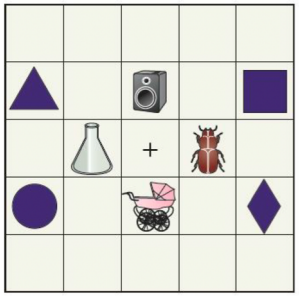
Experimental Methods



Standardized Testing

(3) On a scale from zero to ten please select your *level of* ***proficiency*** in speaking, understanding, and reading from the scroll-down menus:

Speaking	(click here for scale)	Understanding spoken language	(click here for scale)	Reading	(click here for scale

(click here for scale)
0 - none
1 - very low
2 - low
3 - fair
4 - slightly less than adequate
5 - adequate
6 - slightly more than adequate
7 - good
8 - very good
9 - excellent
10 - perfect

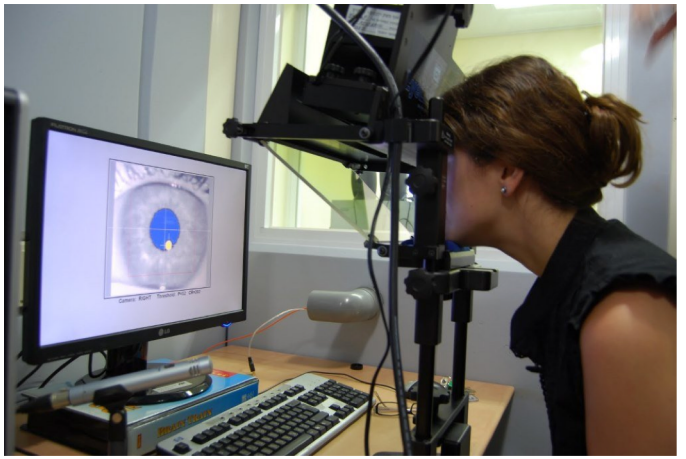
(4) On a scale from zero to ten, please select how much the fo ou learning :

Interacting with friends	(click here for pull-down scale)	Language	pull-down scale)
Interacting with family	(click here for pull-down scale)	Watching	pull-down scale)
Reading	(click here for pull-down scale)	Listening	pull-down scale)

(5) Please rate to what extent you are currently exposed to

Interacting with friends	(click here for pull-down scale)	Listening	pull-down scale)
Interacting with family	(click here for pull-down scale)	Reading	pull-down scale)
Watching TV	(click here for pull-down scale)	Languag	pull-down scale)

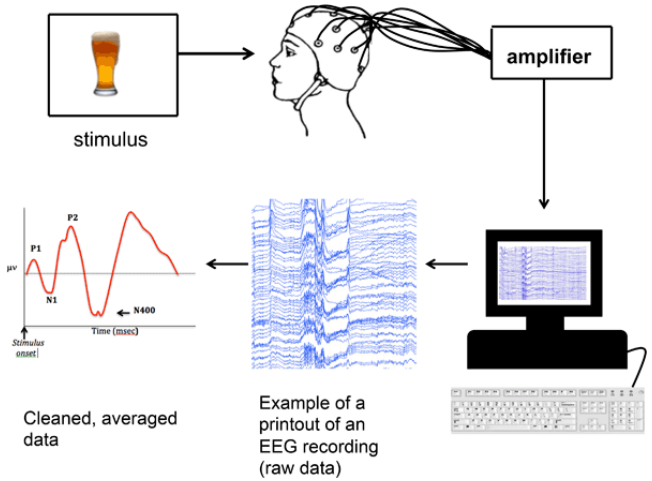
Eye-Tracking



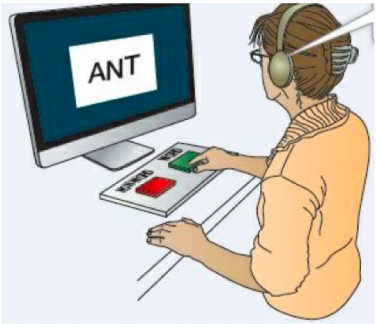
Self-paced Reading

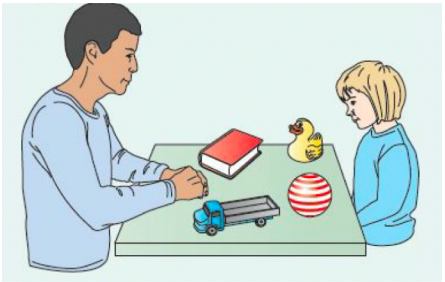
```
2.    --- ------- presented copies -- --- ------ --- ----- -- --- -------.
3.    --- ------- --------- ------ of the report --- ----- -- --- -------.
4.    --- ------- --------- ------ -- --- ------ was aware -- --- -------.
5.    --- ------- --------- ------ -- --- ------ --- ----- of the problem.
```

Event-related Potentials (ERPs)



Behavioral Measures





Road Map

- Fundamentals
- Research ethics and human subjects approval
- Experimental design
- Data collection techniques

Fundamentals

- **Validity**
  - the observed effects are triggered by your manipulated factors
  - rather than by coincidence or other factors
- **Reliability**
  - Consistency
    - Experimental material reliability
    - Instrument reliability
- **Replicability**
  - among another group of participants
  - by other researchers / labs

Before Experimental Design…

- To choose a certain language/linguistic phenomenon
  - observation / literature review
  - list your research questions (variables)

- To ensure validity and reliability, we should be very careful in all aspect of experimental design
  - control extraneous variables
  - create and modify stimuli
  - acceptability and grammaticality judgement
  - run ‘enough’ participants

Before Experimental Design…

- Ask yourself about:
  - Do I need more than one group of participants?
  - If yes, how many groups?
  - How do I assign participants to different groups?

