Blackwell
Reference Online

11. Discourse Coherence

ANDREW KEHLER

Subject	Theoretical Linguistics » Pragmatics
Key-Topics	discourse
DOI:	10.1111/b.9780631225485.2005.00013.x

1 Introduction

While there are many aspects of discourse understanding that are poorly understood, there is one thing that we can be sure of: Discourses are not simply arbitrary collections of utterances. A felicitous discourse must instead meet a rather strong criterion, that of being COHERENT.[1] Passages (1a–b) provide evidence for such a coherence constraint.

(1) a. George W. Bush wanted to satisfy the right wing of his party. He introduced an initiative to allow government funding for faith-based charitable organizations.

   b. ?George W. Bush wanted to satisfy the right wing of his party. He smirked a lot.

Hearers do not generally interpret the two statements in passage (1a) as independent facts about Bush; they identify a causal relationship between the two that I, following Hobbs (1990), will call RESULT. The inference of Result requires that a presupposition be satisfied, specifically that government funding for faith-based charities is something that the right wing of Bush's party wants. Although this relationship is not actually asserted anywhere in passage (1a), a hearer would be well within his rights to question it if it did not accord with his beliefs about the world, say, with a response of the sort shown in (2).

(2) Actually, many on the right are against the initiative because they worry that government interference will affect the independence of religious organizations.

While passage (1a) does not explicitly contradict this statement, the inferences required to establish its coherence do, hence the felicity of the response.

The coherence of passage (1a) contrasts with the more marginal coherence of passage (1b). In this case, a hearer may attempt to establish a similar Result relationship, but it is less obvious how he could accommodate the presupposition that smirking a lot would please the Republican right into his beliefs about the world. Of course, he might nonetheless attempt to construct an explanation that would make passage (1b) coherent. For example, he might reason that smirking is a sign of confidence about winning the election – a form of rubbing it in to the previous Democratic administration – and that the right wing of the Republican party would appreciate such an outward show of confidence. The fact that hearers are driven to try to identify such explanations is itself evidence that coherence establishment is an inescapable component of discourse interpretation.

Of course, Result is not the only type of relation that can connect propositions in coherent discourse.

Passage (3a) is coherent by virtue of a PARALLEL relation, licensed by the fact that similar properties are attributed to parallel entities Dick and George.

>(3) a. Dick is worried about defense spending. George is concerned with education policy.
>b. ?Dick is worried about defense spending. George smirks a lot.

In contrast, passage (3b) seems to be less coherent due to the lack of a similar degree of parallelism. However, in a context that makes it clear that *Dick* refers to Vice President Dick Cheney and *George* refers to President George W. Bush, then the passage might become more coherent under the common topic of (roughly) *what high government officers are doing*. In fact, with this interpretation passage (3b) comes across as a joke at the expense of Bush, since the identification of parallelism between the clauses highlights the contrast between the importance and positive contribution of the activities attributed to the two men. A third type of relation that can connect clauses in a coherent discourse has been called OCCASION, exemplified in passage (4a).

>(4) a. George delivered his tax plan to Congress. The Senate scheduled a debate for next week.
>b. ?George delivered his tax plan to Congress. The Senate scheduled hearings into former President Clinton's pardon of Marc Rich.

Occasion allows one to describe a complex situation in a multi-utterance discourse by using intermediate states of affairs as points of connection between partial descriptions of that situation. As with the other examples discussed thus far, a hearer will normally make certain inferences upon interpreting passage (4a), for instance, that the scheduled Senate debate will center around George's tax proposal. On the other hand, it is harder to determine how the event described by the second sentence of (4b) can be seen as a natural follow-up to the event described by the first, which results in a less coherent passage under this relation. However, if the hearer already knew there to be an external set of factors that required the Senate to deal with the Marc Rich pardon before the tax proposal, then the passage would become more coherent. In this case, the second sentence can be interpreted as a precursor to debating the tax plan, placing the two events in a connected sequence.

In sum, what passages (1a, b), (3a, b), and (4a, b) all have in common is that they each contain two clauses which are independently well-formed and readily understood. The coherence of the (a) passages and relative incoherence of the (b) passages show that interpretation continues beyond this, however, as the hearer is further inclined to assume unstated information necessary to analyze the passage as coherent. These facts demonstrate that the need to establish coherence is a central facet of discourse understanding: Just as hearers attempt to recover the implicit syntactic structure of a string of words to compute sentence meaning, they attempt to recover the implicit coherence structure of a series of utterances to compute discourse meaning.

2 Perspectives on Coherence

In many respects, discourse coherence remains a relatively understudied area of language interpretation. This notwithstanding, it has received some degree of attention within several largely separate strands of research in the language sciences, a sample of which I briefly discuss here.

2.1 Theoretical linguistics perspectives

Theoretical linguists approaching coherence from a variety of perspectives have sought to categorize the different types of coherence relations that can serve to connect clauses, and in fact many of the resulting classifications bear a strong similarity to one another. Halliday and Hasan (1976), for instance, classify relations into four main categories: ADDITIVE, TEMPORAL, CAUSAL, and ADVERSATIVE. Longacre (1983) also distinguishes four categories, CONJOINING, TEMPORAL, IMPLICATION, and ALTERNATION, as does Martin (1992), in his case ADDITION, TEMPORAL, CONSEQUENTIAL, and COMPARISON. The first three categories in each analysis are quite similar, so the main difference lies with respect to the fourth category. Halliday and Hasan's Adversative category separates out relations based on contrast, Longacre's Alternation category distinguishes passages conjoined with *or*, and Martin's Comparison category differentiates comparative constructions. A case could be made that all of these are actually special cases of the Additive/Conjoining category, an idea that I endorse. Indeed, the

agreement among the approaches with respect to the first three categories foreshadows the categorization that I will advocate later in this chapter.

Of course, other sets of relations (and categorizations thereof) have also been proposed, which leads us to the question of how competing proposals should be evaluated and compared.[2] Sanders et al. (1992) propose two criteria: DESCRIPTIVE ADEQUACY, the extent to which a relation set covers the diversity of naturally occurring data, and PSYCHOLOGICAL PLAUSIBILITY, the extent to which the relations are based on cognitively plausible principles. Whereas all proposals are undoubtedly informed by data analysis to some degree, some pursue the goal of descriptive adequacy to a greater extent than others. (One that considers it to be the primary motivating factor is Rhetorical Structure Theory, discussed briefly in the next section.) As pointed out by Knott and Dale (1994), however, without a priori constraints on relation definition one could easily define relations that describe incoherent texts. They suggest, for instance, the possibility of defining an INFORM-ACCIDENT-AND-MENTION-FRUIT relation that would cover example (5):

(5) ?John broke his leg. I like plums.

