Coherence, Reference, and the Theory of Grammar

Andrew Kehler
SRI International*

1 Coherence Establishment

The meaning of a discourse is greater than the sum of the meanings of its parts ⇒ additional meaning comes from establishing *coherence*.

(1) a. Bill is a politician. He's bound to act irresponsibly. (Result)

b. ? Bill is a politician. He's bound to eat at McDonalds a lot.

(2) a. Bill likes to play golf. Al enjoys surfing the net. (Parallel)

b. ? Bill likes to play golf. George hates broccoli.

(3) a. Ken Starr convened his grand jury this morning. Vernon Jordan was subsequently called to testify. (Narration)

b. ? Ken Starr convened his grand jury this morning. Vernon Jordan subsequently scheduled a press conference.

- Establishing coherence is an integral part of language understanding, yet is rarely acknowledged in theories of various intersentential linguistic phenomena.
- One instead finds:
  - an implicit assumption that the resolution process is uniform and can be explained from within a single module of language processing
  - a set of data that appears to defy this assumption

*Artificial Intelligence Center, 333 Ravenswood Avenue, Menlo Park, CA 94025, kehler@ai.sri.com. Presented at the 7th Annual CSLI Workshop on Logic, Language, and Computation, Stanford, May 31, 1998.

- In this talk, we:
  - argue that the data pattern with an independently motivated categorization of coherence relations
  - sketch an account in which the properties of the linguistic phenomenon interact with the process of establishing coherence relations to explain the data

2 Hume's Three Types of Coherence

Hume [1748]:

> "Though it be too obvious to escape observation that different ideas are connected together, I do not find that any philosopher has attempted to enumerate or class all the principles of association—a subject, however, that seems worthy of curiosity. To me there appear to be only three principles of connection among ideas, namely *Resemblance*, *Contiguity* in time or place, and *Cause* or *Effect*."

Categorize previously proposed relations [Hobbs, 1979; Hobbs, 1990] into this trichotomy. Consider the coherence resolution procedure as a two-step process:

1. *Identify* and *retrieve* the arguments to the coherence relation, and
2. *Apply* the constraints of the relation to those arguments

2.1 Resemblance

Parallel: Infer $p(a_1, a_2, ...)$ from the assertion of S_0 and $p(b_1, b_2, ...)$ from the assertion of S_1, where for some property vector q, $q_i(a_i)$ and $q_i(b_i)$

(4) Bill likes to play golf. Al enjoys surfing the net.

2.2 Cause-Effect

Result: Infer P from the assertion of S_0 and Q from the assertion of S_1, where normally $P \rightarrow Q$.

(5) Bill is a politician. He's bound to act irresponsibly.

Relation	Constraints	Conjunctions
Parallel	$p_0 = p_1$, $q_i(a_i)$ and $q_i(b_i)$	and
Contrast	(1) $p_0 = \neg p_1$, $q_i(a_i)$ and $q_i(b_i)$ (2) $p_0 = p_1$, $q_i(a_i)$ and $\neg q_i(b_i)$	but
Exemplification	$p_0 = p_1$; $b_i \in a_i$ or $b_i \subset a_i$	for example
Generalization	$p_0 = p_1$; $a_i \in b_i$ or $a_i \subset b_i$	in general
Elaboration	$p_0 = p_1$, $a_i = b_i$	in other words

Table 1: Resemblance Relations

Relation	Presuppose	Conjunctions
Result	$P \rightarrow Q$	and (as a result) therefore
Explanation	$Q \rightarrow P$	because
Violated Expectation	$P \rightarrow \neg Q$	but and still
Denial of Preventer	$Q \rightarrow \neg P$	even though despite

Table 2: Cause-Effect Relations

2.3 Contiguity

Narration: Infer a change of state for a system of entities from S_1, inferring the initial state for this system from S_0.

(6) Ken Starr convened his grand jury this morning. Vernon Jordan was subsequently called to testify.

3 VP Ellipsis

(7) Bill plays golf, and George does too.

- Much debate in the literature about whether resolution reconstructs a syntactic or a semantic representation in the target

Kehler [1993; 1994a; 1995]:

- Resemblance Relations: *Syntactic* representation reconstructed
- Cause-Effect Relations: *Semantic* representation reconstructed

3.1 Passive/Active Voice Mismatches

Cause-Effect:

(8) In March, four fireworks manufacturers asked that the decision be reversed, and on Monday the ICC did. [reverse the decision] (cited in Dalrymple [1991])

(9) This problem was to have been looked into, but obviously nobody did. [look into the problem] (Vincent Della Pietra, in conversation)

(10) Actually I have implemented it with a manager, but it doesn't have to be. [implemented with a manager] (Steven Ketchpel, in conversation)

Resemblance:

(11) # This problem was looked into by John, and Bob did too.