  - How do I collect participants' data?
    - **Reaction time**
    - **Accuracy**
    - Eye movement
    - Brain activity

    …

Experimental Design

- Hypothesis & Predictions
- Variables
- Stimuli
  - Targets
  - Fillers
- Time constraints
- Feedback
- Participants
- Randomization

Experimental Design

- By the number of **independent variables** (IV) and **dependent variables** (DV) presents in the experiment
  - By the number of IV
    - single-factor vs. multiple-factor
  - By the number of DV
    - univariate vs. multivariate

Experimental Design

- By how we group participants
  - Between-subjects
  - Within-subjects
  - Mixed-design

Single-factor Design

- A single-factor, univariate, between-subjects design
  - IV: native language (native vs. non-native)
  - DV: score in grammaticality judgments

- A single-factor, multivariate, between-subjects study
  - IV: Parkinson's disease (with vs. without)
  - DV: two different verbal fluency tasks
    - common noun
    - proper noun

(Fine et al. 2011)

(Multi-)factorial Design

- Having more than one independent variable
- Allowing researchers to examine ***the effect of each independent variable separately*** (main effects) and to look at ***possible interactions*** between the independent variables (interaction effects)
  - 2 x 2
  - 2 x 3
  - 2 x 2 x 2
  - …

(Multi-)factorial Design

- DV: the perceived processing difficulty of English sentences by English-Chinese bilinguals
- IV: relative clause type
  - SRC vs. ORC
- IV: form of the embedded noun
  - fullNP vs. pronoun
- A set of sample stimuli
  - The reporter [that attacked the senator] admitted the error. (SRC, fullNP)
  - The reporter [that attacked you] admitted the error. (SRC, pronoun)
  - The reporter [that the senator attacked] admitted the error. (ORC, fullNP)
  - The reporter [that you attacked] admitted the error. (ORC, pronoun)

(Gibson 1998; Roland et al. 2007; Staub et al. 2017)

A one-shot survey

- The woman knew the photographer of the singer who won the prize.
    - Who won the prize?

(Multi-)factorial Design

Stimuli Sample I

a. *NP1 Attachment (of)*

- The dean liked / **the secretary of the professors** / who / **was** / reading a letter.

b. *NP2 Attachment (of)*

- The dean liked **/ the secretary of** **the professors** / who / **were** / reading a letter.

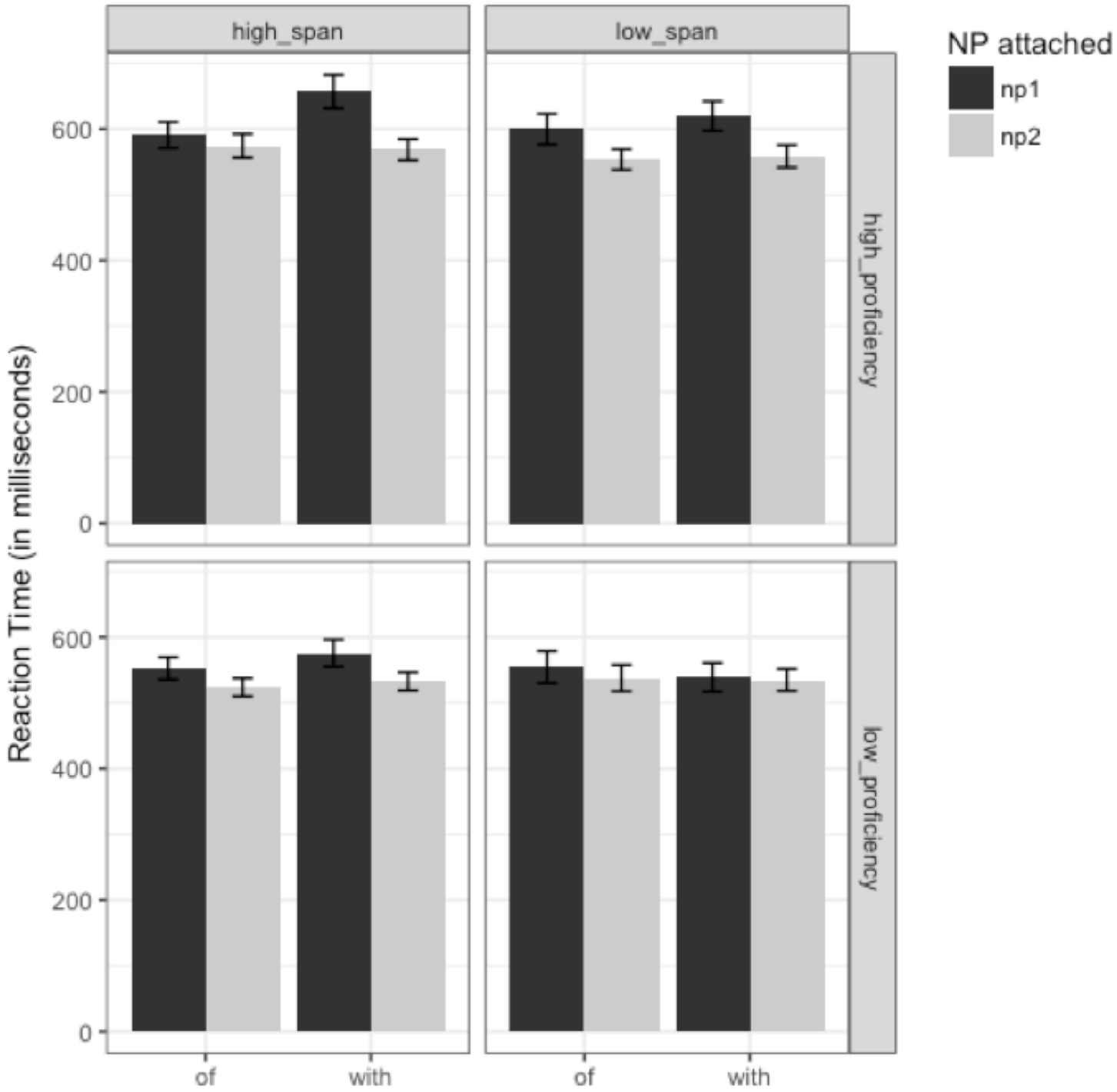
c. *NP1 Attachment (with)*

- The dean liked / **the professors with the secretary** / who / **were** / reading a letter.

d. *NP2 Attachment (with)*

- The dean liked / **the professors with** **the secretary** / who / **was** / reading a letter.

