

# A Probabilistic Reconciliation of Coherence-Driven and Centering-Driven Theories of Pronoun Interpretation

Andrew Kehler  
UCSD Linguistics

(Contains joint work with Hannah Rohde, Jeffrey Elman, and Staci Osborn)

Workshop On Context Dependence In Language And Communication, Stanford University, November 3, 2011

1

# What's the Problem?

Subject Assignment (Crawley et al, 1990)

- a. George narrowly defeated John, and special interests promptly began lobbying him. [ him = George ]
- b. John was narrowly defeated by George, and special interests promptly began lobbying him. [ him = John ]
- c. George narrowly defeated John, and Mitt absolutely trounced him. [ him = John ]
- d. George narrowly defeated John, and he quickly demanded a recount. [ he = John ]

Grammatical Role Parallelism  
(Kamayama, 1986; Smyth, 1994)

Reasoning/World Knowledge  
(Hobbs, 1979)

2

# The SMASH Approach

- \* **S**earch: Collect possible referents (within some contextual window)
- \* **M**atch: Filter out those referents that fail 'hard' morphosyntactic constraints (number, gender, person, binding)
- \* **A**nd **S**elect using **H**euristics: Select a referent based on some combination of 'soft' constraints (grammatical role, grammatical parallelism, thematic role, referential form, ...)

3

# The Big Question

- \* Why would anybody ever use a pronoun?
- \* Speaker elects to use an ambiguous expression in lieu of an unambiguous one, seemingly without hindering interpretation
- \* A theory should tell us why we find evidence for different 'preferences', and why they prevail in different contextual circumstances
- \* We ask: *What would the discourse processing architecture have to look like to allow for a simple theory of pronoun interpretation?*

4

# Two Approaches to Discourse Coherence

- \* Centering Theory (Grosz et al. 1986; 1995):

*"Certain entities in an utterance are more central than others and this property imposes constraints on a speaker's use of different types of referring expressions... The coherence of a discourse is affected by the compatibility between centering properties of an utterance and choice of referring expression."*

- \* Define Centering constructs and rules:
  - \* A (single) backward-looking center ( $C_b$ ; the 'topic')
  - \* A list of "forward-looking centers" ( $C_f$ ; ranked by salience)
  - \* Constraints governing the pronominalization of the  $C_b$
  - \* Ranking on transition types defined by the  $C_b$  and the  $C_f$

5

# Centering

- \* A Centering-driven approach could conceivably explain why linguistic form could affect pronoun biases:

*George narrowly defeated John, and special interests promptly began lobbying him. [ him = George ]*

*John was narrowly defeated by George, and special interests promptly began lobbying him. [ him = John ]*

- \* Semantics and world knowledge do not come into play

6

## Coherence and Coreference

---

- ✦ Hobbs' (1979) Coherence-Driven Approach
- ✦ Pronoun interpretation occurs as a by-product of general, semantically-driven reasoning processes
- ✦ Pronouns are modeled as free variables which get bound during inferencing (e.g., coherence establishment)
  - The city council denied the demonstrators a permit because*
    - a. *they feared violence*
    - b. *they advocated violence* (adapted from Winograd 1972)
- ✦ Choice of linguistic form does not come into play

7

## Agenda

---

- ✦ Discuss the Hobbsian approach to discourse coherence a bit more
- ✦ Describe a series of experiments demonstrating that pronoun interpretation is influenced by coherence relations
- ✦ Present other evidence that suggests a role for a Centering-driven theory
- ✦ Present a model that integrates aspects of both approaches
- ✦ Describe new experiments that examine predictions of the model

8

## The Case for Coherence

---

- ✦ The meaning of a discourse is greater than the sum of the meanings of its parts
- ✦ Hearers will generally not interpret juxtaposed statements independently:
  - I'm traveling to Stanford this week. They've invited me to present at a workshop on Context Dependence in Language and Communication.*
- ✦ Explanation: Infer P from the assertion of  $S_1$ , and Q from the assertion of  $S_2$ , where normally  $Q \rightarrow P$ .
  - ?? *I'm traveling to Stanford this week. I love Trojan football.*

9

## Selected Other Relations

---

- ✦ Occasion: Infer a change of state for a system of entities from the assertion of  $S_2$ , establishing the initial state for this system from the end state of  $S_1$ .
  - Sally sent Sue a book. She read it a few days later.*
- ✦ Elaboration: Infer  $p(a_1, a_2, \dots, a_n)$  from the assertions of  $S_1$  and  $S_2$ .
  - Sally sent Sue a book. She mailed it via Federal Express.*

