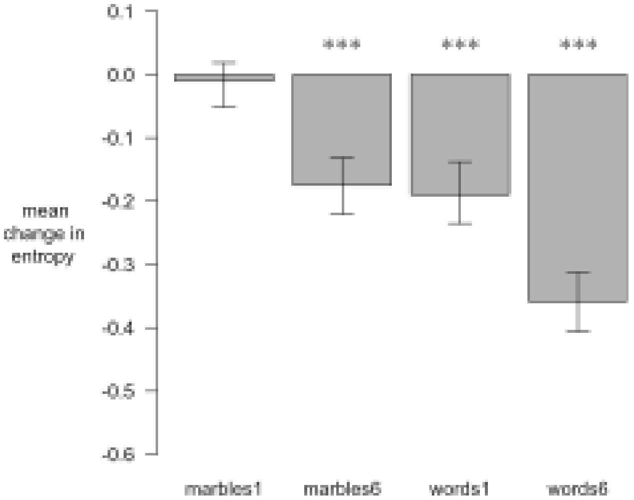


Fig. 6. Entropy drops when learners regularize. Each bar shows the average change in Shannon entropy over all pairs of input-output ratios, per condition. Stars indicate significant difference from zero. Error bars indicate the 95% confidence intervals computed with the bootstrap percentile method (Efron, 1979). A significant drop in entropy means that participants regularized in that condition. Non-significant differences from zero are obtained when participants probability match. The lower and upper bounds on mean entropy change for this experiment are − 0.67 and + 0.33 bits.

From Ferdinand, Kirby & Smith (2019) - week 5

Any questions/concerns so far?

Next up

Wednesday, 9am, Appleton Tower M2: first lab!

- Week 1 practical, linked from the course page on github
- Bring a laptop
- You'll get more out of the lab if you take a look at the materials beforehand!

Next lecture: crowdsourcing experimental data

- Monday 23rd September
- **Do the reading beforehand!**