Results



Sheng & Qiao (2021)

(Multi-)factorial Design

Stimuli Sample II

- The conductor heard the violins were not properly in tune.
- The conductor heard that the violins were not properly in tune.
- The conductor assumed the violins were not properly in tune.
- The conductor assumed that the violins were not properly in tune.

Complementizer *that*

- A sentence comprehension task
- Design: 2 x 2

- The conductor **heard** the violins were not properly in tune.
- The conductor **heard** ***that*** the violins were not properly in tune.
- The conductor **assumed** the violins were not properly in tune.
- The conductor **assumed** ***that*** the violins were not properly in tune.

Exercise: Complementizer *that*

- 法语老师重复这首诗应该在周五之前完成。
- 法语老师决定这首诗应该在周五之前完成。

- A sentence production task

(Multi-)factorial Design

- DV: the amount of speech directed toward infants
- IV: Mother-native language (Japanese or English)
- IV: Infant age (6, 12 or 19 months)

Randomlization

- Random assignment of stimuli sequence
  - rand() function in Excel
  - www.random.org
  - pseudo-random

Counterbalance

Latin squares design

A	B	C
B	C	A
C	A	B

3 x 3 Latin squares

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

4 x 4 Latin squares

Data collection techniques

- Behavioral
  - Acceptability judgment tasks
  - Sentence/discourse completion tasks
  - Multiple-choice questionnaires

off-line

  - Lexical decision tasks
  - Priming
  - Self-paced reading experiments
  - Eye-tracking

on-line

Reaction time (RT)

Reaction time (RT)

- Reaction time (RT) --- by far the most prevalent type of method
  - Time it takes to make a decision, initiate an utterance etc.
  - Errors may be recorded
  - Jiang, N. (2012). Conducting reaction time research in second language studies. Routledge: NY.
- Characterizing RT research (Jiang, 2012: 3-7)
  - Accurate timing (ms): onset of the stimulus --- a response is made
  - Rigorous variable manipulation: RT can be affected by many factors!
  - Time-sensitive assessment of behavior: emphasis on fast response
  - Step-by-step progression: multiple experiments (2-6)

RT research #1
--- Lexical and phonological tasks

- Lexical Decision Task (LDT) (Meyer & Schvaneveldt, 1971)

  - Visual LDT: Is *empty* a word? *mepty*? *brane*? *workly*?

    - frequency, length, neighborhood (size and frequency) most often controlled

    - familiarity, AoA, concreteness, imaginability, regularity etc.

  - Auditory LDT: hear and decide; studying spoken word recognition; children-friendly; uniqueness point (唯一点） rebuild /ri'bild/ , build /bild/

  - Go/no-go procedure: press the button only when you see a word

RT research #1
--- Lexical and phonological tasks

- Word naming / naming task

  - Digit naming (diagnosing reading disabilities) , picture naming (semantic and conceptual), color naming (Stroop effect)

  - Phonology is overtly involved; both a lexical access component and a lexical production component; no decision component; less "lexical" (*Brane*? (faster, slower in LDT)

  - Related factors: word onset (/d/ is slower than /t/ because of the VOT 嗓音/浊音/发声/声带震动 起始时间); spelling-sound regular words are named faster than irregular words (e.g., *mint vs. pint*)

  - The sensitivity of voice key should be properly adjusted

RT research #1
--- Lexical and phonological tasks

- The priming paradigm (3 ways)
  - separate-trial priming;
  - single-trial unmasked priming (doctor-nurse // teacher-student // expectation for semantic relations & strategies effect in priming)
  - single-trial masked priming (Forster & Davis, 1984) demo

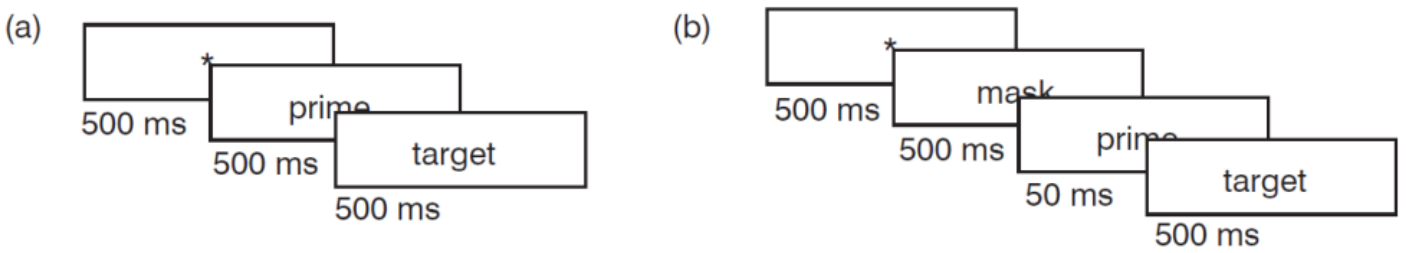


FIGURE 3.3 The sequence of a test trial in (a) an unmasked and (b) a masked single-trial priming experiment.

RT research #1
--- Lexical and phonological tasks

- The priming paradigm
  - prime-target relationship:
    - repetition priming (*book-BOOK*), semantic priming (*doctor-nurse*), morphological priming (*taught-teach*), form priming (*aptitude-attitude*), translation priming (*house-casa*), mediated priming (*lion-stripe*, via *tiger*)
  - Modality: within-modal (auditory-auditory, visual-visual), cross-modal (?)
  - Stimulus onset asynchrony (SOA, 刺激启动异步；刺激启动不同步): the interval between the onset of the prime and of the target (100ms or longer for unmasked priming, 50-60ms for masked priming

RT research #1
--- Lexical and phonological tasks

- The priming paradigm

- At least 2 conditions are necessary: a related condition & a control condition

(a)

Prime	Target	(Condition)
doctor	→ *nurse*	(related)
	→ *police*	(unrelated)

(b)

Prime	Target	(Condition)
doctor →	*nurse*	(related)
police →		(unrelated)

FIGURE 3.4 Two types of priming stimuli: (a) same-prime trials and (b) same-target trials.

