

# Coherence and the Coordinate Structure Constraint

Andrew Kehler  
SRI International

## 1 The Coordinate Structure Constraint

As is well known, the Coordinate Structure Constraint (CSC) was first proposed by 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.

Following Grosu (1973) and Pollard and Sag (1994), we differentiate between two components of the CSC: 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 (1).

(1) \* This is the magazine which John bought the book and.

The Conjunct Constraint has been shown to result from independently motivated constraints in several theories of grammar, for example, Ross (1967) notes that it results independently from his A-over-A Principle, and Pollard and Sag (1994) from their Trace Principle.

We address facts concerning the Element Constraint, which bars the movement of elements contained within a conjunct as opposed to the conjunct itself. The Element Constraint rules out sentences such as (2), because extraction has taken place out of a conjoined verb phrase (VP).

(2) \* What book did John buy and read the magazine? (Parallel)

Several counterexamples to the Element Constraint have been discussed in the literature. First, Ross (1967) himself notes that extraction out of coordinate structures is possible when the same element is extracted from all the conjuncts (so-called “across-the-board” application), as in sentence (3).

(3) What book did John buy and read? (Parallel)

The type of relation that holds between the two conjoined VPs in sentences (2) and (3) is what has been called *Parallel*, in which the *and* can be roughly paraphrased by *and too* or *and also*.

Second, Goldsmith (1985) notes that extraction out of a single conjunct can occur when the “nonetheless” use of *and* is operative between the conjuncts (paraphraseable by *and still*, *and nonetheless*, or *and yet*), as in example (4); this relation has been called *Violated Expectation*.

- (4) How much can you drink and still stay sober? (Violated Expectation)

Lakoff (1986) notes a similar type of case, in which the operative relation between the conjuncts is *Result* (in which *and* is paraphraseable by *and therefore* or *and as a result*), shown in example (5). In both of these cases, extraction has taken place out of the first conjunct but not the second.

- (5) That’s the stuff that the guys in the Caucasus drink and live to be a hundred. (Result) [Lakoff (1986), attributed to Farley]

Finally, extraction can also take place out of coordinate structures when the operative relation is *Narration* as shown in example (6a), in which *and* is paraphraseable by *and then*. However, Lakoff notes that unlike examples (4) and (5), in these cases extraction must take place out of the final conjunct, possibly along with certain others that do not serve a scene-setting function. For instance, if the final conjunct of sentence (6a) is removed to form sentence (6b), the result is unacceptable.

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

To summarize, we can categorize the CSC data into three classes. The first is exemplified examples (2) and (3), in which a *Parallel* relation is extant between the two conjuncts. For reasons that will become clear in Section 4, we categorize this relation as a *Resemblance* relation. The second class is exemplified by examples (4) and (5), in which the relations *Violated Expectation* and *Result* hold, respectively. Because these relations are rooted in cause and effect, we categorize them as instances of *Cause-Effect* relations. The final class of cases is exemplified by example (6), in which the *Narration* relation is operative. We categorize this relation as a *Contiguity* relation.

Viewing the data in light of this categorization, there doesn’t appear to be much support left for the CSC. In fact, in none of the three categories is extraction from a coordinate construct barred entirely. Instead, there appear to be weaker constraints at play that differ with the type of relation: (1) when

a Resemblance relation is operative between the conjuncts, extraction must occur out of each conjunct, (2) when a Contiguity relation is operative, extraction must take place out of the final conjunct (and perhaps certain others), and (3) when a Cause-Effect relation is operative, no constraints appear to be at play.

In the remainder of this paper, we show that the correlation of data with respect to this trichotomy of relation types is not unique to that for extraction from coordinate constructs. We summarize previous work that shows similar patterns in data for VP-ellipsis and gapping. We then discuss how the syntactic operations in question interact with discourse-level coherence resolution mechanisms to account for the data.

## 2 VP-Ellipsis

There has been an ongoing debate in the literature concerning the level of language processing at which VP-ellipsis is resolved. VP-ellipsis is exemplified in sentence (7).

(7) Ross likes his mother, and Bill does too.

The stranded auxiliary in the second clause (the *target* clause) indicates the elision of a verb phrase, a representation for which must be recovered from the representation of another clause, in this case the first clause (the *source* clause). Sentence (7) displays a strict/sloppy ambiguity; Bill may like *Ross's* mother (the strict reading) or *his own* mother (the sloppy reading).

