Structuring Expectation

Matt Wagers
University of California, Santa Cruz

WCCFL 33
Simon Fraser University, Vancouver
March 27-29, 2015

Modes of explanation

- How do we negotiate the incremental uncertainty of language?

- Explanations for what is easy and what is difficult in language processing come in three (overlapping) flavors

  - properties of the representation/representation-building

  - what's likely

  - what's useful in context

What is likely?

- Language processing is probabilistic in at least the sense that, for some unfolding expression, we have an ordering of possible analyses based on likelihood or confidence

- Difficulties occur in language processing when incoming information dramatically shifts our pre-existing allocation of confidence or likelihood (Entropy Reduction, Hale, 2001; Surprisal, Levy, 2008)

  - The horse **raced past the barn** … S → NP **VP**

  - The horse <u>raced past the barn</u> **fell**. S_1 → NP **VP**; NP → NP $\underline{S_2}$

- Very general, approximate and useful formalization

  - *conditional probability:* $P(w_{n+1}|w_0w_1\ldots w_n)$

Three case studies

RC Animacy and English relative clauses

Agr Agreement inside English DPs

Wh Wh-Agreement and Person in Chamorro

Three case studies

RC **Animacy and English relative clauses**

Agr Agreement inside English DPs

Wh Wh-Agreement and Person in Chamorro

QUESTION

Why are object relative clauses easier to understand when the relativized argument is inanimate?

Relative clauses— sources of difficulty

- Generally speaking, subject relative clauses are easier to process and understand than object relative clauses

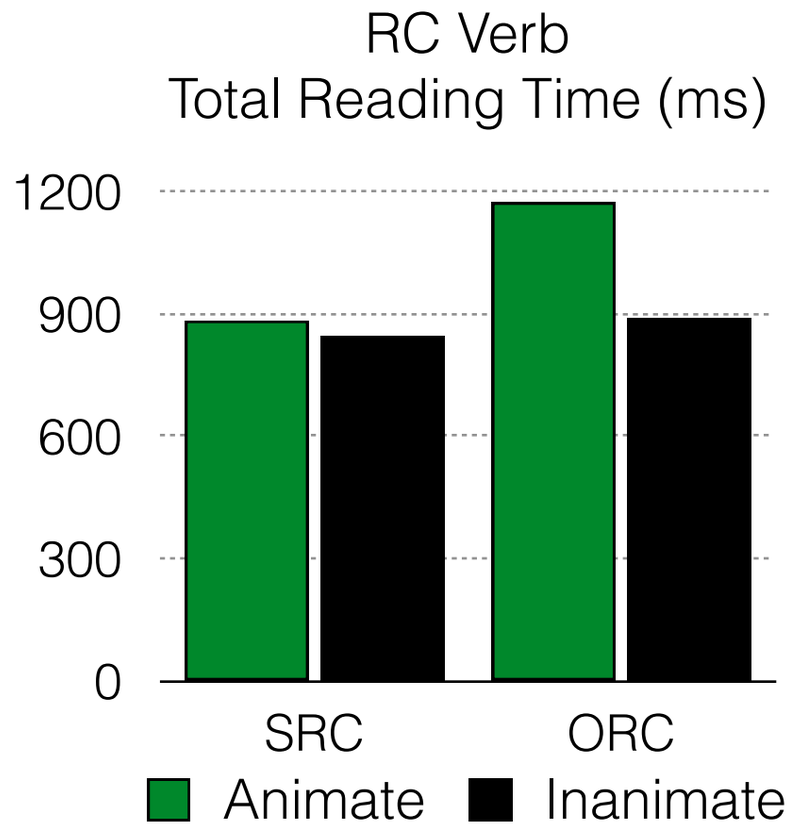
- but … the SRC > ORC advantage can be neutralized under a variety of conditions

- for example, if the RC subject is a pronoun
  (Bever, 1974, Gordon et al. 2001)

- or, if the relativized argument is inanimate
  (Mak et al. 2002, Traxler et al. 2002, Gennari & MacDonald, 2005, i.a.)

Relative clauses— role of animacy

- E.g. Traxler, Morris & Seely (2002):
  - **ANIMATE**
    The director [$_{SRC}$ that __ **watched** the movie]
    The director [$_{ORC}$ that the movie **pleased** __]
  - **INANIMATE**
    The movie [$_{SRC}$ that __ **pleased** the director]
    The movie [$_{ORC}$ that the director **watched** __]



HYPOTHESIS

Animate RC heads are predictively linked to a subject gap.

Inanimate heads are not.
(cf. Active Filler Strategy).

Relative clause expectations: filled gap design

RC

The kindergarten teacher pointed out …

the friendly child who

SUBJECT GAP POSSIBLE

<u>the young girl</u> has played with __ incessantly.

the friendly child with whom <u>the young girl</u> has played __ incessantly.

NO SUBJECT GAP POSSIBLE

Stowe (1986), cf. M.-W. Lee (2004)

Relative clause expectations: filled gap design

RC

The kindergarten teacher pointed out …

the friendly child who, as of yesterday, the young girl has played with incessantly.

SUBJECT GAP POSSIBLE

the friendly child with whom, as of yesterday, the young girl has played incessantly.

NO SUBJECT GAP POSSIBLE

Stowe (1986), cf. M.-W. Lee (2004)

Relative clause expectations: filled gap design

RC

The kindergarten teacher pointed out …

ANIMATE

SUBJECT GAP POSSIBLE

the friendly child who, as of yesterday, <u>the young girl</u> has played with incessantly.

the friendly child with whom, as of yesterday, <u>the young girl</u> has played incessantly.

NO SUBJECT GAP POSSIBLE

INANIMATE

the colorful toy which, as of yesterday, <u>the young girl</u> has played with incessantly.

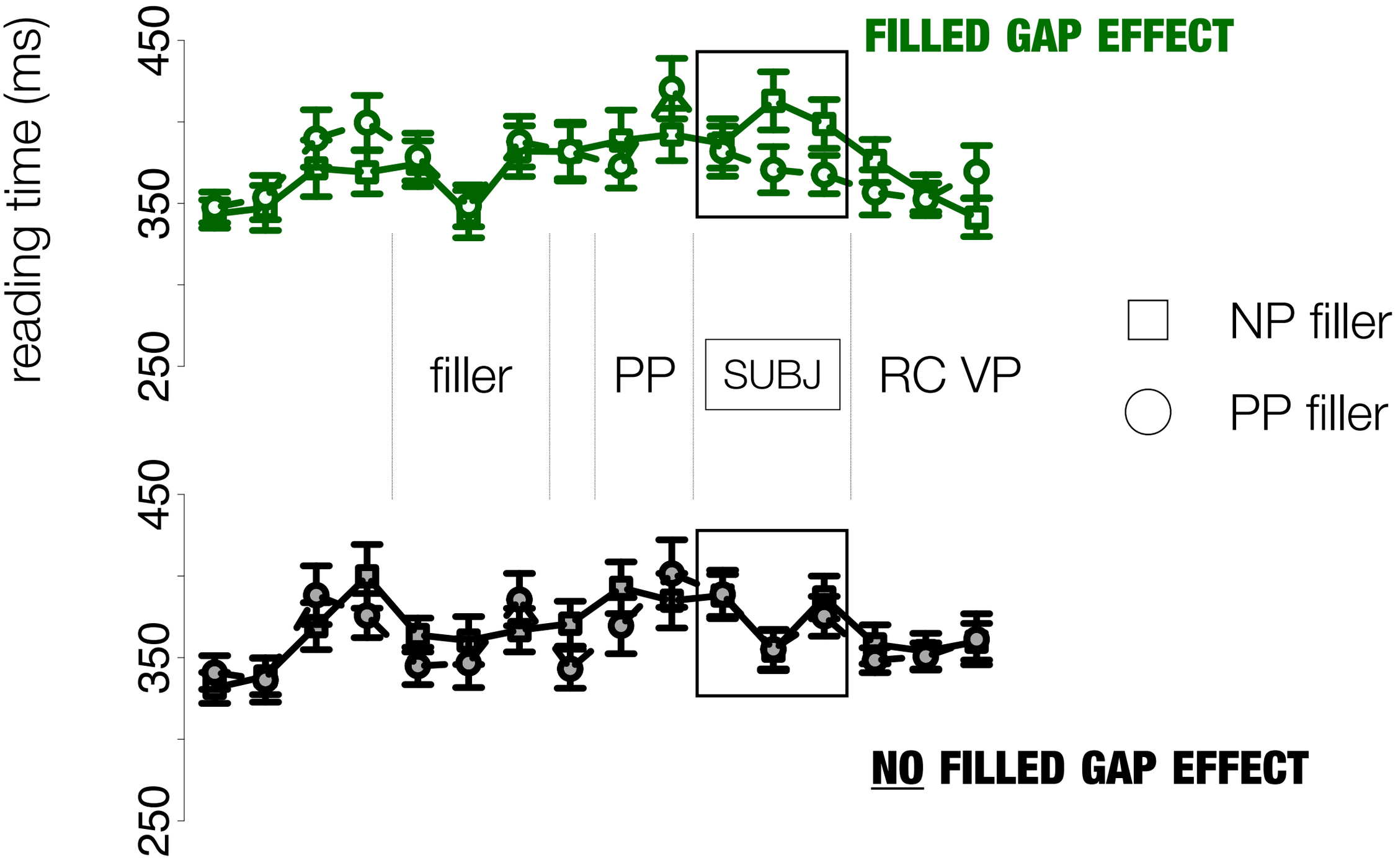
the colorful toy with which, as of yesterday, <u>the young girl</u> has played incessantly.