Thus, an explanatory theory of coherence requires a set of externally driven principles to motivate and ultimately constrain the relation set.

A series of papers by Sanders and colleagues (Sanders et al. 1992, 1993, Sanders 1997) pursues a theory in which psychological plausibility is the primary motivating factor. They analyze relations as composites of more fine-grained features, of which they posit four: BASIC OPERATION (causal or additive), ORDER OF SEGMENTS (basic or non-basic), POLARITY (positive or negative), and SOURCE OF COHERENCE (semantic or pragmatic). By breaking down relations into more primitive features, Sanders et al. take a step toward a more principled and explanatory account of coherence than can be captured by simple lists of relations derived from corpus analysis. Although such an approach will not necessarily offer an exhaustive accounting of all the different coherence relations that researchers have proposed, the resulting set of relations is elegant and economic, and leaves open the possibility that other factors interact with these features to yield a more comprehensive set of distinctions. The more top-down derivational character to Sanders et al.'s analyses has received a more empirically grounded, bottom-up evaluation in several studies by Knott and colleagues (Knott and Dale 1994, Knott and Mellish 1996, Knott and Sanders 1998), which have examined the use and distribution of cue phrases in order to derive hierarchies of relations.

2.2 Computational linguistics perspectives

Computational linguists have also set out to characterize the set of coherence relations that can connect clauses, motivated by the need for computational models of both discourse interpretation and production. From the interpretation side, Hobbs (1979,1990) provides definitions for a set of relations that are rooted in the operations of a computational inference system. In subsequent work, Hobbs et al. (1993) show how a proof procedure based on the unsound inference rule of ABDUCTION can be used to identify coherence in texts. See Hobbs (this volume) for more details on the abductive approach. An alternative proof procedure based on non-monotonic deduction is used for establishing coherence in the DICE system of Asher and Lascarides (Lascarides and Asher 1993, Asher and Lascarides 1994, Asher and Lascarides 1995, Asher and Lascarides 1998a, inter alia); consult those works for further details.

On the discourse production side, analyses of coherence have been used as a basis for the automated generation of coherent text. The Rhetorical Structure Theory (RST) of Mann and Thompson (1986) has been a popular framework for this purpose. RST posits a set of 23 relations that can hold between two adjacent spans of text, termed the NUCLEUS (the more central text span) and SATELLITE (the span containing less central, supportive information).[3] RST relation definitions are made up of five fields: CONSTRAINTS ON NUCLEUS, CONSTRAINTS ON SATELLITE, CONSTRAINTS ON THE COMBINATION OF NUCLEUS AND SATELLITE, THE EFFECT, and LOCUS OF THE EFFECT. While RST is oriented more toward text description than interpretation, it has proven to be useful for developing natural language generation systems, since its relation definitions can be cast as operators in a text planning system that associates speaker intentions with the manner in which they can be achieved. In particular, a high-level communicative goal can be matched against the effect of an RST relation so as to break the problem down into the subgoals necessary to meet the constraints on the nucleus and satellite, which can be iterated until the level is reached at which these constraints can be met by generating single clauses. For further

discussion of using RST for generation and for some of the obstacles such an approach presents, see Hovy (1991, 1993) and Moore and Paris (1993), inter alia. For a discussion of automated parsing of texts in terms of RST relations and its use for discourse summarization, see Marcu (2000).

2.3 Psycholinguistics perspectives

The processes that people use to establish coherence have also been studied from a psycholinguistic perspective; here I briefly mention just a few examples. One line of work has sought to identify which of a potentially infinite number of possible inferences are actually made during interpretation (Garnham 1985, McKoon and Ratcliff 1992, Singer 1994, Garrod and Sanford 1994, inter alia). Inferences are categorized in terms of being NECESSARY to establish coherence versus merely ELABORATIVE, the latter including those suggested by the text but not necessary for establishing coherence. These studies have yielded potentially contradictory results, as they appear to depend to a large degree on the experimental setup and paradigm (Keenan et al. 1990). One of the better known lines of psycholinguistic research into these questions and coherence in general is that of Kintsch and colleagues, who have proposed and analyzed a "construction–integration" model of discourse comprehension (Kintsch and van Dijk 1978, van Dijk and Kintsch 1983, Kintsch 1988, inter alia). They defined the concept of a TEXT MACROSTRUCTURE, which is a hierarchical network of propositions that provides an abstract, semantic description of the global content of the text. Guindon and Kintsch (1984) evaluated whether the elaborative inferences necessary to construct the macrostructure accompany comprehension processes in this framework. Consult these works for further details.

3 A Neo–Humean Analysis of Coherence and Its Application to Linguistic Theory

As the foregoing discussion might suggest, the majority of previous work on coherence relations has operated within the confines of the field of text coherence itself. As it may be tempting to believe that coherence establishment can only occur after all sentence–level interpretation issues have been resolved, theories of coherence rarely play a role in accounts of particular linguistic forms.[4] In this section, I briefly describe several of my own attempts to show that an analysis of coherence is in fact necessary to address outstanding problems in linguistics. I start by presenting my own categorization of a set of coherence relations, and then briefly summarize four linguistic analyses that rely on this categorization as a crucial component. Due to space limitations, I will not attempt to address all of the issues that these brief sketches might raise, and instead refer the reader to Kehler (2002) for more in-depth treatments.

3.1 A neo–Humean classification of coherence relations

In the introduction, I argued for the existence of coherence establishment processes by appealing to three pairs of examples – passages (1a, b), (3a, b), and (4a, b) – which are instances of the coherence relations Result, Parallel, and Occasion, respectively. In Kehler (2002), I argue that these relations can be seen as the canonical instances of three general classes of "connection among ideas," first articulated by David Hume in his *Inquiry Concerning Human Understanding*:

> 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*.

> (Hume 1955: 32 [1748])

In the subsections that follow I analyze a set of coherence relations, many of which are taken and adapted from Hobbs (1990), as belonging to these three general categories.[5] I show that these categories differ systematically in two respects: in the type of arguments over which the coherence relation constraints are applied, and in the central type of inference process underlying this application. Although the details differ in several respects from the Sanders et al. (1992) classification, the two categorizations share the property that the relations are composites of more primitive, cognitively inspired features. The three classes of relations also, at least at a superficial level, show

considerable overlap with the three categories that were common to the classifications of Halliday and Hasan, Longacre, and Martin discussed in section 2.

3.1.1 Cause–Effect relations

Establishing a Cause–Effect relation requires that a path of implication be identified between the propositions denoted by the utterances in a passage. The canonical case of a Cause–Effect relation is Result, which was exemplified in passage (1a).

Result: Infer P from the assertion of S_1 and Q from the assertion of S_2, where normally $P \rightarrow Q$.

The variables S_1 and S_2 represent the first and second sentences being related, respectively.