Inherent in *syntactic* accounts is the claim that VP-ellipsis is resolved at a level of syntactic structure (either surface structure or LF). On the other hand, inherent in *semantic* accounts is the claim that VP-ellipsis is resolved at a purely semantic level of representation. This question remains a point of contention, primarily because there are data to support both views.

In previous work (Kehler, 1993; Kehler, 1994a; Kehler, 1995), we have shown that the seemingly contradictory VP-ellipsis data exhibit a pattern that correlates with the type of coherence relation operative between the source and target clauses. Specifically, we show that the data support a syntactic account when a Resemblance relation is operative between the clauses, whereas the data support a semantic account when a Cause-Effect relation is operative. (The data for the third type of relation, Contiguity, requires more attention than we have space for here. See Kehler (1995) for further discussion.)

To show this pattern we examined two types of data. The first type includes examples for which a suitable semantic representation for the source clause is available, but for which there is a mismatch of surface-syntactic form between the syntactic representations of the source and the target clauses. In

such cases, if ellipsis shows a sensitivity to syntactic form, then the syntactic approaches are evidenced; otherwise, semantic approaches are evidenced. The second type of data includes examples for which there is no syntactic form mismatch, but for which reconstruction of the source VP within the target would possibly result in a syntactic constraint violation in the target. In these cases, if the source and target pair is unacceptable because of such a constraint violation, then the syntactic theories are supported; otherwise, semantic approaches are supported. We examined six types of elliptical contexts: two in the first category (*voice alternation* and *non-VP antecedents*), and four in the second (*Condition A-C violations* and *subjacency violations*). We provide an example of each case here to illustrate the pattern.

**Voice Alternation** In sentence (8), an elided target VP in the active voice receives its interpretation from a source clause in the passive voice, despite the mismatch of syntactic form. Contra syntactic accounts, the result is acceptable (here the Violated Expectation meaning of *but* is operative). However, in accord with syntactic accounts, sentence (9), which is otherwise similar to sentence (8), is unacceptable (here the Parallel relation holds).

- (8) This problem was to have been looked into, but obviously nobody did.  
 [look into the problem] (Vincent Della Pietra, in conversation)  
 (Violated-Expectation)
- (9) # This problem was looked into by John, and Bob did too. (Parallel)

**Non-VP Antecedents** In example (10), the semantic representation for the source is evoked by a nominalization, and not a syntactic VP. While this example is acceptable, the otherwise similar sentence (11) is not.

- (10) This letter deserves a response, but before you do, .... [respond]  
 (Gregory Ward, personal communication) (Violated-Expectation)
- (11) # This letter provoked a response from Bush, and Clinton did too.  
 [respond] (Parallel)

**Condition A Violations** As predicted by Condition A of Binding Theory, it is generally difficult to obtain a strict reading when the source clause contains a reflexive pronoun, as shown in sentence (12). However, in sentence (13), the strict reading is readily available and perhaps preferred.

- (12) ?? Fred<sub>i</sub> voted for himself<sub>i</sub>, and Gary<sub>j</sub> did too. [vote for Fred<sub>i</sub>] (Parallel)
- (13) John<sub>i</sub> voted for himself<sub>i</sub> even though no one else<sub>j</sub> did. [vote for John<sub>i</sub>]  
 (Denial of Preventer)

**Condition B Violations** Sentence (14) is also odd, as predicted by Condition B of Binding Theory. However, sentence (15) seems to be acceptable despite the predicted Condition B violation.

(14) ?? John<sub>i</sub>'s lawyer defended him<sub>i</sub>, and he<sub>i</sub> did too. [defend himself<sub>i</sub>]  
(Parallel)

(15) John<sub>i</sub>'s lawyer defended him<sub>i</sub> because he<sub>i</sub> couldn't. [defend himself<sub>i</sub>]  
(Explanation)

**Condition C Violations** The unacceptability of examples such as (16) is predicted by Condition C of Binding Theory. On the other hand, sentence (17) is also predicted to be unacceptable, but it is instead acceptable.

