Don't measure height with a stopwatch

:
What 'laboratory linguistics'
is(n't) good for

Colin Phillips
Department of Linguistics
Neuroscience & Cognitive Science Program
University of Maryland

languagescience.umd.edu

How I became a psycholinguist

Fall 1991: semantics seems cool!

Spring 1993: morphosyntax of obscure Papuan language

Spring 1995: learning to use slides

Summer 1996: (mostly armchair) psycholinguistics

Linguistics + Brain & Cognitive Science partnership

NSF-funded Research Training Grant ('RTG Program'), 1992-8

Students from each department added ~1 year of training in the other dept.

NSF supported US students
MIT supported international students to do the same (thank you!)

What we learned then is more routine nowadays for linguistics students.
Linguistics looks quite different now than 20 years ago.
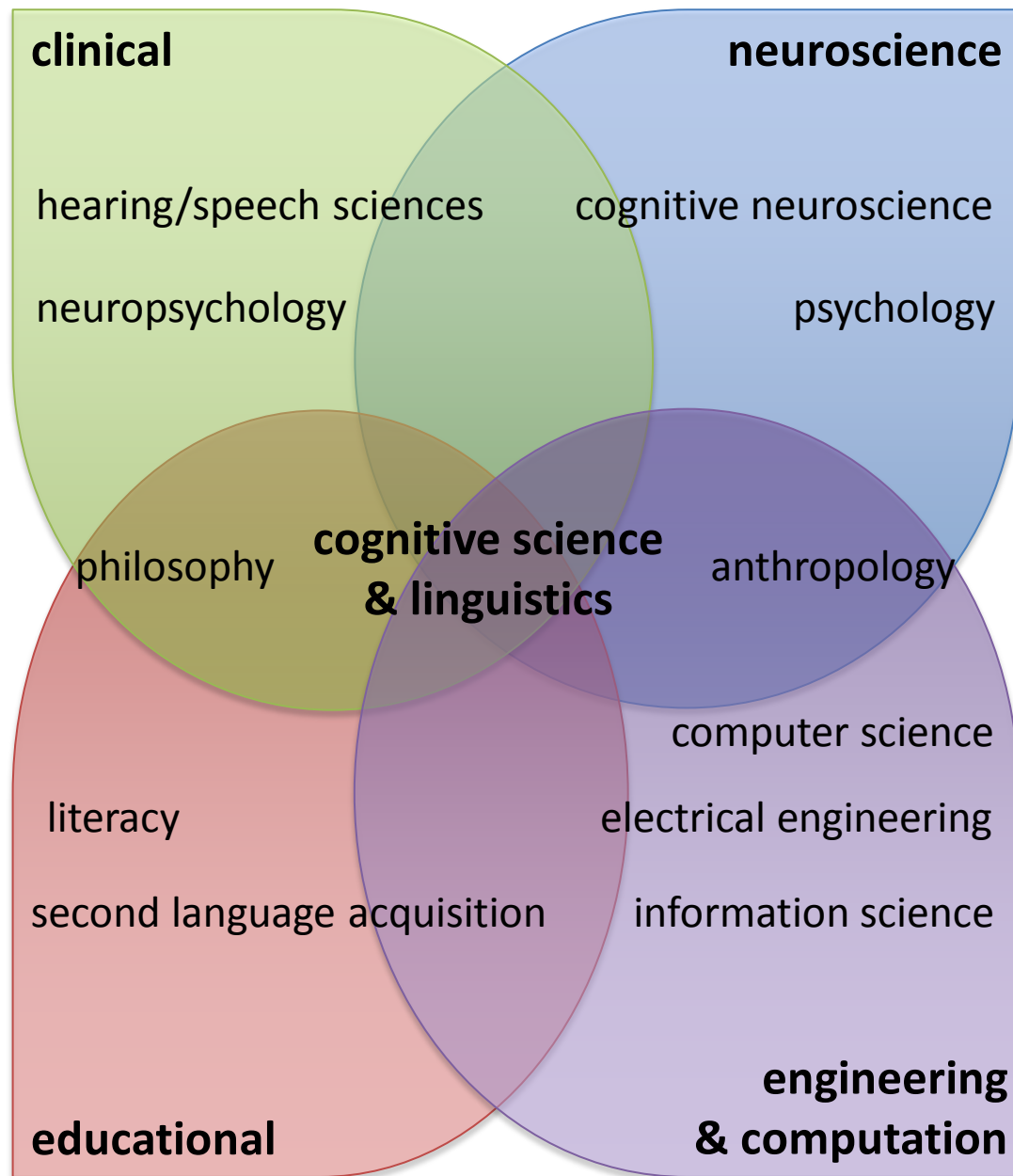… but other fields don't look at linguistics much differently than 20 years ago.
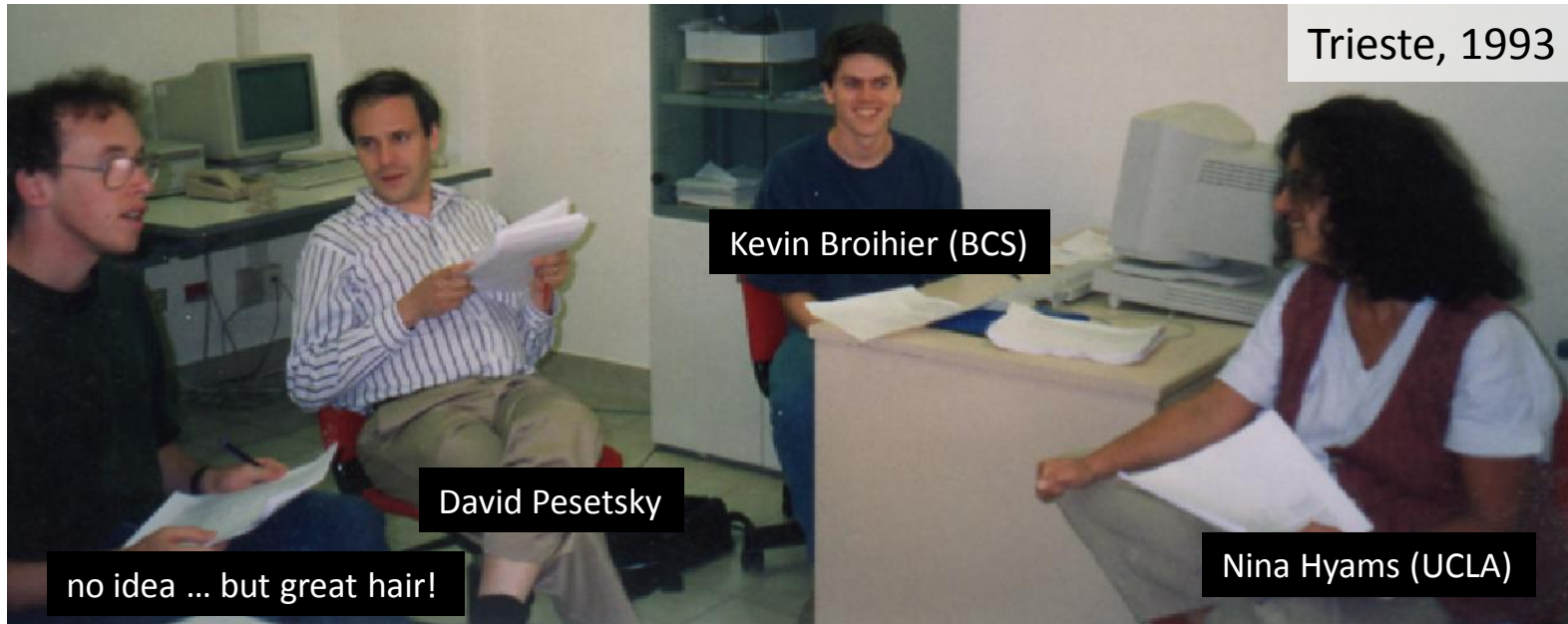
Trieste, 1993

Kevin Broihier (BCS)

David Pesetsky

no idea … but great hair!