(12) # John implemented the indexing program with a manager, and the retrieval program was too.

3.2 Non-VP Antecedents

Cause-Effect:

(13) This letter deserves a response, but before you do, ... [respond] (Gregory Ward, p.c.)

(14) First person pronouns aren't very shiftable, although the plural ones can be. [shifted] (Barbara Partee, in conversation)

Resemblance:

(15) # This letter provoked a response from Bush, and Clinton did too.

(16) # First person pronouns aren't very shiftable, and the plural ones also don't.

3.3 Binding Theory Constraints

3.3.1 Condition A Violations

Resemblance:

(17) ?? John$_i$ defended himself$_i$, and Bob$_j$ did too. [defended John$_i$]

(18) ?? Fred$_i$ voted for himself$_i$, and Gary$_j$ did too. [voted for Fred$_i$]

Cause-Effect:

(19) Bill$_i$ defended himself$_i$ against the accusations because his lawyer$_j$ wouldn't.

(20) John$_i$ voted for himself$_i$ even though no one else$_j$ did.

3.3.2 Condition B Violations

Resemblance:

(21) # John's$_j$ mother introduced him$_j$ to everyone, and he$_j$ did too. [introduced himself$_j$]

(22) # John$_i$'s lawyer defended him$_i$, and he$_i$ did too. [defended himself$_i$]

Cause-Effect:

(23) John's$_j$ mother introduced him$_j$ to everyone because he$_j$ wouldn't.

(24) John$_i$'s lawyer defended him$_i$ because he$_i$ wouldn't.

3.3.3 Condition C Violations

Resemblance:

(25) # I hit Bill$_i$, and he$_i$ did too. [hit Bill$_i$] (Lappin [1993])

(26) # Mary introduced John$_j$ to everyone, and he$_j$ did too. [introduce John$_j$ to everyone]

Cause-Effect:

(27) I expected Bill$_i$ to win even when he$_i$ didn't. (Dalrymple [1991])

(28) The lawyer defended Bill$_j$ against the accusations because he$_j$ wouldn't.

3.4 Subjacency Violations

Antecedent Contained Deletion:

(29) John read everything which Bill believes he did.

Resemblance:

(30) # John read everything which Bill believes the claim that he did. [Haïk, 1987]

(31) # John read everything which Bill wonders why he did. [Haïk, 1987]

Cause-Effect:

(32) Which problem did you think John would solve because of the fact that Susan did? [Rooth, 1981]

Compare (32) with unelided form containing gap:

(33) * Which problem did you think John would solve because of the fact that Susan solved?

Solution [Kehler, 1994a; Kehler, 1995]:

- Can only elide when all subsequently needed info is recoverable:
  - Semantics of VP
  - Arguments to the coherence relation
- VP-ellipsis satisfies the tests for referentiality – it can refer more than one clause back, be cataphoric, etc. – and thus ellipsis succeeds given a semantically suitable antecedent
- Resemblance relations utilize syntax to align parallel, subsentential semantic arguments. Eliding a VP indicates that these parallel elements are shared and recoverable by reconstruction.
- Cause-Effect relations access sentential level antecedents, so no reconstruction is necessary to recover arguments

4 Gapping

(34) Bill ate the Big Mac, and Al the fries.

Gapping only felicitous with 'symmetric' (i.e., Resemblance) *and*, and not 'asymmetric' (i.e., Cause-Effect) *and* [Levin and Prince, 1982]:

(35) a. Sue became upset and Nan became downright angry.

b. Sue became upset and Nan $\emptyset$ downright angry.

(36) Sue and Nan had worked long and hard for Carter. When Reagan was declared the winner, Sue became upset and Nan became/$\emptyset$ downright angry. (Parallel $\in$ *Resemblance*)

(37) Susan's histrionics in public have always gotten on Nan's nerves, but it's getting worse. Yesterday, when she couldn't have her daily Egg McMuffin because they were all out, Sue became upset and Nan became/#$\emptyset$ downright angry. (Result $\in$ *Cause-Effect*)

Correlates with Resemblance vs. Cause-Effect [Kehler, 1994b; Kehler, 1995]:

(38) a. John will vote for Clinton or Tom will vote for Dole.

b. John will vote for Clinton or (alternatively) Tom $\emptyset$ for Dole.

c. # John will vote for Clinton or (else!) Tom $\emptyset$ for Dole.

