

Anti-local contexts improve the *overall* speed of dependency completion

Matthew Wagers

Department of Linguistics
University of California, Santa Cruz

CUNY2010
New York University

 **LINGUISTICS
RESEARCH CENTER**
UNIVERSITY OF CALIFORNIA SANTA CRUZ



Goal

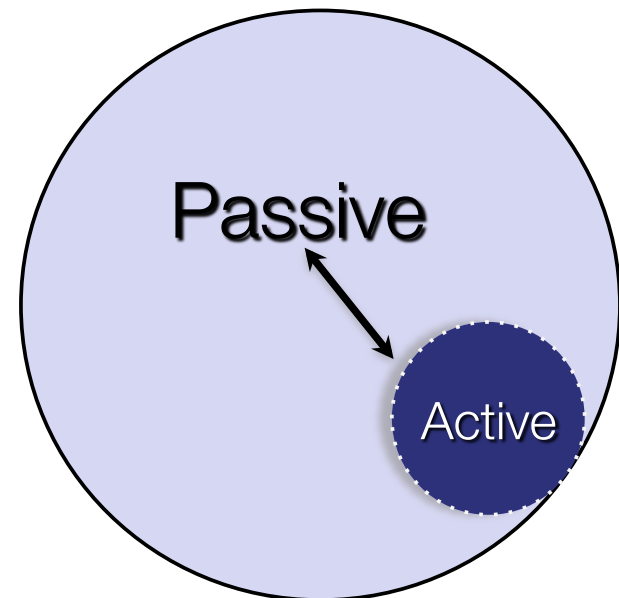
- What is the relation between predictive forces in language comprehension and the concept of focal attention?

Focal attention

- ... is extremely limited

For sequentially-presented information, the capacity of focal attention appears limited to the last “unit” processed (Wickelgren et al., 1980; Garavan, 1998; Cowan, 2001; McElree, 2006; Jonides et al., 2008).

Building structured representations for sequentially-presented input will often require shunting information between memory and focal attention



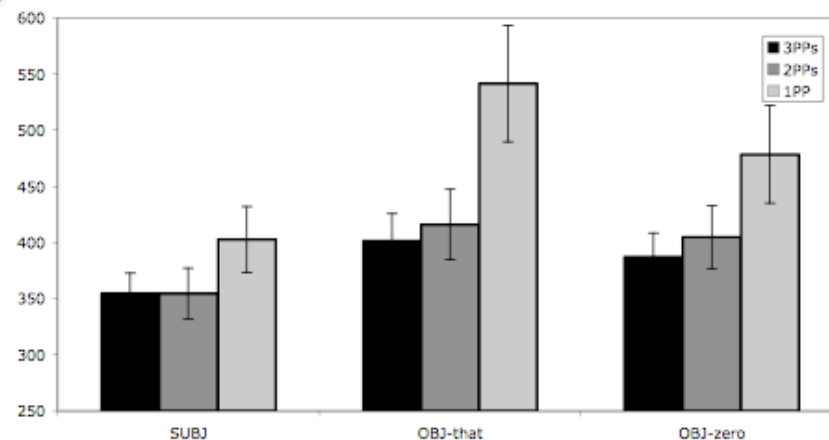
Longer == easier

- Anti-local contexts

Longer can be easier.

E.g., Jaeger, Fedorenko & Gibson, *submitted*

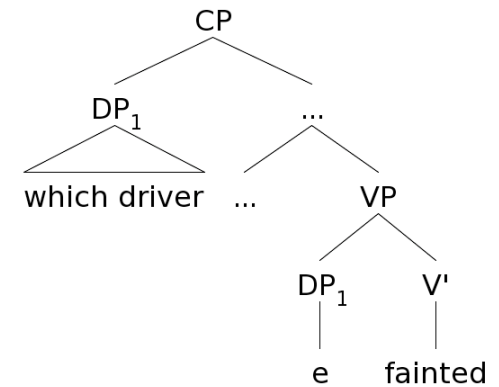
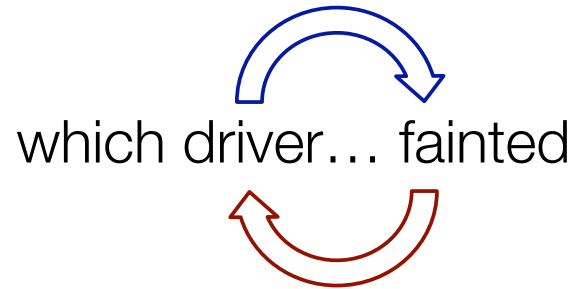
The understudy that the agent telephoned
about the job in Los Angeles
shared the story ...



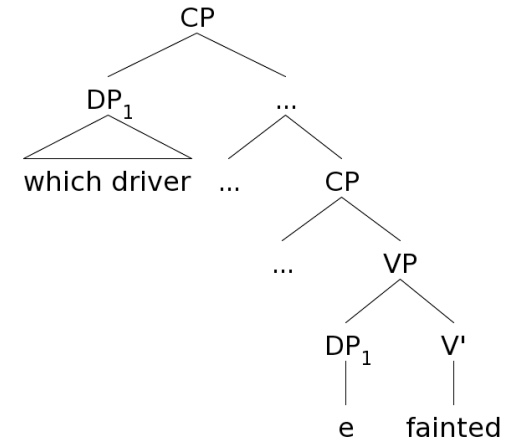
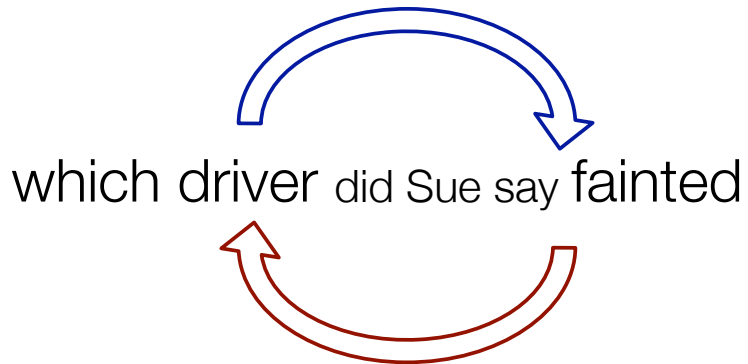
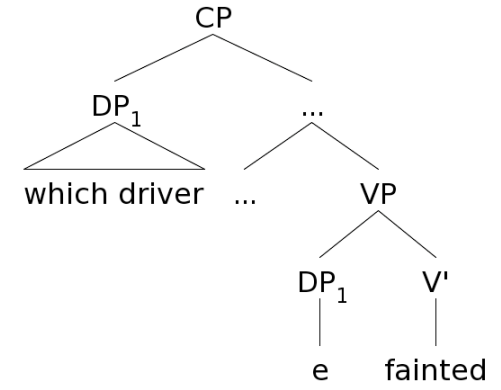
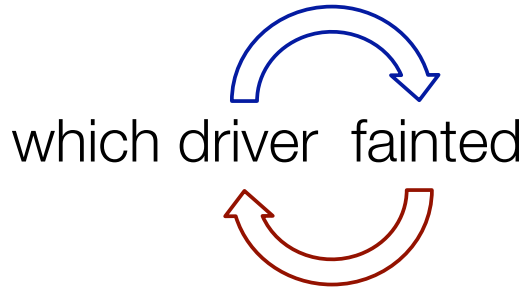
Question

- Today's investigation
 - What is the nature of anti-local facilitation?
 - RT facilitation is fed by many factors: strength of underlying encoding, speed of processing, etc. etc.
 - Which one of these changes?
 - Measure directly with S.A.T. response-signal method.

Locality in comprehension



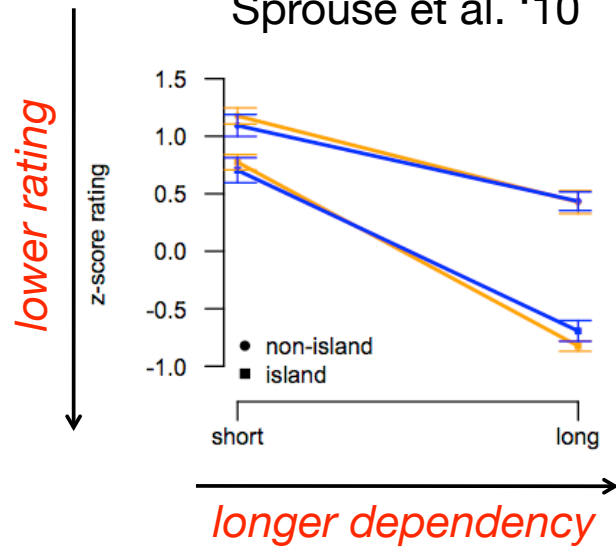
Locality in comprehension



Local is easier

ACCEPTABILITY RATINGS

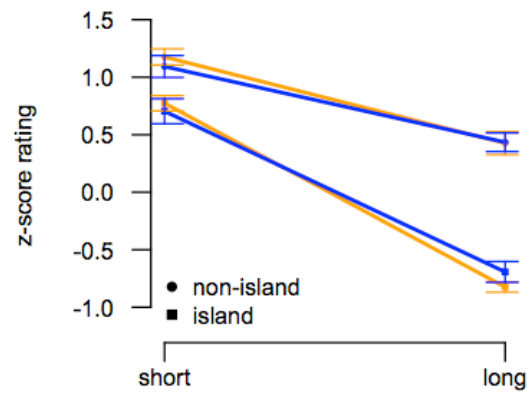
Sprouse et al. '10



Local is easier

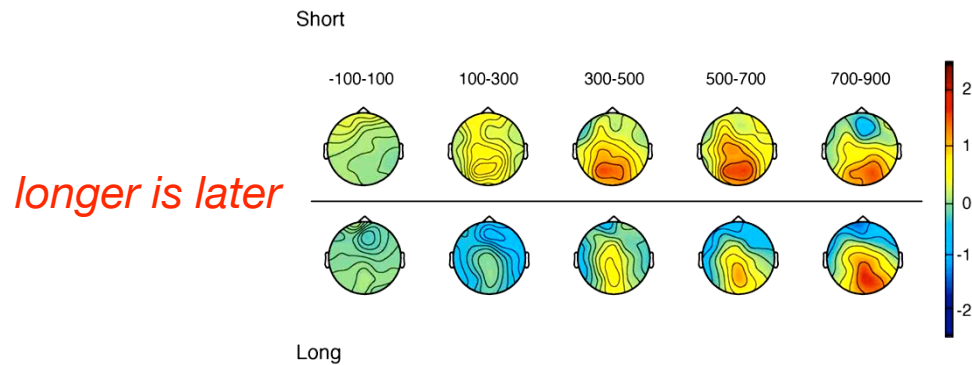
ACCEPTABILITY RATINGS

Sprouse et al. '10



ERP TIMECOURSE

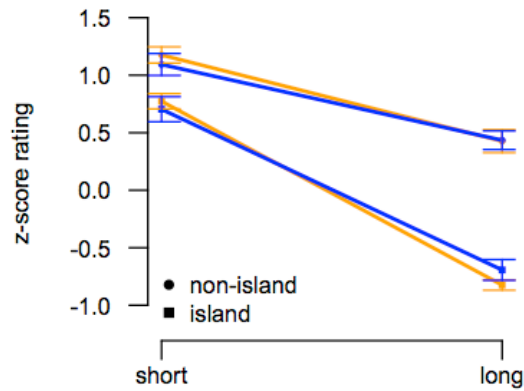
Phillips et al. '05



Local is easier

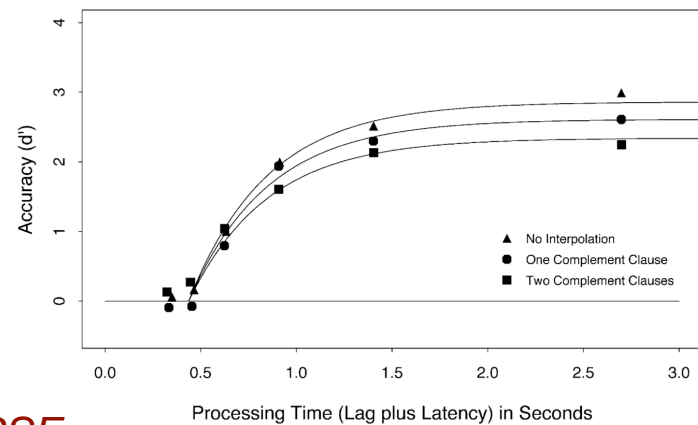
ACCEPTABILITY RATINGS

Sprouse et al. '10



SAT ASYMPTOTIC ACCURACY

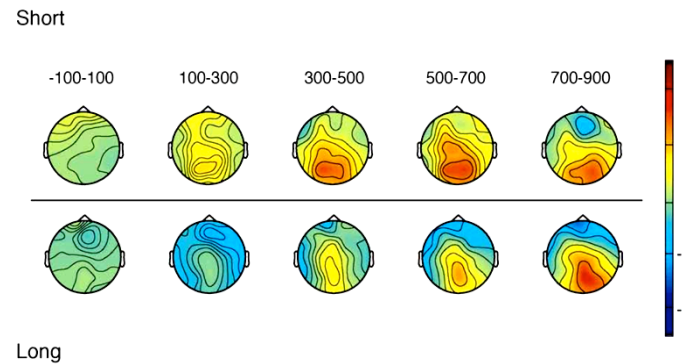
McElree et al. '03



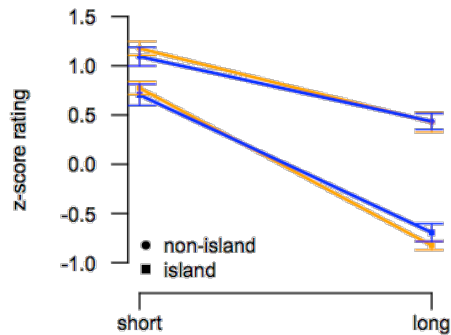
less likely to succeed

ERP TIMECOURSE

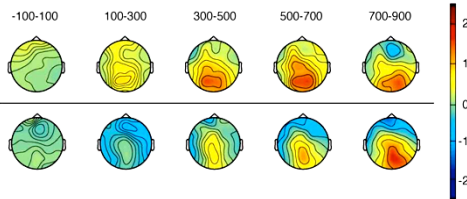
Phillips et al. '05



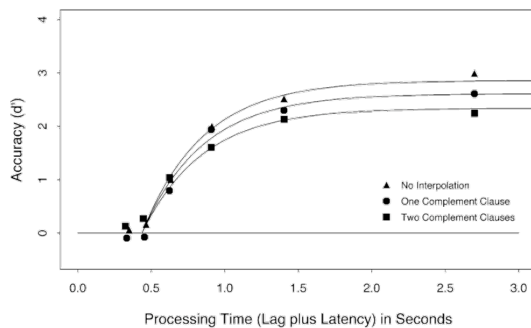
Local is easier



Short



Long



Gibson (2000)

Dependency Locality Theory

Fodor (1978)

Gap-finding is hard

Frazier (1987)

Wanner & Maratsos (1978)

Storing incomplete dependencies is hard

Longer ≠ harder

Konieczny (2000)

German RC preverbal intraposition

Vasishth & Lewis (2006)

Hindi center embedding & object relative clauses

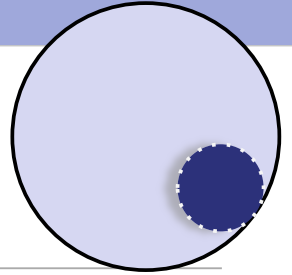
Jaeger, Fedorenko, Gibson (2005,2008,submitted)

RC-modified subjects

Mechanism of facilitation

RT ~ X

Why?



