Agenda

Today

- Finish up discussion from last time (Van Berkum et al., 2008)
- Research methods (experimental paradigms) in psycholinguistics

Monday

- Read: Student paper
- Fleshing out the course experiment

Next Wednesday

- Revising the course experiment

Van Berkum et al. (2008)

The Neural Integration of Speaker and Message

1. What was the theoretical basis/motivation for the study?

   Debate between Gricean vs. contextual models

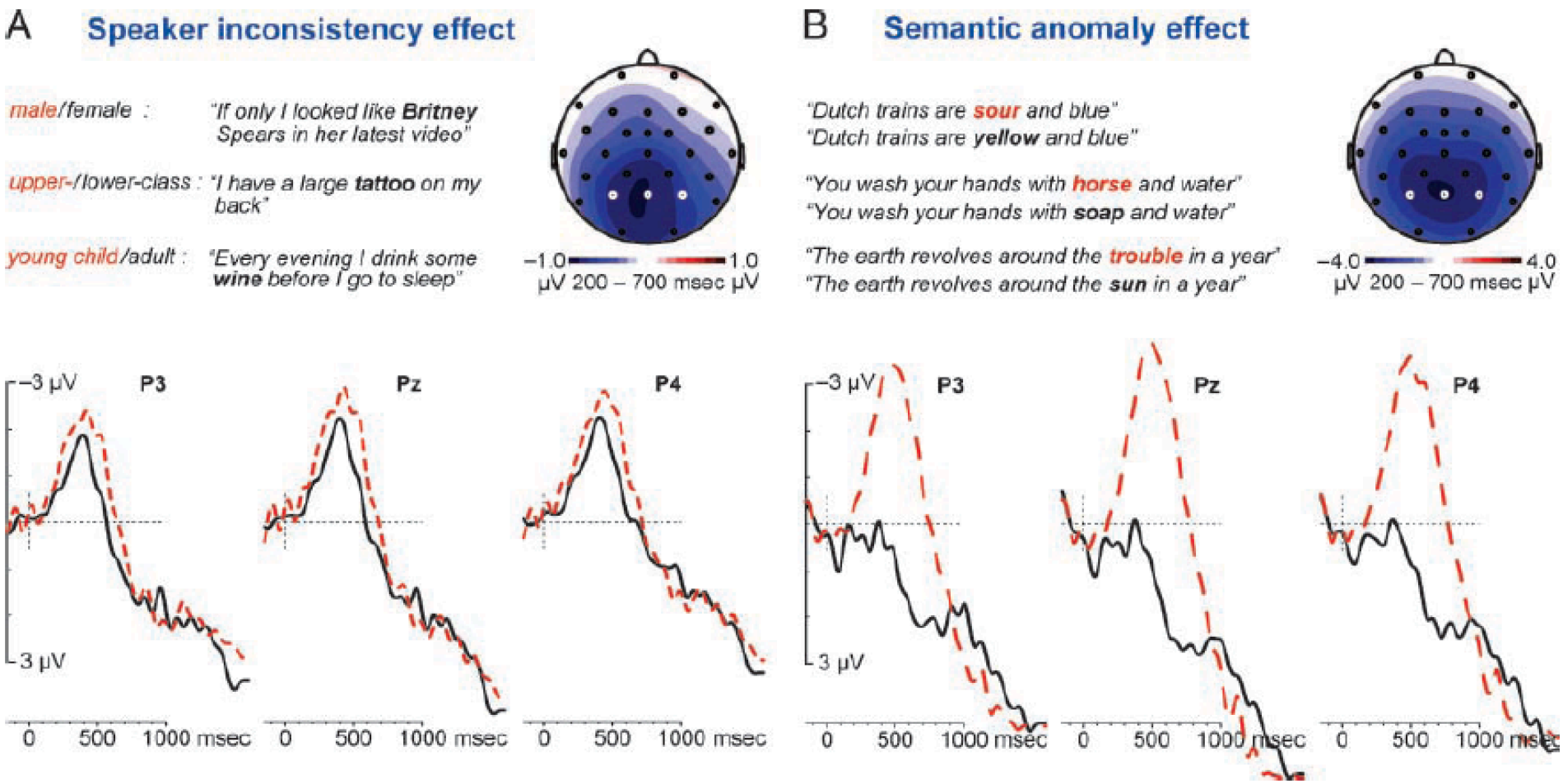
2. What was the general research question?