Which one is better, a or b?

- Upshot: Keep the target constant for different conditions whenever possible.

- Matching primes: related and unrelated primes match in frequency, length etc.

Softwares

- DMDX
  - http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm
- Linger
  - http://tedlab.mit.edu/~dr/Linger/
- Eprime
  - https://pstnet.com/products/e-prime/
- PsychoPy
  - http://www.psychopy.org/

DMDX

THE UNIVERSITY OF ARIZONA
TUCSON ARIZONA

DMDX Display Software

DMDX is a Win 32-based display system used in psychological laboratories around the world to measure reaction times to visual and a Forster at the University of Arizona.

DMDX is a member of the DMASTR family, and represents an extension of the original DOS programs (DM and DMTG) to a Windows

- General Overview
- Hardware and Software Requirements
- Official disclaimer
- Tutorials on DMDX
- Examples of scripts
- Useful Links
- Programmer's notes about Updates to DMDX
- .
- Downloads
- .
- DMDX Users List Serv
- .
- Article on Timing Accuracy

Linger

a flexible platform for language processing experiments
version 2.94

Welcome to the home of Linger, a software package for performing reading, listening, and other sentence processing experiments
experiments. However, the code is flexible and can be adapted to support many other types of experiments.

Because Linger is written in Tcl/Tk, it can run on Unix, Windows, or Macintosh systems. Linger is also able to handle non-English te
Japanese, and other languages.

The Linger package is actually composed of four main programs: Linger, Lingalyzer, Lingrapher, and Subjector. Linger runs the exp
used to quickly generate graphs of averaged reading-time data. The Subjector program is a general purpose application for maint
performed.