cf. M.-W. Lee (2004)

Animacy and the expectation for a subject gap

RC

ANIMATE



NO FILLED GAP EFFECT

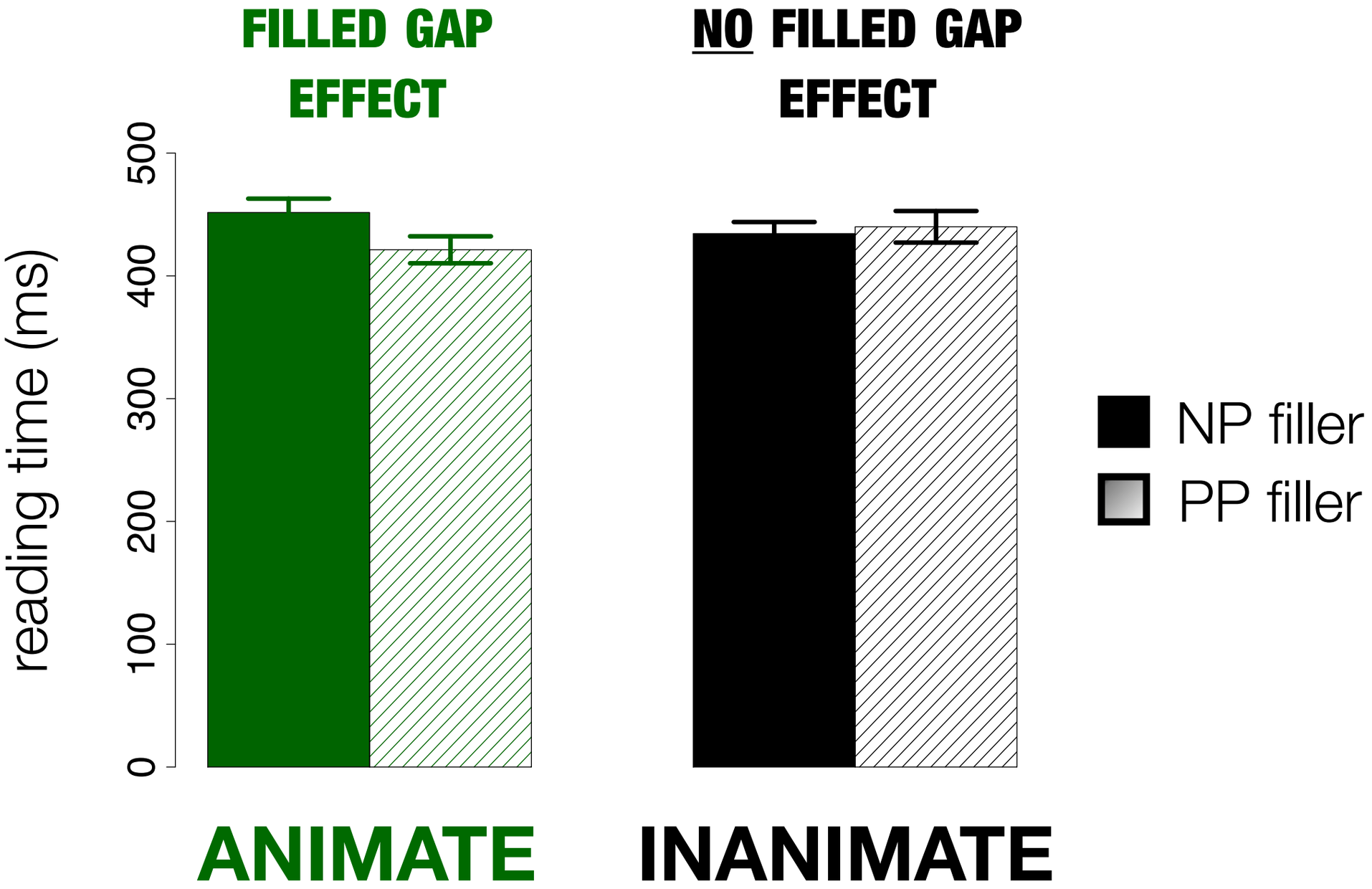
INANIMATE

Pendleton & Wagers (2014)

Animacy and the expectation for a subject gap

RC

Replication attempt



Pendleton & Wagers (2014)

Relative clauses—role of animacy

- Filled-gap design results:
  - **ANIMATE** relative clause heads generate an expectation for a subject gap
  - **INANIMATE** relative clause heads <u>do not</u>

✅

HYPOTHESIS

Animate RC heads are predictively linked to a subject gap.

Inanimate heads are not.

… how does animacy exert this influence?

- What expectations should individuals hold based on their language experience?

Roland, Dick & Elman (2007)

% subject gap

	Brown	Switchboard
Animate	75%	91%
Inanimate	47%	31%

… how does animacy exert this influence?

- Replicated in COCA and Gigaword (parsed NYT subsection)
- … but, what about that RC-initial adjunct?

Roland, Dick & Elman (2007)

% subject gap

	Brown	Switchboard
Animate	75%	91%
Inanimate	47%	31%

% subject gap

+ [XP …]	Gigaword NYT subsection
Animate	99% n = 195
Inanimate	94% n = 449

RC

… how does animacy exert this influence?

- Cloze task (Amazon Mechanical Turk) using actual experimental materials (n=400)

% subject gap completions

	no RC-initial adjunct	adjunct
Animate	100%	
Inanimate	88%	

… how does animacy exert this influence?

- Mediated view
  - $P(w_{n+1}=\text{“the”}|w_0w_1\ldots w_n)$
    - Pre-RC adjunct ⇨ ORC analysis very unlikely
- Direct view, I
  - P(gap:SUBJ | hd:**+**ANIM) >> P(gap:OBJ | hd:**+**ANIM)
  - P(gap:OBJ | hd:**-**ANIM) >> P(gap:SUBJ | hd:**-**ANIM)

… how does animacy exert this influence?

- Direct view, I
  - P(gap:SUBJ | hd:**+**ANIM) >> P(gap:OBJ | hd:**+**ANIM)
  - P(gap:OBJ | hd:**-**ANIM) >> P(gap:SUBJ | hd:**-**ANIM)
- Direct view, II
  - **U**(gap:SUBJ | hd:**+**Anim) >> **U**(gap:OBJ | hd:**+**Anim)
  - **U**(gap:OBJ | hd:**-**Anim) >> **U**(gap:SUBJ | hd:**-**Anim)

… how does animacy exert this influence?

- <u>Mediated view</u>:
  predictions reflect any (known) contingencies

- <u>Direct view, I</u>:
  predictions reflect contingencies on grammatically-active features (or some distinguished set)