   Goal is to investigate:

   1. When during auditory language processing is speaker identity conveyed by voice taken into account, and
   2. Whether it is processed by the same brain system that combines the meanings of individual words into a larger whole

Van Berkum et al. (2008)

Results



Van Berkum et al. (2008)

Conclusions

- Evidence supports the constraint-based (one-step) account of language comprehension
- Construction of meaning based on linguistic information cannot be separated from social aspects of language use

→ The linguistic brain relates the message to the speaker immediately

EXPERIMENTAL PARADIGMS IN PSYCHOLINGUISTICS

Experimental Paradigm (aka, method)

"An experimental paradigm is a model of research that is copied by many researchers who all tend to use the same variables, start from the same assumptions, and use similar special procedures."

— Levine & Parkinson (1994, p. 352)

Experimental Paradigms in Psycholinguistics

Offline methods

- Measure variables that reflect the outcome of processing
  (e.g., Comprehension questions → Final Interpretation)

Online methods

- Measure variables that tap into processing "as it happens"
  (e.g., Brain activity → If and when interpretations are revised)

Online and offline methods can complement each other

- Time-course of processing + outcome of processing

Offline Methods

Common offline methods used in language comprehension:

- Comprehension questions
- Interpretation questions
- Cloze paradigm
- Judgment / Ratings studies

→ Often used to pre-test ("norm") materials created for an on-line experiment

- Check intuitions about phenomena of interest (which are often wrong!)
- Control for potential confounds

Comprehension Questions

DV: Accuracy (number/proportion of questions answered correctly)

While Anna dressed the baby that was small and cute spit up on the bed.

Q: Did Anna dress the baby? 56% incorrect "yes" responses

Q: Did the baby spit up? 9% incorrect "no" responses

While Anna dressed, the baby that was small and cute spit up on the bed.

Q: Did Anna dress the baby? 11% incorrect "yes" responses

(Christianson, Hollingworth, Halliwell & Ferreira, 2001)

Interpretation Questions

DV: Number/proportion of responses of a particular type

John said Fred went to Europe and Mary did too.

Q: What did Mary do?

…went to Europe	60%
…said Fred went to Europe	40%

John said Fred went to Europe. Mary did too.

Q: What did Mary do?

…went to Europe	45%
…said Fred went to Europe	55%

(Frazier & Clifton, 2005)

Judgment/Rating Studies

DV: Judgment or rating

Dimension: Acceptability, grammaticality, naturalness, comprehensibility, plausibility, predictability, felicity, appropriateness, conventionality, coherence, consistency, etc.

Crime is a virus that is infecting the city.

Completely unacceptable	1	2	3	4	5	**6**	7	Highly acceptable
Non-metaphoric	1	2	3	4	5	6	**7**	Highly metaphoric
Highly unusual	1	2	3	4	**5**	6	7	Very common

Implementation Considerations

Type of measurement

- Forced choice (binary or multiple choice)
- Rating scale/continuum
- Provide paraphrase

Time constraint

- Speeded vs. non-speeded (“paper and pencil”)
- Delayed response vs “over-thinking”

Speeded Acceptability Judgment

While Anna dressed the baby that was small and cute spit up on the bed.

Acceptable Unacceptable

Cloze paradigm

DV: Proportion of participants that "fill in the blank" with a certain word

She likes her coffee with cream and _____.

dog	0%
sweetner	14%
sugar	82%

- **Cloze probability** (for a word) – proportion of people who provide that particular word as the completion of a specific sentence frame
- **Constraint** (of a sentence frame) – cloze probability of the best completion for a particular sentence frame

Offline Methods

Strengths

- Useful for pre-testing materials created for an on-line experiment
  - Provides indication of whether or not a variable has expected effect, and (depending on the DV) its size
- Useful for keeping participants "on task" during an on-line experiment
  - Can provide quantifiable measure for excluding "bad" participants

Weakness

- Allows more analysis time than normal situations (may lead to "over thinking")
- Does not provide (much) information about when an effect operates

Online Methods

Seek to answer one or more of three related questions:

- **When** is something happening?
  e.g., at the moment stimulus is perceived vs. after stimulus is perceived