(39) a. John voted for Clinton but Tom voted for Dole.

b. John voted for Clinton but (in contrast) Tom $\emptyset$ for Bush

c. # John voted for Clinton but (despite this fact) Tom $\emptyset$ for Bush.

(40) # John voted for Clinton $\left\{\begin{array}{c}\text{because}\\ \text{even though}\\ \text{despite the fact that}\\ \text{although}\end{array}\right\}$ Tom $\emptyset$ for Bush.

Solution:

- Unlike VP-ellipsis, gapping does not satisfy the tests for referentiality – it cannot refer more than one clause back, cannot be cataphoric, etc. – and thus gapping can only succeed if something causes syntactic reconstruction
- Resemblance relations force such reconstruction, Cause-Effect relations do not

5 Extraction from Multiple Clauses

The Coordinate Structure Constraint [Ross, 1967]:

> In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

The *Element* constraint:

(41) a. * What book did John buy and read the magazine? (Parallel $\in$ *Resemblance*)

b. What book did John buy and read?

Counterexamples to the CSC [Lakoff, 1986]:

(42) a. How much can you drink and still stay sober? [Goldsmith, 1985] (Violated Expectation $\in$ *Cause-Effect*)

b. That's the stuff that the guys in the Caucasus drink and live to be a hundred. [Lakoff [1986], attributed to Farley] (Result $\in$ *Cause-Effect*)

(43) a. What did Harry buy, come home, and devour in thirty seconds? [Adapted from Ross [1967]] (Narration $\in$ *Contiguity*)

b. * What did Harry buy and come home?

Solution [Kehler, 1996]:

> *Topichood Condition for Extraction:* Only those constituents in a sentence that qualify as the topic of the sentence can undergo extraction processes (i.e., *WH*-Q Movement, *Wh*-Relative Movement, Topicalization, and *It*-Clefting). [Kuno, 1987, page 23]

- The topic of a Resemblance construction must be common to each of the conjuncts at some level of generality (i.e., a *Common Topic* [Lakoff, 1971]). Thus, extraction is only possible when the clauses share an element and extraction is across-the-board.
- A topic of a Cause-Effect passage may appear in only one conjunct – no extraction constraints.

- The topic of Contiguity need not appear in every clause (e.g., scene-setting clauses), but must remain the topic throughout the passage (e.g., compare (43a) with (43b)). Thus, extraction will generally be required out of the final clause, but not all clauses.
- Conclusion: There is no purely syntactic Coordinate Structure Constraint. Unacceptable cases of extraction out of coordinate structures result from constraints at the syntax/discourse interface, arising from the interaction between topichood and coherence relations.

6 Pronoun Interpretation

Hobbs (1979): reference resolution is a by-product of coherence resolution

(44) The city council denied the demonstrators a permit because

a. they *feared* violence. [they=the city council]
(Explanation $\in$ *Cause-Effect*)

b. they *advocated* violence. [they=the demonstrators]
(Explanation $\in$ *Cause-Effect*)

Focus- and Centering-based approaches: model discourse state (including salience) and heuristics for resolving referential expressions against it

(45) a. Terry really goofs sometimes.

b. Yesterday was a beautiful day and he was excited about trying out his new sailboat. [He=Terry]

c. He wanted Tony to join him on a sailing expedition. [He=Terry]

d. He called him at 6AM. [He=Terry]

e. He was sick and furious at being woken up so early. [He=Tony (!)]

Effects of parallelism:

(46) Margaret Thatcher admires Hillary Clinton, and Newt Gingrich absolutely worships her. [her=Hillary Clinton (!)]

Kameyama [1986]: property sharing constraint encodes preference for grammatical role parallelism.

Kehler [1997]:

(47) a. The three candidates had a debate today.

b. Bob Dole began by bashing Bill Clinton.

c. He criticized him on his opposition to tobacco.

d. Then Ross Perot reminded him that most Americans are also anti-tobacco. [him=Dole, Narration]

d'. Then Ross Perot slammed him on his tax policies.
[him=Clinton, Parallel]

Competition models: correlations between preference and coherence are lost

Solution [Kehler, 1995]:

- In Resemblance, pronoun is aligned with its parallel element during coherence establishment
  - preference for grammatical role parallelism is likely epiphenomenal
  - does not necessarily result in most semantically plausible referent
- In Cause-Effect, pronoun is aligned with antecedent which allows relation to be established (although salience still plays a role)
- In Contiguity, pronoun is aligned with the passage topic, typically expressed in a salient position (e.g., subject)

7 Conclusions

- We have discussed four areas in which there is an implicit assumption that the resolution process is uniform and can be explained from within a single module of language processing, despite a set of data that appears to defy this assumption
- Many past works have either punted or modified otherwise elegant theories in unintuitive and inelegant ways
- We argued that the data in each case pattern with an independently motivated categorization of coherence relations, and once this is accounted for, simpler and more intuitive theories can be articulated

- For another example, tense interpretation, see Kehler [1994c]
- Note that we have not added new machinery to address these problems: This machinery already exists!