For example (1a), P corresponds to the meaning of the first clause, Q corresponds to the meaning of the second, and the implication that needs to be established is *if someone wants to satisfy the right wing of the Republican party, then it plausibly follows that that person would introduce an initiative to allow government funding for faith-based charitable organizations*. This constraint gives rise to the corresponding presupposition previously cited for example (1a), as well as the analogous one that is less readily satisfied in example (1b).

The definitions of other coherence relations in this category can be derived by simply reversing the clause order and optionally negating the second proposition in the conditional. All of the following examples require that the same presupposition cited above be met:

Explanation: Infer P from the assertion of S_1 and Q from the assertion of S_2, where normally $Q \rightarrow P$.

> (6) George introduced an initiative to allow government funding for faith-based charitable organizations. He wanted to satisfy the right wing of his party.

Violated expectation: Infer P from the assertion of S_1 and Q from the assertion of S_2, where normally $P \rightarrow \neg Q$.

> (7) George wanted to satisfy the right wing of his party, but he refused to introduce an initiative to allow government funding for faith-based charitable organizations.

Denial of preventer: Infer P from the assertion of S_1 and Q from the assertion of S_2, where normally $Q \rightarrow \neg P$.

> (8) George refused to introduce an initiative to allow government funding for faith-based charitable organizations, even though he wanted to satisfy the right wing of his party.

To sum, to establish a Cause–Effect relation the hearer identifies a path of implication between the propositions P and Q denoted by the utterances.

3.1.2 Resemblance relations

Establishing a Resemblance relation is a fundamentally different process. Resemblance requires that commonalities and contrasts among corresponding sets of parallel relations and entities be recognized, using operations based on comparison, analogy, and generalization. The canonical case of a Resemblance relation is Parallel, which was exemplified in passage (3a).[6]

Parallel: Infer $p(a_1, a_2, ...)$ from the assertion of S_1 and $p(b_1, b_2, ...)$ from the assertion of S_2, where for some vector of sets of properties $\vec{i}$, $q_i(a)$ and $q_i(b)$ for all i.

The phrase "vector of sets of properties" simply means that for each i, there is a set of properties q_i representing the similarities among the corresponding pair of arguments a_i and b_i. In example (3a), the parallel entities a_1 and b_1 are Dick and George, respectively, the parallel entities a_2 and b_2 correspond to defense spending and education policy, and the common relation p is roughly *what high government officers are concerned about*. Note that p is typically a generalization of the parallel relations expressed in the utterances.

Two versions of the CONTRAST relation can be derived by contrasting either the relation inferred or a set of properties of one or more of the sets of parallel entities.

Contrast (i): Infer $p(a_1, a_2,...)$ from the assertion of S_1 and $p(b_1, b_2,...)$ from the assertion of S_2, where for some vector of sets of properties $\vec{i}$, $q_i(a)$ and $q_i(b)$ for all i.

(9) Dick supports a raise in defense spending, but George opposes it.

Contrast (ii): Infer $p(a_1, a_2,...)$ from the assertion of S_1 and $p(b_1, b_2,...)$ from the assertion of S_2, where for some vector of sets of properties $\vec{i}$, $q_i(a)$ and $q_i(b)$ for some i.

(10) Dick supports a raise in defense spending, but George wants a raise in education investment.

The EXEMPLIFICATION relation holds between a general statement followed by an example of it.[7]

Exemplification: Infer $p(a_1, a_2 ...)$ from the assertion of S_1 and $p(b_1, b_2, ...)$ from the assertion of S_2, where b_i is a member or subset of a_i for some i.

(11) Republican presidents often seek to put limits on federal funding of abortion. In his first week of office, George W. Bush signed a ban on contributing money to international agencies which offer abortion as one of their services.

The GENERALIZATION relation is similar to Exemplification, except that the ordering of the clauses is reversed.

Generalization: Infer $p(a_1, a_2,...)$ from the assertion of S_1 and $p(b_1, b_2,...)$ from the assertion of S_2, where a_i is a member or subset of b_i for some i.

(12) In his first week of office, George W. Bush signed a ban on contributing money to international agencies which offer abortion as one of their services. Republican presidents often seek to put limits on federal funding of abortion.

From the Exemplification and Generalization relations, negation can be added to derive two definitions for EXCEPTION, depending on the clause order.

Exception (i): Infer $p(a_1, a_2,...)$ from the assertion of S_1 and $\rightarrow p(b_1, b_2,...)$ from the assertion of S_2, where b_i is a member or subset of a_i for some i.

Exception (ii): Infer $p(a_1, a_2,...)$ from the assertion of S_1 and $p(b_1, b_2,...)$ from the assertion of S_2, where a_i is a member or subset of b_i for some i.

Examples in which these two definitions apply are given in (13) and (14) respectively:

(13) Republican presidents do not usually put limits on federal funding of abortion immediately upon entering office. Nonetheless, in his first week, George W. Bush signed a ban on contributing money to international agencies which offer abortion as one of their services.

(14) In his first week, George W. Bush signed a ban on contributing money to international agencies which offer abortion as one of their services. Nonetheless, Republican presidents do not usually put limits on federal funding of abortion immediately upon entering office.

Finally, the ELABORATION relation can be seen as a limiting case of Parallel, in which the two eventualities described are in fact the same.

Elaboration: Infer $p(a_1, a_2, ...)$ from the assertions of S_1 and S_2.

(15) The new Republican president took a swipe at abortion in his first week of office. In a White

House ceremony yesterday, George W. Bush signed an executive order banning support to international agencies which offer abortion as one of their services.

To sum, to establish a Resemblance relation the hearer identifies a relation p that applies over a set of entities $a_1,.., a_n$ from the first sentence and a set of entities $b_1 ..., b_n$ from the second sentence, and performs comparison and generalization operations on each pair of parallel elements to determine points of similarity and contrast. These relations are therefore different than Cause–Effect relations, the arguments of which are simply the sentence–level propositions denoted by each utterance. Indeed, identifying the arguments to a Resemblance relation is considerably less straightforward, since it is not known a priori how many arguments there are; the common relation p to be inferred can have any number of arguments, including zero. Furthermore, in addition to identifying the appropriate sets of arguments from their respective utterances, it must also be determined which members of the first set are parallel to which members of the second. While the inference processes that underlie the establishment of Resemblance ultimately operate on semantic–level constructs, the process of argument identification and alignment likely utilizes the syntactic structure of the utterances, which would explain why many (but not all) passages standing in a Resemblance relation also display some degree of syntactic parallelism.