- **What** is it that is happening?
  e.g., decreased vs. increased reading time; N400 vs. P600 effect in ERP; regression vs. spillover in eye-tracking

- **Where** in the brain is it happening?
  e.g., frontal lobe vs. temporal lobe

Online Methods

Behavioral methods

- Lexical decision, priming, self-paced reading, maze paradigm, eye-tracking, visual world paradigm, etc.
- Pupillometry

Electrophysiological methods

- Electroencephalography (EEG, ERP)
- Magnetoencephalography (MEG)

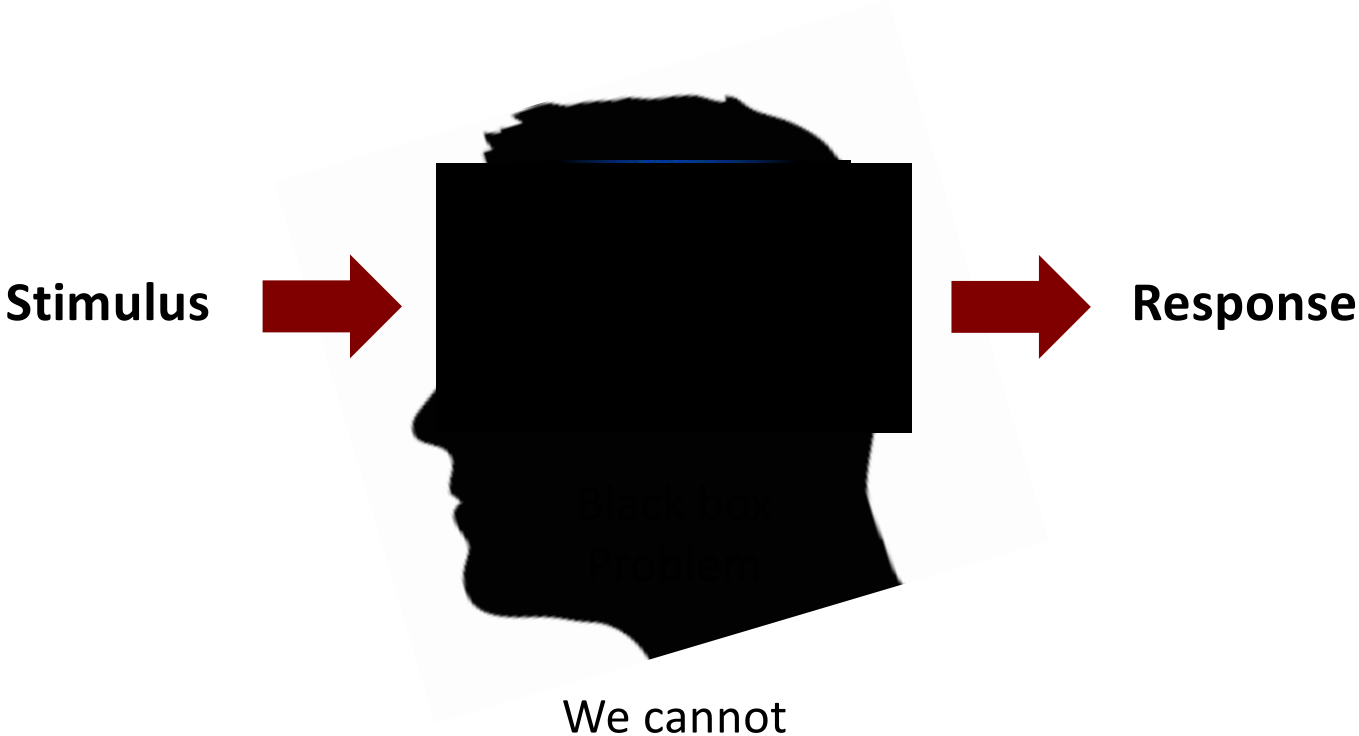
Hemodynamic measures

- Functional Magnetic Resonance Imaging (fMRI)
- Positron Emission Tomography (PET)

Neuroimaging

Behavioral Methods

Why are behavioral studies called "behavioral"?