Nina Hyams (UCLA)

Language Science @ UMd

clinical

hearing/speech sciences

neuropsychology

neuroscience

cognitive neuroscience

psychology

philosophy

cognitive science & linguistics

anthropology

literacy

second language acquisition

computer science

electrical engineering

information science

educational

engineering & computation

200 faculty, PhD students, researchers
12 departments & centers
Collaborative research networks
Courses, talks, outreach
Annual 2-week 'Winter Storm' event
Links to technology, education, clinical

Different than 20 years ago:
Students have broader skill-set
Broader network, goals

Same as 20 years ago:
Seeking reciprocal links, inclusiveness
Linguistics is a hard sell

Trying to put linguistics at center of conversations about language

languagescience.umd.edu

A couple of things that I have learned

1. Don't say "I will never work on X"
2. Success and speed are overrated (in language use)

Rule Learning by
Seven-Month-Old Infants

Linguistic Experience Alters Phonetic Perception in Infants by 6 Months of Age

PATRICIA K. KUHL,* KAREN A. WILLIAMS, FRANCISCO LACERDA, KEN...

Sentence Perception as an Interactive Parallel Process

Abstract. *The restoration of disrupted words to their original form in a sentence shadowing a...*

Brain Activity During Speaking: From Syntax to Phonology in 40 Milliseconds

Miranda van Turennout,* Peter Hagoort, Colin M. Brown

In normal conversation, speakers translate thoughts into words at high speed. To enable this speed, the retrieval of distinct types of linguistic knowledge has to be orchestrated with millisecond precision. The nature of this orchestration is still largely unknown. This report presents dynamic measures of the real-time activation of two basic types of linguistic knowledge: syntax and phonology. Electrophysiological data demonstrate that during noun-phrase production speakers retrieve the syntactic gender of a noun before its abstract phonological properties. This two-step process operates at high speed: the data show that phonological information is already available 40 milliseconds after syntactic properties have been retrieved.

van Turennout et al., *Science,* 1998

Study Reveals: Babies are Stupid

May 21st, 1997

LOS ANGELES – A surprising new study released Monday by UCLA's Institute For Child Development revealed that human babies, long thought by psychologists to be highly inquisitive and adaptable, are actually extraordinarily stupid.

The study, an 18-month battery of intelligence tests administered to over 3,500 babies, concluded categorically that babies are "so stupid, it's not even funny."

According to Institute president Molly Bentley, in an effort to determine infant survival instincts when attacked, the babies were prodded in an aggressive manner with a broken broom handle. Over 90 percent of them, when poked, failed to make even rudimentary attempts to defend themselves. The remaining 10 percent responded by vacating their bowels.

It is unlikely that the presence of the babies' fecal matter, however foul-smelling, would have a measurable defensive effect against an attacker in a real-world situation," Bentley said.

Another test, in which the infants were placed on a mound of dirt outdoors during a torrential downpour, produced similarly bleak results.

"The chicken, dog and even worm babies that we submitted to the test as a control group all had enough sense to come in from the rain or, at least, seek shelter under a leafy clump of vegetation or outcropping of rock," test supervisor Thomas Howell said. "The human babies, on the other hand, could not grasp even this incredibly basic concept, instead merely lying on the ground and making gurgling noises."

The --- --- -- …
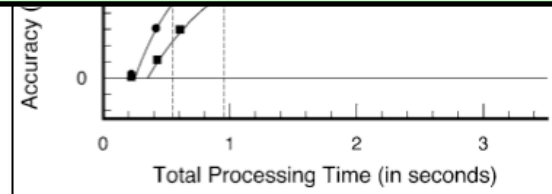--- cat --- -- …
--- --- sat -- …

Self-paced reading

1 2 3 4 5 6 7
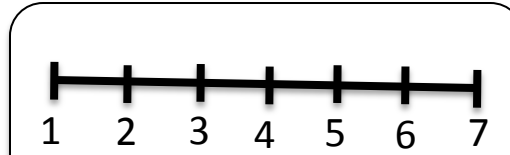
Functional M   **ment**

We are limited less by the sensitivity of our tools

… than by the sensitivity of our theories

**Electroencephalography
(EEG/ERP)**

Total Processing Time (in seconds)

Speed-accuracy tradeoff

Infant language learning

Eye-tracking

# traditional questions

process neutral account of ±grammatical sentences

traditional tools

acceptability judgments (on steroids)

'Experimental Syntax'

- acceptability judgments: lots of them!

- yes/no
  scale ratings
  magnitude estimation

- we do *lots* of these
  ... but the results are rarely surprising

- most results of careful armchair linguistics supported

  Sprouse & Almeida 2011ab: ~98% replication* of cases in Adger's *Core Syntax* textbook and in *Linguistic Inquiry 2001-2010*

- statistical significance is overrated
  subtle contrasts are subtle
  so statistics don't make the contrasts any more categorical
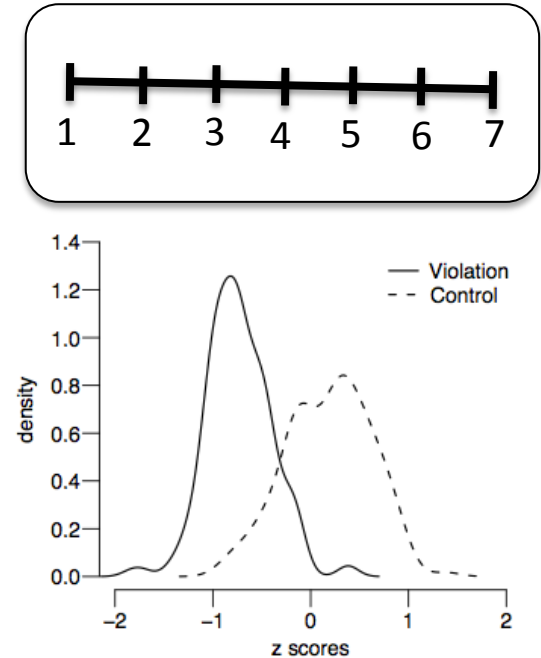
n > 1

- Easy to introspect about individual examples or minimal pairs

- Introspection not feasible for more complex patterns, within & across individuals

- Example #1: large scale judgments as tests of abstract theories (Gibson et al.)