3.1.3 Contiguity relations

The third class of relation in my categorization is Contiguity. I place only one relation in this category, Occasion. Recall that Occasion allows one to express a situation centered around a system of entities by using intermediate states of affairs as points of connection between partial descriptions of that situation. An example of Occasion was given in passage (4a). Two definitions of Occasion, from Hobbs (1990), are given below:

Occasion (i): Infer a change of state for a system of entities from S_1 inferring the final state for this system from S_2.

Occasion (ii): Infer a change of state for a system of entities from S_2, inferring the initial state for this system from S_1.

Whereas the constraints for the other two types of relation and the types of inferential processes required for their establishment are at least somewhat understood, it is less straightforward to state the constraints imposed by Occasion explicitly. Much of what makes for a coherent Occasion is based on knowledge gained from human experience and the granularity with which people conceptualize events and change resulting from them. Certain past treatments (e.g. Halliday and Hasan 1976, Longacre 1983) have equated this relation with temporal progression, the only constraint being that the events described in the discourse display forward movement in time. However, the additional information inferred in order to connect the events in passages like (4a), and the incoherence of passages such as (4b), show that temporal progression is not enough (Hobbs 1990: 86).

3.2 Linguistic case studies

Given the foregoing analysis of coherence relations and the inference processes used to establish them, one might ask whether coherence establishment could potentially affect the behavior and distribution of a variety of linguistic phenomena that operate at least in part at the discourse level. Using three case studies – VP–ellipsis, extraction from conjoined clauses, and pronominal reference – I will now argue that it in fact does.

3.2.1 VP–ellipsis

The first phenomenon I consider is verb phrase (VP) ellipsis, exemplified in sentence (16):

(16) George claimed he won the election, and Al did too.

The stranded auxiliary in the second clause (henceforth, the TARGET clause, following terminology introduced by Dalrymple et al. 1991) marks a vestigial verb phrase, a meaning for which must be determined from some contextually provided material – in this case, the first clause (henceforth, the SOURCE clause). Past theories of VP–ellipsis interpretation can be largely classified into one of two categories: syntactic or semantic (Kehler 2000a). Inherent in syntactic accounts (Sag 1976, Williams

1977, Haïk 1987, Lappin 1993, Fiengo and May 1994, Hestvik 1995, Lobeck 1995, Lappin 1996, Kennedy 1997, inter alia) is the claim that VP-ellipsis is resolved at some level of syntactic structure. The evidence that proponents offer for this view includes the unacceptability of examples such as (24a–c).[8]

> (17) a. #New York was won by Al, and Hillary did too. [won New York]
> b. #Al$_i$ blamed himself$_i$, and George did too. [blamed him$_i$]
> c. #George blamed Al$_i$ for losing, and he$_i$ did too. [blame Al$_i$ for losing]

The unacceptability of sentence (17a) is predicted by a syntactic account due to the mismatch in voice between the clauses. In particular, assuming that a process of SYNTACTIC RECONSTRUCTION copies the source VP to the site of the empty VP in the target representation, the syntactic structure representing the active voice VP *won New York* required in the target is not present in the source. Likewise, binding conditions (Chomsky 1981) predict that (17b–c) are unacceptable on the indicated readings. Specifically, CONDITION A, which requires that a reflexive pronoun have a c-commanding antecedent, predicts that (17b) does not have the strict reading in which George blamed Al. Likewise, CONDITION C, which prohibits coreference between a full NP and a c-commanding NP, predicts that sentence (17c) does not have the reading in which Al blamed himself.

In semantic accounts of VP-ellipsis interpretation (Dalrymple et al. 1991, Hardt 1992, Kehler 1993b, Hobbs and Kehler 1997, Hardt 1999, inter alia), on the other hand, VP-ellipsis is resolved at a purely semantic level of representation. The acceptability of sentences such as (18a–c) has been cited in support of this view:

> (18) a. In November, the citizens of Florida asked that the election results be overturned, but the election commission refused to. [overturn the election results] (adapted from Dalrymple 1991)
> b. Bill is still a great campaigner, but he hasn't been allowed to this year, because Al doesn't want him to. [campaign] (adapted from Hardt 1993)
> c. George expected Al$_i$ to win the election even when he$_i$ didn't. [expect Al$_i$ to win] (adapted from Dalrymple 1991)

VP-ellipsis is felicitous in sentence (18a), unlike (17a), even though the source clause is passivized. This is unanticipated in a syntactic account, since the syntactic structure for the active voice VP that would be required for syntactic reconstruction – *overturn the election results* – is not provided by the source. Likewise, sentence (18b) is at least marginally acceptable, even though the referent is evoked by a nominalization. Finally, sentence (18c) is felicitous despite the fact that Condition C predicts it to be unacceptable (cf. 17c).

The contrast between examples (17a–c) and (18a–c) is puzzling, since each set seems to offer strong evidence for its respective approach. There is an important difference between them, however. The examples that support syntactic theories (17a–c) participate in Resemblance relations, particularly Parallel. On the other hand, the examples that support semantic theories (18a–c) participate in Cause-Effect relations, particularly Violated Expectation (18a–b) and Denial of Preventer (18c). This suggests that the difference between (17a–c) and (18a–c) corresponds to a difference in the way that the inference processes used to establish these two types of coherence relation interact with the process of VP-ellipsis interpretation itself.[9]

There is plenty of evidence to suggest that VP-ellipsis interpretation is fundamentally an anaphoric process. For instance, it behaves like other forms of anaphora (e.g. pronominal reference) with respect to the circumstances in which it can be cataphoric (e.g. *If he$_i$ wants to* φ$_j$, *Al$_i$* [*will contest the election*]$_j$) and the fact that it can access referents evoked from clauses other than the most immediate one (see Hardt 1993 for attested examples). Anaphora resolution is known to be a purely semantically mediated process, whereby referents are identified with respect to the hearer's knowledge state and mental model of the discourse. As such, it would be quite unexpected a priori if VP-ellipsis interpretation was associated with a process of syntactic reconstruction.

I claim that syntactic reconstruction is instead triggered by an aspect of the coherence establishment process, specifically, the need to recover the arguments to the coherence relation being established. In particular, reconstruction is triggered when the constituents corresponding to semantic

representations necessary for coherence establishment have been elided from the syntactic structure. Whether or not reconstruction will be necessary therefore depends on what coherence relation is operative, because Cause–Effect and Resemblance relations differ with respect to the type of arguments they take. In the case of a Cause–Effect relation, such as those operative in (18a–c), reconstruction will not be required. As was indicated in section 3.1, the arguments to these relations are merely the sentence–level semantics of each utterance. Since the top–level sentence node is never elided (obviously), and its semantics will be complete once the meaning of the missing VP has been recovered through anaphora resolution, there will be no missing arguments that require reconstruction. On the other hand, the process of identifying the arguments to a Resemblance relation, as well as the correct parallel pairing among them, requires access to the structure and semantics of subsentential constituents within each utterance. Thus, for many cases in which a Resemblance relation is operative,[10] including (17a–c), certain of these arguments will have been elided in the target clause and reconstruction will be required. As such, the account explains why syntactic constraint violations appear for VP-ellipsis in the context of Resemblance relations, but not Cause–Effect relations.[11]

Other types of syntactic constraints that I have not yet addressed are also observed with VP-ellipsis, including those involving traces in antecedent-contained ellipsis (ACE). Haïk (1987) points out that the unacceptability of examples like (19) is predicted by the subjacency constraint in a syntactic analysis.