We cannot directly observe the mental operation

Behavioral Methods

We need to infer relationship between stimulus (input) and response (output)

To test two candidate processes (A, B)

- Design experiment so that A and B lead to different responses

Behavioral Methods

Stimulus: Sensory event(s)

- Visual or auditory stimuli (e.g., words, pictures, a combination)

Task: The way participants are asked to respond to stimuli

- Press ‘yes’ if the string is a word and ‘no’ if not (lexical decision)
- Read each word and press a button to continue (self-paced reading)

Response: A bodily change that can be physically measured

- Pressing a button
- Eye movements while reading

DV: Aspect of response that is actually measured

- Response (or reaction) time (RT) – time between stimulus and response
- Accuracy – proportion/number of correct responses
- Eye-tracking measures: Fixation duration, position, proportion

Behavioral Methods

What can be inferred?

- To infer something about the mental processes, we need a ***linking hypothesis***
- How do mental processes map onto observed behavior?

Assumptions in behavioral research

- Mental operations take time
- Complexity of a mental process is reflected in response latency
- More difficult mental operations take more time
- Mental operations that have more steps take more time

To infer something more specific about the mental processes involved, we would need a more explicit linking hypothesis

→ Link the behavioral measure to a theoretical construct

Behavioral Methods

Common methods used in language comprehension:

- Lexical Decision
- Priming techniques

Investigate the mental lexicon

- Self-paced reading
- Eye-tracking

Investigate syntactic, semantic, and/or pragmatic processing

- Pupillometry

Complexity, cognitive load

Lexical Decision

Stimuli: A string of characters flashes on computer screen

Task: Decide whether the string is a word or a non-word

+

DVs: RT, Accuracy

Demo:

http://www.intro2psycholing.net/experiments/visual/word_nonword/wnw_index.php

Lexical Decision

What things did you notice that might affect RT and accuracy?

Research questions

- How is the mental lexicon structured?
- How do we access lexical entries (serial or parallel search)?

IVs can be:

- Word vs non-word (basic manipulation)
- Word frequency (e.g., low to high)
- Context (what other entries were just accessed)
  - Phonological relatedness
  - Semantic relatedness

} Priming techniques

Lexical Decision

Frequency effect

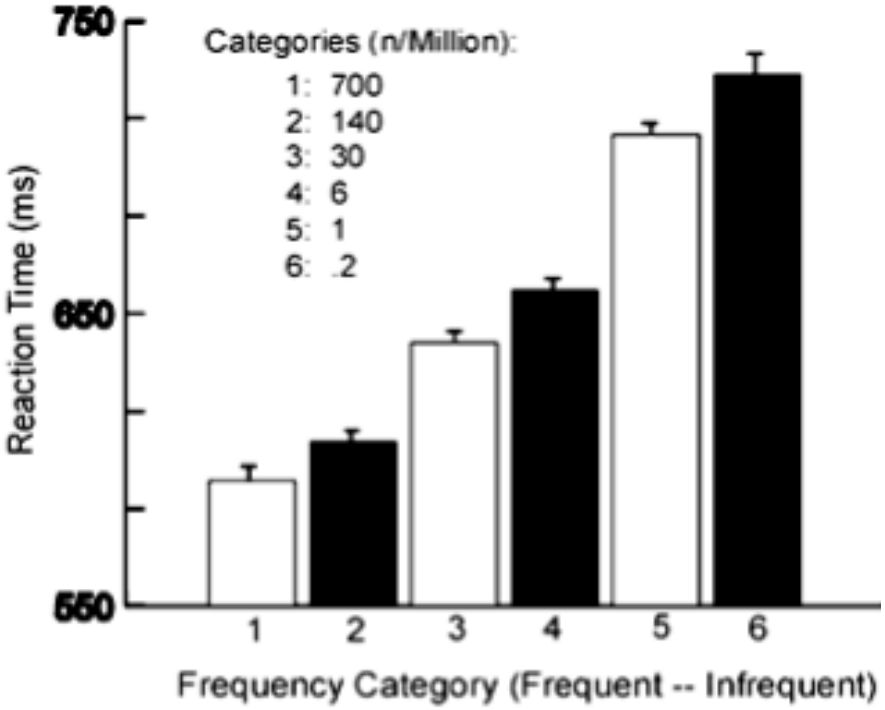


Fig. 1. Reaction time by frequency category (Mean and S.E.).