10

## Selected Other Relations

---

- ✦ Result: Infer P from the assertion of  $S_1$ , and Q from the assertion of  $S_2$ , where normally  $P \rightarrow Q$ .
  - Sally sent Sue a book. She thanked her for the present.*
- ✦ Violated Expectation: Infer P from the assertion of  $S_1$ , and Q from the assertion of  $S_2$ , where normally  $P \rightarrow Q$ .
  - Sally sent Sue a book. She never received it.*

11

## Transfer of Possession (Rohde, Kehler, and Elman 2006)

---

- ✦ Goal/Source preferences (Stevenson et al., 1994):
  - Bush seized the speech from Cheney. He... [Bush]*
  - Bush passed the speech to Cheney. He... [Bush/Cheney]*
- ✦ Possible explanations:
  - ✦ Thematic role preferences ('superficial')
  - ✦ Focus on end states of events ('deep')
- ✦ Latter is what one would expect for Occasion relations
  - Occasion: Infer a change of state for a system of entities from  $S_2$ , establishing the initial state for this system from the end state of  $S_1$

12

## Rohde, Kehler, and Elman (2006)

- ✦ Ran an experiment to distinguish these, comparing the perfective and imperfective forms for Source/Goal verbs

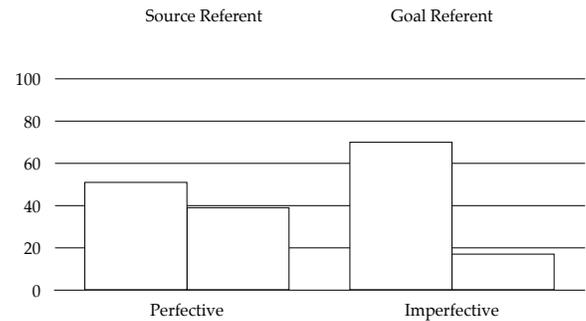
*Bush passed the speech to Cheney. He...*

*Bush was passing the speech to Cheney. He...*

- ✦ More references to the Source/Subject in the imperfective case would support the event structure/coherence analysis

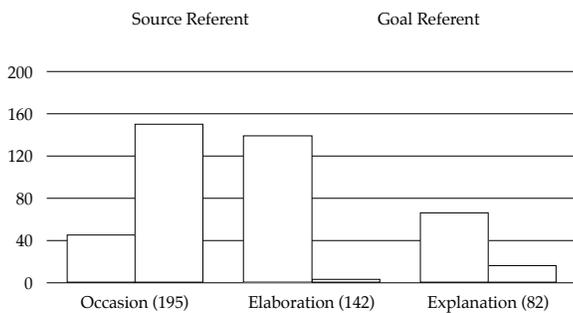
13

## Results



14

## Breakdown by Coherence Type (Perfective Only)



15

## Conditioning on Coherence

- ✦ Expectations are conditioned on coherence relations:

$$P(\text{pronoun}=\text{referent}) = \sum_{\text{CR} \in \{\text{Occasion, Explanation, Elaboration, ...}\}} P(\text{CR}) * P(\text{pronoun}=\text{referent} | \text{CR})$$

- ✦ Different factors in the context could influence these probabilities

- ✦ e.g., connective placement (cf. Kehler et al. 2008)

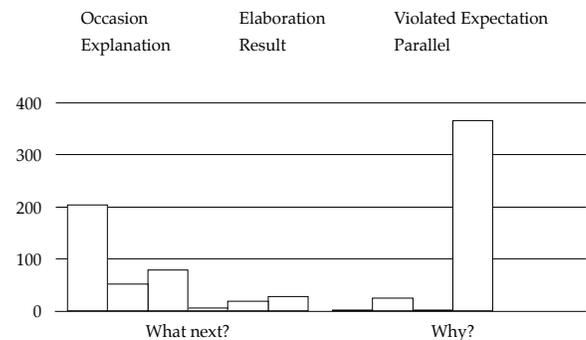
16

## Manipulating Coherence (Rohde, Kehler, and Elman 2007)

- ✦ If coherence matters, a shift in the distribution of coherence relations should induce a shift in the distribution of pronoun interpretations
- ✦ Run the previous experiment again, except with one difference in the instructions for how to continue the passage:
  - ✦ What happened next? (Occasion)
  - ✦ Why? (Explanation)
- ✦ Stimuli kept identical across conditions