- <u>Direct view, II:</u>
  predictions optimize well-formedness

Three case studies

RC Animacy and English relative clauses

Agr **Agreement inside English DPs**

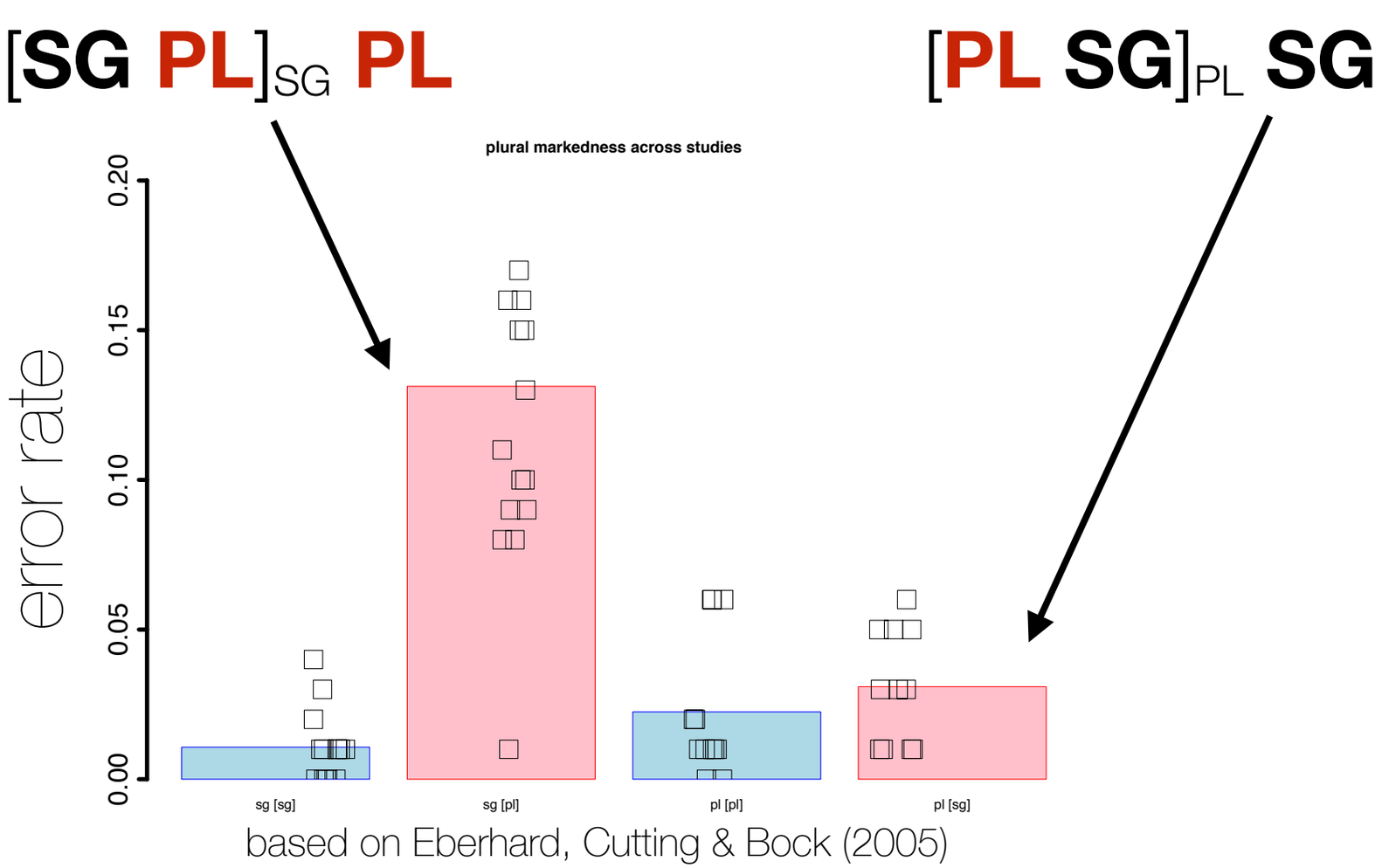
Wh Wh-Agreement and Person in Chamorro

QUESTION

Why are we susceptible to (erroneously) misagree with singular DPs but not with plural DPs?

Agreement attraction

- The **path** to the **monuments** **is**/?**are** littered with bottles.
- The **paths** to the **monument** **are**/***is** littered …



based on Eberhard, Cutting & Bock (2005)

HYPOTHESIS

Plural features are maintained in working memory more durably than **singular** features (e.g., because they are marked).

Bi-partite model of memory

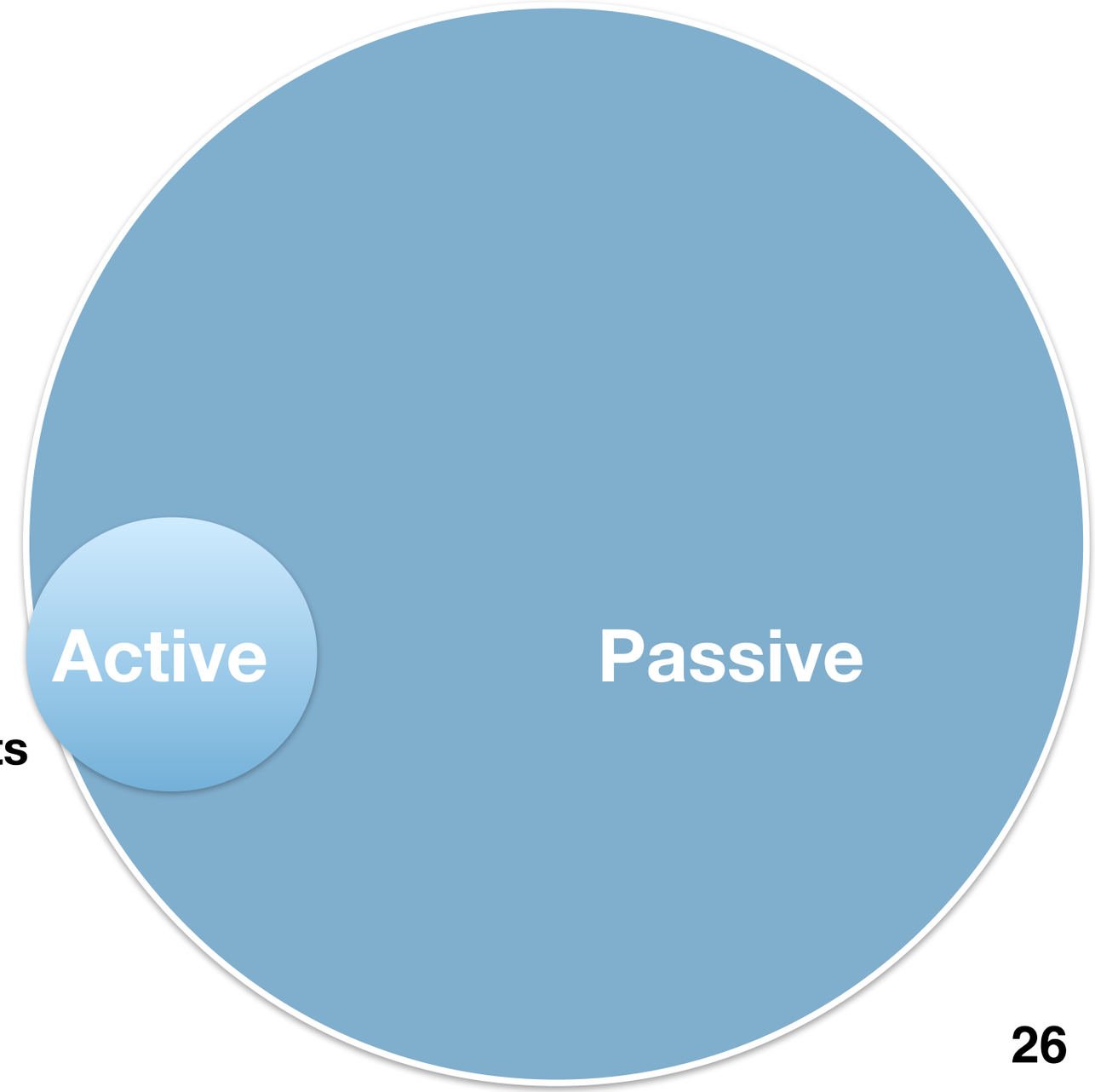
Agr

- Two basic memory states:
  - **active**/focal - stringent capacity limitations - fast processing
  - **passive** - virtually unlimited - requires retrieval/slower

Direct evidence
Broadbent 1958
Wickelgren et al., 1980
Garavan, 1998; Cowan, 2001
McElree, 2006
Verhaegen & Basak, 2007
Jonides et al., 2008

Architecture
ACT-R: Lewis & Vasishth, 2005
Full/reduced representations Hinton, 1990

Similarity-based retrieval interference effects
Gordon et al., 2001, et seq
Van Dyke & Lewis, 2003, Van Dyke, 2007, *et seq.*
Drenhaus et al., 2008
Badecker & Kuminiak, 2007
Wagers, Lau & Phillips, 2009



HYPOTHESIS v.2

Plural features are more likely than **singular** features to be maintained in the active (fast) state of memory.