Wagers & McElree (AMLaP 2009)

Adjacent

The officer was informed that **the driver fainted**/*drained.

+Adverb

... **the driver abruptly fainted**

+PP

... **the driver of the ambulance fainted**

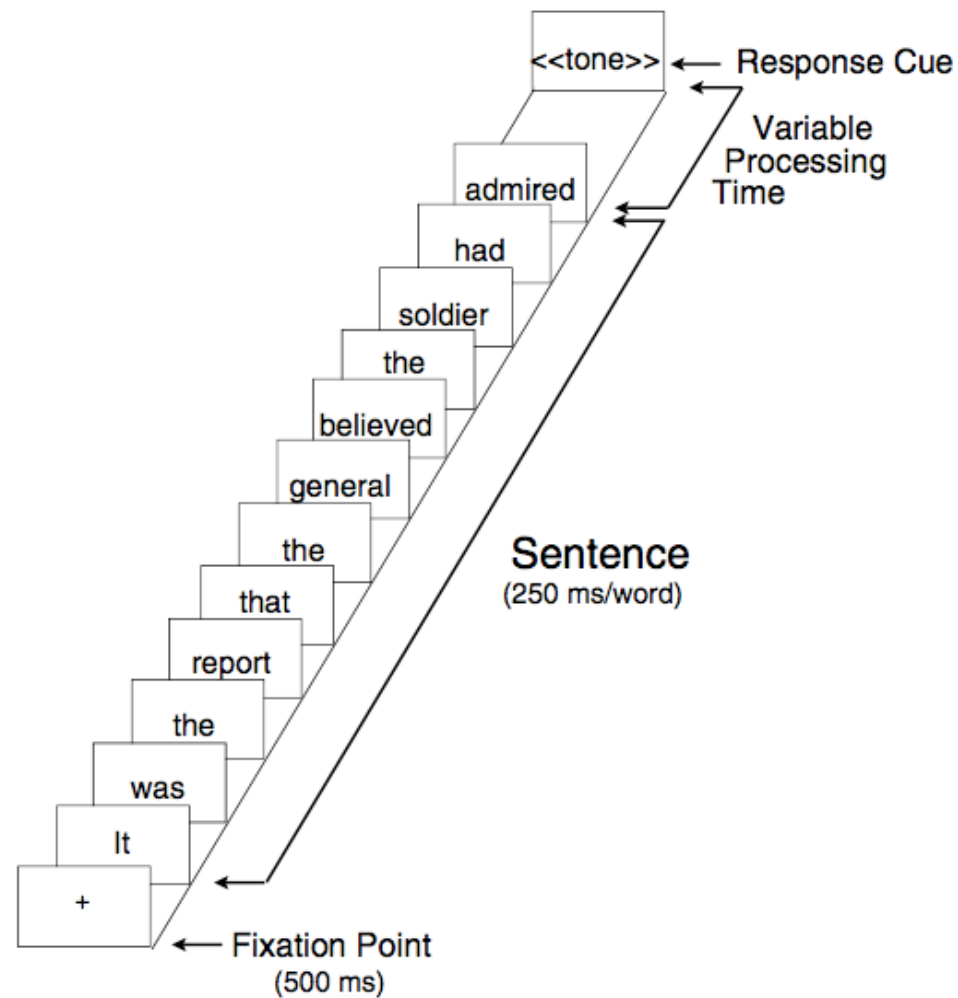
+Subject Relative Clause

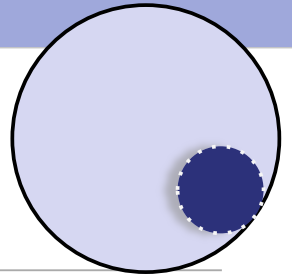
... **the driver who wrecked the ambulance fainted**

+Object Relative Clause

... **the driver who the ambulance hit fainted**

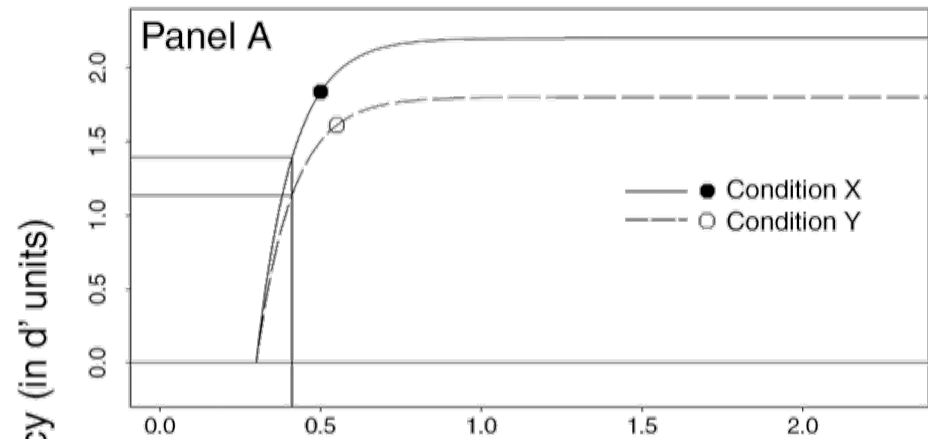
Speed-accuracy tradeoff



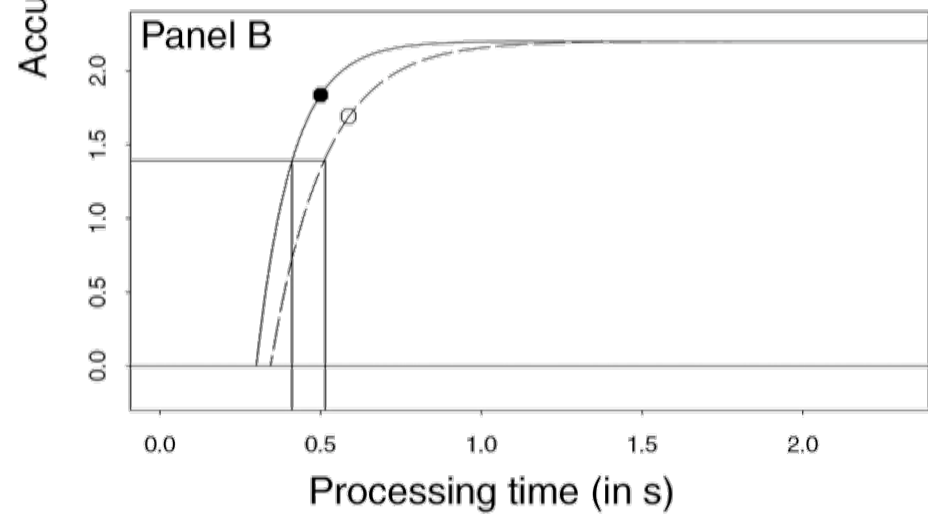


Theoretical outcomes

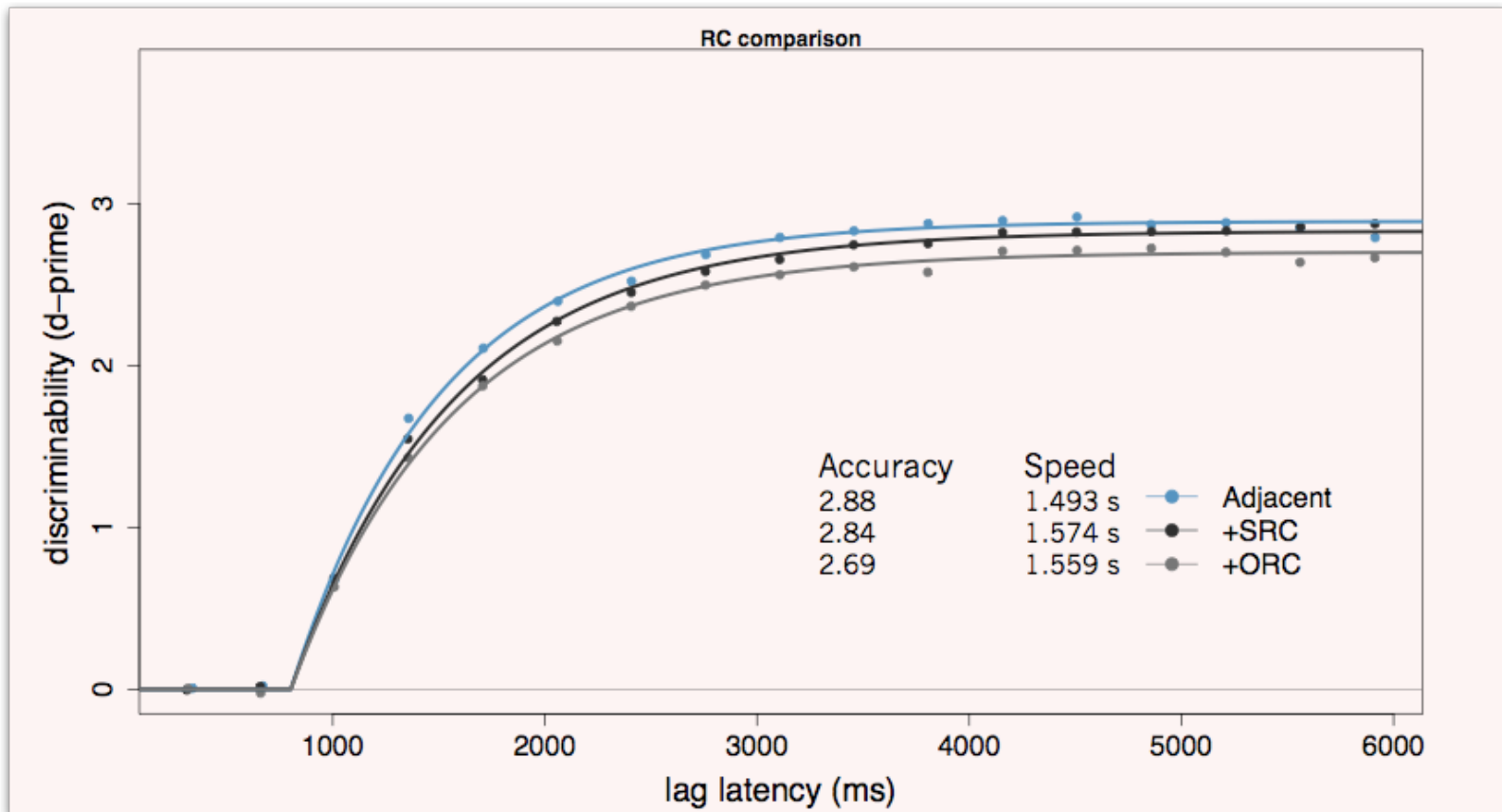
Accuracy difference



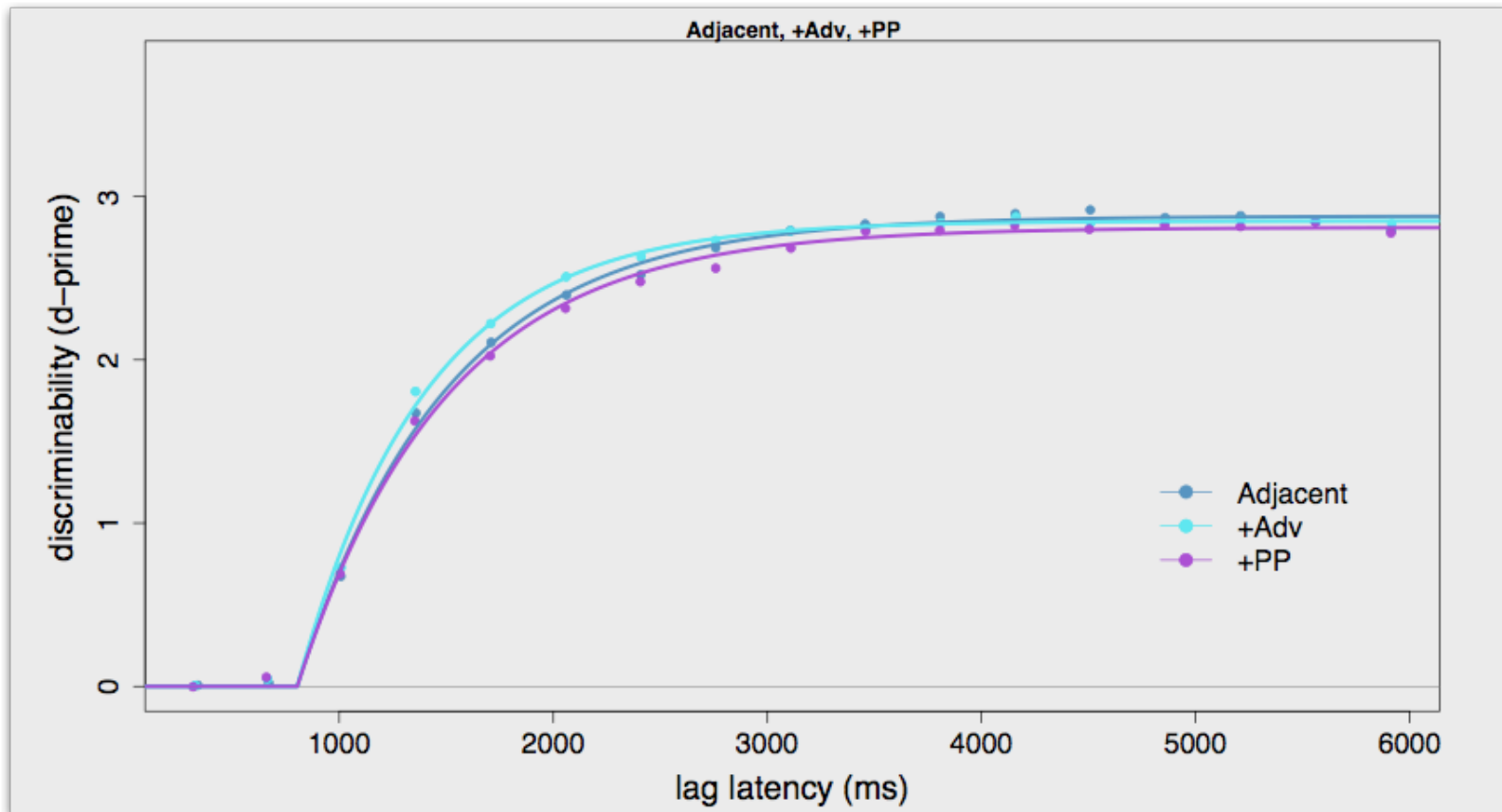
Rate difference

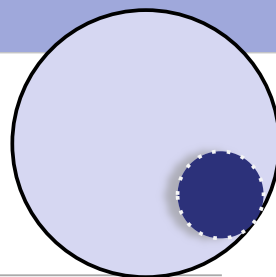


Wagers & McElree (2009) actual data



Wagers & McElree (2009) actual data





Wagers & McElree (2009) summary

- Facilitation only observed in the +ADVERB conditions
- However, RCs were simple
- Give anti-locality a better chance by extending the RCs
- Follow the Jaeger et al. materials design

Exp. 1: Materials and Methods

- **ADJACENT**

Wounded by the rebel in the trenches near the border,

the soldier twitched/*snagged.

