Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives University of Tübingen 4 February 2006



# LOCALITY AND SLOPPY IDENTITY: EVIDENCE FROM A WEB SURVEY

#### Daniel Hardt

CMOL/Computational Linguistics Copenhagen Business School

## 1 A Controversial Judgment

- (1) John said Mary hit him. Harry did too. (said Mary hit John/Harry)
- (2) John said Mary hit him. Harry said she did too. (hit John/\*Harry)

## Sag/Williams Account

- In (1), VPE can mean Mary hit John (strict) or Mary hit Harry (sloppy)
- In (2), the sloppy reading is ruled out
- In general: sloppy readings in VPE must be *local* the sloppy pronoun must co-vary with the subject of the elided VP.

(Sag, 1976; Williams, 1977)

## An Alternative View

- If you try, you can get the sloppy reading for (2)
- The strict/sloppy alternation is not restricted to VP Ellipsis, so it's better to explain in terms of some general discourse principles (like a general preference for parallelism)

## 2 Plan

- Locality and Sloppy Identity: Two Views
- Web Survey
  - Results: Strict blocks Sloppy
  - Who is Right?
- Proposal

# 3 Locality and Sloppy Identity: Two Views

### Sag/Williams

- Elided VP must have the same meaning as its antecedent
- Lambda binder with scope over VP
  - John<sub>1</sub>  $\lambda$  x.x said Mary hit x
  - Harry\_2 did too. $\lambda$ x. x said Mary hit x

- With lambda-bound sloppy pronoun, VPE has same meaning as antecedent
  - John\_1 said Mary\_3 hit him\_1
  - Harry<sub>2</sub> said she<sub>3</sub> did too.  $\lambda$  x. x hit him<sub>2</sub>
- With non-local sloppy (also called *re-binding*), VPE can't satisfy same-meaning condition. Lambda binding doesn't help

### Non-Local Parallelism

- Subsequent authors have argued (Dalrymple *et al.*, 1991; Rooth, 1992; Hardt, 1993; Fiengo and May, 1994; Asher *et al.*, 2001) that there is no locality constraint on sloppy readings.
- VPE facts are a special case of a more general preference for Parallelism

## 4 Web Survey

- Linguist-GRID.org: web-based tool for interactive linguistic surveys
  - Developed by Matthias Kromann at Copenhagen Business School
- Survey advertised on LINGUIST and elsewhere in Nov/Dec 2004
- 29 subjects rated 30 examples of VP ellipsis, with reading indicated
- Four-point scale
  - 0 Fully Acceptable
  - 1 Closer to Acceptable
  - 2 Closer to Unacceptable
  - 3 Fully Unacceptable
- Examples presented together with the reading to be evaluated (parenthesized in italics)

## Results

• Four categories of examples:

Example Type	Mean Judgment
Strict	.66
John said Mary hit him. Harry did too. (said Mary hit John)	
Local Sloppy	.61
John said Mary hit him. Harry did too. (said Mary hit Harry)	
Non-local Sloppy	1.16
John said Mary hit him. Harry said she did too. (hit Harry)	
Non-local Sloppy (Strict Blocked)	.39
Nearly every student felt I ought to tutor her.	
But Susan didn't feel I should (tutor Susan)	

- Strict: the embedded pronoun in the antecedent receives the same reading at the ellipsis site
- Local Sloppy: embedded pronoun covaries with the local subject.
- *Non-local Sloppy*: embedded pronoun covaries with something other than the local subject. In this category, the strict reading is also a potential reading.
- Non-local Sloppy (Strict Blocked) the strict reading is unavailable (or degraded), because the pronoun is bound by a quantifier in the first sentence.



Figure 1: Survey Results

#### **Overall Results**

- Non-local sloppy readings are normally degraded
- They become completely acceptable when the strict reading is blocked.

#### Who is Right?

- Nobody: No previous account predicts these results
  - **Sag/Williams:** non-local sloppy readings should always be unacceptable; thus, their acceptability when strict is blocked is unaccounted for.
  - Non-Local Theories: (Dalrymple et al., 1991; Rooth, 1992; Hardt, 1993; Fiengo and May, 1994; Asher et al., 2001) the degraded quality of non-local sloppy readings is unexplained, as is the fact that they become acceptable when strict is blocked.

#### MaxElide

- MaxElide requires maximal ellipsis within a given Parallelism Domain (PD) (Takahashi and Fox, 2005)
- PD is a constituent that can be semantically identified with some antecedent, modulo focused elements
- This captures contrast between (1) and (2)
- In (2) on the sloppy reading, the only PD is [*Bill said she did too. (hit him)*] within this domain, a larger constituent could have been elided.
- On the strict reading of (2) there is a smaller PD: [(hit him)]. Within this PD ellipsis is maximal.
- But MaxElide doesn't capture the fact that non-local sloppy is ok when strict is blocked:
- (3) Nearly EVERY boy said Mary hit him. But BILL didn't say she did. hit Bill
- This violates MaxElide just like (2) does.
- In other cases, MaxElide incorrectly *permits* re-binding:
- (4) Bill<sub>1</sub> believes that Sally<sub>2</sub> will marry him<sub>1</sub>, but everyone knows that she won't. [marry him] (Bach and Partee, 1980)
- With the following variant of (4), the sloppy reading becomes acceptable.
- (5) Everyone<sub>1</sub> HOPES that Sally<sub>2</sub> will marry him<sub>1</sub>, but Bill KNOWS that she will. [marry him]

# 5 Proposal

- As soon as a PD can be identified during a bottom-up derivation, *indexation* and *focus licensing* takes place, and cannot be modified later in the derivation.
- A constituent E is a PD if there is an antecedent A such that there is a *valid indexing* E' of E such that E' is semantically identical to A, modulo focus marked constituents.
- A valid indexing must respect Familiarity (Heim 82): a pronoun must have an index *i*, such that  $i \in Dom(F)$ , where *F* is the File representing the current state of the discourse. (This is equivalent to the *accessibility* requirement in the DRT literature (Kamp and Reyle 93)).