  Theories claim that some superficially similar phenomena are unrelated
  Theories claim that some superficially different phenomena are related

 Test: greater co-variance in judgments for related phenomena

- Example #2: testing reductionist accounts of island effects …

n > 1

- Example #2: testing 'reductionist' accounts of islands

- Theoretical claim: island violations are not ungrammatical, just difficult

  (e.g., Kluender & Kutas 1993; Hofmeister & Sag 2010)

- **if** island effects reflect processing capacity overload,
  then severity of island effects should covary with individual capacity

- Island: interaction of (i) long extraction, (ii) island-inducing structures

  Subject island
  a. Who ____ thinks the speech interrupted the TV show?
  b. Who do you think ____ interrupted the TV show?
  c. Who ____ thinks the speech about global warming interrupted the TV show?
  d. *Who do you think the speech about ____ interrupted the TV show?

  ±Length   ±Complex Subject   ±Both

Sprouse, Wagers, & Phillips 2012, *Language*
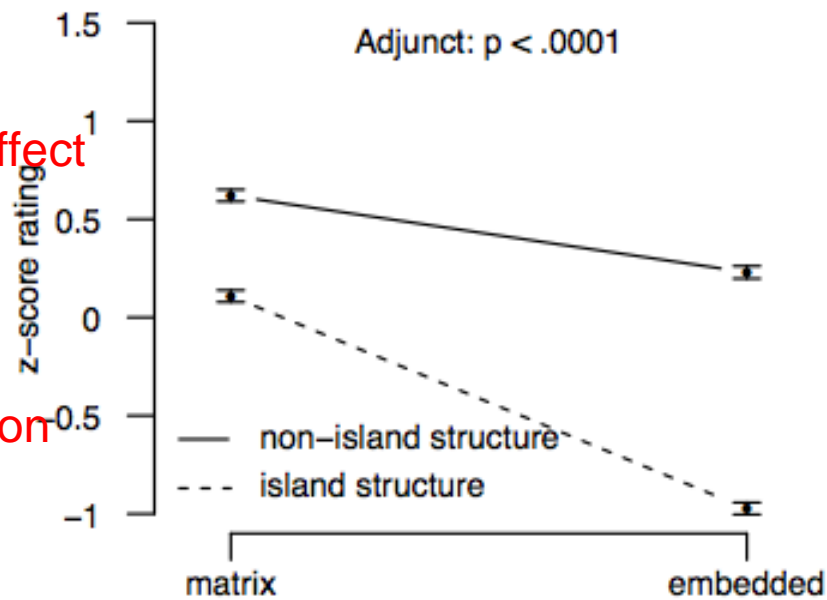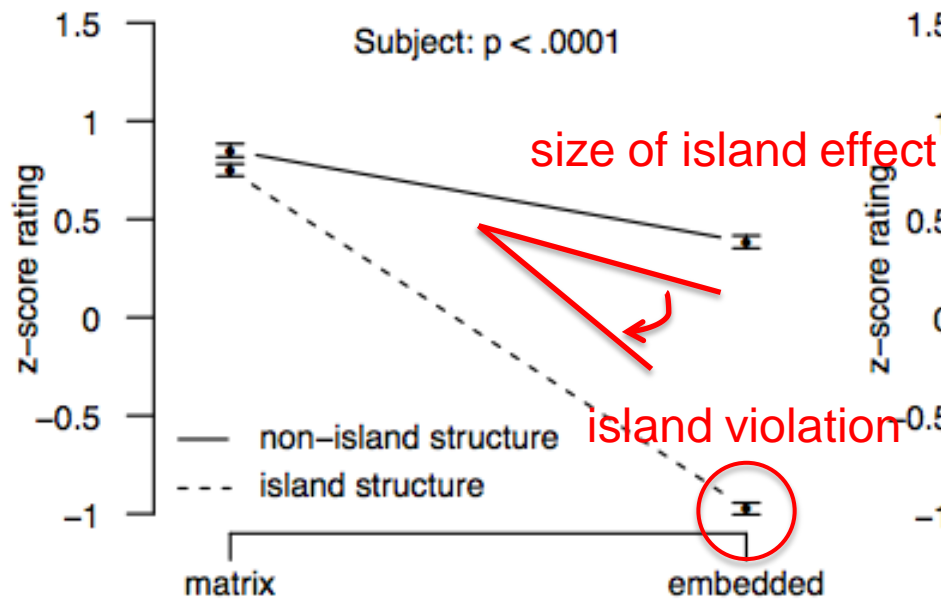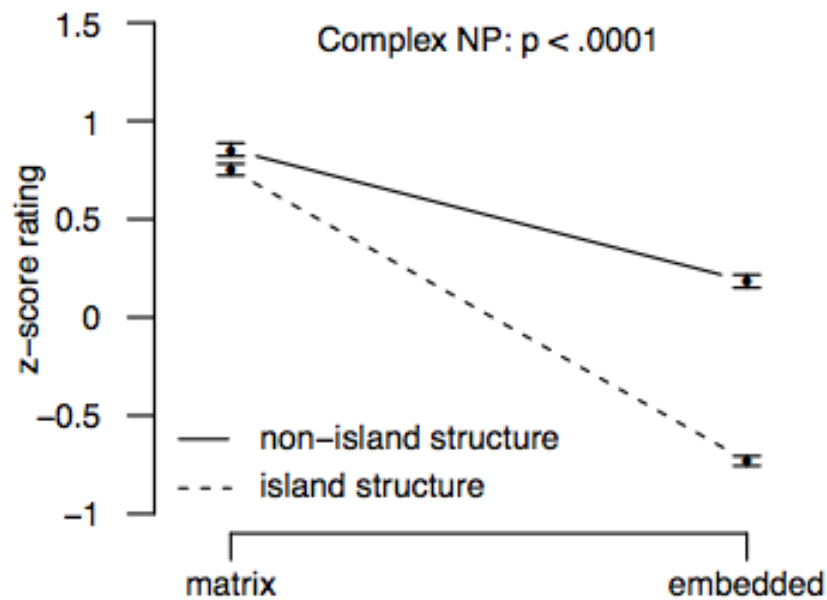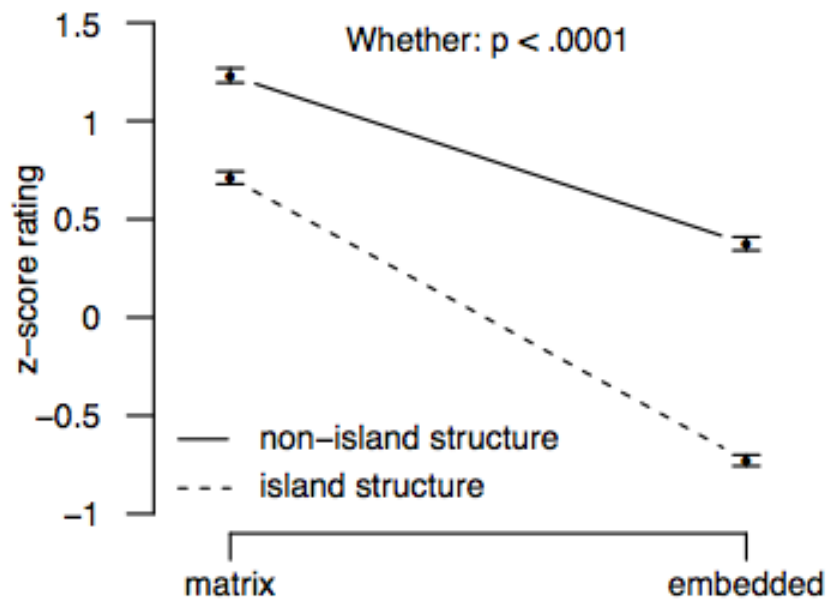
4 island types, 2 memory tasks (serial recall, n-back), n = 315
island severity vs. memory capacity
capacity differences account for 0% - 3% of variance