- **+Object Relative Clause/NoPP**

In the trenches near the border,

the soldier **that the rebel wounded** twitched/*snagged.

- **+Object Relative Clause/+1PP**

In the trenches,

the soldier **that the rebel wounded near the border** twitched/*snagged.

- **+Object Relative Clause/+2PP**

The soldier **that the rebel wounded in the trenches near the border** twitched/*snagged.

Materials and Methods

- Acceptability × Length

ADJACENT

+OBJECT RC (ORC.noPP)

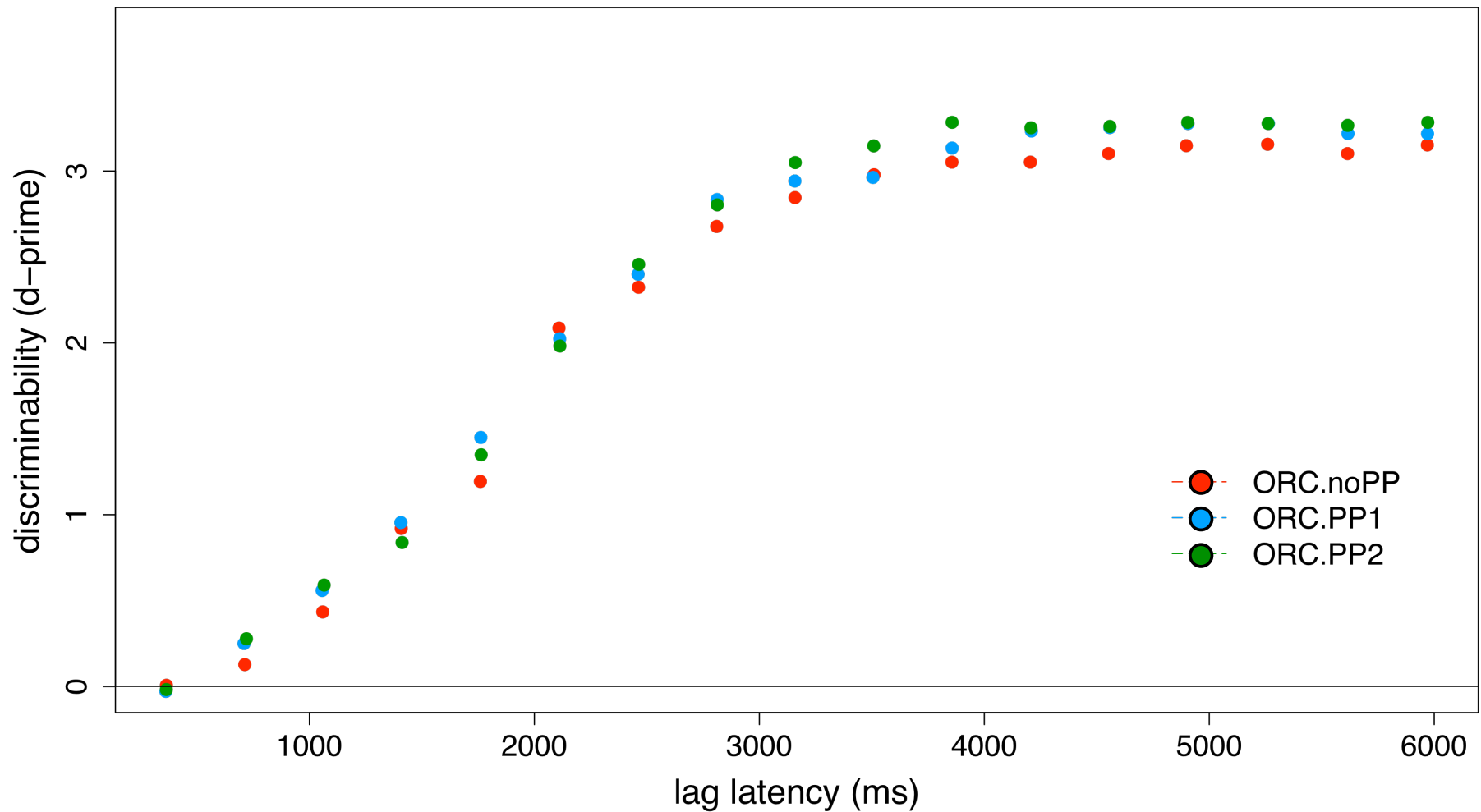
+OBJECT RC/1PP (ORC.2PP)

+OBJECT RC/2PP (ORC.1PP)

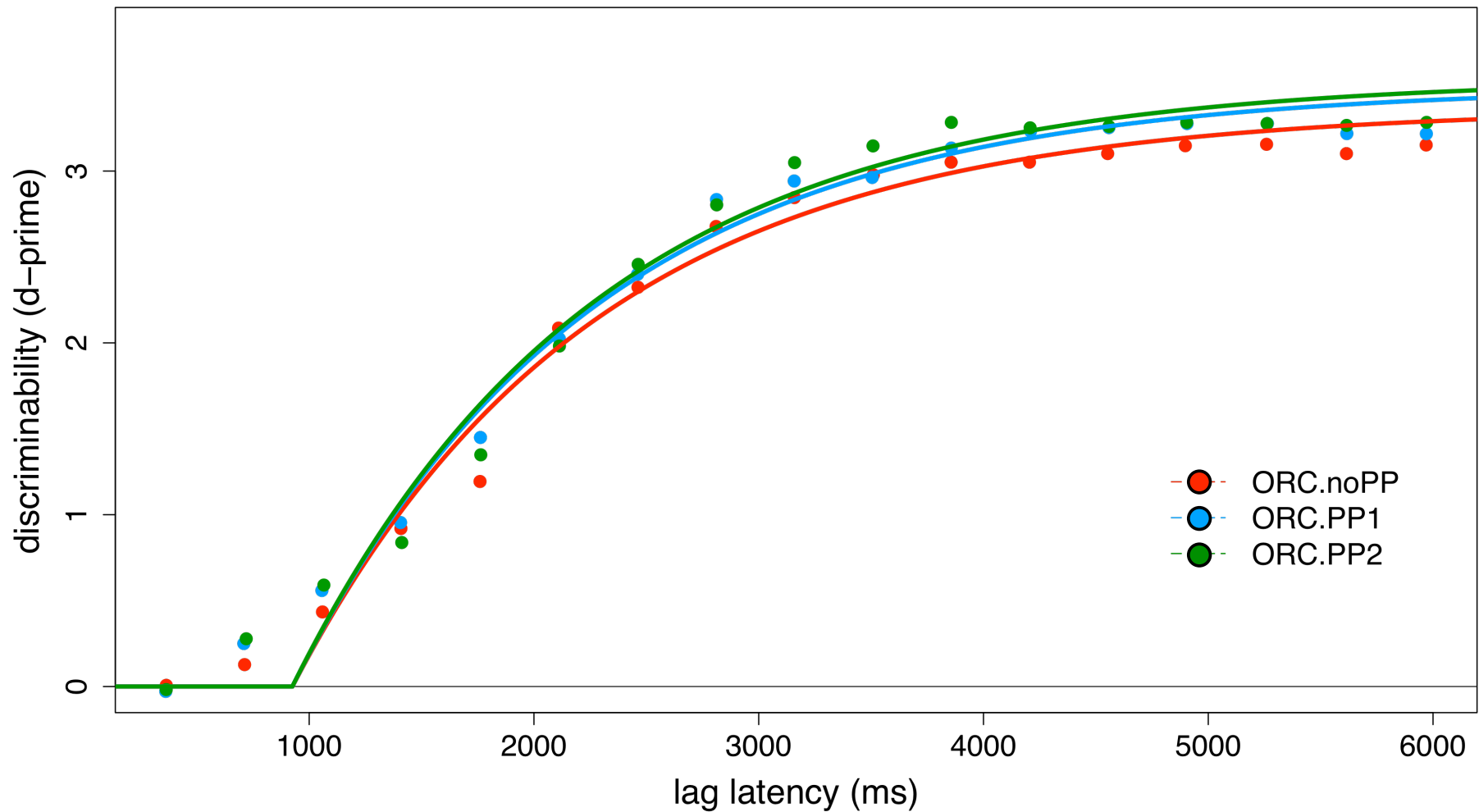
- 36 item sets
- MR-SAT
 - n = 10, compensated
 - Fillers with sentence-medial errors
 - Three sessions + with a practice session
 - Liu et al. (2009): Competitive model analysis

$$d' = \lambda \cdot \left(1 - e^{-\beta \cdot (t - \delta)}\right)$$

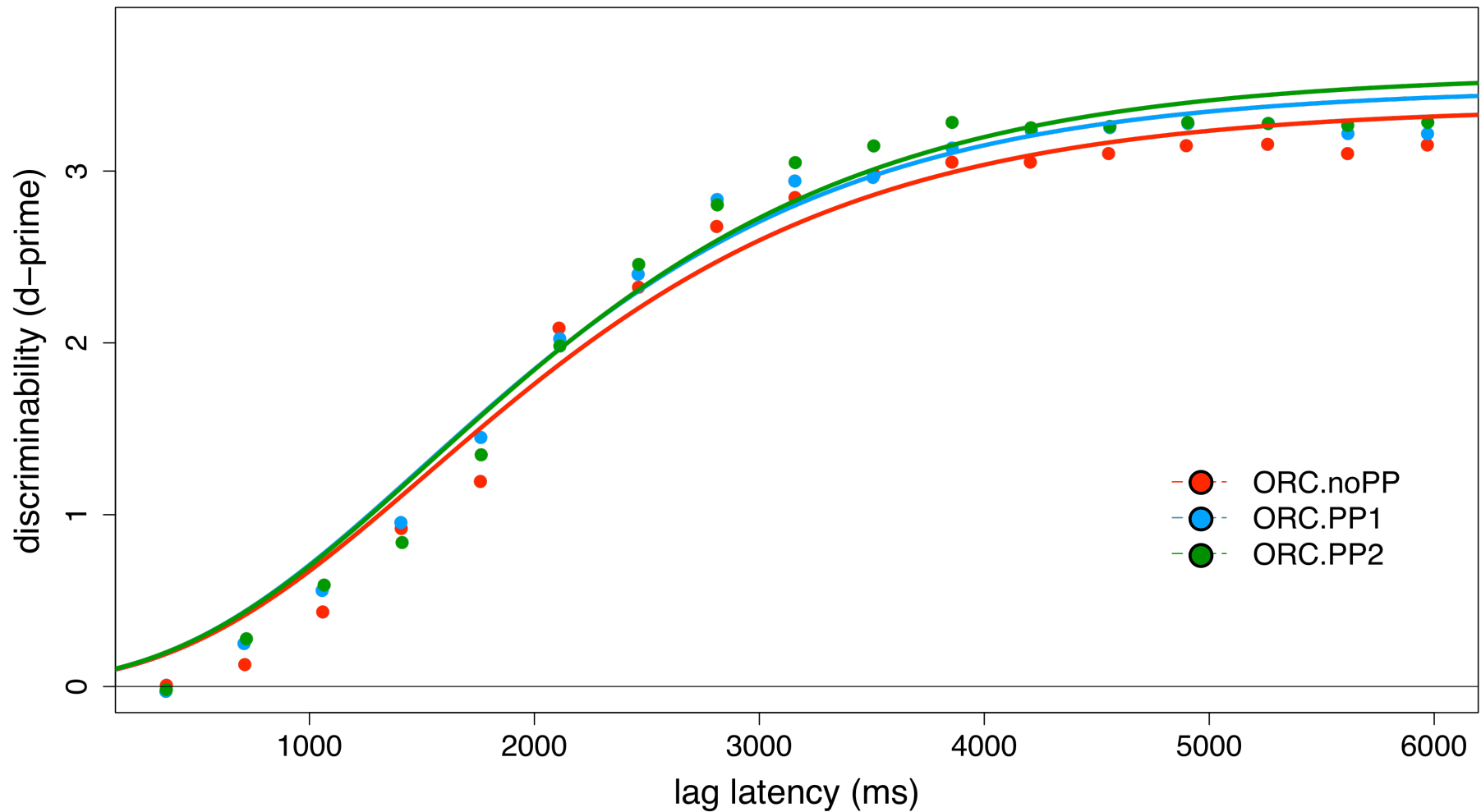
Results: Object Relative Clauses



Results: Object Relative Clauses



Results: Object Relative Clauses



Results: Best-fit parameters

	noPP	PP+1	PP+2
Asymptote λ	3.37	3.49	3.54
Rate β (sec ⁻¹)	0.746		
Intercept δ (sec)	0.724		