(16) ?? Mary introduced John<sub>j</sub> to everyone, and he<sub>j</sub> did too. [introduced John<sub>j</sub> to everyone] (Parallel)

(17) The lawyer defended Bill<sub>j</sub> against the accusations because he<sub>j</sub> couldn't.  
[defend Bill<sub>j</sub> against the accusations] (Explanation)

**Subjacency Violations** Haik (1987) gives examples of apparent subjacency violations in cases of antecedent-contained deletion (ACD), exemplified by sentence (18). However, Rooth notes that one would expect subjacency to apply in sentence (19), but it appears not to; in this case the Cause-Effect relation *Explanation* holds between the clauses. Furthermore, in the corresponding case without ellipsis shown in sentence (20), the gap remains and a subjacency violation results, seemingly precluding the possibility that syntactic material is reconstructed in interpreting sentence (19).

(18) # John read everything which Bill believes the claim that he did. [read  $\phi$ ]

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

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

The data given throughout this section suggest that VP-ellipsis resolution requires a syntactically suitable antecedent in examples in which a Resemblance relation is operative, but only a suitable semantic representation in examples in which a Cause-Effect relation is operative. (Again, for reasons of space, we omit discussion of the third category, Contiguity, which appears to have more subtle constraints at work.)

### 3 Gapping

We now move on to consider the gapping construction. Gapping is characterized by an initial *source* clause and the elision of all but two (and in constrained circumstances, more than two) constituents in one or more subsequent *target* clauses, as exemplified in sentence (21a).

- (21) a. Sue became upset and Nan  $\emptyset$  downright angry.  
b. Sue became upset and Nan became downright angry.

We address a particular phenomenon noticed by Levin and Prince (1982), who note that pairs of conjoined sentences such as (21b) have what they call *symmetric* and *asymmetric* readings. That is, sentence (21b) could have a *symmetric* reading in which the two events are understood as independent, or an *asymmetric* reading in which the first event is interpreted as the cause of the second event. In our terms, the symmetric readings correspond to the Resemblance relation *Parallel*, whereas the asymmetric readings correspond to the Cause-Effect relation *Result*. Levin and Prince contrast the sentence (21b) with its gapped counterpart, given in (21a), which unlike sentence (21b) has only the symmetric reading. That is, whereas sentence (21b) can have a reading in which Nan became angry *because* Sue became upset, this reading is unavailable in (21a). This can be seen by the following contexts, again due to Levin and Prince, in which gapping is acceptable in the context favoring the symmetric reading in (22), but not in the context favoring the asymmetric (causal) reading given in (23), although in both cases the non-gapped versions are acceptable.

- (22) 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.  
(23) 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.

The causal interpretation of the two final clauses in example (23), supported by the given context, is unavailable when gapping has applied.

While Levin and Prince limit their discussion to sentences conjoined with *and*, in a previous paper (Kehler, 1994b) we have shown that this pattern is quite robust with respect to the Resemblance versus Cause-Effect relation distinction. For instance, like *and*, *or* also has Resemblance and Cause-Effect uses. Consider example (24a).

- (24) a. John will go to New York, or Bill will go to Boston.  
 b. John will go to New York, or Bill  $\emptyset$  to Boston.

Sentence (24a) has two readings: a symmetric (disjunctive) reading, and an asymmetric causal reading (e.g., to express a threat of the form *If A doesn't happen then B will!*). Like the case with *and*, the gapped counterpart of sentence (24a) shown in sentence (24b) has only the symmetric reading. Likewise, the sentence (25) is only felicitous with the pure Contrast meaning of *but*, and not with the Violated Expectation meaning.

- (25) John voted for Clinton but (# nonetheless) Tom  $\emptyset$  for Bush.

Finally, gapping is unacceptable with subordinating conjunctions that indicate Cause-Effect relations, as in examples (26a-c).

- (26) John voted for Clinton,  $\left\{ \begin{array}{l} \# \quad \text{because} \\ \# \quad \text{even though} \\ \# \quad \text{despite the fact that} \end{array} \right\}$  Tom for Bush.

A stipulation commonly seen in the syntactic literature is that gapping may apply only in coordinate structures, and not in subordinate ones (such as those in example (26)). However, the facts concerning *and*, *or*, and *but* suggest that the Resemblance versus Cause-Effect distinction may instead be the correct one to make.

Relation Type	Across-the-Board Extraction Required?	VPE Requires Syntactic Parallelism?	Gapping Felicitous?
Resemblance	yes	yes	yes
Cause-Effect	no	no	no
Contiguity	no*	no*	no*

Table 1: The Pattern for Extraction, VP-Ellipsis, and Gapping

The pattern we have established for extraction, gapping, and VP-ellipsis is summarized in Table 1. The asterisks in the *Contiguity* row indicate the more subtle constraints at play in that case, which we are not addressing here.