(Embick et al., 2001)

Linking hypothesis:

RTs reflect the time needed to access the mental lexicon

The 'ease' of accessing a lexical entry depends on the level of activation in that entry prior to access

Higher frequency words are more active than lower frequency words, and are therefore easier to access

Priming Techniques

Stimuli: A ***prime*** string and a ***target*** string are presented serially

Task: Make a lexical decision on the target

+

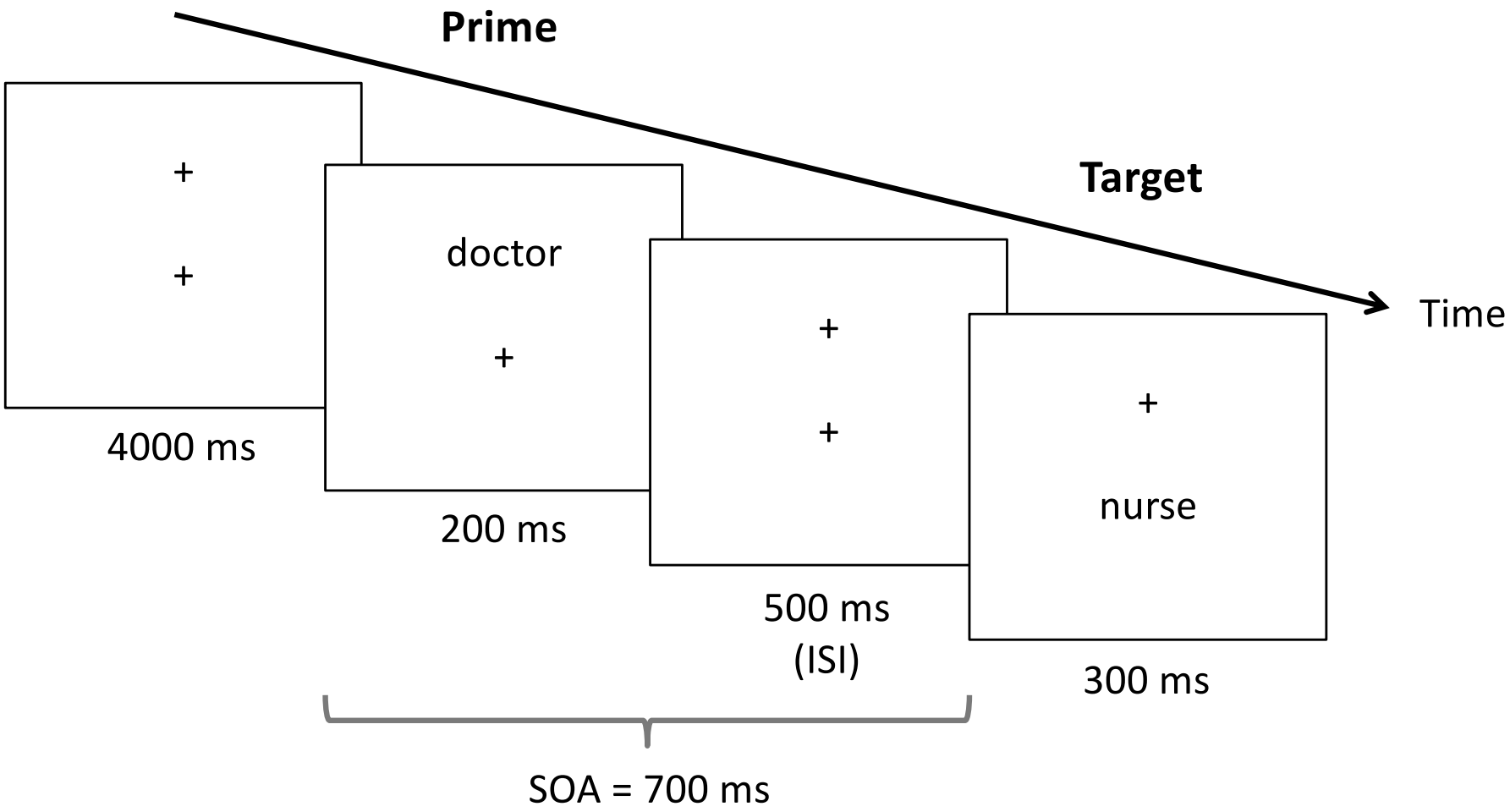
+

IVs can be:

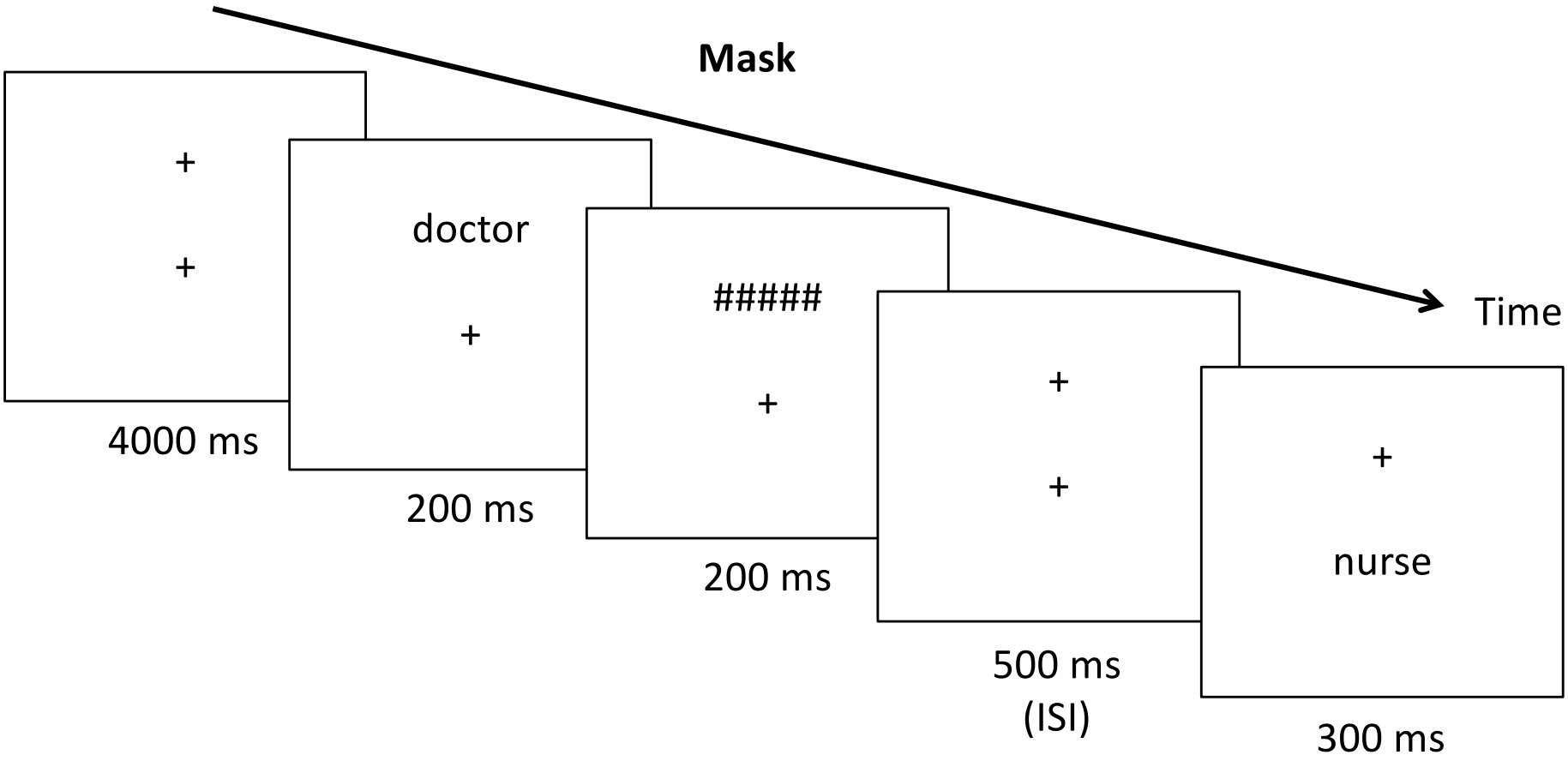
- Relationship between prime and target (semantic, phonological)
- Stimulus Onset Asynchrony (SOA)
- Degree to which prime reaches conscious awareness (masked priming)