2064 ms

$R^2 = 0.98$

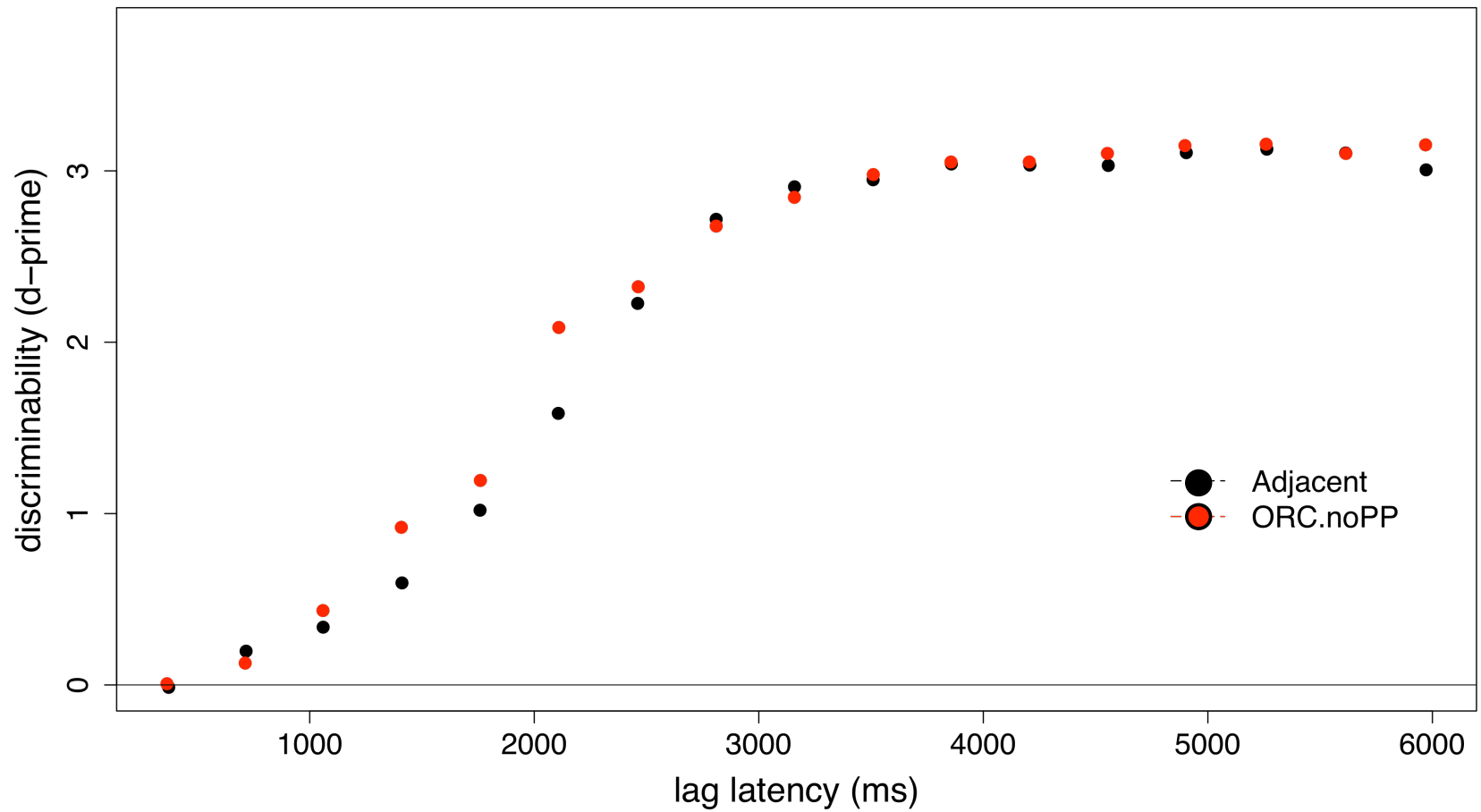
Results: Best-fit parameters

	noPP	PP+1	PP+2
Asymptote λ	3.37	3.49	3.54
Rate β (sec ⁻¹)		0.746	
Intercept δ (sec)		0.724	

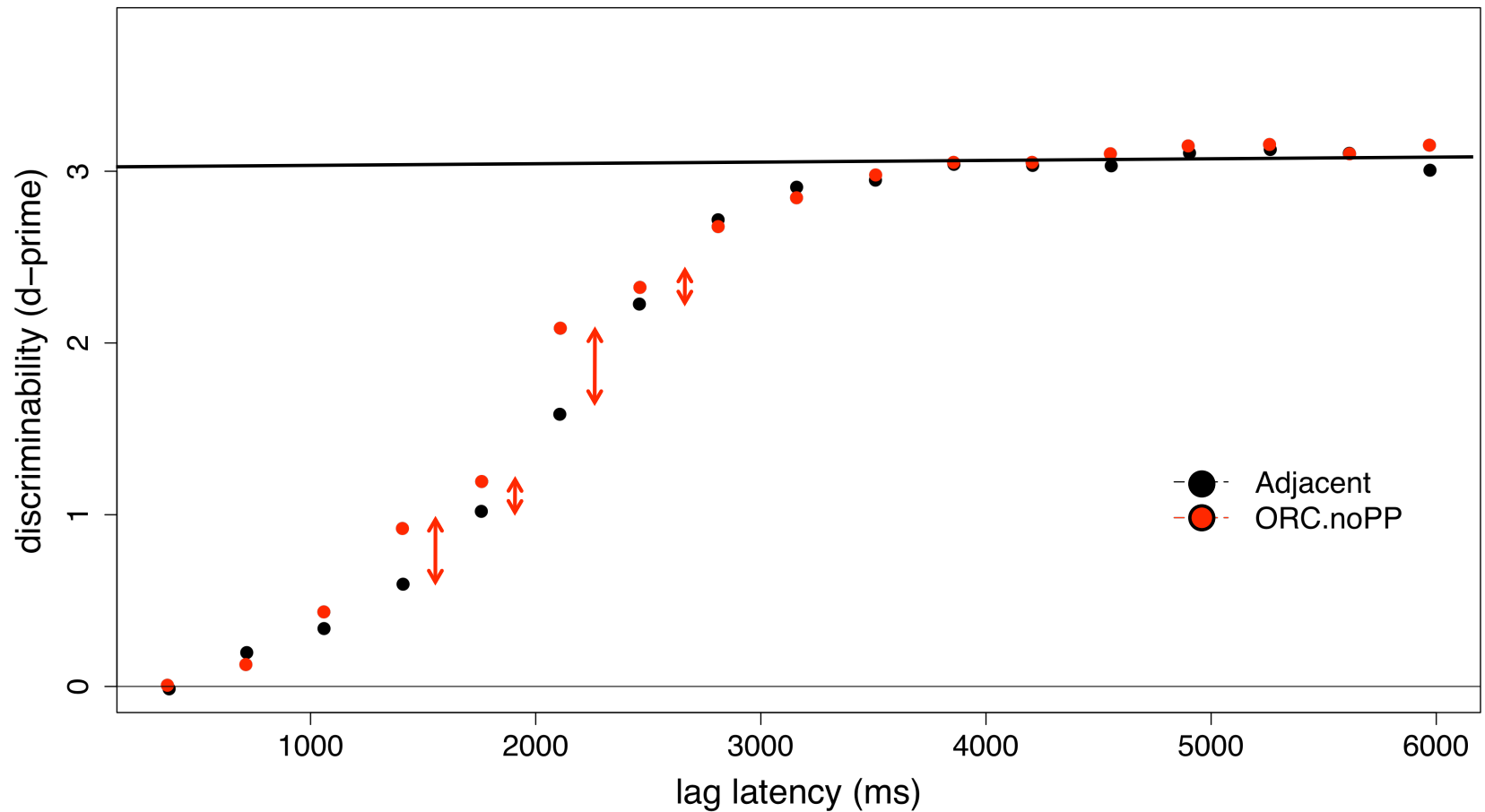
2064 ms

$R^2 = 0.98$

Results: Adjacent v. ORC



Results: Adjacent v. ORC



Summary

- For RC-modified subjects, there was no rate effect of adding more PPs
- There was an increase in asymptotic accuracy as more PPs were added:
[3.37 d' < 3.49 d' < 3.54 d']
consistent across subjects
- **Concern**: overall processing was slow and adjacent subject-verb dependencies were slowest of all

Experiment 2

- The preposed XPs that controlled for ordinal position may have significantly dampened/swamped the subject-verb relevant processing
 - XPs were attachable to either subject or verb
 - ... and sometimes ambiguously
- Experiment 2 uses local environments that are identical to Experiment 1, but with an unambiguous embedding context to control for ordinal position

Materials and Methods

- **ADJACENT**

The medic who was tending wounds in the trenches near the border observed that

the soldier twitched/*snagged.

- **+Object Relative Clause/NoPP**

The medic in the trenches near the border observed that

the soldier that the rebel wounded twitched/*snagged.

- **+Object Relative Clause/+2PP**

The medic observed that

the soldier that the rebel wounded in the trenches near the border twitched/*snagged.

- **+ADVERB**

The medic who tended wounds in the trenches near the border observed that

the soldier slightly twitched/*snagged.

Materials and Methods

- Acceptability × Length

ADJACENT

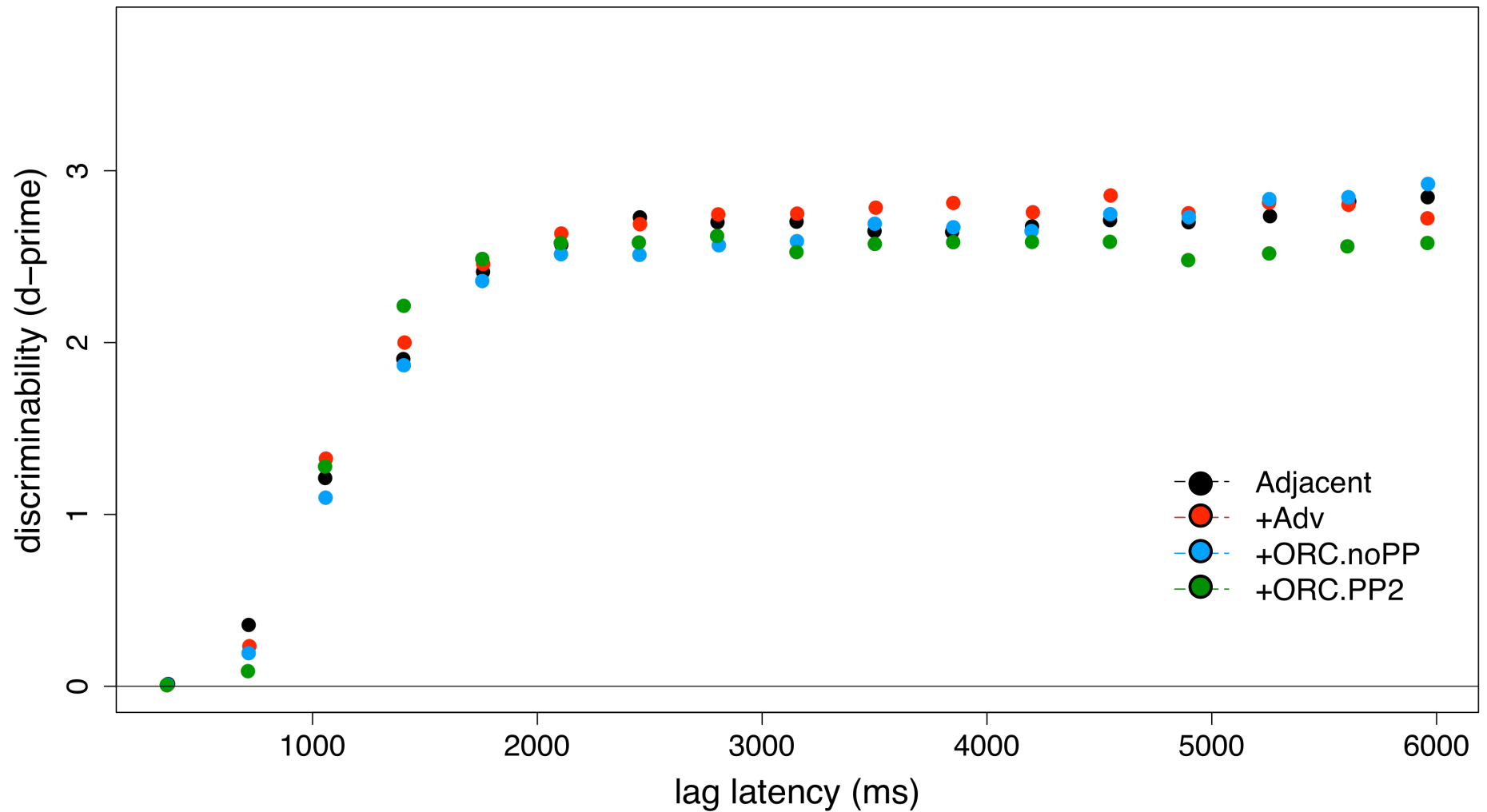
+OBJECT RC (ORC.noPP)

+OBJECT RC/2PP (ORC.2PP)