Testing for the maintenance of [PL]

PL demonstrative DP

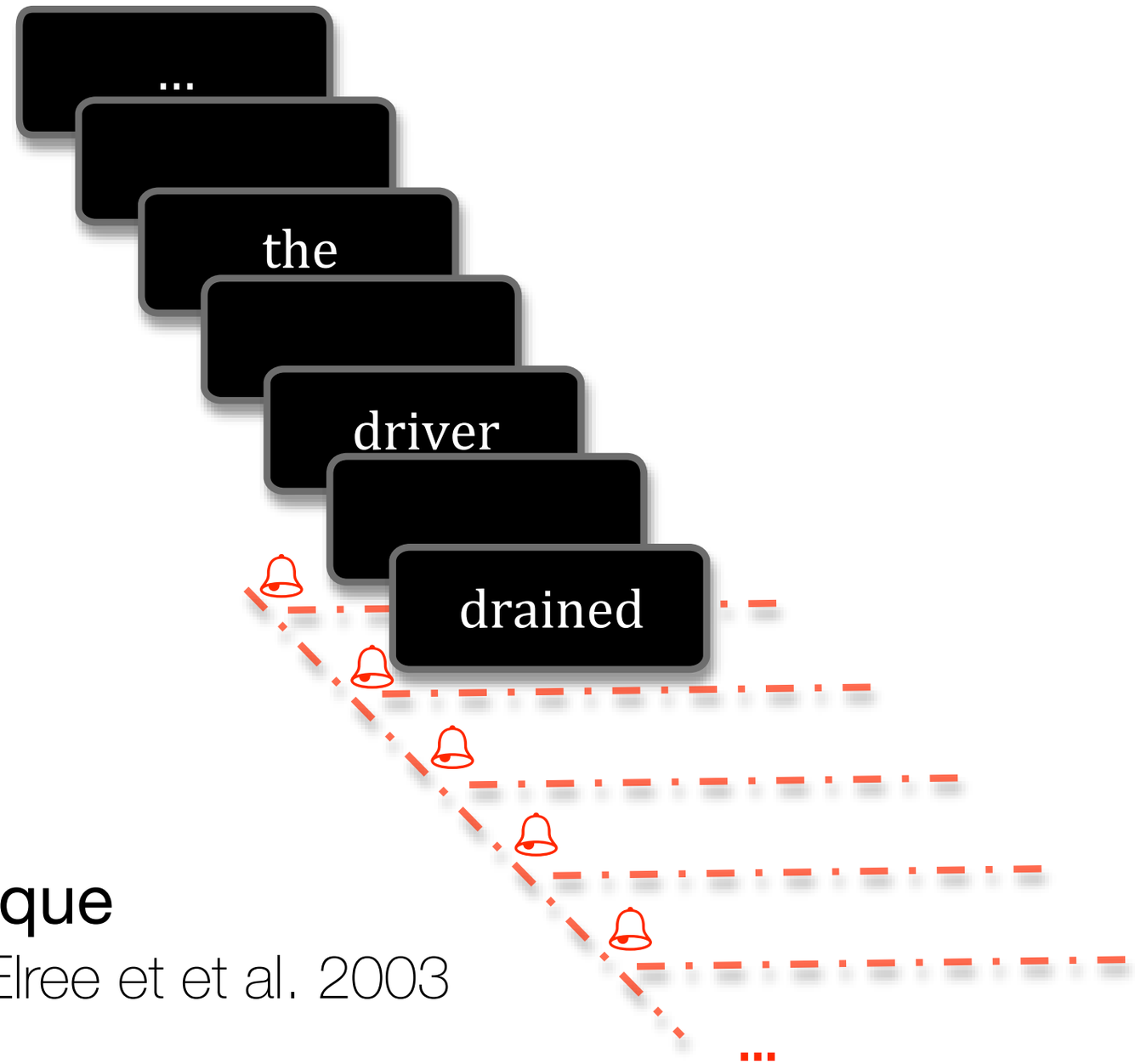
those monkeys (*monkey)
those face-making **monkeys** (*monkey)
those mischievous, face-making **monkeys** (*monkey)

SG demonstrative DP

that monkey (*monkeys)
that face-making **monkey** (*monkeys)
that mischievous, face-making **monkey** (*monkeys)

Testing for the maintenance of [PL]

Method



Speed-accuracy tradeoff

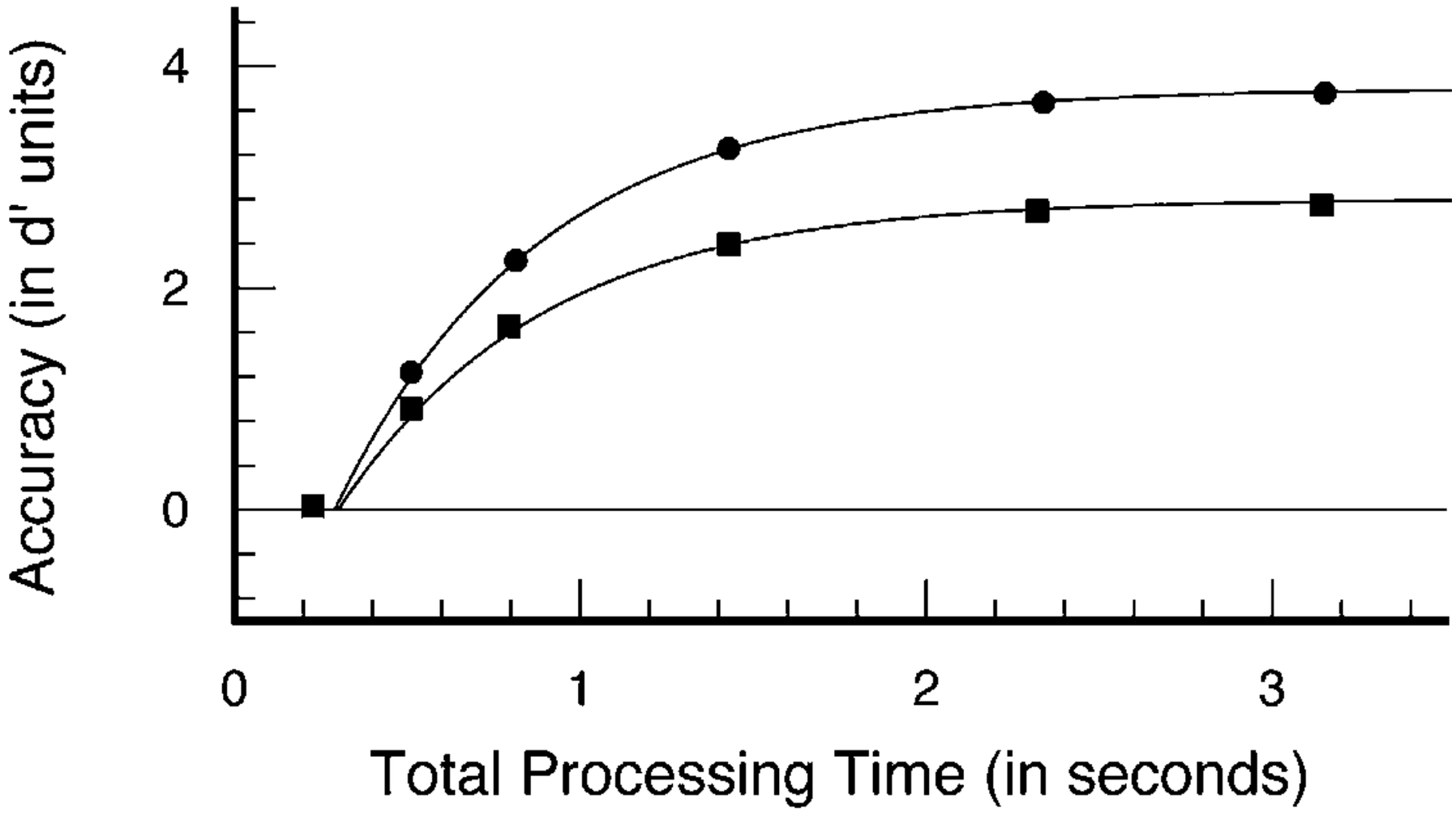
(multiple-response)