Sprouse, Wagers, & Phillips 2012, *Language*
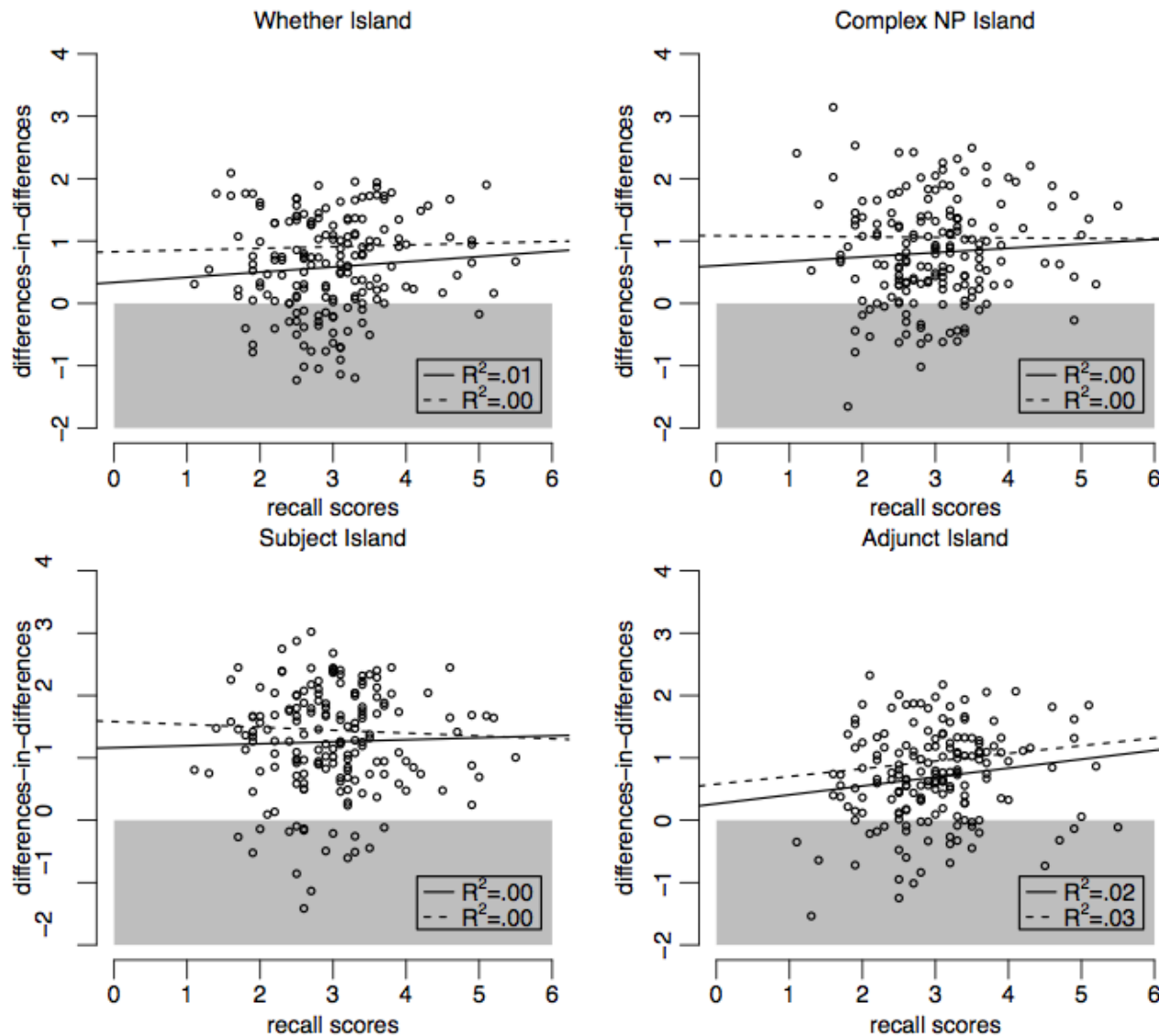
traditional questions