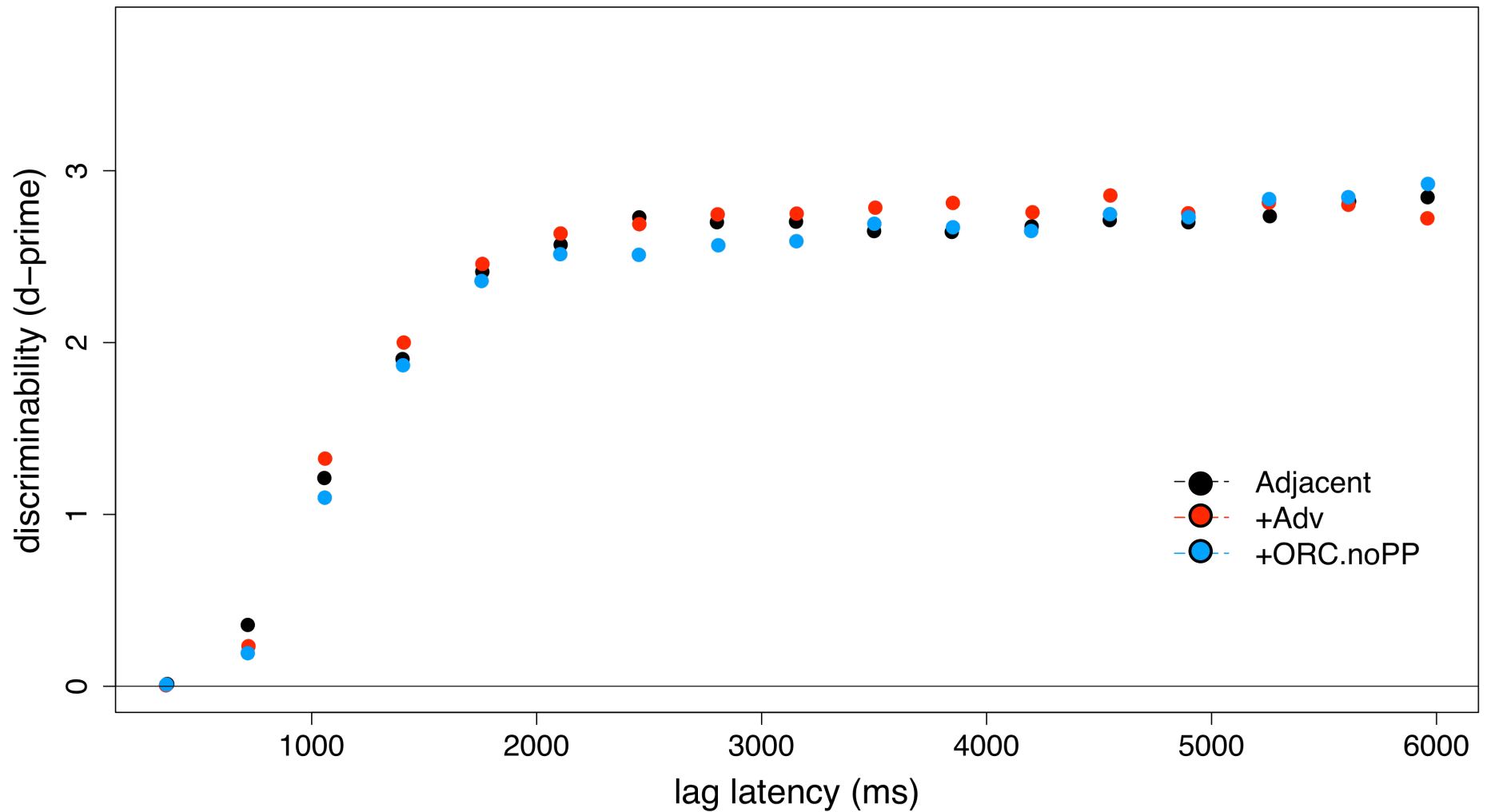
+ADVERB

- 36 item sets
- Fillers identical to experiment 1
- MR-SAT
 - n = 10, course credit for a Semantics course
 - Five sessions + 1 practice session

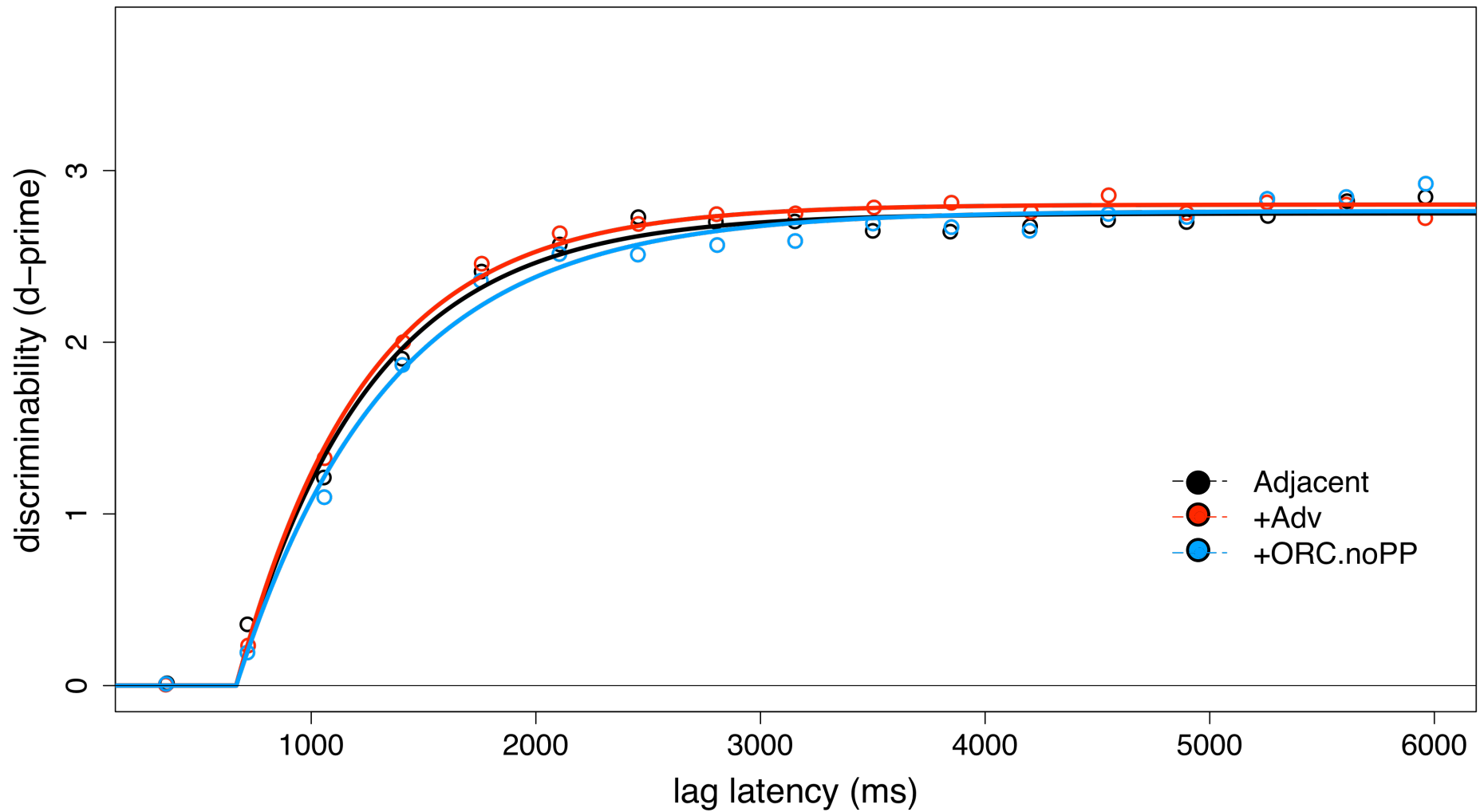
Results: all data



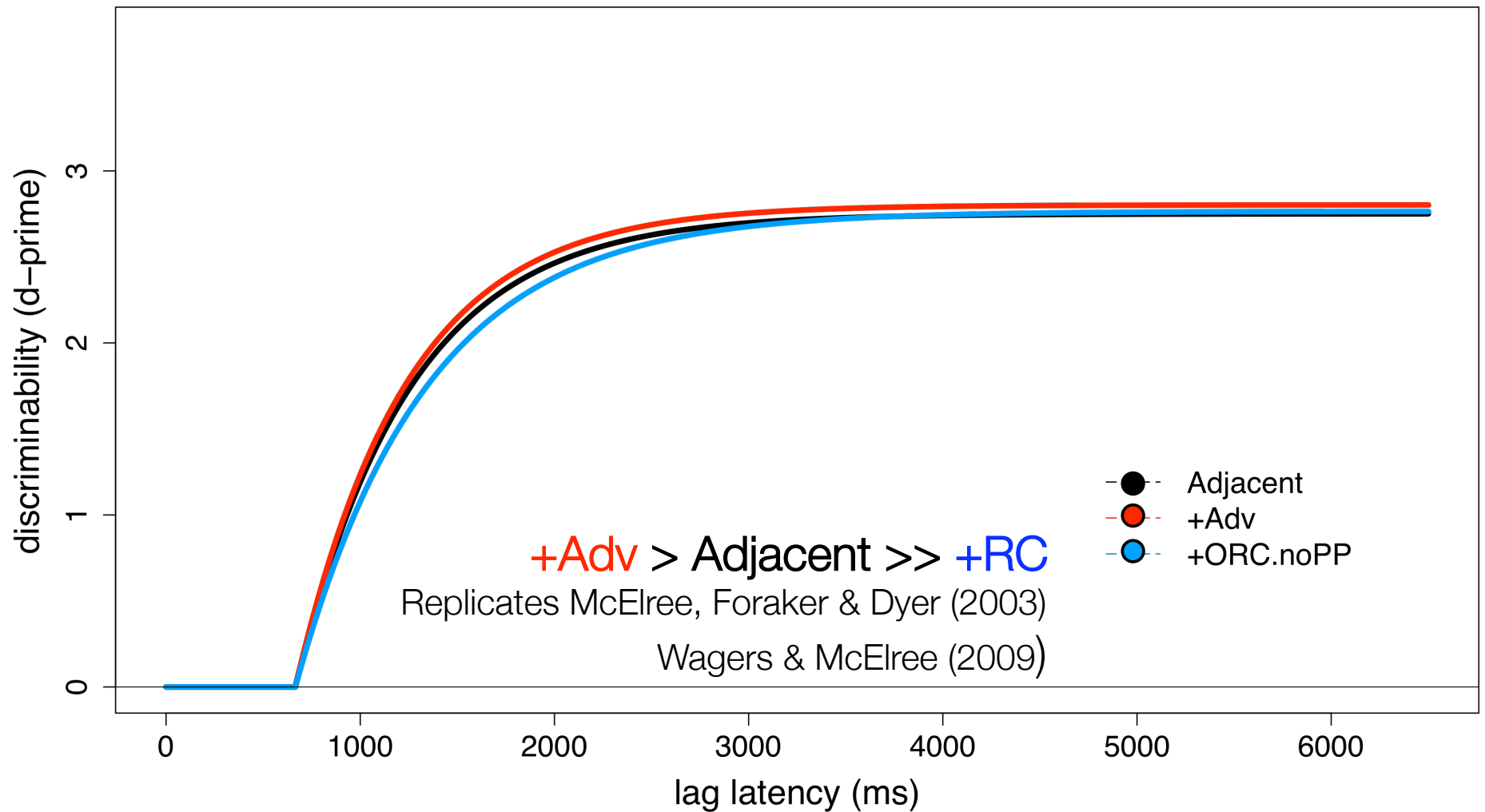
Results: Adjacent, +Adverb, +ORC



Results: Best-fit curve



Results: Best-fit curve



Results: Best-fit parameters

	Adjacent	+ADV	+ORC <i>noPP</i>	
Asymptote λ	2.75	2.80	2.76	
Rate β (sec ⁻¹)	1.69	1.74	1.48	
Intercept δ (sec)		0.466		

Results: Best-fit parameters

	Adjacent	+ADV	+ORC <i>noPP</i>	
Asymptote λ	2.75	2.80	2.76	
Rate β (sec ⁻¹)	1.69	1.74	1.48	
Intercept δ (sec)	0.466			

Average 'speed': $1/\beta + \delta$

Results: Best-fit parameters

	Adjacent	+ADV	+ORC <i>noPP</i>	
Asymptote λ	2.75	2.80	2.76	
Rate β (sec ⁻¹)	1.69	1.74	1.48	
Intercept δ (sec)		0.466		