(19) #John read everything which Bill believes the claim that he did. [read ϕ]

In Kehler (2000a, 2002), I suggest several ways that such violations could be addressed within the current proposal. First instance, an analysis in the spirit of Chao (1988) could be posited, in which the need to satisfy *wh-trace* dependencies in the target can also force the reconstruction of missing syntactic material. Alternatively, assuming a lexicalist theory of syntax capable of representing trace dependencies without movement or reconstruction (e.g. HPSG, LFG), a trace dependency represented at the elided VP node could be coordinated with a variable within the anaphorically resolved semantic representation (Mark Gawron, p.c., cf. Lappin 1999). Interestingly, cases like (19) contrast with ones involving traces in parasitic gap configurations (Fiengo and May 1994, Kennedy 1997, Lappin 1999); consider the unacceptability of (20a) versus the acceptability of its elided counterpart (20b), both from Rooth (1981).

(20) a. *Which problem did you think John would solve because of the fact that Susan solved?
b. Which problem did you think John would solve because of the fact that Susan did? [solved the problem]

This example differs from examples (19) in that there is no dependency within the sentence that requires there to be a trace within the elided VP. Thus, assuming a semantic theory in which the representation of the missing VP in (20b) contains a bound variable (Dalrymple et al. 1991), the fact that a Cause–Effect relationship is operative in (20b) is consistent with its acceptability.[12]

3.2.2 Extraction from conjoined clauses

The second phenomenon I consider is extraction from conjoined clauses. As is well known, Ross (1967: 89) first proposed the Coordinate Structure Constraint (CSC) as a basic constraint in universal grammar:

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

Grosu (1973) makes a convincing case that two components of the CSC should be differentiated: the CONJUNCT CONSTRAINT and the ELEMENT CONSTRAINT. The Conjunct Constraint bars the movement of whole conjuncts out of coordinate structures, ruling out sentences such as (21).

(21) This is the energy policy that George proposed the tax cut and.

The Conjunct Constraint is extremely robust, and has been argued to result from independently motivated constraints in several theories of grammar (but cf. Johannessen 1998).

The Element Constraint, which bars the movement of elements contained within a conjunct as opposed to the entire conjunct itself, has been more controversial, and hence will be my focus here. The Element Constraint rules out sentences such as (22a–b), because extraction of an NP has taken place out of a conjoined VP:

(22) a. #What energy policy did George support and propose the tax cut?
b. #What energy policy did George propose the tax cut and support?

A very general class of counterexamples to the Element Constraint was first noticed by Ross (1967) himself, whereby extraction out of coordinate structures is possible when the same element is extracted "across the board" from all the conjuncts, as in sentence (23):

(23) What energy policy did George propose and support?

Since Ross, the Element Constraint and the across-the-board exception has been taken by many syntactic theorists to be a valid generalization of the facts, and they have thus typically sought to explain it solely at the level of syntax (Schachter 1977b, Gazdar et al. 1985, Steedman 1985, Goodall 1987, Postal 1998). However, several other counterexamples that violate this generalization have been discussed in the literature.

For instance, Goldsmith (1985) points out that extraction out of a single conjunct can occur when the "nonetheless" use of *and* is operative between the conjuncts (which can be paraphrased as *and still, and nonetheless*, or *and yet*), as in example (24):

(24) How much can you drink and still stay sober?

G. Lakoff (1986) discusses the similar example given in (25), which he attributes to Peter Farley. In both (24) and (25), extraction has taken place out of the first conjunct but not the second.

(25) That's the stuff that the guys in the Caucasus drink and live to be a hundred.

Finally, Ross (1967) discusses examples of the sort shown in (26a), in which extraction has occurred only from the first and third conjuncts. Lakoff notes that unlike examples (24) and (25), extraction must take place out of the final conjunct in such cases, possibly along with certain other conjuncts that do not serve a scene-setting function. For instance, sentence (26b), which is the same as sentence (26a) but without the final (gap-containing) conjunct, is unacceptable.

(26) a. What did Harry buy, come home, and devour in thirty seconds?
b. #What did Harry buy and come home?

The foregoing data pattern directly with the three classes of coherence relations I have proposed. First, the Resemblance relation Parallel holds between the conjoined VPs in sentences (22a, b) and (23). In these cases, any extraction must be across-the-board. Second, the Cause–Effect relations Violated Expectation and Result are operative in examples (24) and (25) respectively. In these cases, extraction from only one conjunct is permitted. Third, example (26a) is related by the Contiguity relation Occasion. These cases also allow extraction to occur out of only a subset of conjuncts, but appear to require extraction out of the final one (per 26b).

Viewing the data in light of this pattern, there does not appear to be much ground left on which to argue for the CSC as a purely syntactic constraint. In none of the three categories is extraction from within a conjunct barred entirely; instead, there appear only to be weaker constraints at play that differ with respect to the type of coherence relation. This fact suggests that the data would be better explained at the level of the syntax/pragmatics interface, rather than stipulated as a purely grammatical constraint.[13]

The factor that appears to be coming into play in these data is what Kuno (1976b, 1987) calls a TOPICHOOD CONDITION on extraction. Kuno cites sentences (27a–d) to support such a constraint:

(27) a. I read a book about John Irving.
b. I lost a book about John Irving.
c. Who did you buy a book about?
d. ??Who did you lose a book about?

