

The Long and the Short of Names and Pops in Discourse

Nicholas ASHER

CNRS, laboratoire IRIT

(in collaboration with P. Denis, J. Hunter and B. Reese of UT Austin)

May 22, 2006

Overview

- Many different types of referential devices: ellipsis, pronouns, definite descriptions, demonstrative NPs, noun phrases
- What, if any, are the differing conditions of use of these expressions?
- Ellipsis, pronouns and definite descriptions differ markedly in their behavior to retrieve discourse unavailable antecedents (not on the right frontier of the discourse structure's graph).
- What about names?
- Theoretical repercussions

Names in actually occurring texts

- The MUC6 and MUC7 corpora offer many examples of anaphoric chains where we have uses of names of various lengths alternating with uses of pronouns and definite descriptions.
- More referential uses of names and definites in the texts than of pronouns.
- Full names, names with titles like *President Bush* and then short names like *Bush*.

An example of names in a referential chain

- (1) a. 0 :: Chrysler to Create Grievance Process For Non-Union Staff
- b. 1 :: Chrysler Corp. this year will take the highly unusual step of setting up a formal grievance procedure for its non-union white-collar employees.
- c. 2 :: "We understand that (Chrysler's) dramatic restructuring . . . will inevitably lead many employees to question management's decisions and actions,"

Previous Hypotheses

- Reinhardt: the use of a proper name to refer back to some already introduced discourse entity previously referred to by a pronoun marks a shift in the discourse topic.
- The hierarchy of referential expressions (Prince, Ariel, Gundel et al.):
zero pronouns >> pronouns >> definite descriptions >> demonstrative NPs >> short PNs >> full PNs >> full PNs + appositive.

Expressions on the left of the referential hierarchy require more salient antecedents than those on the right.

- Does the use of different referential expressions correlate with attachment facts of new information to existing discourse structure?

Our Approach

- As part of an NSF project on anaphora and discourse structure we are annotating the MUC6, MUC7 and ACE2 corpora with discourse structures. 2 annotators independently produce SDRSs for each story and then compare. We have done about 50 so far.
- MUC texts are annotated with anaphoric chains involving indefinites, proper names, definites, pronouns. (No bridging uses and no event or abstract entity reference here, no ellipsis either.)
- We've checked 15 stories concerning the behavior of definites and proper names and 10 stories for correlations between pops and shifts along the referential heirarchy.

Hypotheses revisited

- Hypothesis 1: a shift “rightward” in the RH corresponds to a discourse pop
- Hypothesis 2: discourse pops are in general signalled by shifts in the RH.
- Hypothesis 3: sizes of jumps in the RH correspond to sizes of discourse pops.

An example of the correlation

- (2) a. 23 :: That view was supported in a tender-offer document filed by USAir with the Securities and Exchange Commission and in another document filed in court...
- b. 47 :: “He doesn’t give up easily
- c. 48 :: and one should never underestimate what he can or will do.”
- d. 49 :: Mr. Icahn couldn’t be reached for comment.

Annotators agreed that EDU 49 attached to EDU 23 with Continuation, although Icahn was mentioned via pronouns in EDUS 47 & 48. The rightward shift in the RH here coincides with a large discourse pop.

What we found

- No significant correlation between the size of jump in the Referential Hierarchy and the size of the discourse pop in the discourse structure.
- No support that pop is always signalled by a shift rightward (*jump*) in the RH.
- Sometimes pops occur with leftward shifts in the RH (*plunge*)
- Correlations between discourse pops and referential shifts

POP & JUMP	POP & NO JUMP	POP & NO REF CH	POP & PLUNGE
17	17	29	13

Cells 2 and 3 are about pops that are triggered by content or structural factors of the discourse and involve either no right shift in the RH or no anaphoric chain continued over the pop.

The fourth cell

- (3)
- a. 30 :: But Chrysler Canada’s chief negotiator, William Fisher, said yesterday
 - b. 31 :: that the two sides had reached “some understandings” on economic issues, including pensions.
 - c. 32 :: “We have a good chance of wrapping something up by the end of the week,”
 - d. 33 :: he said.

EDU 33 is connected in the discourse structure to EDU 30, but the anaphor is to the left on the Reference Hierarchy of its antecedent. All of the items in cell 4 are of this kind. An artefact perhaps of segmentation conventions (also adopted by Marcu et. al.)

Interesting Counterexample

- (4)
- a. 64 :: But sources close to USAir said
 - b. 65 :: they didn't believe there would be any problem obtaining approval for the merger
 - c. 66 :: and pointed out
 - d. 67 :: that the company plans to complete its tender.
 - e. 68 :: Sources close to USAir and Piedmont conceded
 - f. 69 :: that if TWA is permitted to buy more USAir shares,
 - g. 70 :: it certainly has enough cash, an estimated \$900 million, to do so.

In (4e), *USAir* coreferring to *the company* in (4d) marks a jump in the RH, but (4e) *backgrounds* or *comments* on (4d) where a coreferential pronoun is used.