DVs: RT and Accuracy

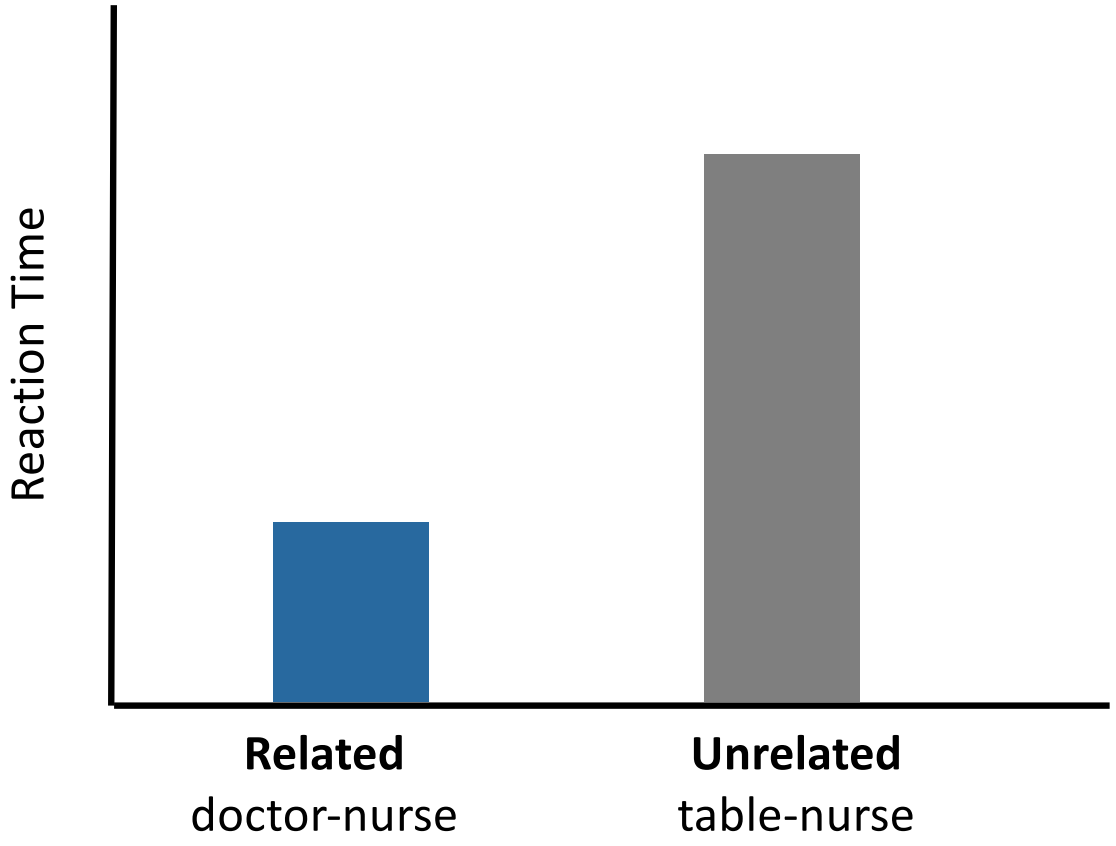
Semantic Priming



Semantic Priming



Semantic Priming Effect



Behavioral Methods

Common methods used in language comprehension:

- Lexical Decision
- Priming techniques

Investigate the mental lexicon

- Self-paced reading
- Eye-tracking

Investigate syntactic, semantic, and/or pragmatic processing

- Pupillometry

Complexity, cognitive load

Self-Paced Reading (or listening)

Stimuli: Sentences presented word-by-word or segment-by-segment

Task: Participant controls rate of presentation by pressing a button

DV: RT for each word or segment

Linking Hypothesis: Longer RTs at particular positions reflect processing difficulties

Moving-Window Technique

Most commonly used version of self-paced reading

- Sentence appears with all words/phrases dashed out
- Participant presses button to see first word/phrase
- Each button press turns current word/phrase to dashes and reveals the next (i.e., non-cumulative)