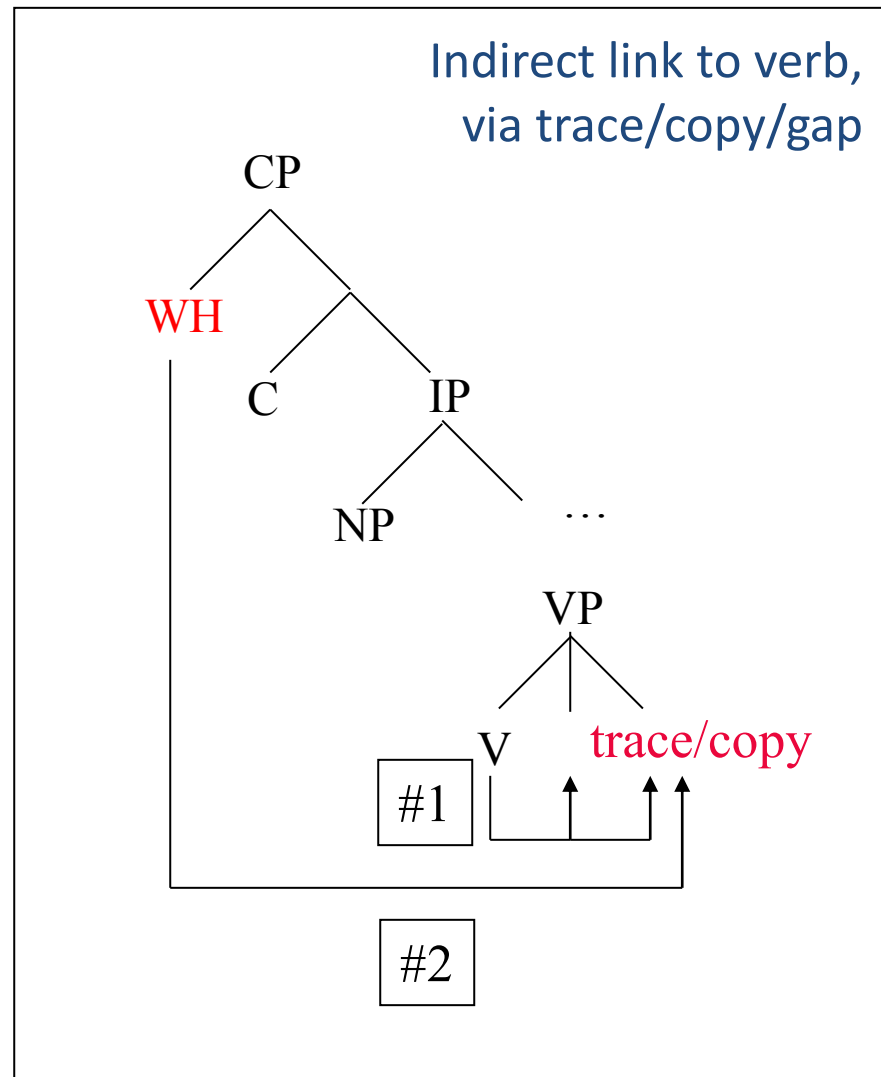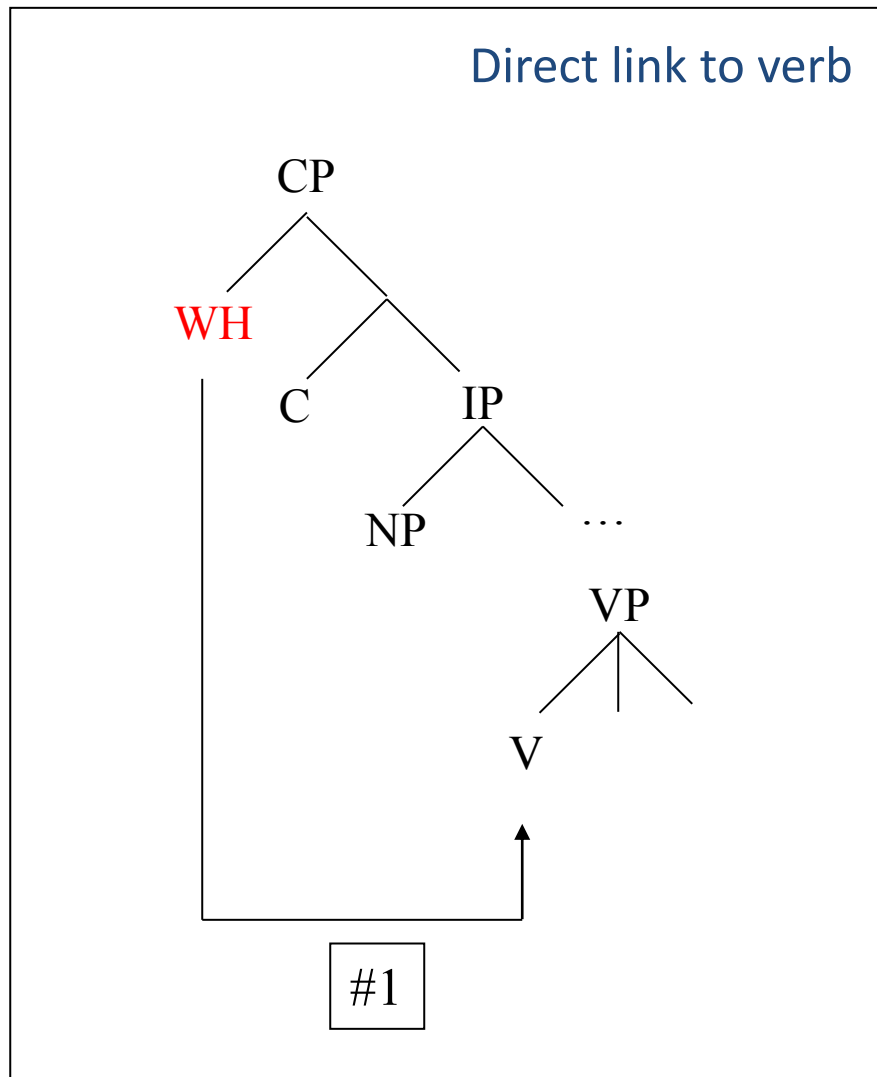process neutral account of ±grammatical sentences

non-traditional tools

time-sensitive measures
developmental dissociations

experiments as arbiters of theoretical disputes

Timing & Traces

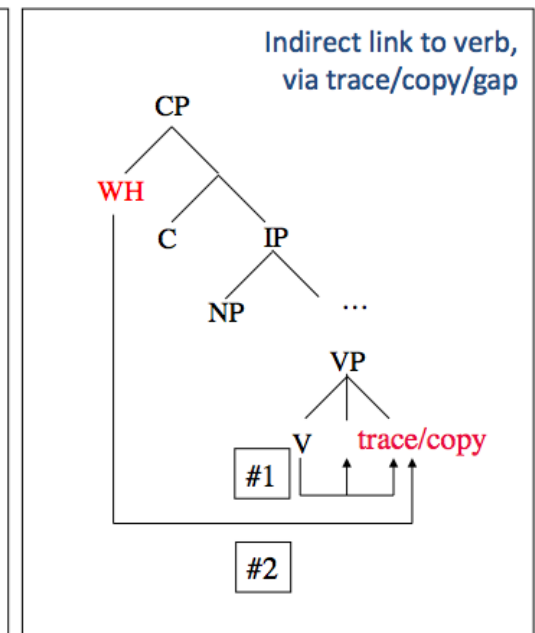Direct link to verb

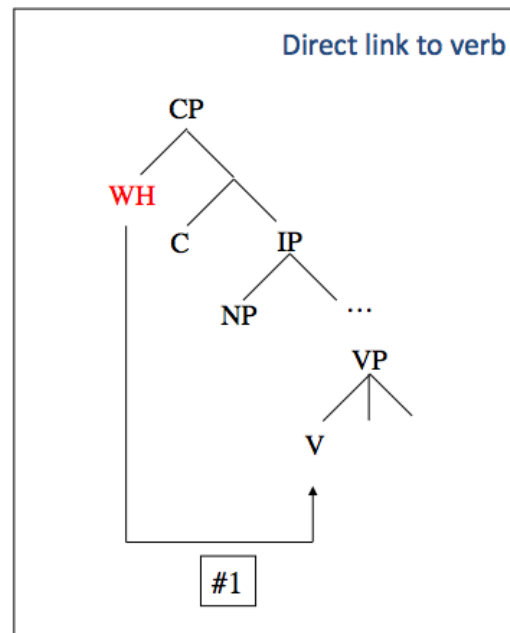Indirect link to verb,
via trace/copy/gap

Wh-phrase forms dependency with different position in different theories.
Suggestion: use timing evidence to decide between competing theories

Timing & Traces

Can timing evidence decide
between competing theories?