Whereas sentences (27a, b) are both perfectly acceptable, the extracted variants in sentences (27c, d) are not equally acceptable. Kuno says:

> In a highly intuitive sense, we feel that the fact that the book under discussion was on John Irving is much more *relevant* in [(27a)] than in [(27b)]. This is undoubtedly due to the fact that one buys books, but does not lose them, because of their content.
>
> (Kuno 1987: 23)

Although the notion of topichood still lacks a concrete definition in the literature,[14] one can still ask how such a constraint might apply when extraction from coordinate clauses is considered. Indeed, as I describe in detail in Kehler (2002), one would expect that the constraints on what can serve as a topic of a conjoined set of clauses will depend on the operative coherence relation. Clauses related by a Resemblance relation are coherent by virtue of the very fact that they manifest a common topic (cf. R. Lakoff 1971b); the possibilities for topichood are thus provided by the properties shared by the corresponding elements over which the relation applies (at some, perhaps inferred, level of generalization). Syntactic constructions that involve extraction are thus felicitous only when a set of parallel elements that denote this topic can be extracted to a common topic-denoting position, which in most cases is possible only when the elements are identical, resulting in the "across-the-board" behavior.[15] In contrast, there is no similar property of Cause–Effect relations nor of the inference processes that underlie their establishment that would prohibit an element in one clause from serving as the topic of both; indeed, one would expect the topic of a clause expressing a cause to be relevant to one expressing its effect. Finally, the topic of a set of clauses related by Occasion need not be mentioned in every clause – in particular, coherent Occasions commonly contain scene-setting and similar types of clauses that do not mention, but also do not distract attention from, the discourse topic – so extraction need not take place out of all conjuncts. On the other hand, extraction must take place out of the conjuncts that do not serve this type of scene-setting function, including the final clause, insofar as a failure to do this would suggest that the extracted element is no longer the topic at that point in the passage.

3.2.3 Pronominal reference

Finally, I consider pronominal reference. Sure enough, three different types of approach can be found in the literature, each motivated by different types of examples. While the data that support each approach are often problematic for the others, a pattern emerges when the operative coherence relation is taken into account.

Hobbs (1979) presents what I will call a COHERENCE-DRIVEN theory, in which pronoun interpretation is not an independent process but instead a by-product of more general reasoning about the most likely interpretation of an utterance. Pronouns are modeled as free variables which become bound during these inference processes; potential referents of pronouns are therefore those which result in valid proofs of coherence. A typical example used to support a coherence-driven theory is given in sentence (28) with follow-ons (a) and (b), adapted from an example from Winograd (1972).

(28) The city council denied the demonstrators a permit because
a. they *feared* violence.
b. they *advocated* violence.

Hearers appear to have little difficulty interpreting the pronoun *they* in each case, despite the fact that

it refers to *the city council* in sentence (28a) and *the demonstrators* in sentence (28b). These different interpretations result even though the two sentence pairs have identical syntactic configurations. Indeed, the only difference between them is the verb in the second clause, suggesting that semantics is the key factor in determining the correct referents. In each case, the pronoun receives the assignment necessary to establish the Cause–Effect relation Explanation.

Contrasting with coherence–driven theories are what I will call ATTENTION–DRIVEN theories. Attention–driven theories treat pronoun interpretation as an independent process associated with its own interpretation mechanisms, rather than as a side–effect of more general interpretation mechanisms. Instances of this approach include Sidner's focusing framework (Sidner 1983) and Centering theory (Kameyama 1986, Brennan et al. 1987, Grosz et al. 1995), among others. Examples such as passage (29), from Grosz et al. (1995), have been used to support this type of approach.

> (29) a. Terry really goofs sometimes.
> b. Yesterday was a beautiful day and he was excited about trying out his new sailboat.
> c. He wanted Tony to join him on a sailing expedition.
> d. He called him at 6 a.m.
> e. He was sick and furious at being woken up so early.

This passage is perfectly acceptable until sentence (29e), which causes the hearer to be misled. Whereas commonsense knowledge would indicate that the intended referent for *He* is Tony, hearers tend to initially assign Terry as its referent, creating a garden–path effect. Such examples therefore provide evidence that there is more to pronoun interpretation than simply reasoning about semantic plausibility. In fact, they suggest that hearers assign referents to pronouns at least in part based on other factors (e.g. the grammatical role of the antecedent; cf. examples 28a, b) before interpreting the remainder of the sentence. In this example, the Contiguity relation Occasion is operative.

Finally, passages such as (30) and (31), from Sidner (1983) and Kameyama (1986) respectively, are potentially problematic for both types of approach:

> (30) a. The green Whitierleaf is most commonly found near the wild rose.
> b. The wild violet is found near it too.

> (31) a. Carl is talking to Tom in the Lab.
> b. Terry wants to talk to him too.

In these cases there is a strong bias for the pronoun to refer to its syntactically parallel referent (*the wild rose* in (30), and *Tom* in (31)). This fact poses a problem for attention–driven approaches that, as is typical, prefer referents evoked from subject position over those evoked from other grammatical positions (cf. example 29). Furthermore, there is no semantic basis to prefer one referent over the other in each case, since coherence could be established assuming either assignment. In these examples the Resemblance relation Parallel is operative.

While these three types of example offer contradictory evidence about the mechanisms that underlie pronoun interpretation, they pattern directly with the neo–Humian categorization of coherence relations I have proposed. Examples that offer specific support for interpretation preferences based on semantics and coherence, such as (28), tend to participate in Cause–Effect relations. Examples that directly support preferences based on a hierarchy of grammatical role, like (29), tend to be instances of Contiguity relations. And examples that support grammatical role parallelism preferences, such as (30) and (31), are typically instances of Resemblance. This pattern suggests that an adequate analysis will have to account for the differences among the inference processes that underlie the establishment of these different types of relations.

In Kehler (2002), I argue that these data can be explained by an analysis that shares characteristics of both coherence–driven and attention–driven approaches. As in attention–driven approaches, pronouns impose the requirement that their referents be highly attended to within the current discourse state. However, in contrast to the clause–by–clause discourse update mechanisms often found in such approaches, the discourse state changes on a rapid time scale as coherence establishment processes redirect the focus of attention as necessary during inferencing. The fact that different types of pronoun

interpretation preferences appear to be in force when different coherence relations are operative is a result of the different types of reasoning underlying their recognition.

Recall, for instance, that the inference process associated with Resemblance first identifies pairs of parallel entities and relation from each clause, and then attempts to identify points of similarity and contrast among the members of each pair. When examples like (30) or (31) are interpreted, this process will pair each pronoun with its parallel element, placing the latter in focus. At that point, similarity can be established between the two by simply assuming coreference. The effect of attending solely to the pronoun's parallel element at that point during inferencing is so strong that it trumps even a strong world knowledge bias toward another referent. For instance, hearers universally assign Clinton as the referent of *her*[16] in (32a) even though world knowledge would strongly suggest Thatcher. Likewise, hearers even get confused by example (32b) due to the gender mismatch between *her* and Reagan, even though Thatcher is an otherwise perfectly suitable referent (cf. Oehrle 1981).

> (32) a. Margaret Thatcher admires Hillary Clinton, and George W. Bush absolutely worships her.
> b. Margaret Thatcher admires Ronald Reagan, and George W. Bush absolutely worships her.

On the other hand, the inference process associated with Cause–Effect relations attempts to identify a chain of implication between the semantics representations of the clauses being related. As such, there would be no reason to expect a bias toward parallelism in such cases. Instead, an axiom used to create the implicational chain may bring into focus the pronominal referent that is necessary for that axiom to apply. As a result, examples that are ambiguous between Resemblance and Cause–Effect readings (for example) may display a corresponding ambiguity with respect to the interpretation of the pronoun, as in (33):

> (33) Colin Powell defied Dick Cheney, and George W. Bush punished him.