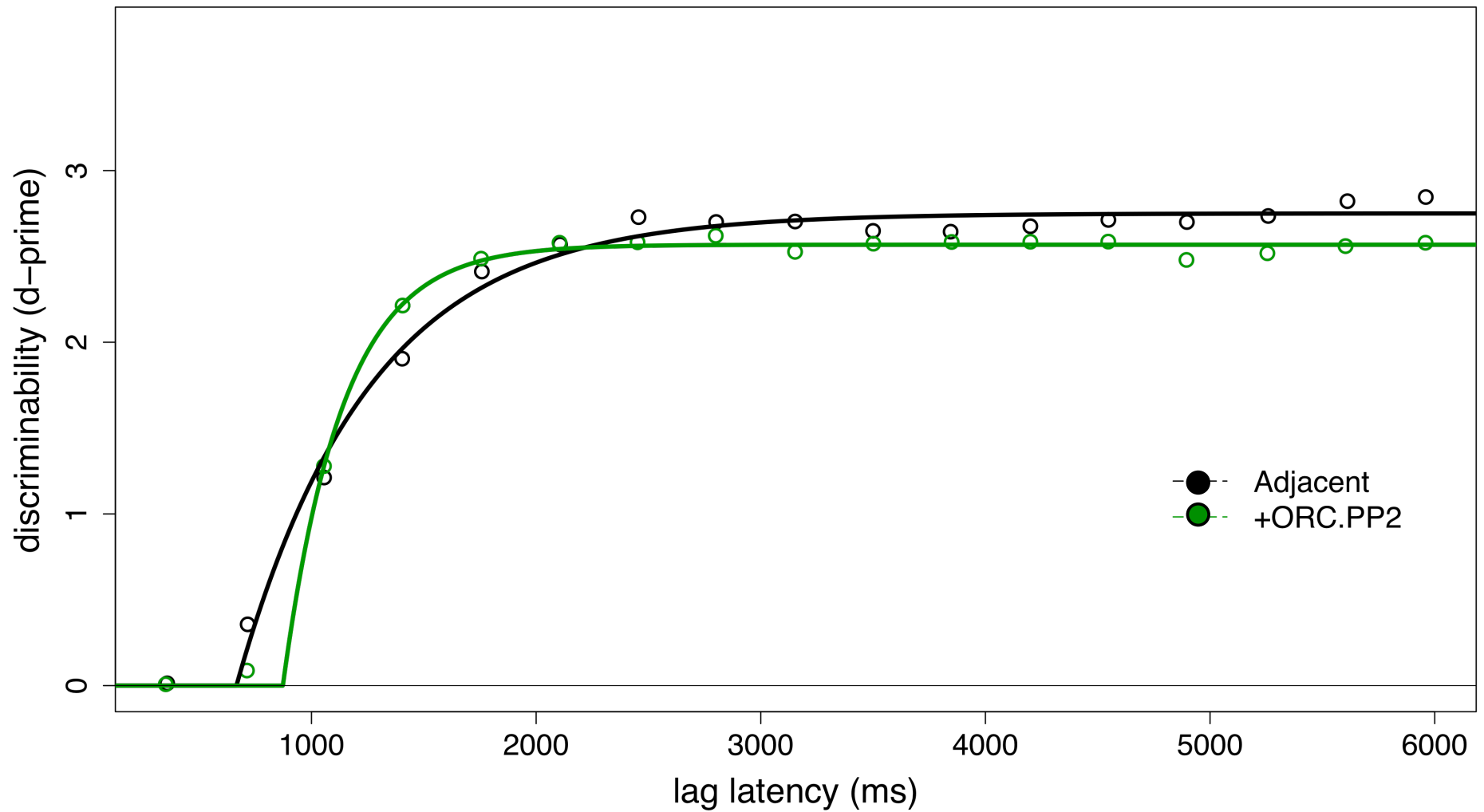
1056 ms 1040 ms 1142 ms

+Adv > Adjacent >> **+RC**

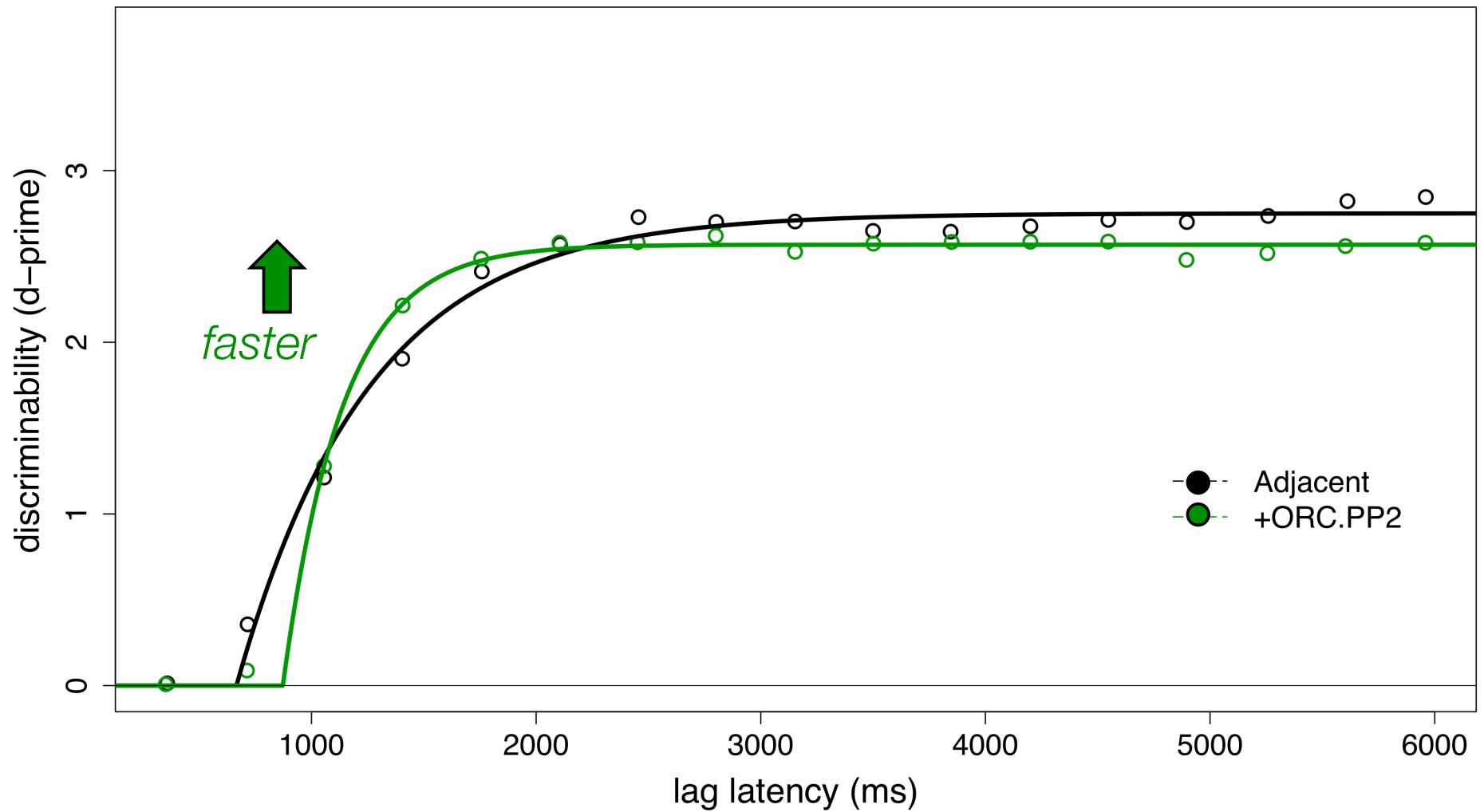
Replicates McElree, Foraker & Dyer (2003)

Wagers & McElree (2009)

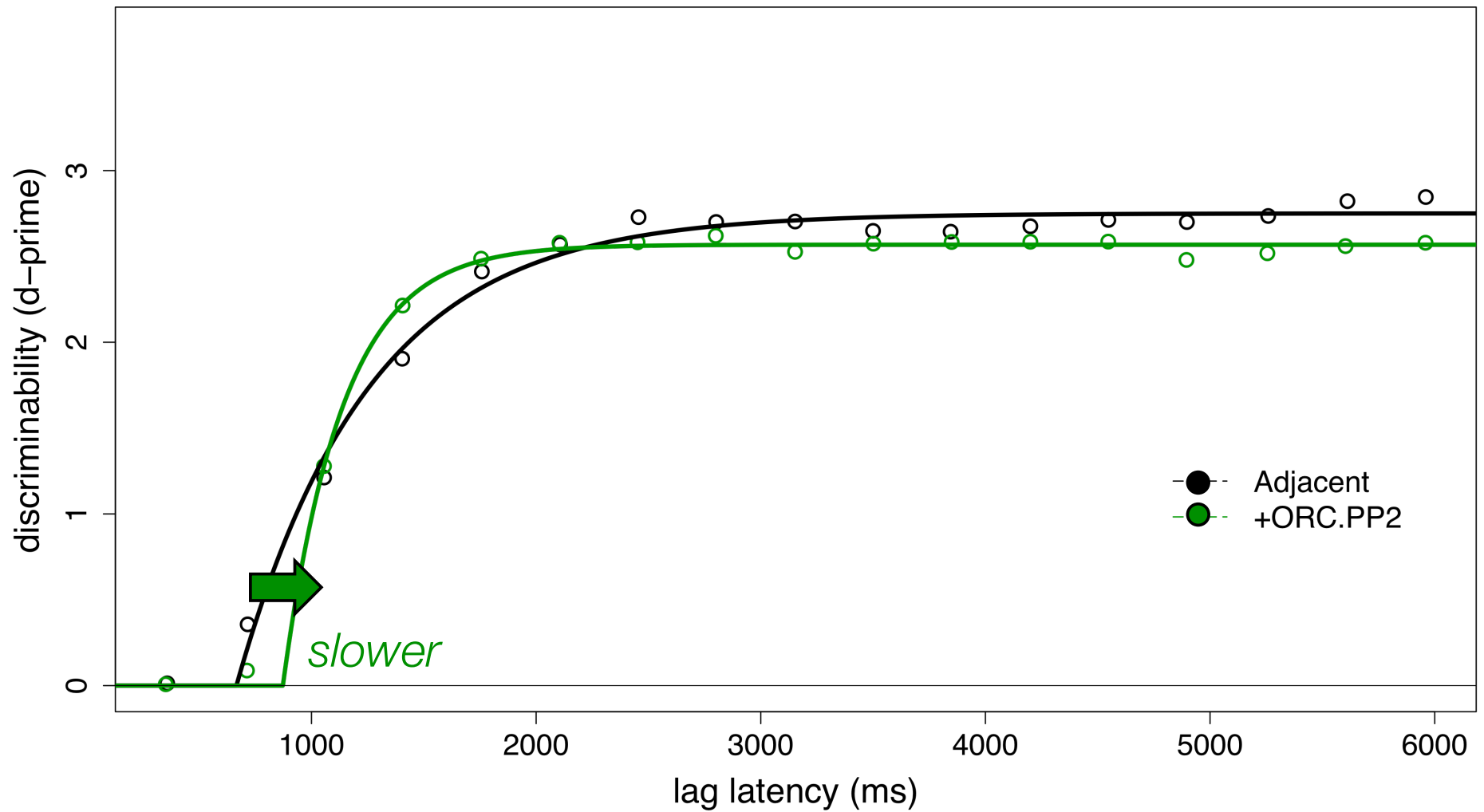
Results: Adjacent v. ORC.PP2



Results: Adjacent v. ORC.PP2



Results: Adjacent v. ORC.PP2



Results: Best-fit parameters

	Adjacent	+ADV	+ORC <i>noPP</i>	+ORC <i>PP2</i>
Asymptote λ	2.75	2.80	2.76	2.58
Rate β (sec ⁻¹)	1.69	1.74	1.48	3.75
Intercept δ (sec)		0.466		0.673
	1056 ms	1040 ms	1142 ms	939 ms