Much evidence collected …

Wh-dependencies in English formed at verb, i.e., earlier than gap site
> (e.g., Crain & Fodor 1985; MacDonald 1989; Pickering & Barry 1991; Kaan et al. 2000)

Wh-dependencies in Japanese formed at gap site, i.e., earlier than verb
> (e.g., Nakano, Felser, & Clahsen 2002; Aoshima, Phillips, & Weinberg 2004)

Problem: timing evidence isn't decisive *because structure-building is not synchronous with spoken/written input*   (Gibson & Hickok 1993; Phillips & Wagers 2007)

In other words: the theories don't really make timing predictions.

So, fine-grained timing measures weren't so useful after all

Developmental Dissociations

Rationale: do children reveal theoretically interesting contrasts that are obscured in adults?

Method: ≈ acceptability test ... but we can't ask kids directly, experimental tricks needed

Suitable? Yes, the measure fits the theoretical question ... but interpreting the experimental findings is not trivial

Theoretically decisive? Sometimes

Binding vs. coreference

Mama Bear$_i$ touched her$_i$
Every bear$_i$ touched her$_i$

Delay of Principle B Quantificational Asymmetry

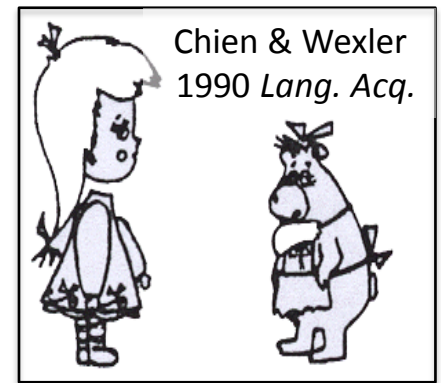Attempt to defend against criticisms in Elbourne (2005) & review of 30+ studies on PrB

~~*Delay of Principle B Quantificational Asymmetry*~~

but fragility is very informative

Imperfective Paradox

Monkey build.imp smurf

Children interpret imperfectives as entailing completion .. unless temporal frame-of-ref given.

Chien & Wexler 1990 *Lang. Acq.*

Conroy, Takahashi, Lidz, & Phillips 2009 *Ling. Inq.*

Kazanina & Phillips 2007 *Cognition*

non-traditional questions

theories of real-time grammatical computation

non-traditional tools

many toys!

Real-time Linguistic Computation

- Traditional view: generative grammars are not process models

- But: sentence representations appear not to be process-independent, i.e., we can't create the same representation in multiple ways

  → the traditional process-neutrality is merely a convenient simplification
  → better to develop process models, claims about real-time operations

- How do we <u>encode</u> and <u>navigate</u> linguistic representations in memory?

  What mental operations underlie relations that we normally draw with indices, lines, copies, etc.?

- Initial strategy
  Try to show that real-time processes are "grammatically infallible"

Selective Fallibility

- Many grammatical constraints have robust & early on-line effects

- Binding Condition C

  While he was reading the book, John ate an apple.

  He ate an apple while John was reading the book.

  Search for pronoun interpretation ignores inappropriate nouns

  (Kazanina et al. 2007; Aoshima et al. 2009)

- Island constraints – does on-line gap search look inside islands?

  No!   (many demonstrations)

  Even Japanese head-final relative clauses (Yoshida, Aoshima & Phillips 2004)

- Implies grammatically rich representations and processes

Selective Fallibility – Grammatical Illusions

Agreement

"The ideas that Cheney are putting forward could find fertile ground with the American people." (*Washington Post,* 5/21/09)

"The word order preference previously observed in [..] type languages are not universal." (*Poster session – today!*)
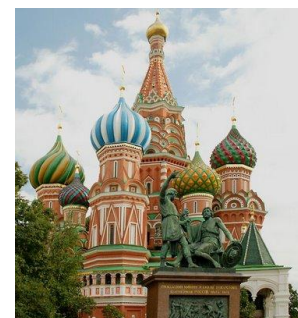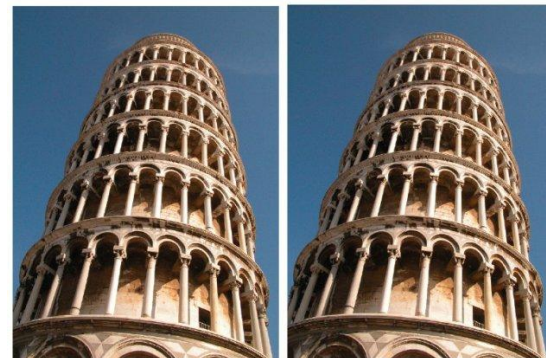

Comparatives

"More people have been to Russia than I have."

Negative Polarity Items (NPIs)

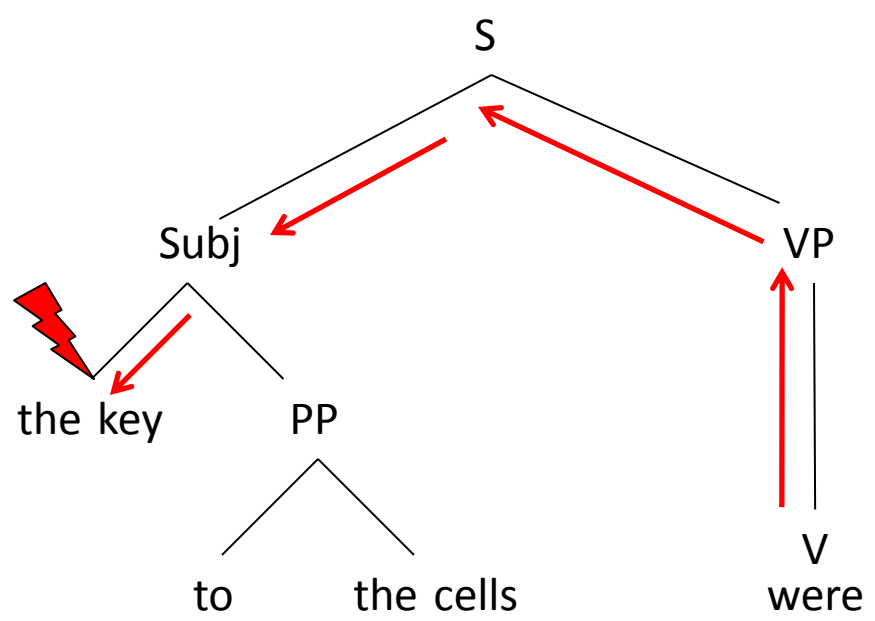"The bills [that no senators supported] will ever become law."