- response-signal technique
  Wickelgren et al. 1980, McElree et et al. 2003

- sensitive measure of retrieval speed
  McElree, 2006

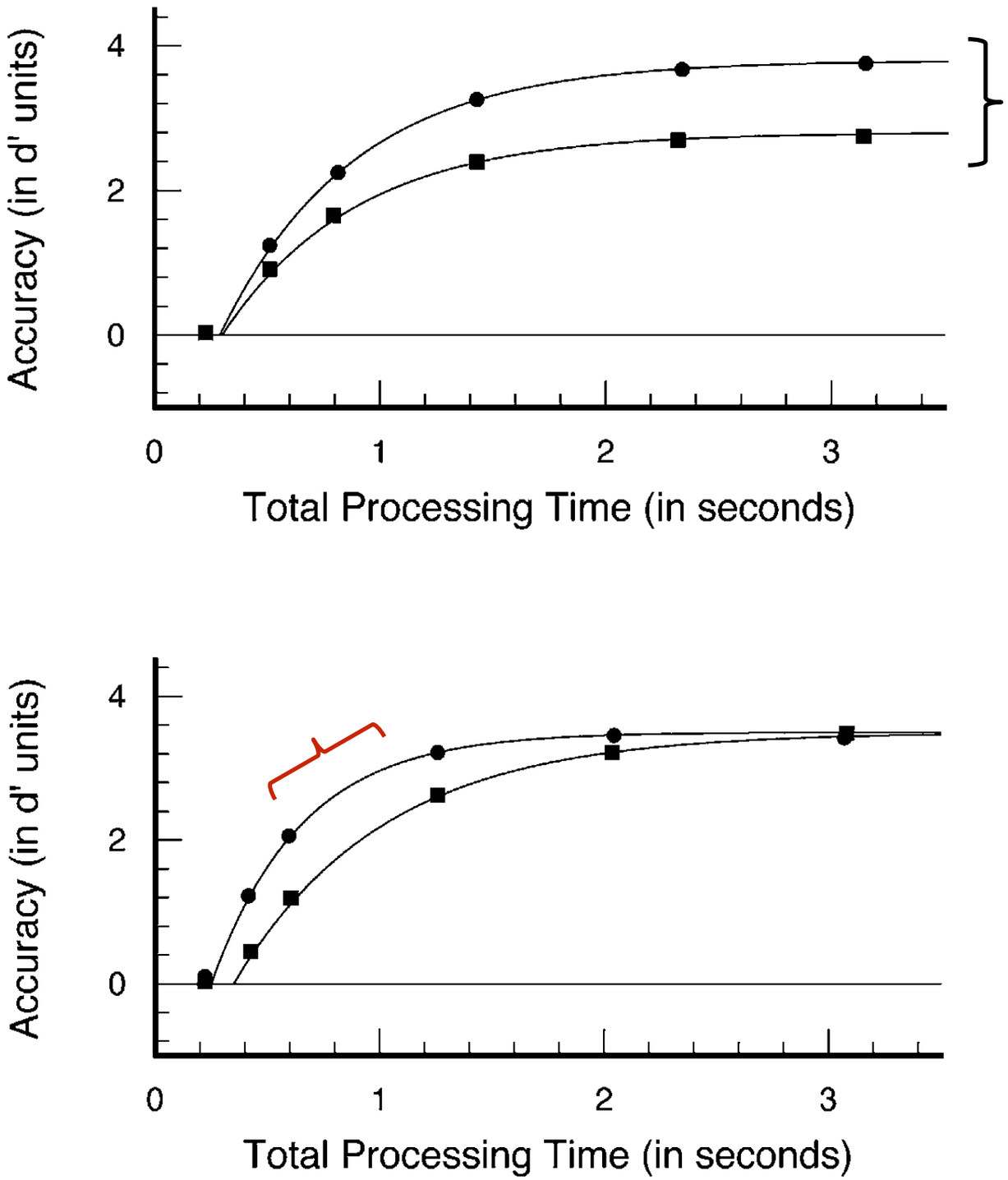
Testing for the maintenance of [PL]

Analysis



Testing for the maintenance of [PL]

Analysis



Asymptotic difference
Reflects the likelihood of completing a parse/process.

Rate difference
Reflects speed of processing
how quickly information accumulates continuously, or the differences in an underlying discrete finishing time distribution

Testing for the maintenance of [PL]

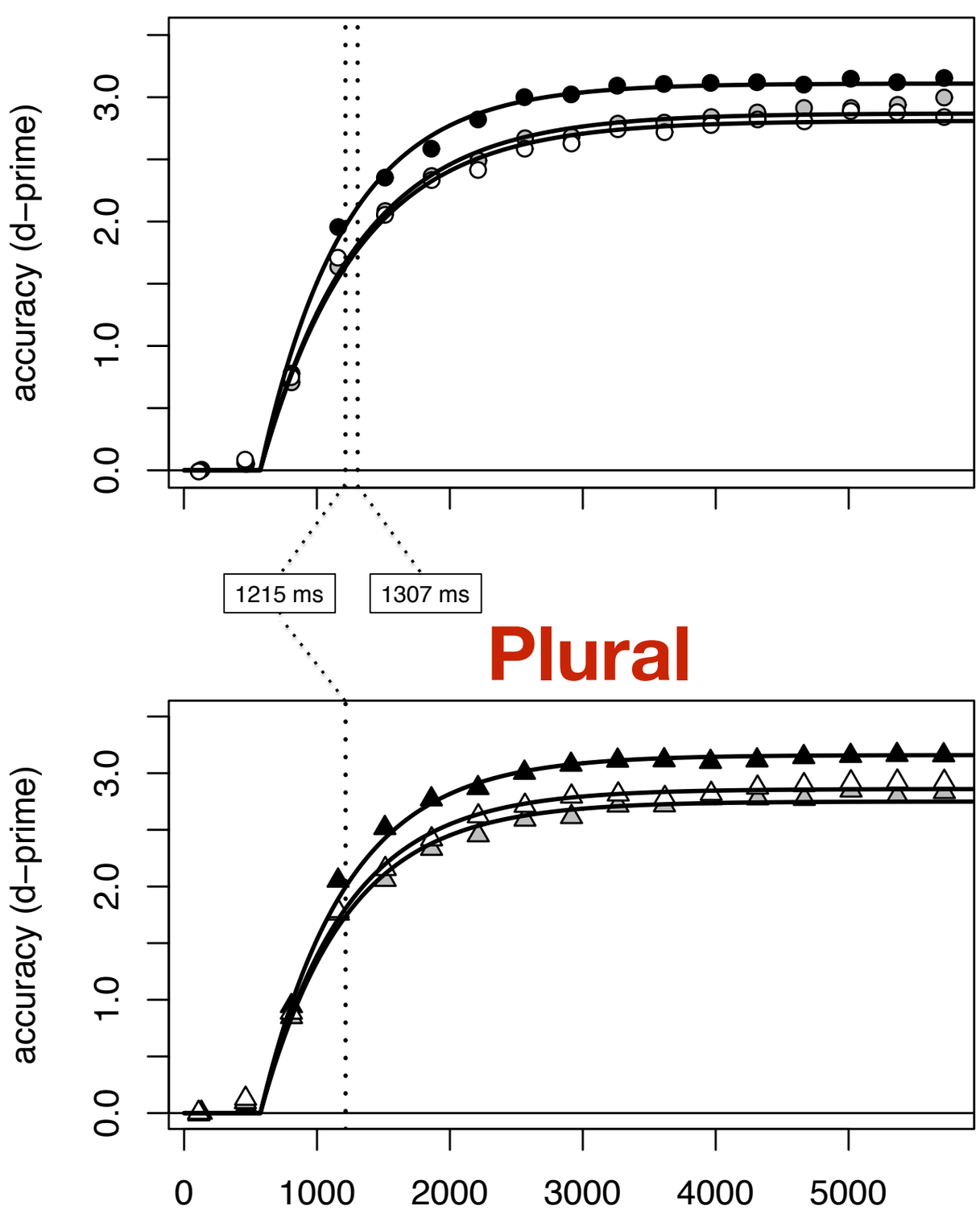
Agr

Results

Computing non-adjacent agreement
versus adjacent agreement

Singular: +92 ms

Plural: no longer



✅

HYPOTHESIS

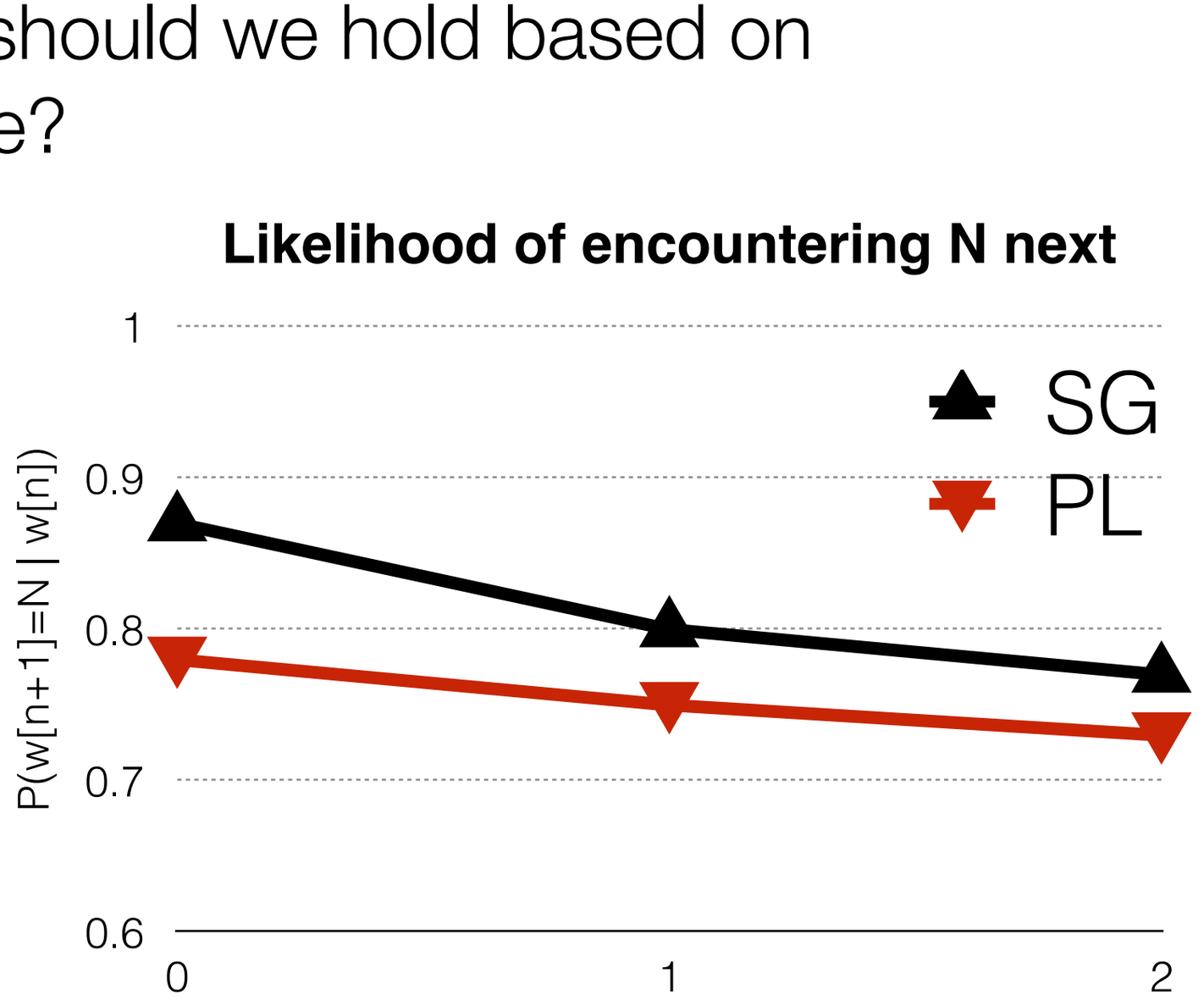
Plural features are more likely than **singular** features to be maintained in the active (fast) state of memory.