## 4 Hume’s Three Types of Coherence

We have categorized the data with respect to the type of relation that holds between the clauses in question; these relations have included *Parallel*, *Result*, *Violated Expectation*, *Explanation*, and *Narration*, among others. Past researchers have posited such *coherence relations* to account for why discourses are interpreted under the assumption that they are coherent, a process which may require that the hearer adopt additional inferences beyond what is explicitly stated. While many such sets of relations have been posited with associated categorizations, our categorization appeals to the sentiment summarized in the following passage, due to the philosopher David Hume (1748):

“To me there appear to be only three principles of connection among ideas, namely *Resemblance*, *Contiguity* in time or place, and *Cause* or *Effect*.”

In this section, we analyze a set of coherence relations as belonging to these three general categories. The three classes are shown to differ systematically in the type of arguments over which the coherence constraints are applied, as well as in the type of inference process underlying this application.

### 4.1 Resemblance Relations

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 2: Resemblance Relations

Establishing a passage as coherent under a Resemblance relation requires that commonalities and contrasts among corresponding sets of properties and entities be recognized. For each relation, the hearer identifies a relation  $p_0$  that applies over a set of entities  $a_1, \dots, a_n$  from the first sentence  $S_0$ , and a corresponding relation  $p_1$  that applies over a corresponding set of entities  $b_1, \dots, b_n$  from the second sentence  $S_1$ . Coherence results from these corresponding components being related; in this case a common (or contrasting) relation  $p$  subsuming  $p_0$  and  $p_1$  is inferred along with common (or contrasting) properties  $q_i$  of the corresponding elements  $a_i$  and  $b_i$ . The Resemblance relations are summarized in Table 2.

## 4.2 Cause-Effect Relations

Establishing a passage as coherent under a Cause-Effect relation requires that a path of implication be established between the propositions denoted by the utterances. (We are using implication in a very loose sense here, meaning roughly “could plausibly follow from”.) For each relation, the hearer identifies a proposition  $P$  from the first sentence  $S_0$  and a proposition  $Q$  from the second sentence  $S_1$ . Coherence results from these two propositions being related; in this case an implication relationship is inferred between the two. The Cause-Effect relations are summarized in Table 3.

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

Table 3: Cause-Effect Relations

## 4.3 Contiguity Relations

The sole relation in the Contiguity category is *Narration*. The Narration relation allows one to express a coherent sequence of events centered around some system of entities.

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

Further discussion of this relation is given in Kehler (1995).

## 5 The Analysis

One of the main differences between the constraints that Resemblance and Cause-Effect relations impose is in the type of arguments over which they are applied. Establishing a Resemblance relation requires access to the semantics of subclausal constituents in the source and target sentences, that is, the relations  $p_i$  and the corresponding elements  $a_i$  and  $b_i$ . In contrast, Cause-Effect relations require access only to the clause-level semantics, that is, the  $P$  and  $Q$ . We now ask how this difference affects the constraints on the syntactic

form used in the two clauses being related, in light of syntactic operations such as elision and extraction.

Put briefly, to explain the VP-ellipsis and gapping data noted earlier, in Kehler (1994a) and Kehler (1995) we posited that elided information in the syntax of the target is recovered in just those cases in which the coherence resolution mechanism needs to access the semantics of syntactic nodes within the elided material. This will be the case when Resemblance relations are operative, because the arguments to the relations (the  $p_i$ ,  $a_i$ , and  $b_i$ ) will generally be associated with nodes within the elided material. In contrast, the elided syntactic material need not be restored when Cause-Effect relations are operative, because the arguments to the relations (the  $P$  and  $Q$ ) are associated with the clause-level sentence nodes, which are always present. Therefore, the missing syntax in VP-ellipsis and gapping constructions is recovered only when a Resemblance relation is operative, accounting for why they require a suitably matching syntactic antecedent in these cases.

However, VP-ellipsis and gapping differ in that VP-ellipsis is least constrained when a Cause-Effect relation is operative (requiring only a suitable semantic antecedent without regard to syntactic parallelism), whereas gapping is most constrained in this case (not being felicitous at all). We argue in Kehler (1994a) and Kehler (1995) that this results from the referential properties of the two forms. Particularly, VP-ellipsis is referential; it behaves like a pronoun in that it can refer cataphorically, as well as have antecedents that are evoked from clauses other than the one immediately prior to it. Therefore, in Cause-Effect relations a purely semantic antecedent can be recovered without regard to syntactic parallelism. By similar reasoning, gapping is shown not to be referential (i.e., it does not allow cataphora nor reference to antecedents more than one clause back), so it is infelicitous in Cause-Effect relations because its meaning cannot be recovered without reconstruction of the missing syntax, which comes about only with Resemblance relations.