Thematic Role Reversals

"I'm not going to solely blame all of man's activities on climate change." (9/30/08)
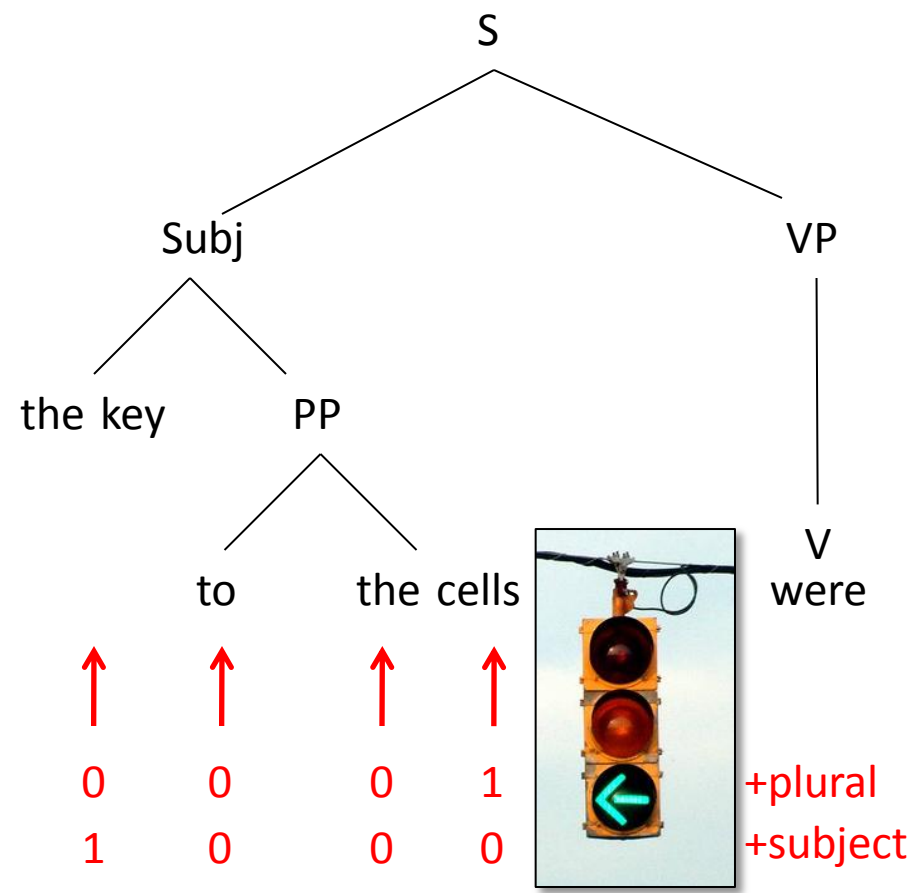
Two ways to search structures in memory

serial, structure-guided search

```
                        S
              /                  \
          Subj                    VP
         /    \                   /  \
   the key     PP                /    \
              /   \             V
            to   the cells    were
```

structure-sensitive, avoids interference

slow, esp. for longer relations

parallel, cue-guided search

```
                        S
              /                  \
          Subj                    VP
         /    \                    |
   the key     PP                 V
              /   \              were
            to   the cells
```

0	0	0	1	+plural
1	0	0	0	+subject

susceptible to interference

fast, even for longer relations

McElree et al. 2003; Lewis et al. 2006; Wagers, Lau, & Phillips, 2009

Same Memory – Different Access

Subject-Verb Agreement

The diva [that accompanied the **harpist** on stage] clearly **was** flawless …
The diva [that accompanied the **harpists** on stage] clearly **was** flawless …

The diva [that accompanied the **harpist** on stage] clearly **were** flawless …
The diva [that accompanied the **harpists** on stage] clearly **were** flawless …

illusion

Subject-Reflexive Agreement

The diva [that accompanied the **harpist** on stage] clearly presented **herself** …
The diva [that accompanied the **harpists** on stage] clearly presented **herself** …

The diva [that accompanied the **harpist** on stage] clearly presented **themselves** …
The diva [that accompanied the **harpists** on stage] clearly presented **themselves** …

no illusion

Both processes require access to identical element
-- the subject of the same clause.

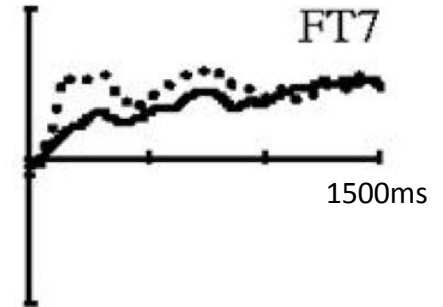Dillon, Mishler, Sloggett, & Phillips, 2011

Timing – surprising speed

Ultra-fast syntactic analysis seen with ERPs

Early Left Anterior Negativity (ELAN) – 150-250ms

John criticized Max's proof <u>of </u>the theory.
John criticized Max's <u>of </u>proof the theory.

(Hahne et al., 2002)

Timing of brain response is as fast or faster than lexical access.
This leaves almost zero time for syntactic analysis.
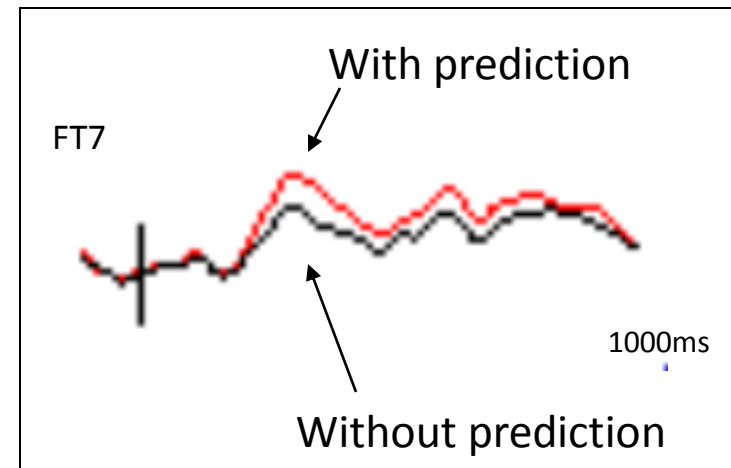
<u>Proposal</u>
Speed is achieved by syntactic pre-computation
Possessor predicts noun → preposition mismatches

<u>Evidence</u>
Ellipsis can "turn off" prediction for noun

Although I like Bill's book, I don't like Max's …

→reduced ELAN effect

Syntactic predictions are very fast

(Lau, Stroud, Plesch, & Phillips, 2006)

Timing – surprising slowness

ERP N400 – reflects lexico-semantic processes
(Kutas & Hillyard 1980)