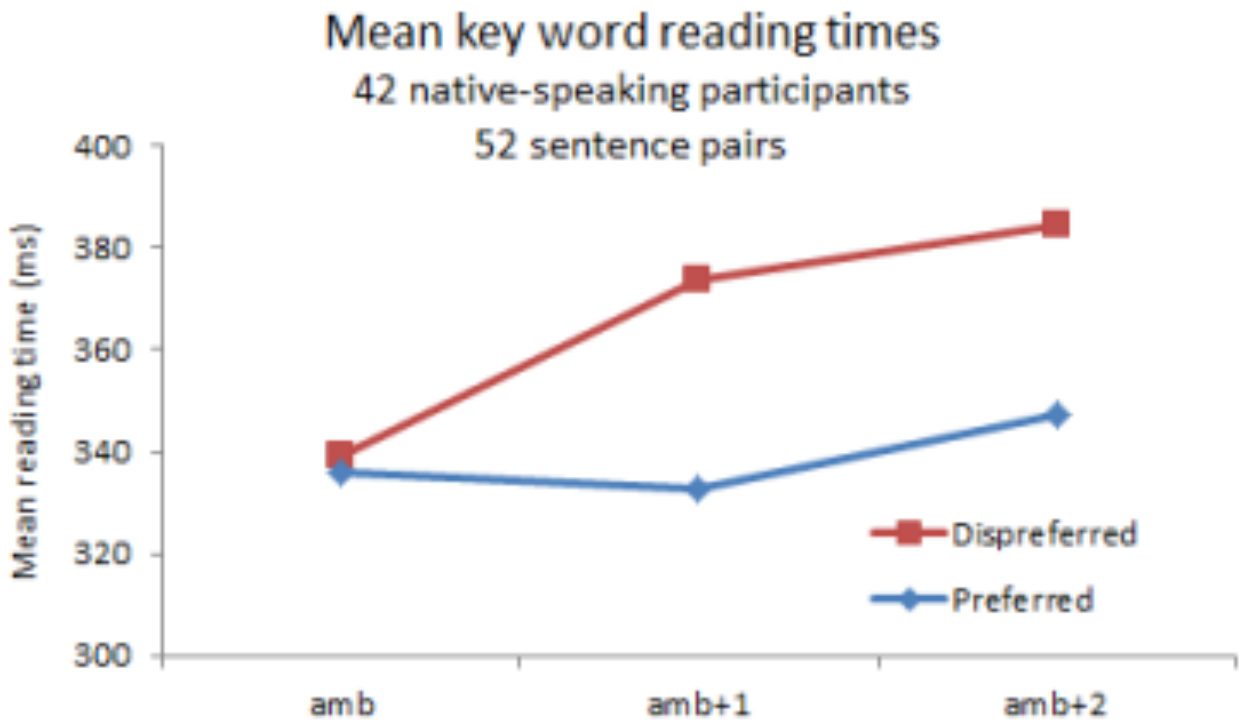
```
James knew that kids like Hannah enjoyed playing …
```

- Additional offline tasks (e.g., plausibility judgments, comprehension questions, etc.) can be used to ensure participant is attending

Demo:

http://www.intro2psycholing.net/experiments/sentence_processing/self_paced/selfpaced_index.php

Results



James knew that kids... like Hannah enjoyed playing....

James knew that kids... like Hannah because she enjoys ...

Self-Paced Reading

Strengths

- Sheds light on processes underlying incremental integration of words into developing sentence representations
- Provides measure for every word (no skipping)
- Cheap and easy to implement (e.g., can be done in web-based environments)

Weaknesses

- Slower than normal reading (about 300-350 vs 175-200 ms/word)
- Encourages early commitments (no parafoveal preview)
- ***Spillover effects*** — readers may click fast and buffer words in memory, leading to effects becoming visible "downstream" of critical word

→ Phrase-by-phrase presentation can help overcome these issues, but loses precision

Course Experiment

What open questions remain after van Berkum et al. (2008)?

Manipulation: Could the speaker-inconsistency effect interact with addressee knowledge?

- Imagine an addressee who knows something about the “unusual” speaker

Measures: Could these results extend to reading?

- Imagine getting an email from someone that you either know or don’t know

Could such speaker inconsistencies be reflected in ICA?

Homework: Read student paper from last semester **(S16_final_paper.pdf)**

- What were the Stimuli, Task, DV and Linking Hypothesis?
- Can you think of any confounds that were missed?
- Problems with the stimuli or procedure? Ideas for fixing?

Agenda

Next Monday

- Read: Student paper for our course experiment
- Discussion of course experiment

Next Wednesday

- Read: Demberg & Sayeed (2016)
- ICA and self-paced reading combined
- Apply to course experiment