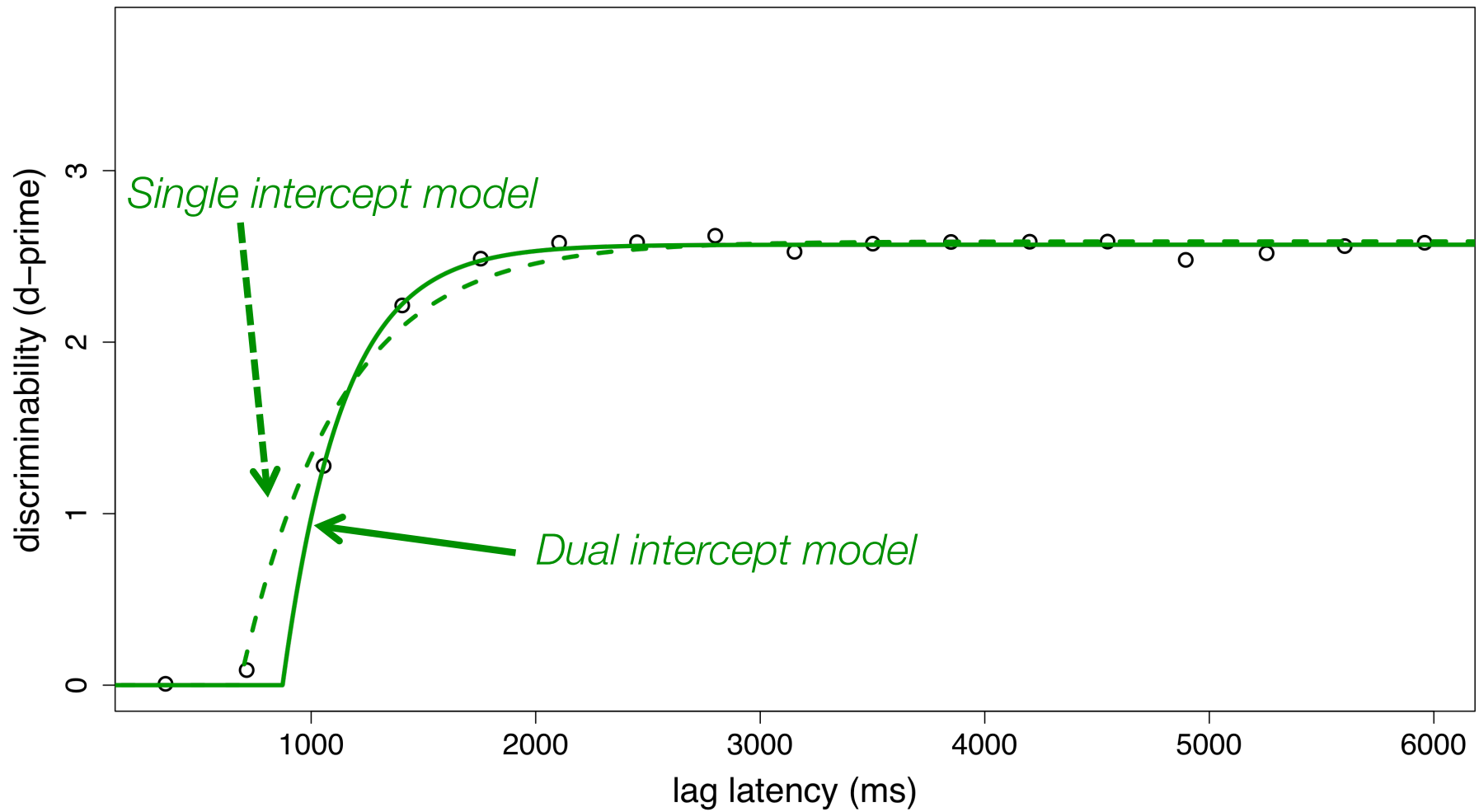
RC.PP2 >> +Adv > Adjacent >> +RC.noPP

Results: Best-fit parameters

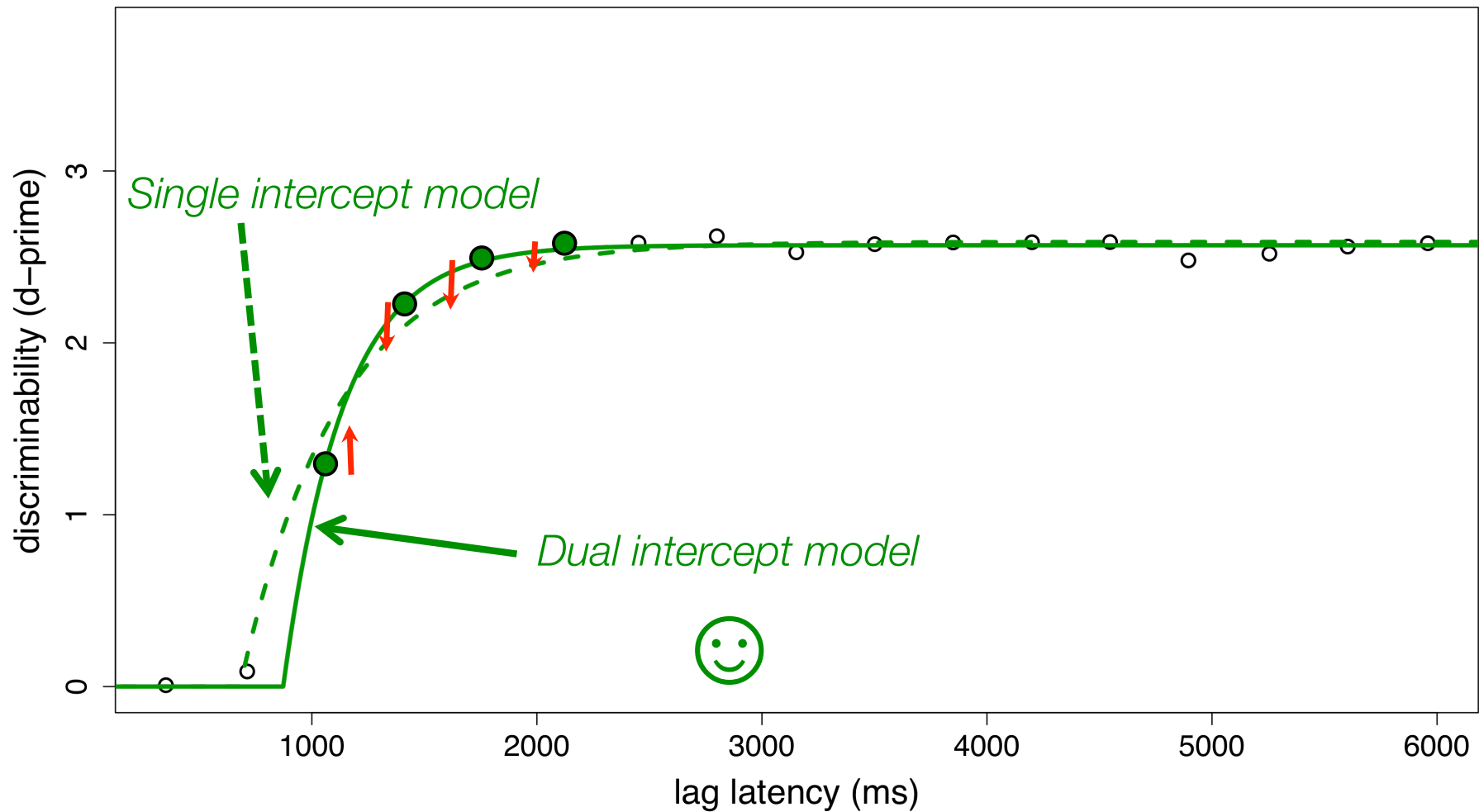
	Adjacent	+ADV	+ORC <i>noPP</i>	+ORC <i>PP2</i>
Asymptote λ	2.75	2.80	2.76	2.58
Rate β (sec ⁻¹)	1.69	1.74	1.48	> 3.75
Intercept δ (sec)		0.466		< 0.673
	1056 ms	1040 ms	1142 ms	939 ms

RC.PP2 >> +Adv > Adjacent >> +RC.noPP

Results: Model comparison



Results: Graphical model comparison



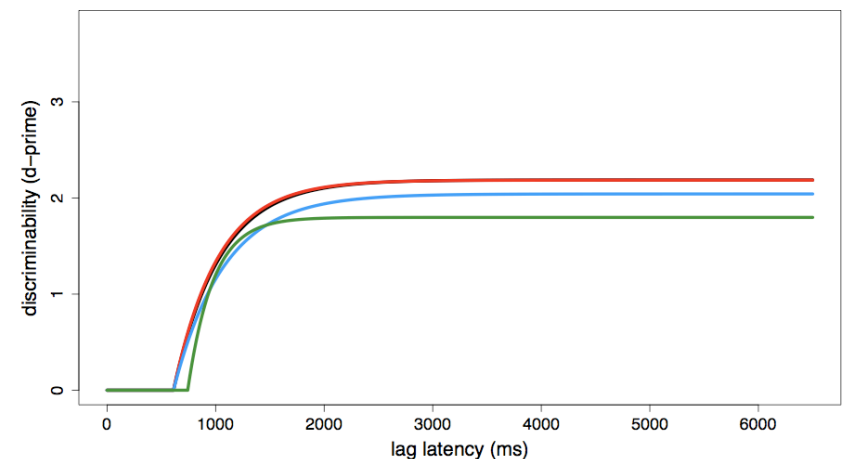
Results: Model comparison quantified

	Adjusted R-squared	Deviance	AIC	BIC
4-4-1 Single intercept	0.9917	-147.41	-129.41	-109.44
4-4-2 Dual intercept	0.9938	-156.47	-136.47	-114.28
4-4-4 Saturated intercept	0.9938	-157.30	-133.31	-106.68

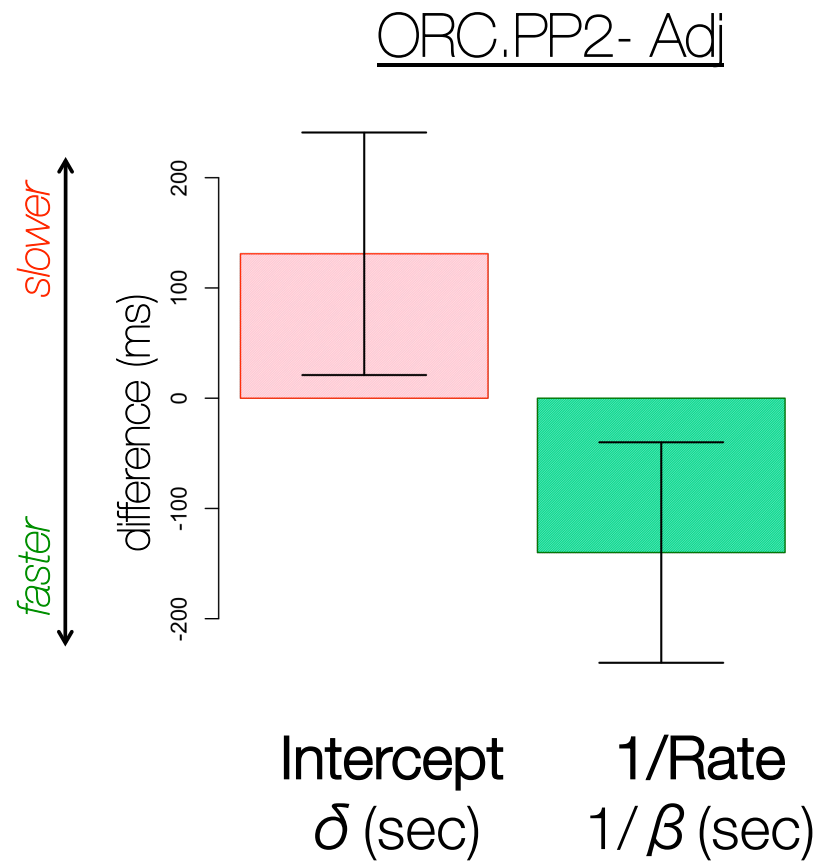
$$G^2(1) = 9.1, p < .005$$

$$G^2(2) = 0.8, n.s.$$

Consistent parameter ranking
across participants ($p < .05$)



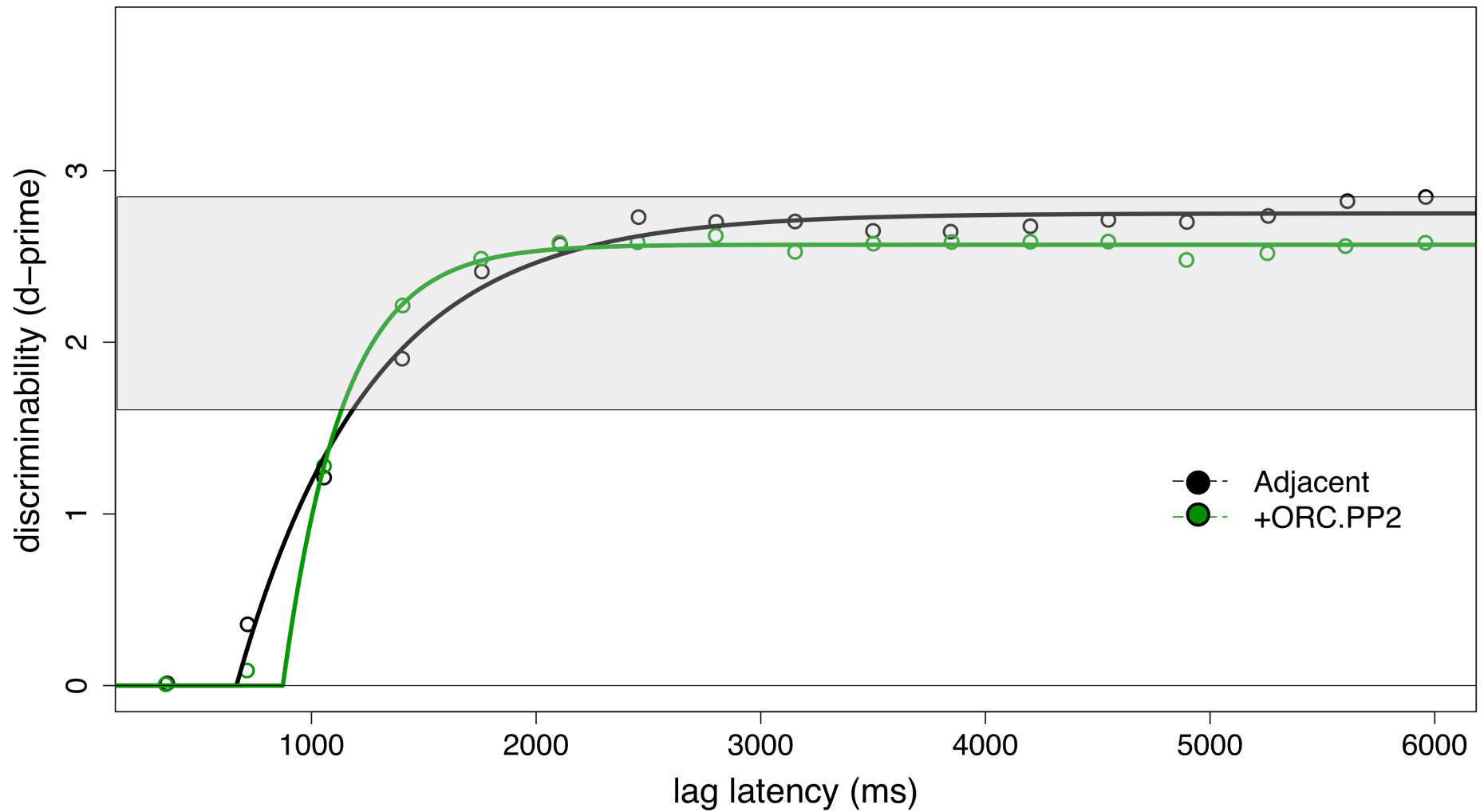
Results: Across participants



Results: Summary

- When the V-dependencies are strictly local, we observe two distinct effects on speed for the strong anti-local context:
 - An intercept shift **SLOWER**
Discriminating information is available much later
 - A rate increase **FASTER**
Information is accrued much faster
 - Overall: a facilitation in speed

Relation to reaction times



Conclusion

- Strongly anti-local S-V relationship formation is associated with faster dynamics
 - Consistent with predictive accounts (sloughing, working ahead)
 - Less consistent with memory-strength accounts
- However, it associated with a cost: discriminative information is available later
- *Facilitation* obtains on balance for modest-to-high accuracy processing

What is the cost?

- Focus of attention costs
 - 85 ms / McElree et al. (2003)
 - 87 ms, 74 ms / Wagers & McElree (2009)
 - 83 ms / this study [ORC.noPP-Adjacent]
- Intercept cost:
 - 207 ms (+44%)

What is the cost?

- Previous studies:

- ...]_{VP}]_S]_{NP} → [NP V]_S

- Current study

- ... NP]_{PP-2}]_{PP-1}]_{VP}]_S]_{NP} → [NP V]_S

- ... NP]_{PP-2}]_{PP-1}] → ... PP-1]_{VP}]_S]_{NP} → [NP V]_S

- Relating the scope of focal attention with the chunking of syntactic category

A diversity of timing measures

- RTs masked two underlying effects
 - Speed and accuracy tradeoffs are not predictable
- Dillon et al., *Thurs*, Binding *ziji*
faster rate was associated with lower accuracy
- Staub, *Fri*, frequency & predictability in fixation times
RT distribution modeling

Collaborators and acknowledgments

- Sarah Napoli
(UCSC Linguistics)
- Shayne Sloggett, Pranav Anand and
LING116 members
- CUNY reviewers
- Office of the Dean of Humanities, UCSC,
and UCSC Academic Senate Committee on Research



(c) Brian Dillon

Thank you.