Wagers & McElree (2012)

… how does number exert its influence?

- What expectations should we hold based on language experience?

Occurrence of N head as the next element in DP is always less surprising for **singular DPs**



Wagers & McElree (2012)

… how does number exert its influence?

- Mediated: if predictions reflect (known) contingencies for DPs, the number-bearing N head is always expected sooner when Dem=**SG**

- Direct: if predictions stem from properties of feature structure, the marked **PL** feature is able to persist in working memory (=facilitating longer DPs?)

Three case studies

RC Animacy and English relative clauses

Agr Agreement inside English DPs

Wh **Wh-Agreement and Person in Chamorro**

QUESTION

How does obligatoriness and optionality in agreement paradigms affect interpretation of cross-indexed dependencies?

Chamorro: Wh-Agreement

(1) **Ha fåhan si Vicente i gima' Antonio.**
AGR buy NM Vicente the house.L Antonio
"Vicente bought Antonio's house."

(2) **Håyi fumåhan ___ i gima'?**
who? WH[SBJ].buy the house
"Who bought the house?"

Chamorro: Wh-Agreement

(3) **Håfa <u>ha fåhan</u> si Maria ___ gi tenda?**
what? AGR buy NM Maria LOC store

"What did Maria buy at the store?"

(4) **Håfa <u>finåhan-ña</u> si Maria ___ gi tenda?**
what? WH[OBJ].buy-AGR NM Maria LOC store

HYPOTHESIS

Object Wh-Agreement will trigger faster interpretation of an object gap dependency compared to ordinary S-V agreement.

Chamorro: Wh-Agreement

Kuåntu na chinina **prinensåm-mu** nigap **…**

how.many L shirt iron[WH:OBJ]-AGR yesterday

“How many shirts did you iron yesterday…?”

OBJECT GAP OBLIGATORY

Kuåntu na patgun låhi **prinensåm-mu** nigap …

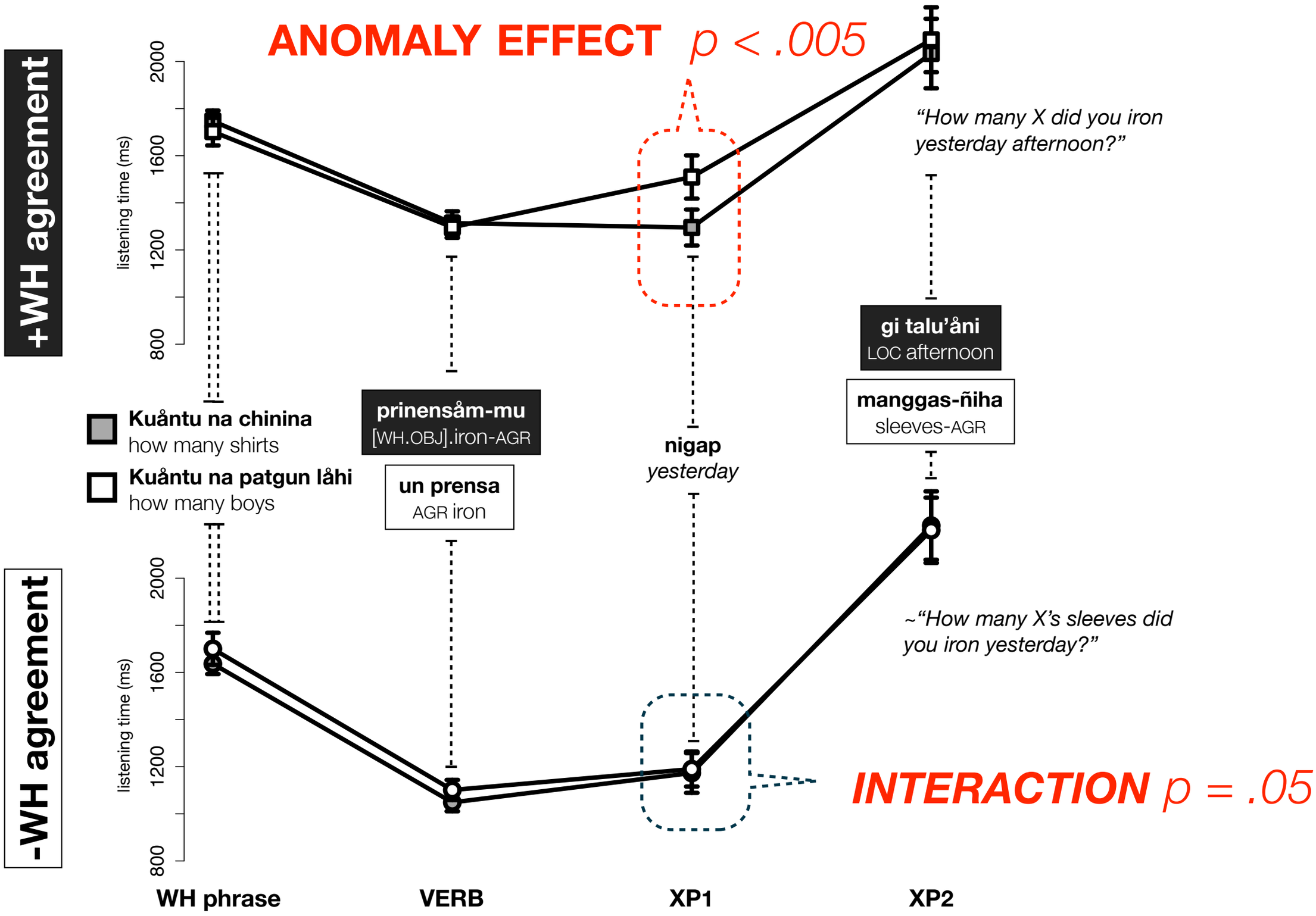
“How many boys did you iron yesterday…?”

OBJECT GAP OPTIONAL

Kuåntu na chinina/patgun låhi **un prensa** nigap **…**

AGR iron

Transitive clauses SELF-PACED LISTENING



Wagers, Borja & Chung (2015)

HYPOTHESIS

✅ Object Wh-Agreement will trigger faster interpretation of an object gap dependency compared to ordinary S-V agreement.

… how does Wh-Agreement exert its influence?

- Why is the anomaly contrast evident earlier for Wh-Agreement?

  - **Presence** of Wh-Agreement ***promotes*** dependency construction because it must be licensed

  - **Absence** of Wh-Agreement ***restrains*** dependency construction because it is compatible with other continuations

Wagers, Borja & Chung (2015)

Informativity of the bare form

- **Absence** of Wh-Agreement ***restrains*** dependency construction because it is compatible with other continuations

  - Possessor extraction → no Wh-Agreement

    Håyi un låksi chininå-ña?

    who? 2SG sew shirt-AGR

    “Whose shirt did you sew?”