Documentation

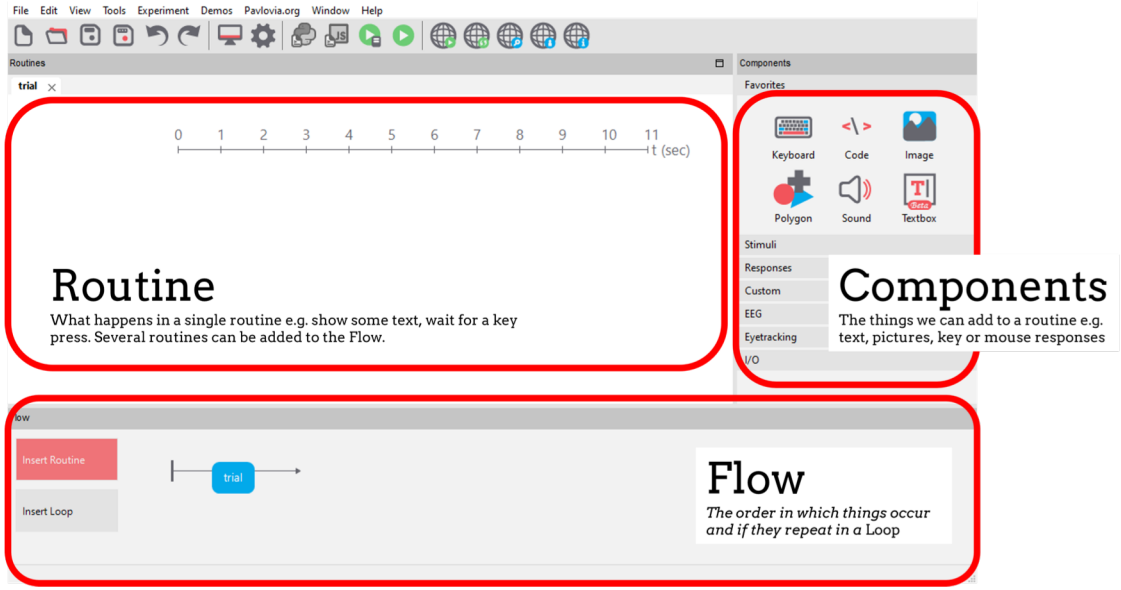
Linger Lingalyzer Lingrapher

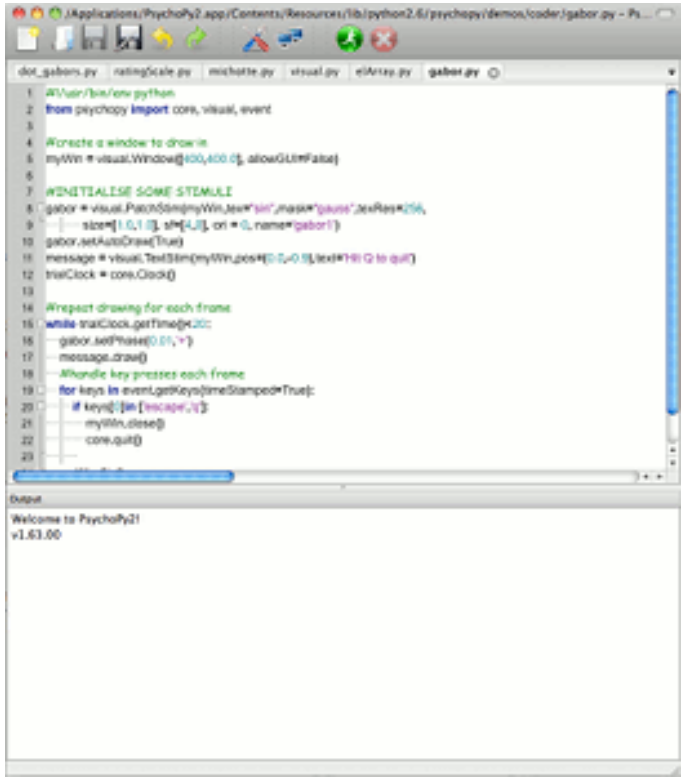
Installing Linger

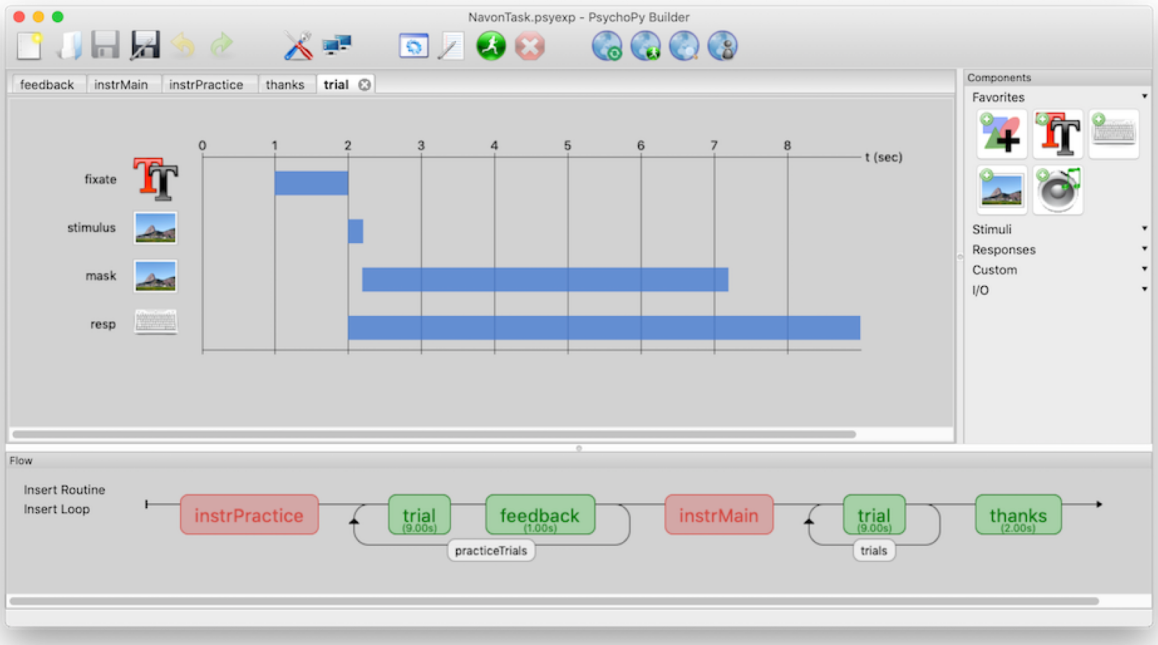
To install Linger, or to upgrade your current installation, you will first need to download the latest version.

Then you will need to install Tcl/Tk as follows...