Appendices

Results: Best-fit parameters

	Adjacent	+ADV	+ORC <i>noPP</i>	+ORC <i>PP2</i>
Asymptote λ	2.75	2.80	2.76	2.58
Rate β (sec ⁻¹)	1.69	1.74	1.48	3.75
Intercept δ (sec)		0.466		0.673
	1056 ms	1040 ms	1142 ms	940 ms

RC.PP2 >> +Adv > Adjacent >> +RC.noPP

Results: 4-4-4 parameters

	Adjacent	+ADV	+ORC <i>noPP</i>	+ORC <i>PP2</i>
Asymptote λ	2.75	2.80	2.76	2.57
Rate β (sec ⁻¹)	1.63	1.77	1.51	3.76
Intercept δ (sec)	0.467	0.476	0.476	0.673
	1057 ms	1039 ms	1140 ms	939 ms

RC.PP2 >> +Adv > Adjacent >> +RC.noPP

Results: 4-4-1 parameters

	Adjacent	+ADV	+ORC <i>noPP</i>	+ORC <i>PP2</i>
Asymptote λ	2.75	2.80	2.76	2.58
Rate β (sec ⁻¹)	1.68	1.72	1.46	1.77
Intercept δ (sec)	0.460			

1056 ms

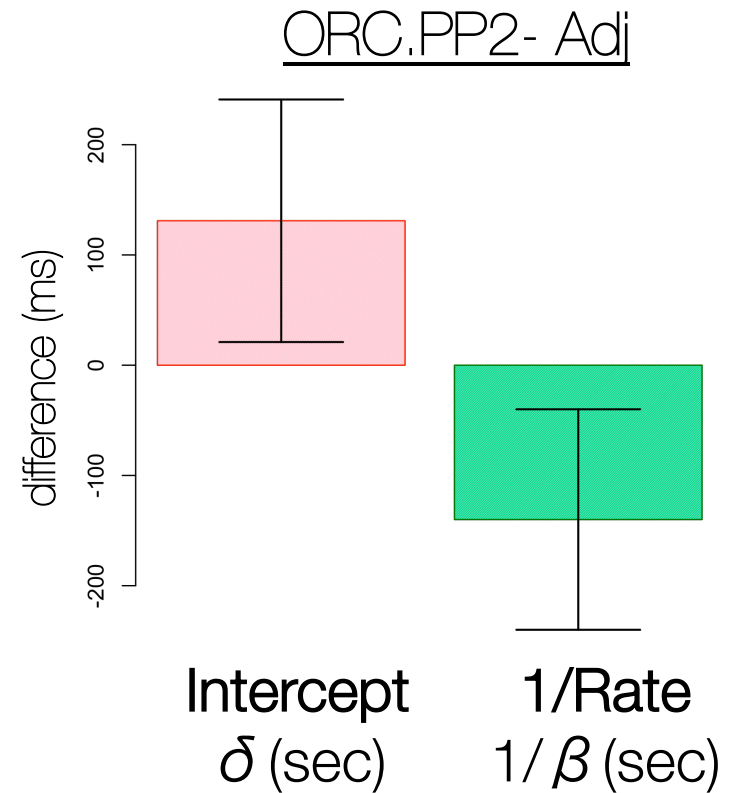
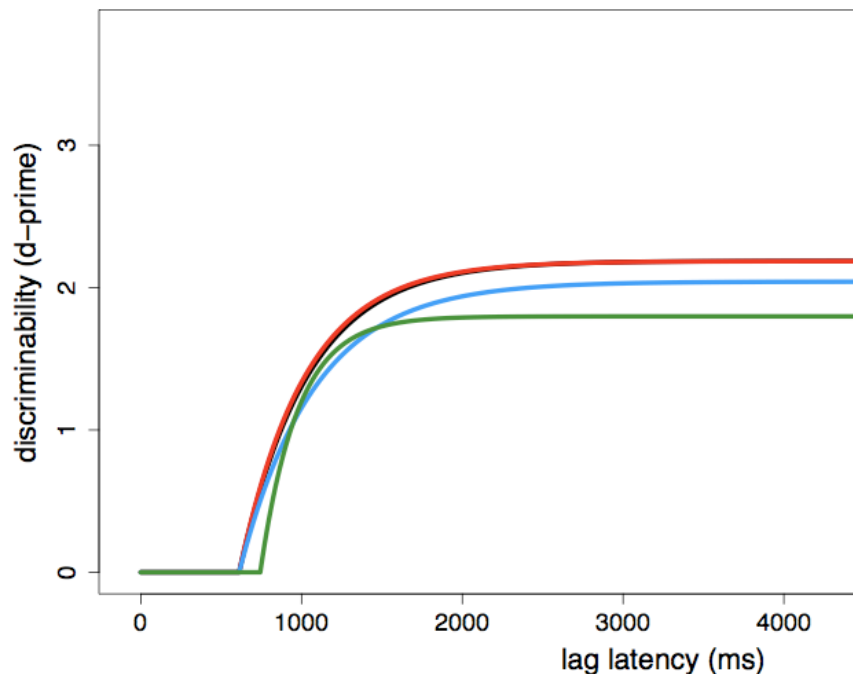
1042 ms

1143 ms

1023 ms

RC.PP2 >> +Adv > Adjacent >> +RC.noPP

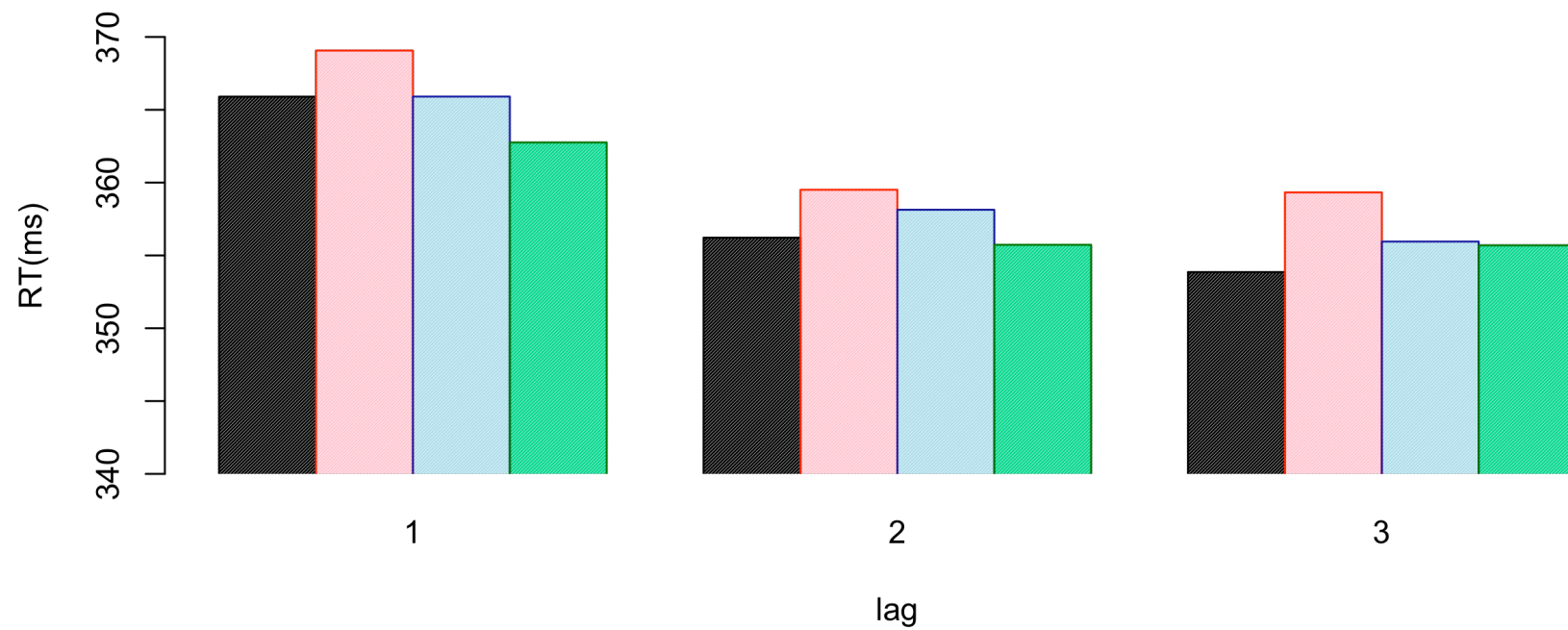
Results: Across participants



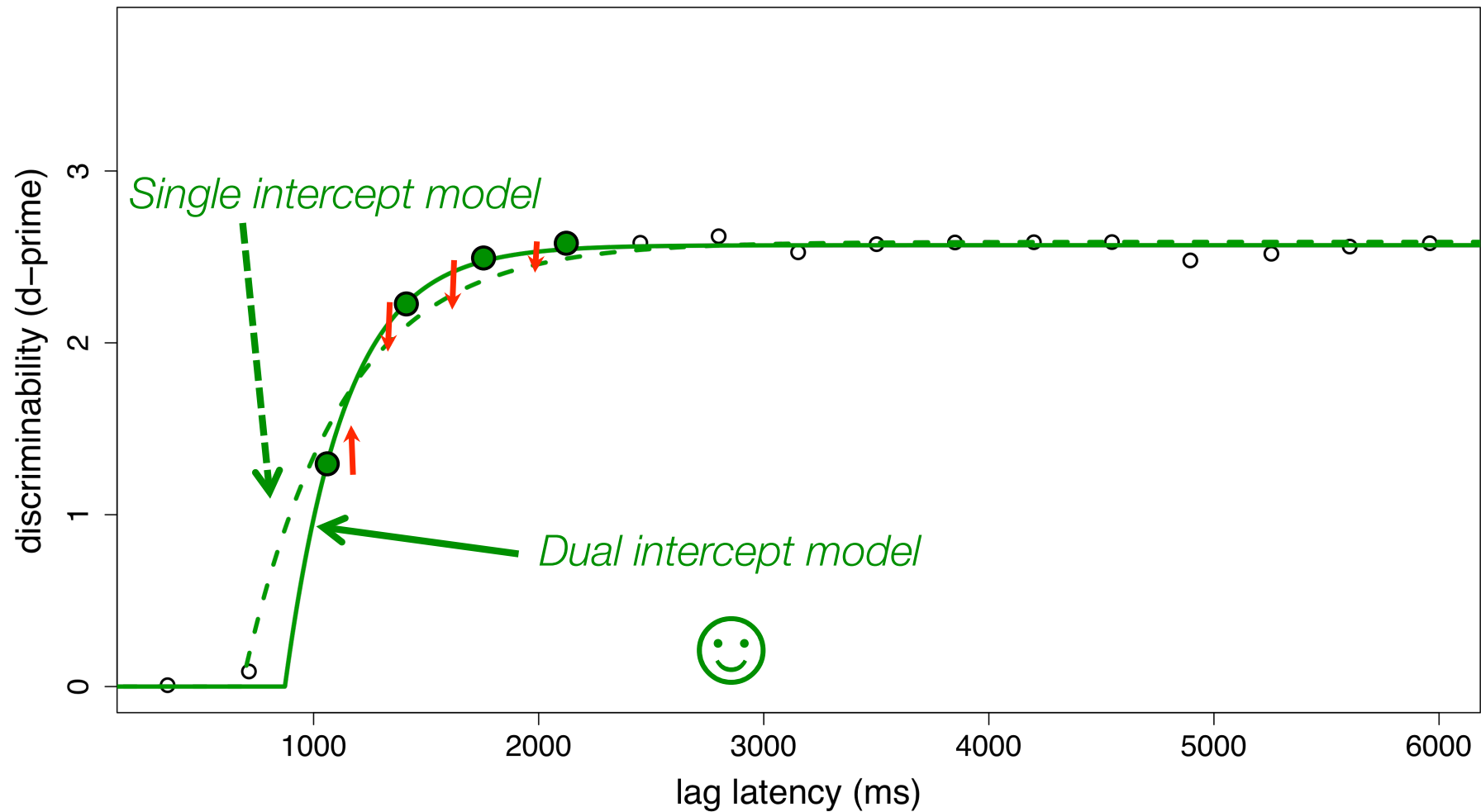
Consistent parameter ranking
across participants ($p < .05$)

Spill-over?

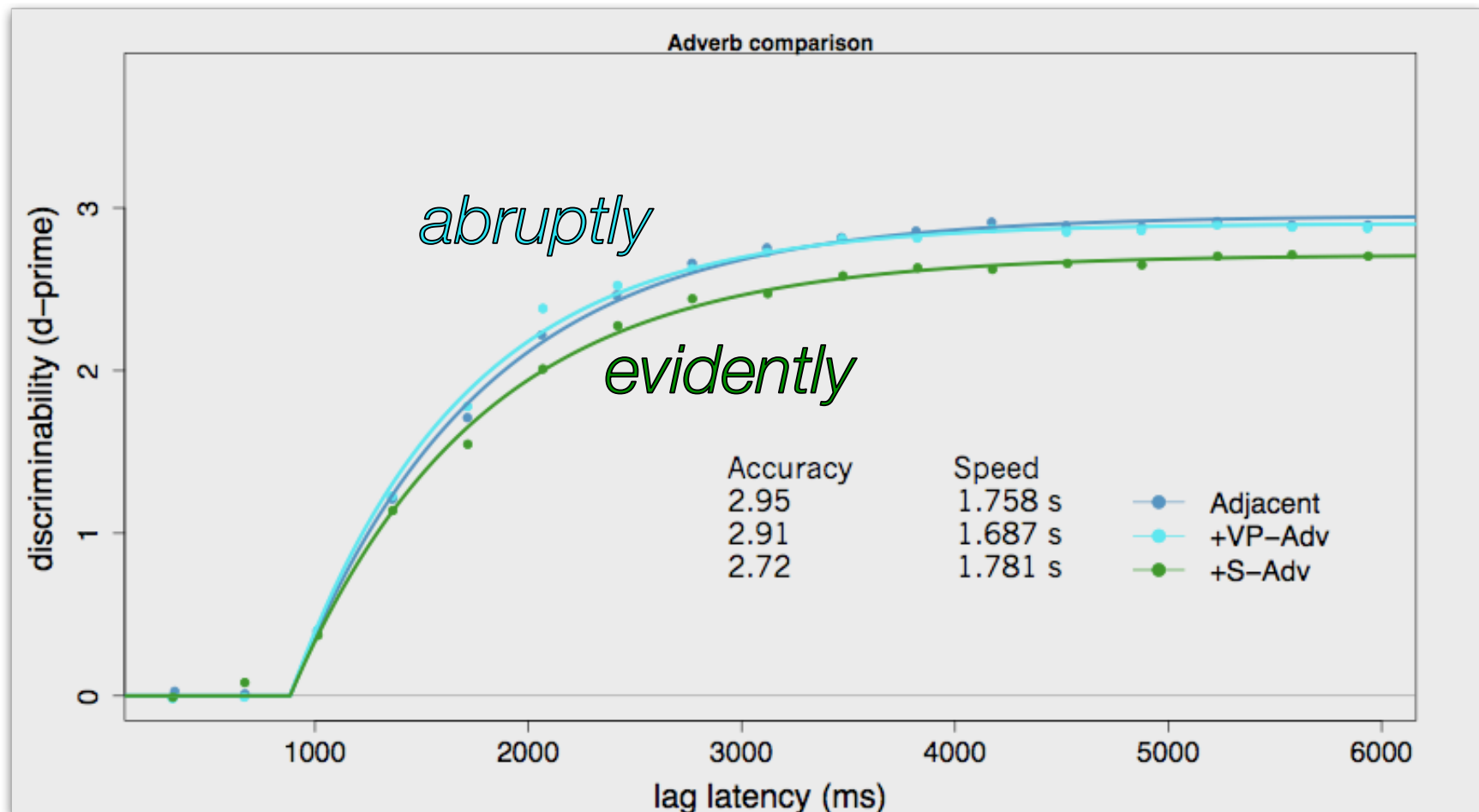
... lag latency plots



Results: Graphical model comparison



Wagers & McElree (2009) Adverb specificity



Exponential equation

$$d' = \lambda \cdot \left(1 - e^{-\beta \cdot (t - \delta)}\right)$$

Accuracy ~ session