Generally shows clear semantic sensitivity

She was stung by a {bee|wasp} (Kutas et al.)
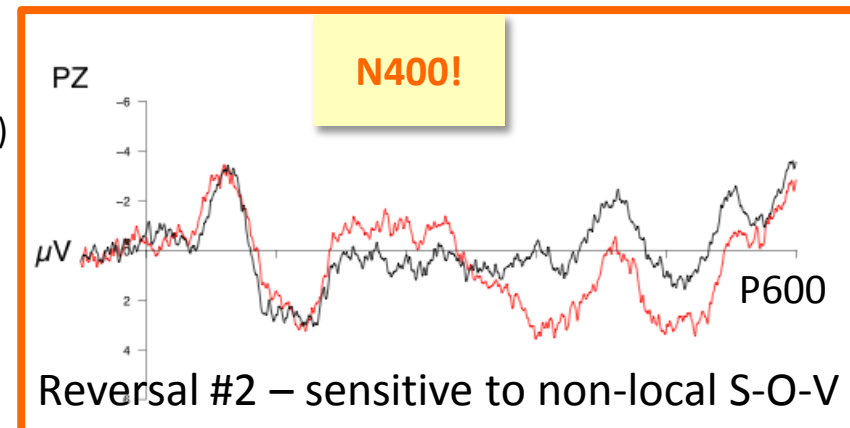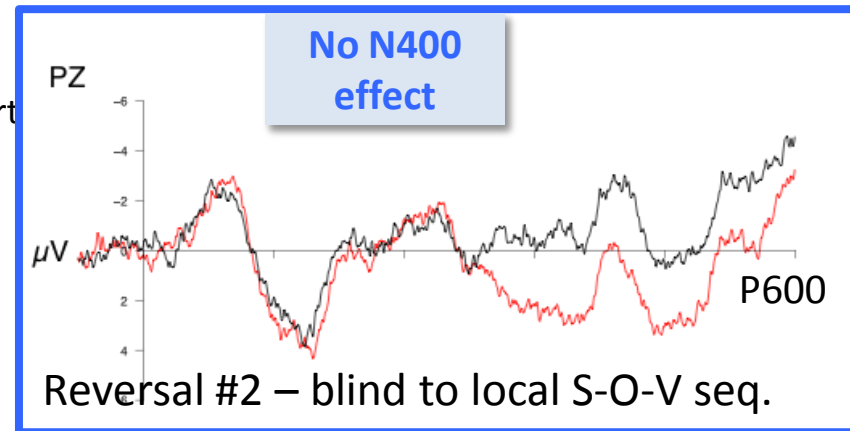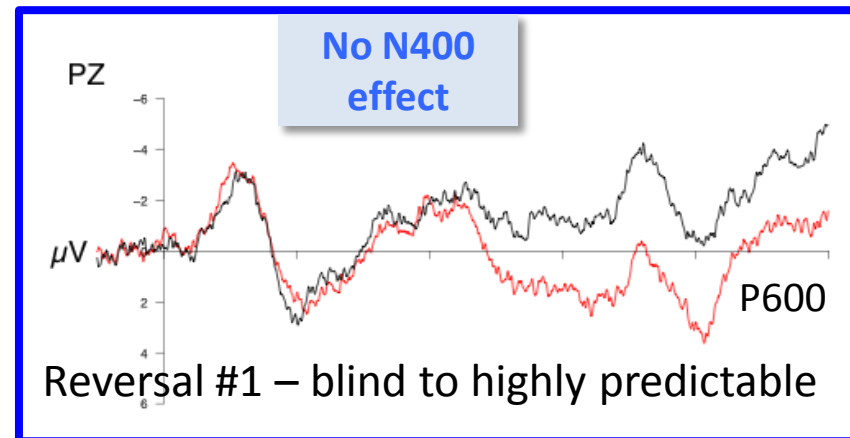
The author {wrote|began} the book. (Kuperberg, Hagoort

I know which book you {read|hoped} that …

Cases of semantic 'blindness'

A robin is not a {bird|tree} (Fischler et al. 1983)

cop BA thief arrest (Chow & Phillips 2011)
thief BA cop arrest

last week thief BA cop arrest
thief BA cop last week arrest

Reversal #1 – blind to highly predictable

Reversal #2 – blind to local S-O-V seq.

Reversal #2 – sensitive to non-local S-O-V

Don't measure height with a stopwatch

- Lab toys are nice

  But they become more useful once we develop theories/models that have the temporal (or spatial) precision of the tools

- High-powered **judgment** studies yield few surprises about simple paradigms
  Statistical precision does not make subtle judgments clearer

  But more complex hypotheses about judgments can be explored

- **Arbitrating** traditional process-neutral theories:
  Timing evidence is good for testing timing predictions
  Developmental dissociations do sometimes work out

- Developing detailed **models/theories** with **time/space predictions**
 Rapidly developing area; sometimes slow and dumb is very informative
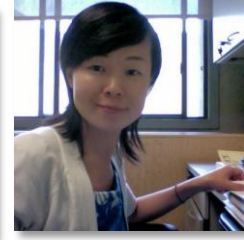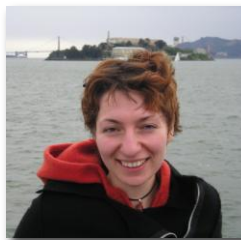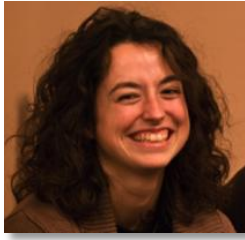
Matt Wagers (UCSC)
Ellen Lau (MGH/UMd.)
Brian Dillon (UMass)
Nina Kazanina (Bristol)
Masaya Yoshida (NWU)
Ming Xiang (Chicago)
Clare Stroud (Natl. Acad.)

Sachiko Aoshima
Akira Omaki (JHU)
Pedro Alcocer
Wing Yee Chow
Jon Sprouse (UC Irvine)
Alexis Wellwood
Dan Parker

Amy Weinberg
Jeff Lidz
Roumi Pancheva (USC)
Valentine Hacquard
Moti Lieberman (McGill)
Shevaun Lewis
Dave Kush

"It has sometimes been argued that linguistic theory must meet the empirical condition that it account for the ease and rapidity of parsing. But parsing does not, in fact, have these properties. [...] In general, it is not the case that language is readily usable or 'designed for use.'" (Chomsky & Lasnik, 1993, p. 18)
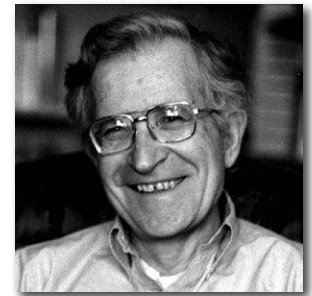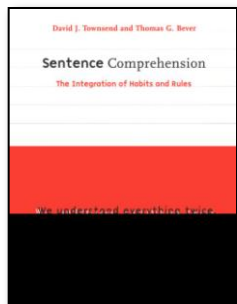

Noam Chomsky, MIT

Howard Lasnik, Maryland

Fernanda Ferreira, Edinburgh

"...the language comprehension system creates representations that are 'good enough' (GE) given the task that the comprehender needs to perform. GE representations contrast with ones that are detailed, complete, and accurate with respect to the input." (Ferreira & Patson, 2007, p. 71)

Tom Bever, Arizona

"we understand everything twice"
(Townsend & Bever, 2001)

At least 4 different uses of the terms:
1. Distinction between mechanism and behavior   [*necessary*]
2. Declarative vs. procedural specification of a formal system   [*necessary*]

3. Distinction between specialized mental systems   [*empirical hypothesis*]
4. Abstraction from resource limitations   [*controversial]*