17

## Results: Coherence Distribution



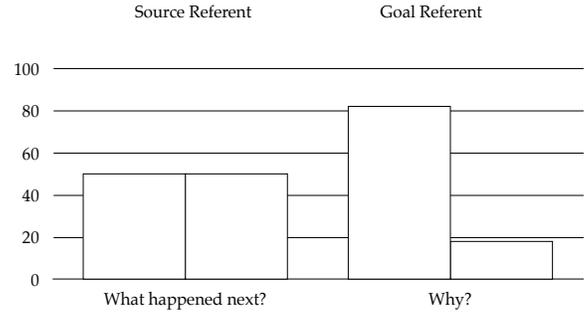
18

## Pronoun Biases

Coherence Relation	Prob(Source   Coh Reln) 1st Exp	Prob(Source   Coh Reln) 2nd Exp
Elaboration	0.99	1.00
Explanation	0.75	0.81
Violated Exp	0.87	0.81
Occasion	0.20	0.28
Result	0.16	0.10

19

## Results



20

## The Subject Preference

- Stevenson et al's (1994) study paired their pronoun-prompt condition with a no-prompt condition:

*Bush passed the speech to Cheney. He \_\_\_\_\_*  
*Bush passed the speech to Cheney. \_\_\_\_\_*

- They found a near 50/50 split in Source vs. Goal interpretations for pronouns in the prompt condition
- But in the no-prompt condition, they found a strong tendency to use a pronoun to refer to the subject and a name to refer to the object

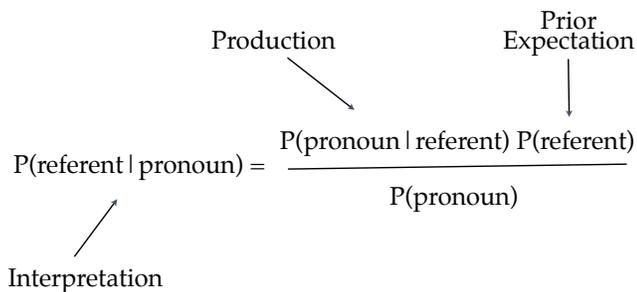
21

## The Subject Preference

- The latter result was replicated by Arnold (2001), and by us
- Arnold's numbers in a no-prompt condition for Source-Goal sentences:
  - 76% of Source-subjects pronominalized
  - 20% of Goal-non-subjects pronominalized
- Why would hearers resolve pronouns to the Goal so often when they don't have a similar production bias?

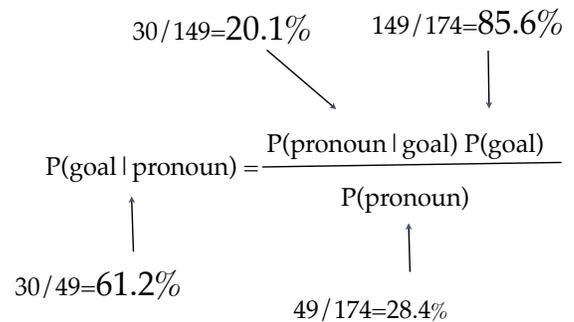
22

## Bayesian Pronoun Interpretation



23

## Arnold (2001)'s Data



24

## Bidirectional Influences (Rohde and Kehler, 2008)

- \* Bidirectional influences between pronoun interpretation and coherence establishment

*Bush passed the speech to Cheney. He \_\_\_\_\_*

*Bush passed the speech to Cheney. \_\_\_\_\_*

$$P(\text{CR}=\text{cr}) = \sum_{\text{referent} \in \text{referents}} P(\text{CR}=\text{cr} \mid \text{referent}) * P(\text{referent})$$

- \* Will the occurrence of an ambiguous pronoun affect the distribution of coherence relations?