Summary

- Support for hypothesis 1 at least for jumps using names. (4) is the only counterexample to hypothesis 1 we found with names.
- No support for hypotheses 2 and 3

Definites vs. Short Names

In 15 stories surveyed with a total of 35 discourse pops,

- we found only 5 pops with definite descriptions,
- 2 with indirect discourse contexts with pronouns (e.g., 3)
- and 25 with short proper names and 3 with long proper names.

An extended example

The definite in EDU 5 is a plunge but in EDU 12 the definite actually marks a jump.

- (5) 0. :: Chrysler to Create Grievance Process For Non-Union Staff
1. :: Chrysler Corp. this year will take the highly unusual step of setting up a formal grievance procedure
 2. :: "We understand that (Chrysler's) dramatic restructuring . . .
 3. :: Anthony P. St. John, vice president for human resources of the auto maker's..., said in a speech ...yesterday.
 4. :: "Employees may not feel that they are being treated fairly."
 5. :: To "strengthen management's credibility," Mr. St. John added, the company will set up an "alternative disputes resolution" system ...
 10. :: Chrysler's move apparently was inspired by the turmoil created by its current cost-cutting and restructuring,
 12. :: The auto maker is significantly changing wage and benefit plans,
 15. :: Chrysler has about 31,000 white-collar employees in the U.S.,

Discourse behavior and semantics

Hypothesis: the discourse pop behavior of names has to do with the treatment of presuppositions

- Different presupposition triggers may provide different presupposition strategies—e.g., *too*:

(6) John lives in New York too.

- Some triggers, e.g. indexicals, prefer their presuppositions to be bound in the widest possible context, but may shift (Hunter and Asher, 2005)

We will encode the preferential behavior of binding for short names—binding to a particular location.

A Sketch of Hunter & Asher 2005

Treat accessible DRSs as a sequence of contexts, K_0, \dots, K_n (where K_0 is the global context and K_n is the local context in which the presuppositional material is introduced).

- \uparrow is an operator introduced by indexicals with scope over the rest of their presupposed content.
- \models is a notion of presupposition resolvability.
- $K_j \triangleright \phi$ means that the presupposition ϕ is incorporated into K_j , subject to some constraints.

Binding and Accomodation

Our definition simplifies binding to \models

- (a) $K_0, \dots, K_{l-1}, [K_l, \dots, K_n] \models \phi$ iff ϕ is a dynamic consequence of K_l, \dots, K_n relative to any assignment to free variables occurring in K_l, \dots, K_n, ϕ that are declared in K_0, \dots, K_{l-1} and satisfy (some of) the conditions in K_0, \dots, K_{l-1} .
- (b) $K_0, \dots, K_i \Vdash \phi$ iff
 $\exists j \leq i, \exists l \geq 0$ such that $K_0, \dots, K_{l-1}, [K_l, \dots, K_j] \models \phi$ or,
for some $k, 0 \leq k \leq j, K_k \triangleright \phi$, for ϕ a normal DRS or DRS condition.

Clause (b) says that a presupposition is resolvable in a sequence of contexts just in case some subsequence entails the presupposition or it is accommodated at some element in the sequence.

Applications in Semantics

- Presupposition triggers like indexicals introduce a presupposition of the form $\uparrow \phi$:
 $K_0, \dots, K_i \Vdash \uparrow \phi$ iff $\exists j \leq i$ such that $K_0, \dots, K_j \Vdash \phi$
and there is no $k < j$ such that $K_0, \dots, K_k \Vdash \phi$.

- Presupposition triggers like *too* introduce a presupposition of the form $B\phi$:
 - $K_0, \dots, K_i \Vdash B\phi$ iff $\exists j \leq i$ and $\exists l \geq 0$ such that $K_0, \dots, K_{l-1}, [K_l, \dots, K_j] \models \phi$

That is, such a presupposition will be resolvable at a context iff the material under its scope can be bound.

- extensions to proper names, complex demonstratives and definite descriptions,

The discourse behavior of presuppositions

Complicate the language of presupposition operators to allow for discourse attachments while preserving the semantic behavior of presuppositions:

- E.g., introduce dynamic sequence ; to first handle semantic instructions, e.g, \uparrow, B and then discourse instructions.
- instructions involving presupposition attachment preferences given by triggers: B^d (attach to a constituent via a discourse relation that captures binding—e.g., *Defeasible-Consequence*).
- more complex instructions like Van der Sandt's proposal (V^d) can also be encoded.

For short names:

- Short names prefer to bind above the previous bit of information attached (defn. of a pop).
- So index the discourse binding operator so that it prefers to bind above *LAST* (the node given by the last bit of information attached in the discourse structure:

$$(7) \quad (\uparrow^s; B_{>LAST}^d, \text{else } V^d)\phi$$

Concluding observations

- In principle, since presuppositions are separate constituents, they could attach to nodes in the discourse structure independently of the preferred content associated with them from the composition process.
- But this doesn't happen. Presupposed constituents given by names, definites, and pronouns and associated preferred constituents attach at the same place (modulo logical operators—not a problem in our corpus).
SDRT's glue logic can axiomatize a preference for this joint attachment
- By figuring out the attachment conditions on presuppositions, we give more content to the theory of discourse attachment