#### Good and Bad Re-Binding.

Bad Re-binding:

- (6) John<sub>1</sub> said Mary<sub>2</sub> hit him<sub>1</sub>. Bill said she did. [hit him]
- The smallest potential PD is [did hit him]
- This is indeed a PD, with valid indexing [did  $hit him_1$ ]
- The strict reading results, and re-binding is not possible.

#### Good Re-binding:

- (7) Nearly EVERY boy<sub>1</sub> said Mary<sub>2</sub> hit him<sub>1</sub>. But BILL didn't say she did.[hit him]
- Again, begin with [did hit him].
- This fails to be a PD, since Familiarity rules out the indexing [did  $hit him_1$ ]
- Derivation continues until we construct [BILL  $\lambda$  x.x didn't say she did *hit him*]
- This constituent is a PD, with the pronoun him bound by the lambda operator
- We correctly derive re-binding in this case
- In a similar way, the proposed account captures all the facts presented above.

## 6 Questions and Prospects

## What is the Status of "Bad" Re-binding?

- Survey suggests that Sag was right, that there is a difference between (1) and (2)
- But: (2) receives an intermediate status, not unacceptable.
- Consider
  - First Mary was told by SUSAN that I ought to tutor her, and then JUDY did. (her=Susan, did=tell Mary that I ought to tutor Judy)
  - Rating: 2.69 (3 is "Fully Unacceptable")
  - First Mary was told by SUSAN that I ought to tutor her, and then JUDY did. (her=Susan, did=tell Mary that I ought to tutor Susan)
  - Rating: 2.72

### How is Indexing Decision Made?

- Result of Survey: If strict indexing is *structurally* ruled out, sloppy becomes fully acceptable
- Things are actually more complicated: there is an alternative "strict-like" reading:
  - (8) Nearly every student felt I ought to tutor her. But Susan didn't feel I should *(tutor them)*
- This *E-type* reading could in principle block sloppy; but it doesn't. Perhaps such E-type readings are more difficult in general.

- Could strict reading be ruled out based on pragmatic factors? Yes.
- This is suggested by data from "admit" examples below A admits P presupposes that P is incriminating for A. In the "pragmatically good" examples, the presupposition is satisfied, while in the "pragmatically bad" examples, it must somehow be accomodated.
- local sloppy, pragmatically good: .29
- (9) MARY might admit that the criminals had been in contact with her, but SUSAN wouldn't (admit that the criminals had been in contact with Susan).
- non-local sloppy, pragmatically good: .59
- (10) MARY might admit that the criminals had been in contact with her, but SUSAN wouldn't admit that they had (been in contact with Susan).
- strict, pragmatically bad: 1.21
- (11) MARY might admit that the criminals had been in contact with her, but SUSAN wouldn't admit that they had (been in contact with Susan).
- strict, pragmatically bad: 1.24
- (12) MARY might admit that the criminals had been in contact with her, but SUSAN wouldn't (admit that the criminals had been in contact with Mary).

## References

- [Asher et al., 2001] Nicholas Asher, Daniel Hardt, and Joan Busquets. Discourse parallelism, ellipsis, and ambiguity. Journal of Semantics, 18(1), 2001.
- [Bach and Partee, 1980] E. Bach and B. Partee. Anaphora and semantic structure. In A. Ojeda, editor, *Papers from the parasession on pronouns and anaphora*, pages 1–28. CLS, Chicago, 1980.
- [Dalrymple *et al.*, 1991] Mary Dalrymple, Stuart Shieber, and Fernando Pereira. Ellipsis and higher-order unification. *Linguistics and Philosophy*, 14(4), August 1991.
- [Fiengo and May, 1994] Robert Fiengo and Robert May. Indices and Identity. MIT Press, Cambridge, MA, 1994.
- [Hardt, 1993] Daniel Hardt. Verb Phrase Ellipsis: Form, Meaning, and Processing. PhD thesis, University of Pennsylvania, 1993.
- [Rooth, 1992] Mats Rooth. Ellipsis redundancy and reduction redundancy. In Proceedings of the Stuttgart Ellipsis Workshop, Stuttgart, Germany, 1992.
- [Sag, 1976] Ivan A. Sag. Deletion and Logical Form. PhD thesis, Massachusetts Institute of Technology, 1976. (Published 1980 by Garland Publishing, New York).
- [Takahashi and Fox, 2005] Shoichi Takahashi and Danny Fox. Maxelide and the re-binding problem. In Proceedings of the Fifteenth Conference on Semantics and Linguistic Theory, Los Angeles, CA, 2005.
- [Williams, 1977] Edwin Williams. Discourse and logical form. Linguistic Inquiry, 8(1):101–139, 1977.