If the Parallel relation is inferred then the pronoun must be interpreted to refer to its parallel element, Cheney. If a Result relation is inferred then Powell is the preferred referent, in accordance with our world knowledge about the relationship between defying and punishing.

Finally, the inference process associated with Contiguity attempts to draw the connections necessary to interpret the final state of one eventuality as being the initial state of the next. As such, at the time that a pronoun is encountered in a sentence like (29e), the referent most attended to should be the one that is most prominent with respect to the hearer's conceptualization of the end state of the previous eventuality. While this will often be the subject of the preceding sentence, this is not always so; consider examples (34a, b), from Stevenson et al. (1994).

> (34) a. John seized the comic from Bill. He …
> b. John passed the comic to Bill. He …

In a set of psycholinguistic experiments, Stevenson et al. found that hearers are more likely to interpret *he* to refer to John in passage (34a) and to Bill in (34b), despite the fact that in (34b) Bill is mentioned from within a sentence-final prepositional phrase, a position which attention-driven theories typically consider to be much less salient than subject position. What these examples have in common is that the preferred referent occupies the Goal thematic role of its respective predication, which is presumably more central to the final state of the eventuality than the entity that occupies the Source thematic role.

To sum, by accounting for the different properties of the inference processes that underlie the establishment of the three types of coherence relations in the neo-Humean classification, facets of both coherence-driven and attention-driven theories can be integrated to account for pronominal reference behavior that is problematic for each in isolation.

4 Informational and Intentional Coherence

Thus far I have described an approach to establishing coherence based on making the inferences necessary to meet the constraints imposed by one of a set of coherence relations. Following the

terminology of Moore and Pollack (1992), I will refer to this view as the INFORMATIONAL approach to coherence. Historically, this approach has been applied predominantly to monologues.

In contrast, other researchers (Grosz and Sidner 1986, inter alia), following work in speech act theory and plan recognition (e.g. Cohen and Perrault 1979, Allen and Perrault 1980), have argued that the role of the utterance in the overall plan underlying the speaker's production of the discourse is the determining factor of coherence. I will likewise follow Moore and Pollack and refer to this view as the INTENTIONAL approach. In this view, a hearer considers utterances as actions and infers the plan-based speaker intentions underlying them to establish coherence. The intentional approach has been applied predominantly to dialogues, such as the following interchange from Cohen et al. (1990):

> (35) *Customer:* Where are the chuck steaks you advertised for 88 cents per pound?
> *Butcher:* How many do you want?

A more appropriate information-level response to the customer's question would be *behind the counter*. However, the butcher recognizes the customer's higher-level goal of purchasing the steaks and responds with a question that is designed to address and ultimately satisfy that goal, hence the coherence of the response with respect to the speaker's intentions.

In the intentional approach of Grosz and Sidner (1986), each discourse segment has a corresponding DISCOURSE SEGMENT PURPOSE, or DSP. In contrast to the large set of relations commonly found in informational analyses, there are only two relations that can hold between discourse segments: DOMINANCE, in which the satisfaction of the DSP of one discourse segment is intended to provide part of the satisfaction of the DSP of another segment, and SATISFACTION-PRECEDENCE, in which the DSP of one segment must be satisfied as a prerequisite to satisfying the DSP of another.

The relationship between these two conceptions of coherence has been a topic of some debate (Moore and Pollack 1992, Moore and Paris 1993, Asher and Lascarides 1994, Hobbs 1997, inter alia). Moore and Pollack (1992) argued that in fact both levels of analysis must co-exist, illustrating the point with passage (36):

> (36) a. George Bush supports big business.
> b. He's sure to veto House Bill 1711.

Passage (36) can be analyzed from either the intentional or informational perspective. At the intentional level, the speaker may be trying to convince the hearer of the claim being made in sentence (36b), and offering sentence (36a) as evidence to support it. At the informational level, she intends that the hearer recognize a Result relationship between the fact expressed in sentence (36a) and the event expressed in sentence (36b). This duality is not surprising, since one way to provide evidence for a proposition is to show that it follows as a consequence of another proposition that the hearer already believes.

Moore and Pollack demonstrate that this connection may allow a hearer to recognize a relation at one level from the recognition of a relation at the other level. For instance, if the hearer knows that House Bill 1711 imposes strong environmental controls on manufacturing processes, but does not know the intentions of the speaker a priori, he can infer the intention of providing evidence from having recognized the informational Result relation. Alternatively, if the hearer has no knowledge of the content of House Bill 1711, but has prior reason to believe that the speaker is attempting to provide evidence for the proposition in (36b), he may infer that a Result relation holds, and from this that House Bill 1711 must place undesirable constraints on businesses. As such there is reason to suggest that both levels co-exist, with links between the two to enable the recognition of relationships on one level from the recognition of relationships on the other.

5 Conclusion

In this chapter, I have attempted to convince the reader that coherence establishment is not only a fundamental aspect of discourse interpretation, but that it needs to be accounted for in analyses of a variety of linguistic phenomena that operate across clauses. I categorized a set of coherence relations with respect to Hume's three types of connection among ideas, and demonstrated how differences

among the inference processes that underlie the establishment of such relations affect the distribution and behavior of three different linguistic phenomena. Given the centrality of coherence establishment to language interpretation, it would perhaps have been surprising if it were found that these processes did not affect the behavior of such phenomena.

I again refer the reader to Kehler (2002) for considerably more detailed treatments. I also address two other linguistic phenomena in that work, gapping and tense interpretation. The treatment of gapping accounts for the fact that conjoined clauses that are compatible with Resemblance and Cause-Effect interpretations lose the latter after gapping has applied (Levin and Prince 1986). This prediction falls out from the theory of VP-ellipsis summarized in this paper, along with the fact that the gapping, unlike VP-ellipsis, is not anaphoric. The treatment of tense shows how the temporal constraints imposed by coherence relations interact with the anaphoric properties of tense to predict data that are problematic for past analyses, including approaches that treat the simple past as anaphoric (Partee 1984b, Hinrichs 1986, Nerbonne 1986, Webber 1988, inter alia), as well as one that resolves temporal relations involving both simple (e.g. the past) and complex (e.g. the past perfect) tenses purely as a by-product of coherence establishment processes (Lascarides and Asher 1993).