References

[Dalrymple, 1991] Mary Dalrymple. Against reconstruction in ellipsis. Technical Report SSL-91-114, Xerox, 1991.

[Goldsmith, 1985] John Goldsmith. A principled exception to the coordinate structure constraint. In *Papers from the Twenty-First Annual Regional Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, Chicago, 1985.

[Haïk, 1987] Isabelle Haïk. Bound variables that need to be. *Linguistics and Philosophy*, 10:503–530, 1987.

[Hobbs, 1979] Jerry Hobbs. Coherence and coreference. *Cognitive Science*, 3:67–90, 1979.

[Hobbs, 1990] Jerry Hobbs. *Literature and Cognition.* CSLI Lecture Notes 21, 1990.

[Hume, 1748] David Hume. *An Inquiry Concerning Human Understanding.* The Liberal Arts Press, New York, 1955 edition, 1748.

[Kameyama, 1986] Megumi Kameyama. A property-sharing constraint in centering. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, pages 200–206, New York, 1986.

[Kehler, 1993] Andrew Kehler. The effect of establishing coherence in ellipsis and anaphora resolution. In *Proceedings of the 31st Conference of the Association for Computational Linguistics (ACL-93)*, pages 62–69, Columbus, Ohio, June 1993.

[Kehler, 1994a] Andrew Kehler. Common topics and coherent situations: Interpreting ellipsis in the context of discourse inference. In *Proceedings of the 32nd Conference of the Association for Computational Linguistics (ACL-94)*, Las Cruces, New Mexico, June 1994.

[Kehler, 1994b] Andrew Kehler. A discourse processing account of gapping and causal implicature. Manuscript presented at the Annual Meeting of the Linguistic Society of America, Boston, January 1994.

[Kehler, 1994c] Andrew Kehler. Temporal relations: Reference or discourse coherence? In *Proceedings of the 32nd Conference of the Association for Computational Linguistics (Student Session)*, Las Cruces, New Mexico, June 1994.

[Kehler, 1995] Andrew Kehler. *Interpreting Cohesive Forms in the Context of Discourse Inference.* PhD thesis, Harvard University, 1995.

[Kehler, 1996] Andrew Kehler. Coherence and the coordinate structure constraint. In *Proceedings of the 22nd Annual Meeting of the Berkeley Linguistics Society (BLS-22)*, Berkeley, CA, February 1996.

[Kehler, 1997] Andrew Kehler. Current theories of centering for pronoun interpretation: A critical evaluation. *Computational Linguistics*, 23(3):467–475, 1997.

[Kuno, 1987] Susumu Kuno. *Functional Syntax - Anaphora, Discourse and Empathy.* The University of Chicago Press, Chicago and London, 1987.

[Lakoff, 1971] Robin Lakoff. If's, and's, and but's about conjunction. In Charles J. Fillmore and D. Terence Langendoen, editors, *Studies in Linguistic Semantics*, pages 114–149. Holt, Rinehart, and Winston, New York, 1971.

[Lakoff, 1986] George Lakoff. Frame semantic control of the coordinate structure constraint. In *Papers from the Twenty-Second Annual Regional Meeting of the Chicago Linguistic Society*, Chicago, 1986. Chicago Linguistic Society.

[Lappin, 1993] Shalom Lappin. The syntactic basis of ellipsis resolution. In S. Berman and A. Hestvik, editors, *Proceedings of the Stuttgart Workshop on Ellipsis.* University of Stuttgart, University of Tuebingen, and IBM Germany, 1993.

[Levin and Prince, 1982] Nancy Levin and Ellen Prince. Gapping and causal implicature. Presented at the Annual Meeting of the Linguistic Society of America, 1982.

[Rooth, 1981] Mats Rooth. A comparison of three theories of verb phrase ellipsis. University of Massachusetts Occasional Papers in Linguistics, Volume 7, 1981.

[Ross, 1967] John R. Ross. *Constraints on variables in syntax.* PhD thesis, MIT, 1967.