PsychoPy

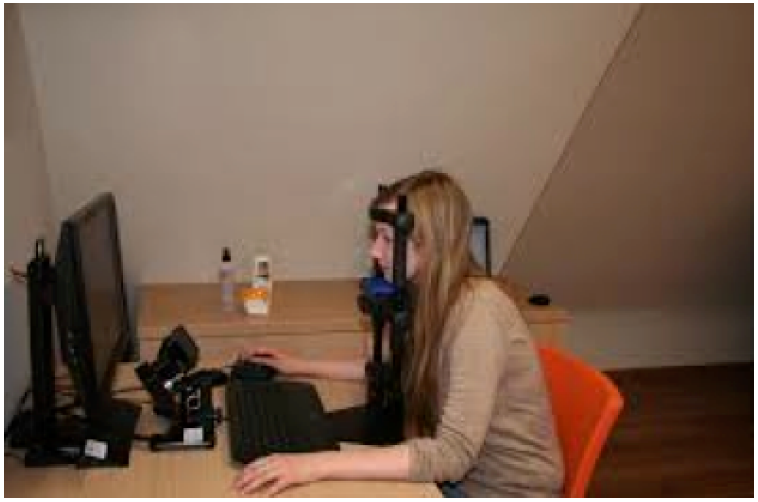


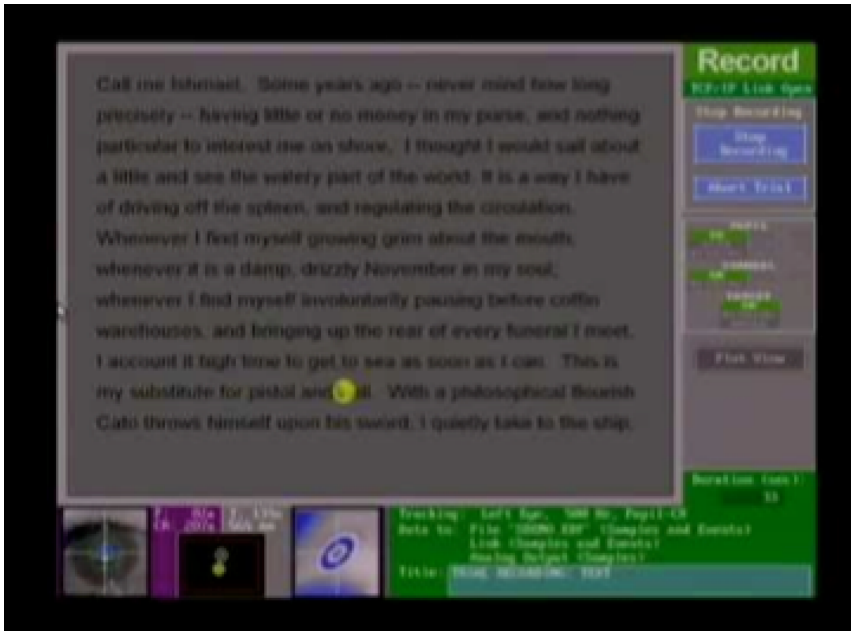


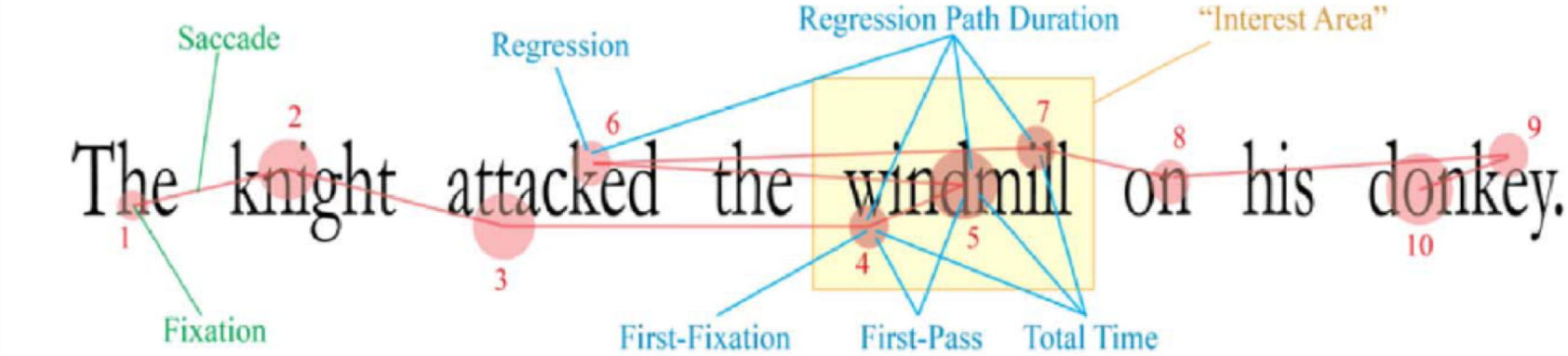


https://workshops.psychopy.org/3hrs/buildingBetter.html

Eye-tracking (reading)

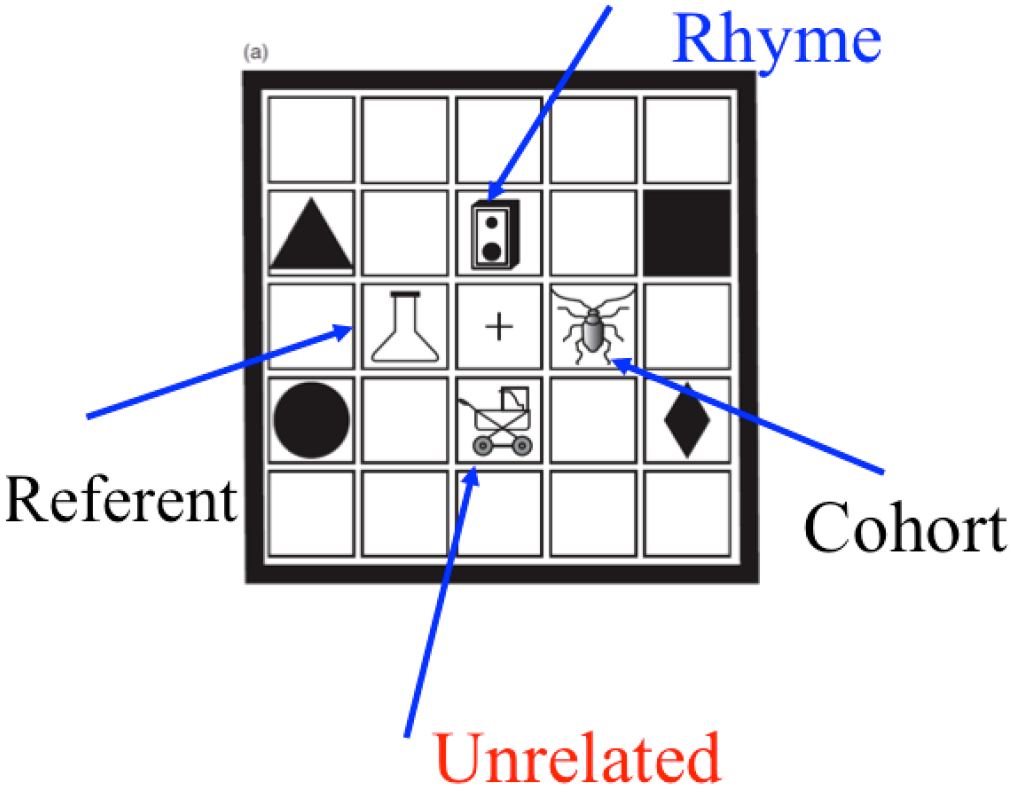


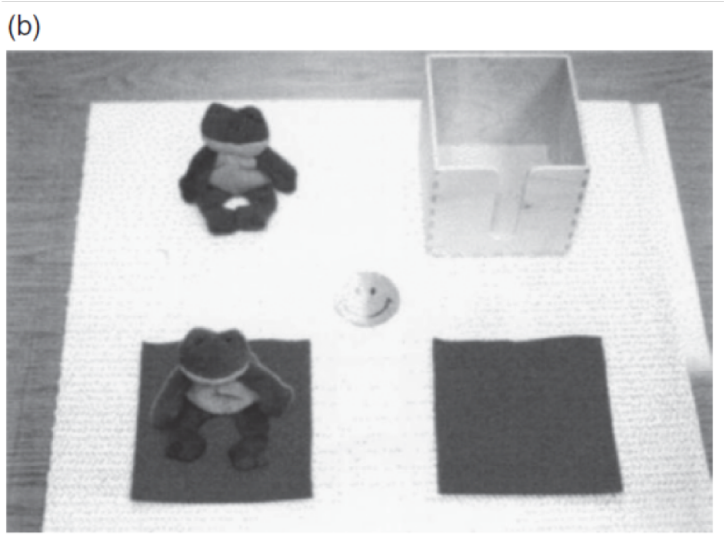




https://www.sr-research.com/eyelink-1000-plus/

Eye-tracking (visual-world paradigm)