Much remains to be accomplished before we have a fully adequate analysis of discourse coherence. We have only scratched the surface with respect to understanding the answers to many questions: the detailed workings of the inference mechanisms that underlie relation establishment; how a preferred interpretation emerges from a set of alternative proofs of coherence; the manner in which these mechanisms operate during on-line, left-to-right processing; in what contexts a discourse connective is necessary, optional, or even redundant; what the basic principles of coherence are and how they interact with the large set of connectives that a language makes available; how deeply embedded these basic principles are not only with respect to discourse but also at the levels of the lexicon and clause; and many, many others. It is hoped that this chapter has inspired some new interest in these questions.

ACKNOWLEDGMENTS

1 In the majority of this chapter, I will focus almost solely on coherence in monologue, specifically with respect to relationships that hold between adjacent clauses and their implications for sentential interpretation. I will not address coherence among larger segments of discourse here, and I speak only briefly about dialogue in section 4.

2 Indeed, a common objection to theories of coherence is that the proposed relations have a laundry-list quality, with no rationale given for why they constitute the correct set. The unwieldiness of the situation is demonstrated in Hovy (1990), who compiled over 350 previously proposed relations from 26 researchers. After merging redundancies, a hierarchy with 63 relations resulted.

3 I am oversimplifying a bit here, since there is a small set of relations which are multi-nuclear and can relate more than two spans of text (e.g. the JOINT relation).

4 Notable exceptions include Hobbs's (1979) approach to resolving pronouns in the context of coherence establishment and Lascarides and Asher's (1993) similar approach to tense interpretation.

5 Hobbs (1990: 101-2) was the first to point out that Hume's principles could be used as a basis for categorizing coherence relations, but he did not pursue such a classification in depth, opting instead for a different basis for categorizing relations.

6 For ease of exposition, I will treat Parallel as if it always relates only two clauses, whereas in reality it can operate over longer sequences.

7 While not directly stated in this definition of Exemplification, the subset relationship can also hold between the relations expressed instead of one or more pairs of entities (or both). The same is true for the Generalization and Exception relations, discussed below.

8 Here and elsewhere in the paper, my analyses suggest that various examples commonly considered to be ungrammatical are actually pragmatically infelicitous. As such, I will typically indicate unacceptability with a

hash mark (#) rather than an asterisk, except in a few instances that I still consider to be ungrammatical.

9 The situation with respect to Contiguity relations is less clear, and will not be discussed further here. See Kehler (2002) for discussion.

10 The situation is more complicated for comparative and temporal subordination constructions, which, despite being instances of Resemblance, readily violate Binding Theory constraints: (i)a. Al$_i$'s lawyer defended him$_i$ better than he$_i$ did. [defend him$_i$]
b. Al claimed that George$_i$ won before he$_i$ did. [claimed that George$_i$ won] However, these constructions still require that the source and target VPs be parallel; consider the infelicity of the following voice mismatches: (ii)a. #Al was defended by Joe more competently than Bill did. [defend Al]
b. #CNN announced Al as the winner before George was. [announced as the winner] I argue in Kehler (2000a, 2002) that the process of establishing the coherence of such constructions only requires VP-level parallelism, which would account for these facts.

11 Kennedy (in press), who advocates a syntactic reconstruction account, criticizes my analysis by stating that: there is a third, more general problem with a mixed approach such as Kehler's. If a purely semantic analysis is available in some examples, then it ought to be in principle available in all examples, even if a syntactic analysis is preferred. In actuality, my analysis specifically predicts that this is not the case. This prediction comes about because the relevant syntactic and semantic constraints originate from different recovery processes. To reiterate, the meaning of the elided VP is always recovered anaphorically (i.e. semantically), regardless of the coherence relation. In no case is syntactic reconstruction performed out of a need to recover the meaning of an elided VP. Syntactic reconstruction is instead necessitated only by the independent need to recover elided arguments to a coherence relation. Thus, while a "semantic analysis" is in fact available for all examples, some examples have the *additional* constraint of requiring a syntactic antecedent that can be reconstructed. Hence there is no respect in which my analysis is "mixed" with respect to the mechanism used for recovering the meaning of an elided VP. There are two other respects in which Kennedy criticizes my account. One pertains to the status of Condition B violations in Cause–Effect relations – e.g. *Joe has to take care of Ali because he$_i$ won't* – which I find acceptable and he does not. The other pertains to certain examples involving trace violations, which space precludes addressing here. I refer the reader to Kennedy's discussion and the relevant discussions in Kehler (2000a, 2002).

12 But cf. Kennedy (1997), who presents a syntactic analysis in which reconstructed target VPs in examples such as these contain a pronoun rather than the expected trace, utilizing the VEHICLE CHANGE proposal of Fiengo and May (1994).

13 Per footnote 8, I have marked the unacceptable extraction examples (except 21) as infelicitous rather than ungrammatical, as my analysis implies that none of these are unacceptable on purely syntactic grounds. Although this may contradict the intuitions of some with respect to examples like (22a, b), it can hardly be any other way. As Lakoff (1986) points out, if one allows an autonomous syntactic module to generate sentences such as (24), (25), and (26a), then it must also be able to generate (22a, b), leaving the task of filtering out (22a, b) to semantic or pragmatic constraints. The converse – keeping the CSC and adding a semantic or pragmatic condition to allow (24), (25), and (26a) – is not an option, since such conditions could not turn an ungrammatical sentence into a grammatical one.

14 In lieu of such a definition, Kuno (1976b) offers a "Speaking of X" test for a potential sentence topic X (see also Reinhart 1981), in which this phrase is placed at the beginning of the sentence and the mentions of X pronominalized: (i)a. Who did you buy a book about?
b. Speaking of John Irving, I just bought a book about him.
c. ??Who did you lose a book about?
d. ??Speaking of John Irving, I just lost a book about him.

15 Actually, it is possible to extract more than one element from a conjoined clause when a RESPECTIVE READING is operative, as in the following example (Postal 1998): (i) What book and what magazine did John buy and Bill read, respectively? In fact, Gawron and Kehler (2000) show that similar examples are felicitous even when the "extracted" elements are denoted as a group by a single noun phrase, posing a challenge for theories that account for extraction by movement or that otherwise require coreference between the gap site and the constituent on which it is dependent. Both predict only an across-the-board reading for example (ii): (ii) I finally met Susan, Marilyn, and Lucille yesterday. They are the three sisters that Bob married, John is engaged to, and Bill is dating (respectively).

16 This assumes that the pronoun remains unaccented, of course.

Cite this article

KEHLER, ANDREW. "Discourse Coherence." *The Handbook of Pragmatics*. Horn, Laurence R. and Gregory Ward (eds). Blackwell Publishing, 2005. Blackwell Reference Online. 28 December 2007 <http://www.blackwellreference.com/subscriber/tocnode?id=g9780631225485_chunk_g978063122548513>

Bibliographic Details

The Handbook of Pragmatics

Edited by: Laurence R. Horn And Gregory Ward
eISBN: 9780631225485
Print publication date: 2005