25

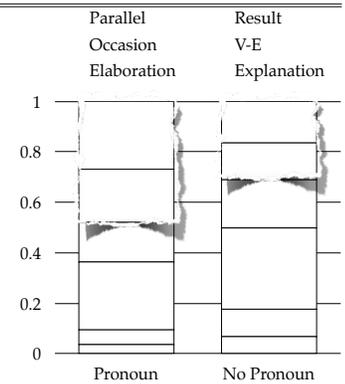
## Results: Coherence Relations

- \* Results replicated previous studies:

- \* Pronoun bias toward Goal: 47.98%

- \* Next-mention bias toward Goal: 84.23%

- \* Significant difference in coherence distribution: More source-biased relations in pronoun condition

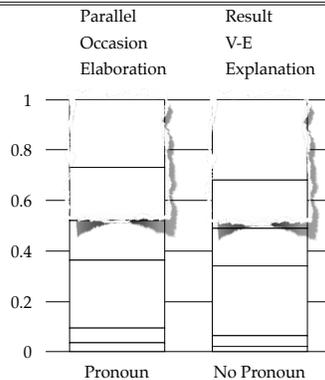


26

## Results: Coherence Relations (Pronoun continuations only)

- \* Coherence distributions are near identical, however, when only the no-pronoun condition continuations that used a pronoun are analyzed

- \* Pronouns are not merely free variables!



27

## Factors Influencing Production

- \* Bayesian formulation again:

$$P(\text{referent} \mid \text{pronoun}) = \frac{P(\text{pronoun} \mid \text{referent}) P(\text{referent})}{P(\text{pronoun})}$$

Production (Subject Bias)
Prior Expectation (Coherence-Driven)  
↓
↓

- \* Our data is consistent with a scenario in which coherence-driven biases primarily affect probability of *next-mention*, whereas grammatical biases affect *choice of referential form*

- \* Fukumura and van Gompel (2010) tested this latter prediction

28

## Implicit Causality

- \* Previous work has shown that so-called *implicit causality* verbs are associated with strong pronoun biases (Garvey and Caramazza, 1974 and many others)

*Amanda amazes Brittany because she \_\_\_\_\_* [subject-biased]

*Amanda detests Brittany because she \_\_\_\_\_* [object-biased]

- \* Also have strong coherence biases toward Explanation relations in full-stop prompts (Kehler et al. 2008)

*Amanda amazes Brittany. \_\_\_\_\_*  
(~60% Explanations in full-stop condition)

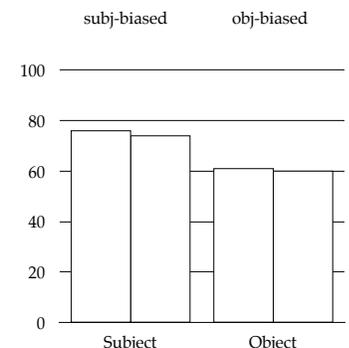
29

## IC Manipulation

- \* Contexts:

- \* Gary *scared* Anna after the long discussion ended in a row. This was because... [subject-biased]

- \* Gary *feared* Anna after the long discussion ended in a row. This was because... [object-biased]

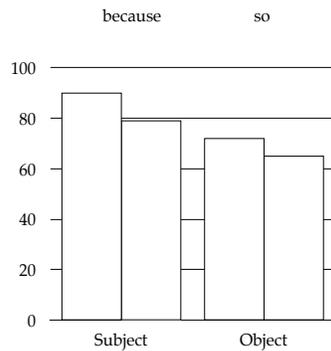


30

## Connective Manipulation

### Contexts:

- ✦ Gary scared Anna after the long discussion ended in a row, *because* \_\_\_\_\_
- ✦ Gary scared Anna after the long discussion ended in a row, *so* \_\_\_\_\_

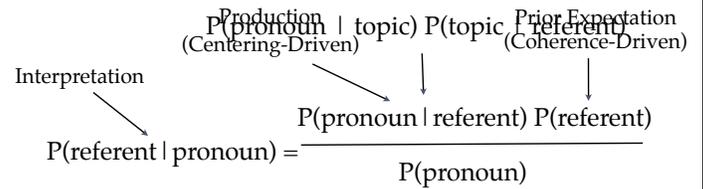


31

## Taking Stock

- ✦ We thus find evidence for a subject bias for pronouns: a *production* bias
- ✦ Or perhaps a bias toward pronominalizing the *topic*. Back to Centering:

... the use of a pronoun to realize the  $C_b$  signals the hearer that the speaker is continuing to talk about the same thing. (GJW 95, p. 214)



32

## Taking Stock

### What happens if we do more to establish the topic?

- ✦ Voice alternations: stronger biases for passives than actives
- ✦ Extended discourses: stronger biases when discourse topic is well-established
- ✦ Cross-linguistic differences: languages with null pronouns and topic marking

33

## IC and Passivization (Rohde and Kehler 2009)

- ✦ We used IC-NP1 verbs to test several predictions:

Amanda amazed Brittany. She \_\_\_\_\_

Brittany was amazed by Amanda. She \_\_\_\_\_

Amanda amazed Brittany. \_\_\_\_\_

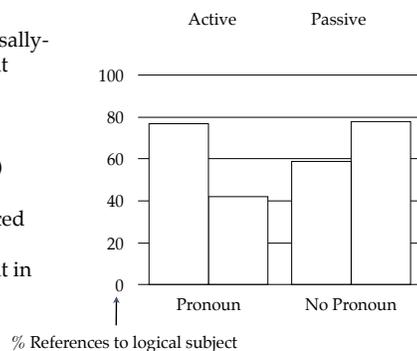
Brittany was amazed by Amanda. \_\_\_\_\_

- ✦ Question 1: Does passivization change the pronoun interpretation bias?
- ✦ Question 2: Does passivization effect the bias toward ensuing Explanation relations?
- ✦ Question 3: Does passivization change the pronoun production bias?

34

## Results: Mentions

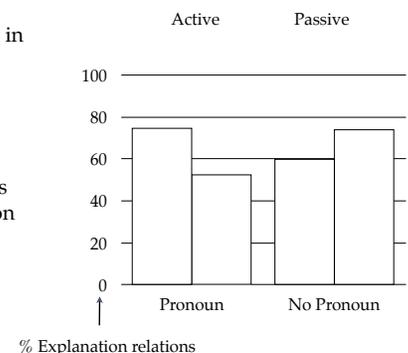
- ✦ Preference for causally-implicated referent ( $p < .001$ )
- ✦ Subject bias for pronouns ( $p < .001$ )
- ✦ Interaction: Reduced bias for causally-implicated referent in passive/pronoun condition ( $p < .05$ )



35

## Results: Coherence

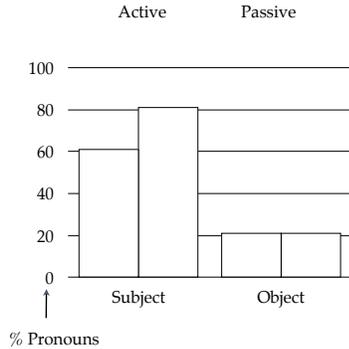
- ✦ Fewest Explanations in Pronoun+Passive condition ( $p < .001$ )
- ✦ Even in strongly-biased contexts, ambiguous pronouns not only shift mention biases, but expectations about ensuing coherence relations as well



36

## Results: Production

- \* Greater rate of pronominalization to the subject than non-subject
- \* Greater rate of pronominalization for passive subjects than active ones
- \* No difference for non-subjects, as expected



37

## Extended Discourses (Kehler and Osborn, 2010)

- \* If the discourse topic is more established (higher probability), then coherence relations shouldn't push biases around as easily

- \* Rohde and Kehler:

*Nick kicked a soccer ball to Justin.*  
(He) \_\_\_\_\_

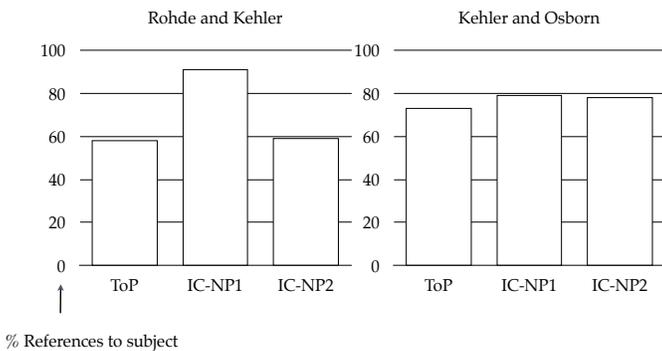
Used subset of ToP, IC-NP1, and IC-NP2 verbs from Rohde and Kehler

- \* Kehler and Osborn (2010):

*Nick is a very good athlete.*  
*He joined Justin on the field.*  
*He kicked a soccer ball to him.*  
(He) \_\_\_\_\_