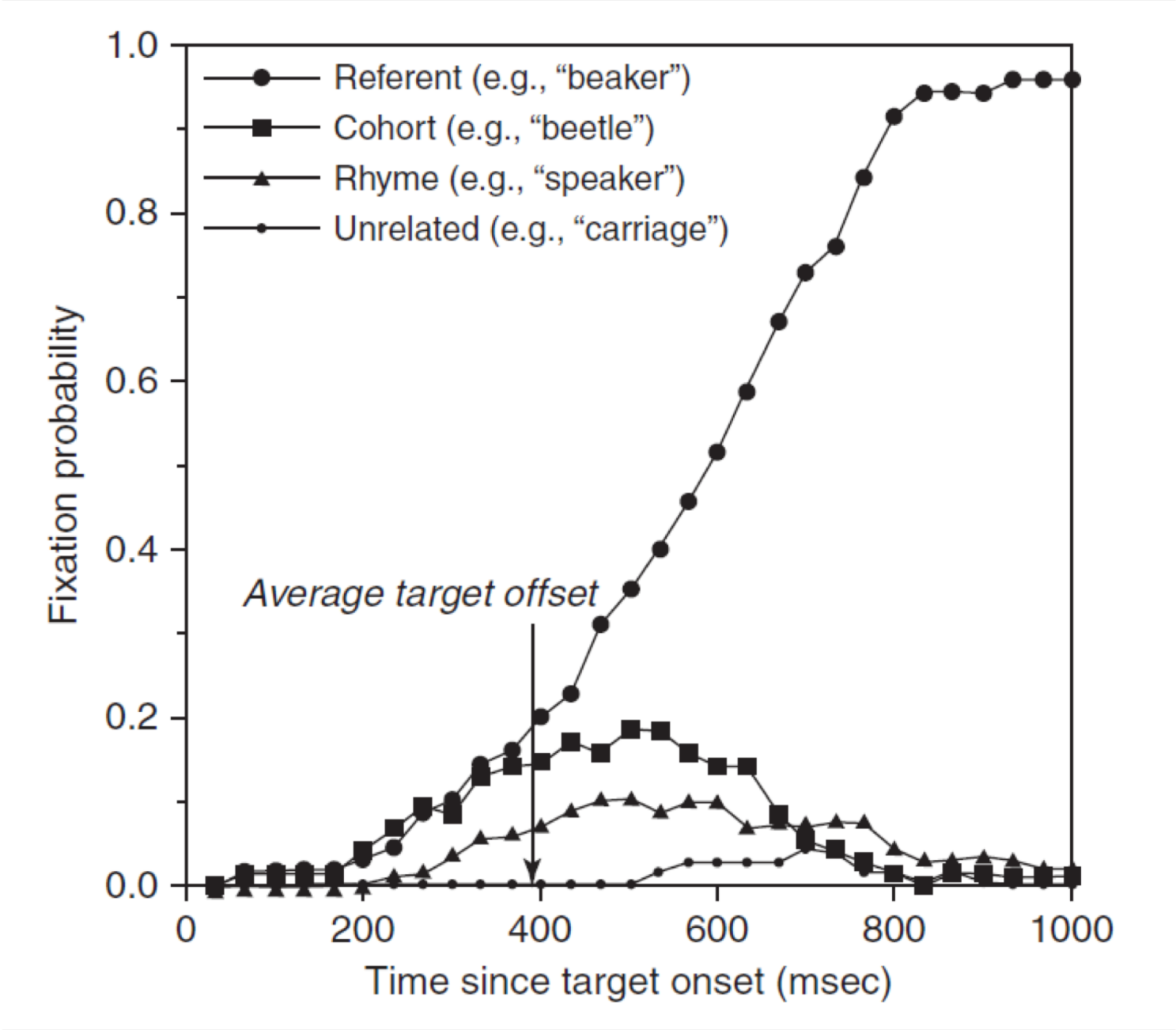


Put the bear on the towel in the box.

a) Allopenna et al. (1998)

b) Trueswell et al. (1999)

Eye-tracking (visual-world paradigm)



Reaction time methods

- Lexical decision
- Self-paced reading
- Eye movements during reading

- Speech-onset latencies (production side)

Reaction time methods

- Measuring reaction times
  - How rapidly people perform different kinds of tasks
    - how quickly people read sentences
    - how quickly people start to produce sentences
    - how quickly people recognize strings of letters (or strings of phonemes) as being real words or nonsense words
- An indication of processing complexity

Reaction time methods

- The reaction time methodology has a number of advantages.
  - Inexpensive to implement
  - A number of software options available
  - Data analysis is fairly straightforward
  - The methodology is technologically very portable
    - all that is needed is a computer and headphones/speakers.

Neurological methods

- Event-related potentials (ERPs)
  - to measure electrical activity in the brain

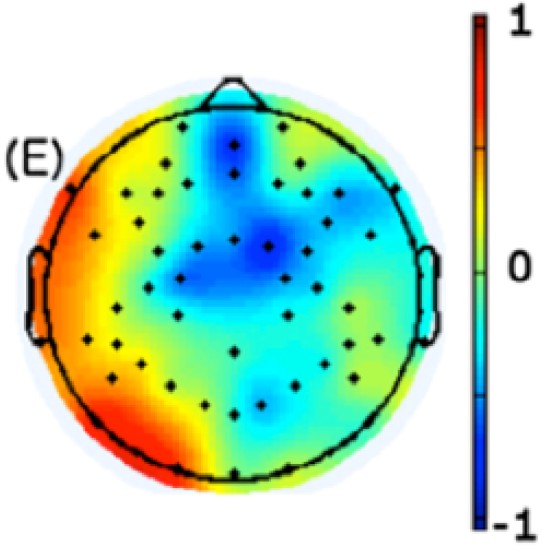
- Functional magnetic resonance imaging (fMRI)
  - to measure blood flow levels in the brain