Wagers, Borja & Chung (2015)

Informativity of the bare form

- Odds in favor of possessor gap:

$$\underbrace{\frac{P(Poss \mid Verb)}{P(Obj \mid Verb)}}_{\text{POSTERIOR ODDS}} = \underbrace{\frac{P(Verb \mid Poss)}{P(Verb \mid Obj)}}_{\text{LIKELIHOOD RATIO}} \times \underbrace{\frac{P(Poss)}{P(Obj)}}_{\text{PRIOR ODDS}}$$

- Preference survey (n=13)
  - Object extraction: **bare** v. Wh-Agreeing **72%**
  - Possessor: **bare** v. periphrastic form **74%**
- Odds in favor of possessor extraction given a bare verb? ~ **Prior odds of a possessor extraction**

Wagers, Borja & Chung (2015)

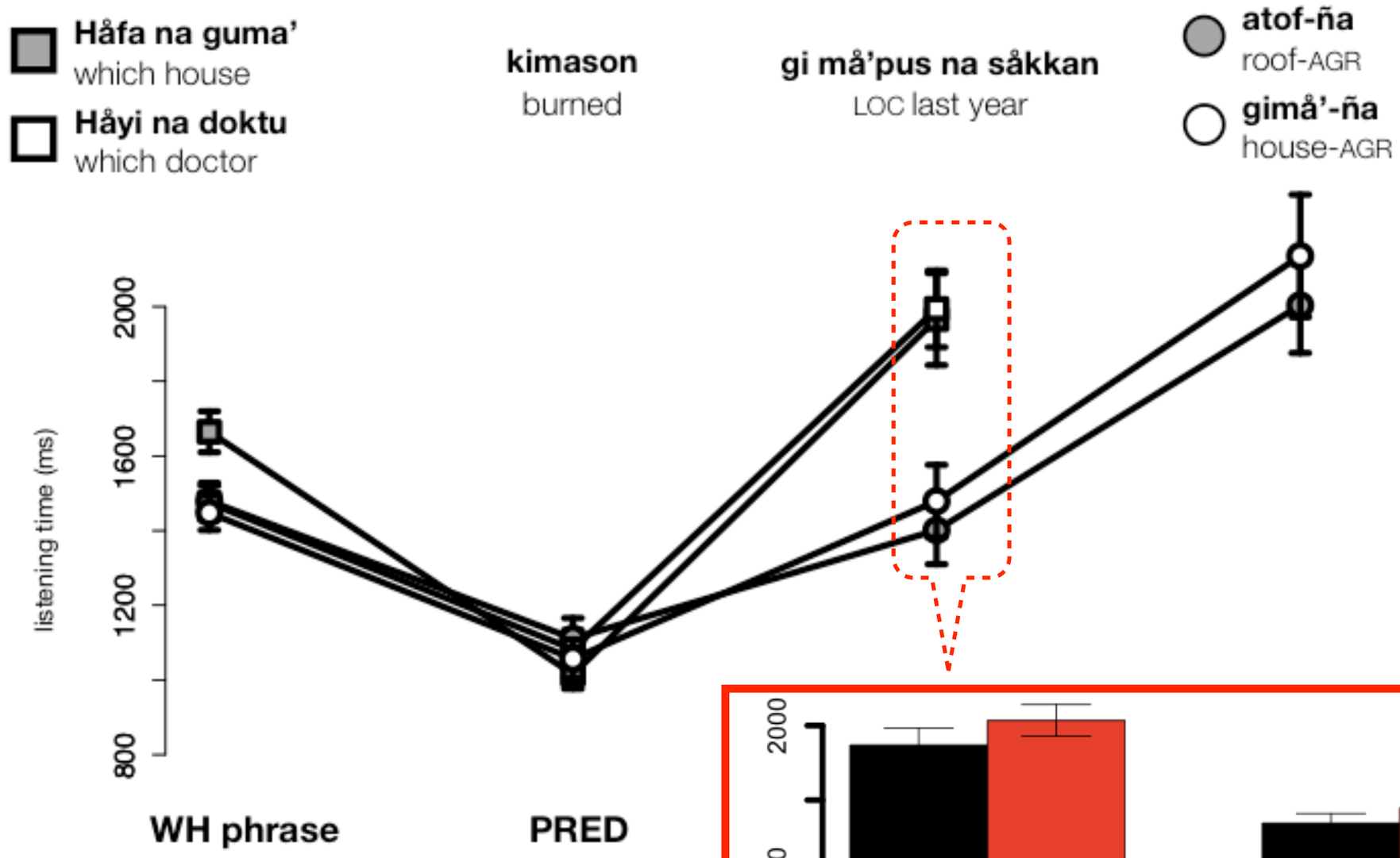
Informativity of the bare form

- **Prior odds of a possessor extraction**
  - Probably pretty low? especially for transitive objects
  - In a recent production study we elicited (at least) 691 gap-containing relative clauses:

	Transitive Subject	Object	Intransitive Subject	Locative	
Argument	210	**10**	308	106	**634**
Possessor	0	**1**	16	40	**58**

For RCs and objects, **prior odds are 1:10** (but, small *n*)

Intransitive clauses SELF-PACED LISTENING



SMALL **ANOMALY EFFECT** $p < .05$

NO INTERACTION

Wagers, Borja & Chung (2015)

Informativity of the bare form

- Absence of Wh-Agreement *restrains* dependency construction because it is compatible with other continuations

- <u>Possessor extraction → no Wh-Agreement</u>

  Håyi un låksi chininå-ña?

  who? 2SG sew shirt-AGR

  “Whose shirt did you sew?”

Informativity of the bare form

- Absence of Wh-Agreement *restrains* dependency construction because it is compatible with other continuations

- <u>Possessor extraction → no Wh-Agreement</u>
  Håyi un låksi chininå-ña?
  who? 2SG sew shirt-AGR
  “Whose shirt did you sew?”

- **<u>No possessor extraction</u>** <u>over 3.PERS DP</u>
  ***Håyi ha låksi si Bedu’ chininå-ña?**
  who? 3SG sew Bedu’ shirt-AGR
  (“Whose shirt did Bedu’ sew?”)

Alignment hierarchies and prediction

- ***Subj=3.pers > Obj=2**

  ***Ha li'i hao si Dolores nigap**

  3SG see 2SG UNM Dolores yesterday

  'Dolores saw you yesterday'

- ***Subj=DP > Obj=3.pers anim. pron.**

  ***Ha li'i' gui' si Maria**

  3SG see 3SG UNM Maria

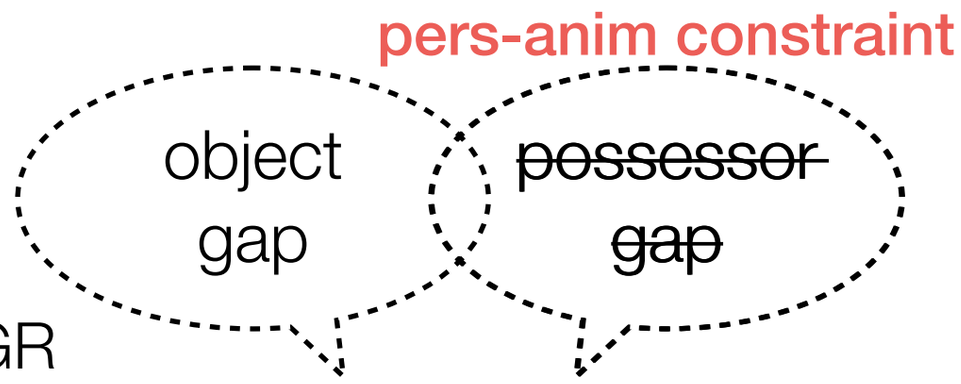
  'Maria saw him'

- **The Chamorro Person-Animacy Hierarchy**

  2.pers > 3.pers anim. pron. > anim. non-pron. > inanimate

cf. Aissen, 1997;
Christianson & Ferreira, 2005, Christianson & Cho, 2009

Compare effect of +WH.AGR to 3.PERS



3PERS, -WH.AGR

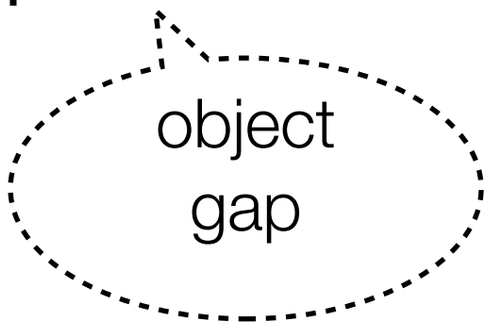
Bula katpinteru [ha apåsi si Bedu’ __ gi lanchu gi ma’pus na mes].

There were a lot of carpenters who Bedu’ paid at the ranch last month.

3PERS, +WH.AGR

Bula katpinteru [inapasi-ña si Bedu’ __ gi lanchu gi ma’pus na mes].

There were a lot of carpenters who Bedu’ paid at the ranch last month.