The constraints on extraction out of coordinate constructions can also be seen to arise out of the interaction between a syntactic operation and coherence resolution processes. Kuno (1987) and others have noted that for extraction to take place, the extracted element must be able to serve as the *topic* of the clause (in some appropriate sense) from which it is extracted:

*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)

An element that is extracted from a conjoined clause must therefore be able to serve as a topic for the conjoined clause. The situation differs for this constraint with respect to the type of coherence relation that is operative between the

clauses. Resemblance relations are coherent by virtue of the fact that they share a common topic, identifiable from the common relation and the common properties of the corresponding elements over which the relation applies (at some, perhaps inferred, level of generalization). Therefore, the topic of a Resemblance construction must be common to each of the conjuncts; that is, if an element in one of the clauses gives rise to a topic, then its corresponding element in the other clause must also. Extraction can thus occur only if both elements are extracted to the topic-denoting position. Of course, this is possible only when the elements are in fact the same, resulting in the “across-the-board” type of extraction. In contrast, there is nothing prohibiting an element in one clause from serving as the topic of a pair of clauses related by a Cause-Effect relation; we therefore do not see similar constraints on extraction in those cases. Finally, the topic of a set of clauses related by Narration need not be mentioned in every clause (in particular, not in scene-setting clauses), so extraction need not take place out of all conjuncts. On the other hand, extraction must take place out of certain of the conjuncts that do not serve a scene-setting function, insofar as an inability to do this would suggest that the extracted element is no longer the topic of that part of the narration.

## 6 Conclusions

The data discussed herein suggest that there is no purely syntactic Coordinate Structure Constraint operative in natural language grammar, and in fact that it is unlikely that any purely syntactic explanation will be able to account for the data involving extraction from coordinate structures. Instead, as with the data involving VP-ellipsis and gapping, any such explanation needs to take into consideration the type of coherence relation operative between the clauses over which the relevant syntactic operation applies. With each of these phenomena, the strongest constraints on choice of syntactic structure appear to be at play when a Resemblance relation is operative between the clauses; such constraints have independent motivation from the perspective of coherence resolution mechanisms and the syntax/discourse interface.

## Acknowledgements

I would like to thank Jerry Hobbs, who brought Lakoff’s article to my attention, and Mary Dalrymple for valuable discussions. This research was funded by the National Science Foundation and the Advanced Research Projects Agency under Grant IRI-9314961 (Integrated Techniques for Generation and Interpretation).

## References

- Goldsmith, John. 1985. 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.
- Grosu, Alexander. 1973. On the nonunitary nature of the coordinate structure constraint. *Linguistic Inquiry*, 4:88–92.
- Haïk, Isabelle. 1987. Bound variables that need to be. *Linguistics and Philosophy*, 10:503–530.
- Hume, David. 1748. *An Inquiry Concerning Human Understanding*. The Liberal Arts Press, New York, 1955 edition.
- Kehler, Andrew. 1993. 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.
- Kehler, Andrew. 1994a. 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.
- Kehler, Andrew. 1994b. A discourse processing account of gapping and causal implicature. Manuscript presented at the Annual Meeting of the Linguistic Society of America, Boston, January.
- Kehler, Andrew. 1995. *Interpreting Cohesive Forms in the Context of Discourse Inference*. Ph.D. thesis, Harvard University.
- Kuno, Susumu. 1987. *Functional Syntax - Anaphora, Discourse and Empathy*. The University of Chicago Press, Chicago and London.
- Lakoff, George. 1986. Frame semantic control of the coordinate structure constraint. In *Papers from the Twenty-Second Annual Regional Meeting of the Chicago Linguistic Society*, Chicago. Chicago Linguistic Society.
- Levin, Nancy and Ellen Prince. 1982. Gapping and causal implicature. Presented at the Annual Meeting of the Linguistic Society of America.
- Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. CSLI and University of Chicago Press.
- Ross, John R. 1967. *Constraints on variables in syntax*. Ph.D. thesis, MIT.