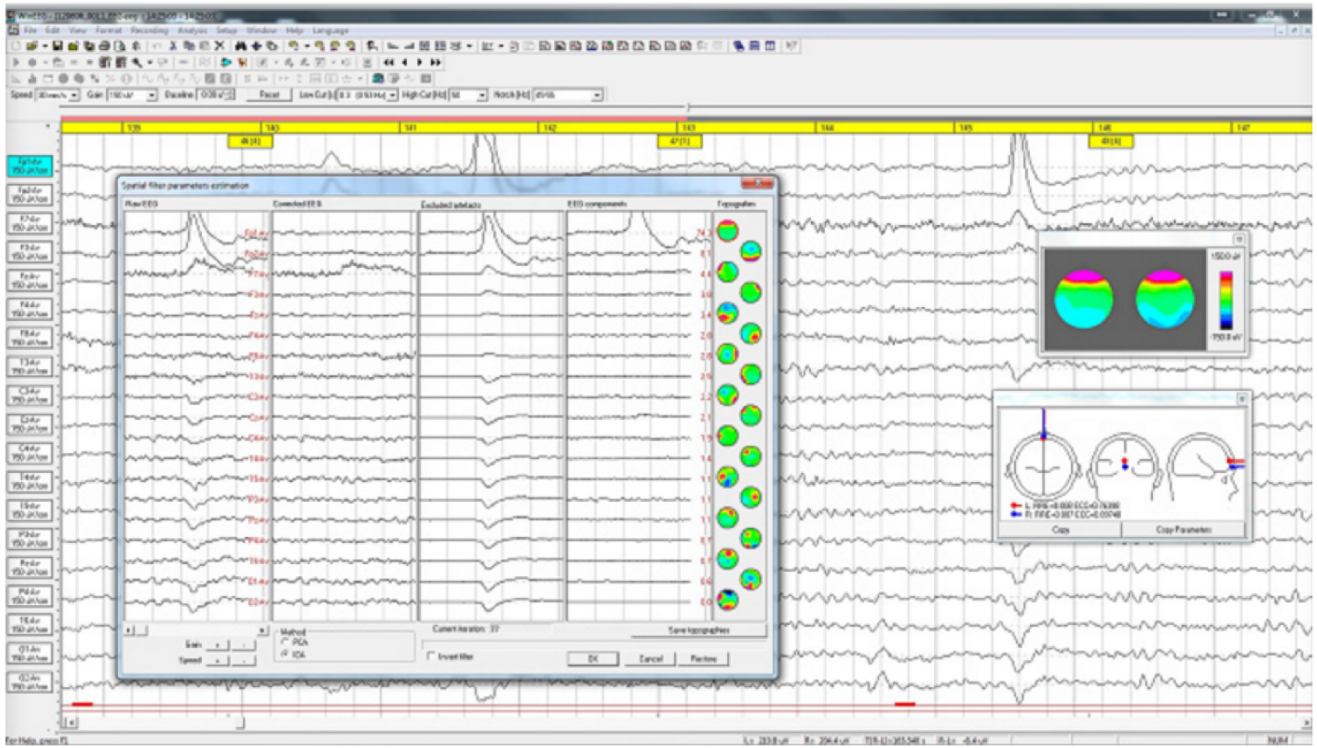
Event-related Potentials (ERP)

Stimulus – Response

High time resolution

Low space resolution

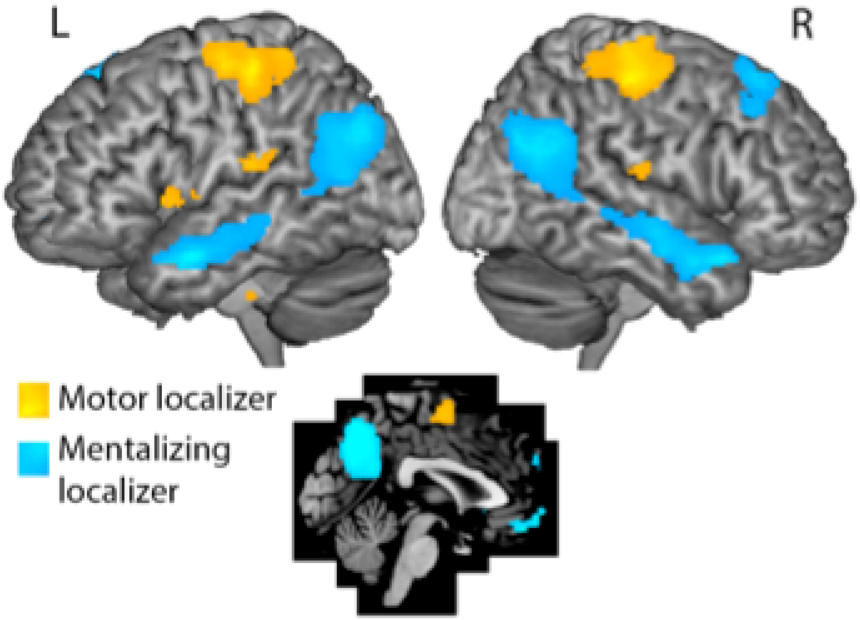


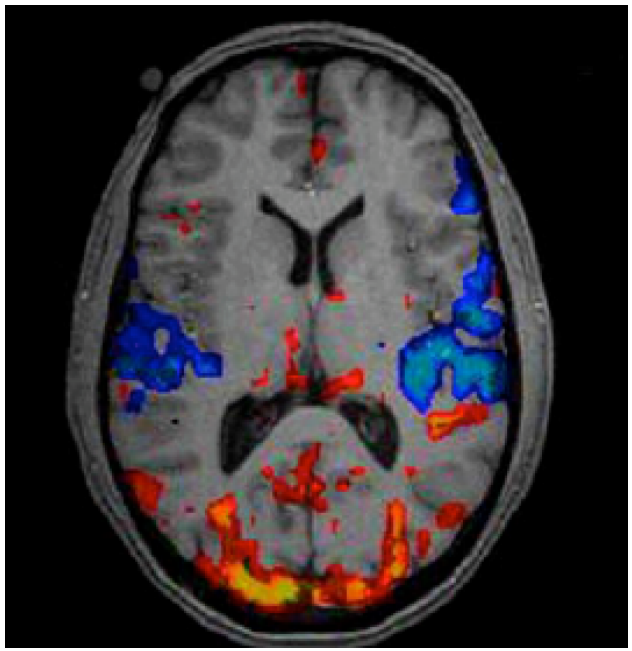


fMRI

High space resolution

Low time resolution





The benefits of on-line methods

- To gain insights into transient effects that are often not explicitly “noticed” by language users

- To make it possible to learn about the time-course of both language production and comprehension

Choosing the ‘right’ method

- Research questions —> the ‘right’ method
- One method is better than the others?
- Each has its own strengths and weaknesses