Compare effect of +WH.AGR to 3.PERS

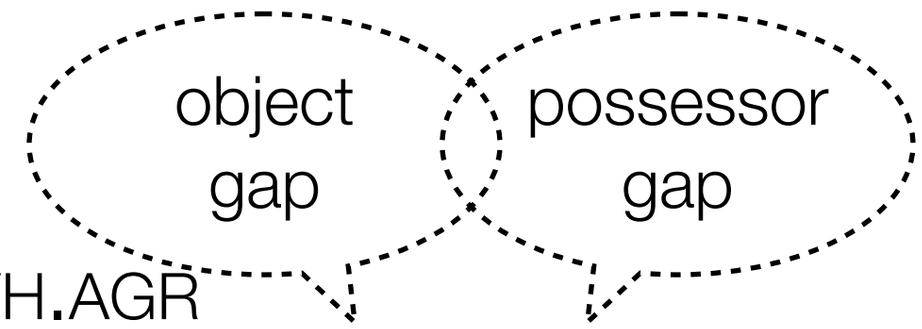
PLAUSIBLE EXTRACTIONS

3PERS, -WH.AGR

Bula katpinteru ha apåsi si Bedu’ gi lanchu gi ma’pus na mes.
Bedu’ paid a lot of carpenters at the ranch last month.

3PERS, +WH.AGR

Bula katpinteru inapasi-ña si Bedu’ gi lanchu gi ma’pus na mes.
Bedu’ paid a lot of carpenters at the ranch last month.



2PERS, -WH.AGR

Bula katpinteru un apåsi gi lanchu gi ma’pus na mes.
You paid a lot of carpenters at the ranch last month.

Compare effect of +WH.AGR to 3.PERS

IMPLAUSIBLE EXTRACTIONS

3PERS, -WH.AGR

Bula katpinteru ha dingding si Bedu' gi lanchu gi ma'pus na mes.
Bedu' rang a lot of carpenters at the ranch last month.

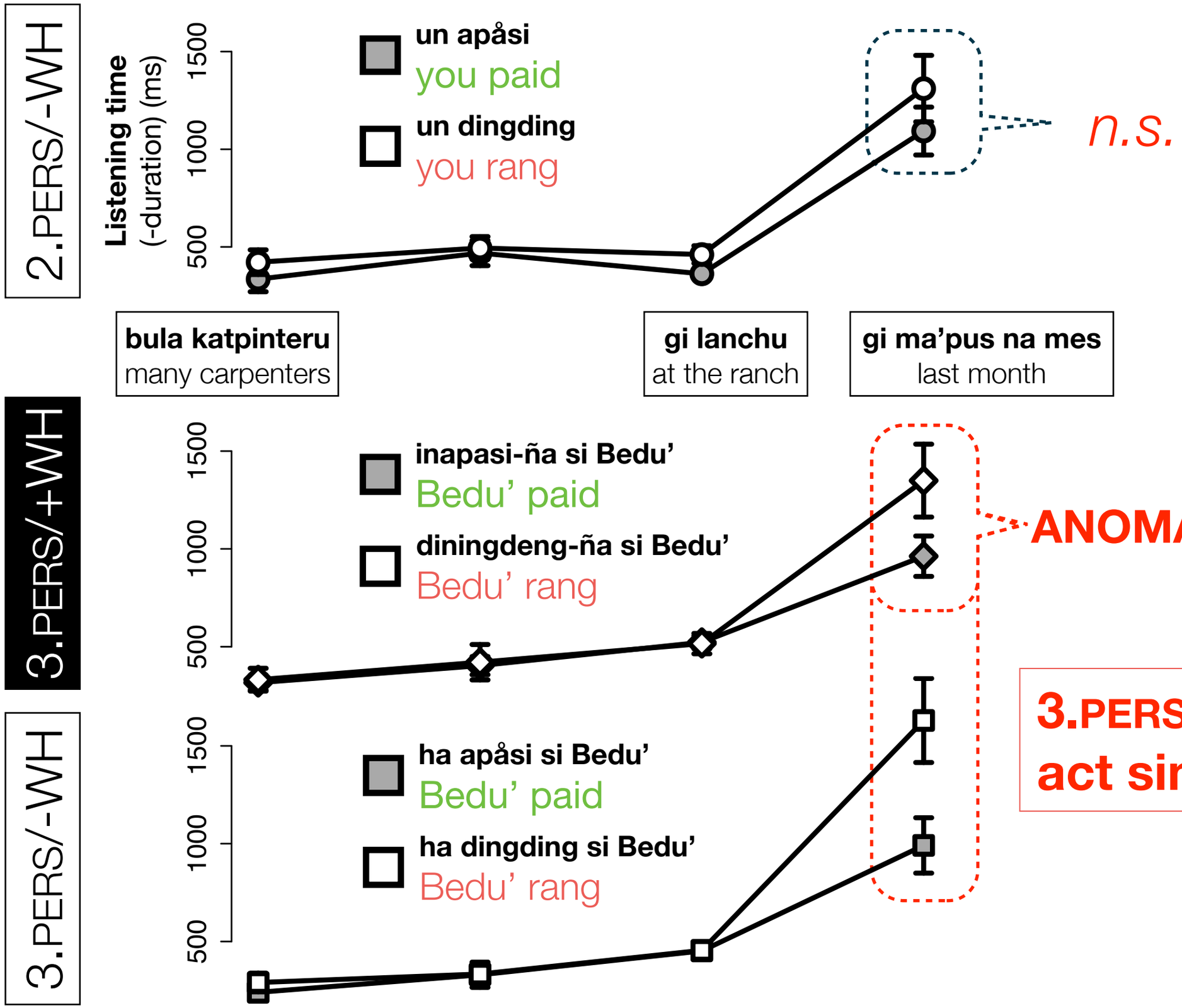
3PERS, +WH.AGR

Bula katpinteru diningdeng-ña si Bedu' gi lanchu gi ma'pus na mes.
Bedu' rang a lot of carpenters at the ranch last month.

2PERS, -WH.AGR

Bula katpinteru un dingding gi lanchu gi ma'pus na mes.
You rang a lot of carpenters at the ranch last month.

Compare effect of +WH.AGR to 3.PERS



3.PERS and +WH.AGR act similarly!

Wagers, Borja & Chung (*in prep*)

Summary

On-line anomaly effect

	2.PERS	3.PERS
+WH.AGR	✅	✅
-WH.AGR	❌❌	✅

Wagers, Borja & Chung (*in prep*)

… how does Wh-Agreement exert its influence?

- Mediated: if predictions reflect (known) contingencies, the chance of an object gap is essentially maximal when Wh-Agreement is present - just like when a 3.PERS DP subject is present

- Direct: if predictions stem from grammatical licensing requirements

  - Object Wh-Agreement predictively extends the representation to include an object gap, triggering interpretation

  - 3.PERS DP subjects require “justification” from PERS-ANIM constraints ↝ gap must be a VP-internal argument

Taking stock

	Finding	Direct Mechanism
RC	Animate RC heads are prospectively linked to subject gap, but inanimate heads are not.	Licensing of animacy? Online “harmonic alignment”?
Agr	Faster processing of complex PL demonstrative DPs.	PL is more durable in memory.
Wh	Wh-Agreement triggers early interpretation of movement dependencies.	Wh-Agreement prospectively agrees with object gap.

Taking stock

	Finding	Mediated Mechanism
RC	Animate RC heads are prospectively linked to subject gap, but inanimate heads are not.	Conditional probabilities, ignorant of pre-RC adjunct.
Agr	Faster processing of complex PL demonstrative DPs.	Ignorance of pre-nominal modifiers' effect on DP size distributions.
Wh	Wh-Agreement triggers early interpretation of movement dependencies.	Knowledge of prior odds of possessor extraction.

Managing our expectations

- Comprehenders exhibit misalignments between predicted dependency elements - as demonstrated in online experiments - and the most probable continuation - as estimated from a corpus, or a Cloze task

- Independent mechanisms affect ordering of predictions -

  - the *value* of one representation over another: which satisfies more constraints? or is incrementally more grammatical? (Pritchett, 1992, Chater, Crocker, Pickering, 1998, Borja, Wagers, & Chung 2015)

  - workspace constraints: how features/constituents are maintained in short-term memory (Wagers & McElree, 2013)

Ignorance is bliss

- Hard to tell whether comprehenders are not using contingencies, or are or ignorant of them.

- Ignorance is not a bad thing.

  - Syntactic probabilities are not universal or absolute - e.g., genre/context dependence. More abstract generalizations may be more stable than “construction-specific” ones (Roland et al. 2007)

- What’s stored is what drove learning and generalization?

  - input/intake distinction of Gagliardi & Lidz (2014)

  - argument/adjunct hypothesis of Boland & Blodgett (2006)/ Tutunjian & Boland (2008)

Thank you

RC

Emily Pendleton

UCSC B.A. ‘13

Wh

Manuel F. Borja

Inetnun Åmut yan Kutturan Natibu, CNMI

Agr

Brian McElree

NYU Psychology

Wh

Sandra Chung

UCSC Linguistics

National Science Foundation BCS #1251429 (MW, SC)