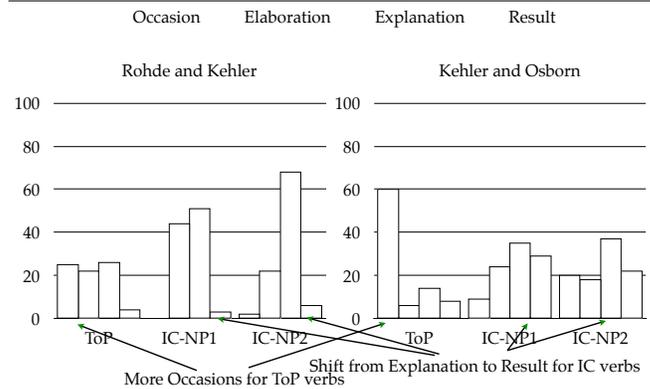
38

## Results: Reference (Pronoun Condition)



39

## Results: Coherence (Pronoun Condition)



40

## Japanese Pronouns

- \* Japanese has both null and overt pronouns, and topic marking
- \* Lore from the literature:
  - \* Japanese null pronouns should behave like English overt pronouns (Kuroda 1965; Kameyama 1985; Gundel et al, 1993)
  - \* Japanese overt pronouns should display a division-of-labor effect, e.g. whereby they prefer non-subjects
  - \* Topic (-wa) marked subjects should be more salient than -ga marked ones (Walker et al., 1994)

41

## Japanese Pronouns (Ueno and Kehler, 2010)

- \* Thus we ask what factors influence null and overt pronoun interpretation (grammatical role, semantic, topichood)

- \* Ran a version of Rohde et al. (2006) using ToP contexts:

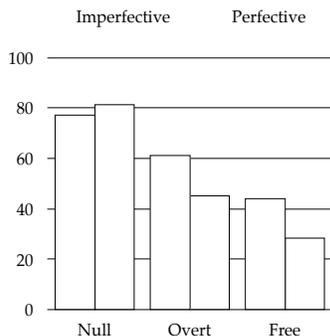
太郎が/は 次郎に本を 渡した / 渡している ところだった。  
Taro-wa/ga Jiro-ni hon-o watashita/watashi-te-iru tokoro-datta.  
Taro-TOP/NOM Jiro-to book-ACC handed/hand-INF-ASP scene-was  
'Taro-TOP/NOM handed/was handing a book to Jiro.'

主語省略/彼は/自由 \_\_\_\_\_  
shugo-shoryaku/kare-wa/jiyu  
subject-omission/he-TOP/free

42

## Results: ToP verbs, Aspect

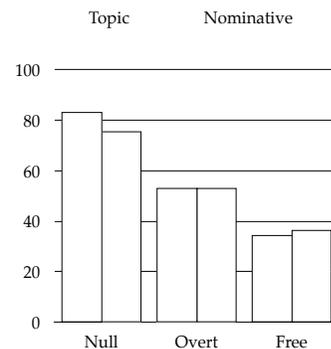
- ♦ Subject bias: null > overt > free ( $p < .0001$ )
- ♦ Aspect: Imperfective > perfective ( $p < .0001$ )
- ♦ But only for the overt and free conditions (null condition n.s.)
- ♦ Null pronouns in Japanese do not behave like overt English pronouns; Japanese overt pronouns do
- ♦ No division-of-labor effect



43

## Results: ToP verbs, Topichood

- ♦ Topic-marking: no main effect, nor pairwise effect in any prompt condition
- ♦ Surprising per Walker et al. (1994)



44

## Conclusions

- ♦ We come a long way from one-size-fits-all views like the SMASH approach
- ♦ The data presented here suggests a potential reconciliation of coherence-relation-driven and Centering-driven theories that accords with this view:
  - ♦ Coherence relations create top-down expectations about next mention
  - ♦ Centering-style constraints yield bottom-up evidence specific to choice of referential form
- ♦ Fits within a modern view in psycholinguistics that casts interpretation as the interaction of “top-down” expectations and “bottom-up” linguistic evidence
- ♦ We have gained insight into why we see evidence for so-called ‘preferences’ and their apparent inconsistency in different contextual circumstances
- ♦ The behavior of pronouns is thus an important source of insight into larger questions concerning the discourse processing architecture

45

## Conclusions

- ♦ I have focused on the results of passage completion studies
- ♦ Unlike many methodologies, passage completions allow us to compare the pronoun patterns against an appropriate baseline, i.e., those biases measured in no-pronoun prompt continuations
- ♦ These data allow us to explain how a production bias toward using pronouns to refer to subjects is compatible with an interpretation bias toward a non-subject
- ♦ It is therefore crucial that psycholinguistic studies control for the operative next mention biases (and coherence relations!) in their stimuli

46

Thank you!

47