



Published in final edited form as:

Linguist Philos. 2006 June ; 29(3): 315–346.

Ellipsis and discourse coherence

Lyn Frazier and Charles Clifton Jr

University of Massachusetts

Abstract

VP-ellipsis generally requires a syntactically matching antecedent. However, many documented examples exist where the antecedent is not appropriate. Kehler (2000, 2002) proposed an elegant theory which predicts a syntactic antecedent for an elided VP is required only for a certain discourse coherence relation (resemblance) not for cause-effect relations. Most of the data Kehler used to motivate his theory come from corpus studies and thus do not consist of true minimal pairs. We report five experiments testing predictions of the coherence theory, using standard minimal pair materials. The results raise questions about the empirical basis for coherence theory because parallelism is preferred for all coherence relations, not just resemblance relations. Further, strict identity readings, which should not be available when a syntactic antecedent is required, are influenced by parallelism per se, holding the discourse coherence relation constant. This draws into question the causal role of coherence relations in processing VP ellipsis.

1. INTRODUCTION

There is a long-standing debate in the linguistic literature about the correct approach to VP ellipsis. One approach claims that only a semantic antecedent is required for an elided VP (Dalrymple et al., 1991, Hardt, 1993, 1999). This approach seems to overgenerate, generally allowing ellipsis without a formally or syntactically appropriate antecedent, e.g., an active VP antecedent if the elided clause calls for an active VP, a passive VP if the elided VP is passive. The other major approach is syntactic and essentially it requires syntactic reconstruction at the ellipsis site (Fiengo and May, 1994, Hestvik, 1995, Sag, 1976, Williams, 1977.). It tends to undergenerate, not accounting for attested examples where no syntactically appropriate antecedent is available for an elided constituent.

In several papers, Kehler (1993, 1995, 2000, 2002) proposed that which conditions are operative for VP ellipsis depends on which particular discourse coherence relations the ellipsis clause instantiates. He argued that three distinct types of discourse coherence relations exist: resemblance relations and cause-effect relations, the two coherence relations on which we will focus, and contiguity relations, which often involve a sequence of events. Resemblance relations allow sentences of a discourse to be connected by emphasizing the similarity or contrast between entities or events. These will be recognized “using comparison and generalization operations” (Kehler, 2002, p. 542). In an example like (1), the similar entities are Bill and Al, and the common relation is participation in a recreational activity.

(1) Bill likes to play golf. Al enjoys surfing the net. (= Kehler, 2000, example 17a) With resemblance relations, an elided VP should be ‘syntactically sensitive.’ The elided VP should be grammatical only if there is a syntactically appropriate antecedent available. Further, if there is an elided VP in a clause instantiating a resemblance coherence relation, its syntactic representation is reconstructed. Therefore, it is natural to expect syntactic principles to apply

to the elided constituent in such cases, e.g., the conditions on binding. Finally, on Kehler's view, discourse coherence relations influence not only the type of ellipsis that is possible, but also the resolution of anaphors.

Turning to Cause-effect relations, the canonical case is the Result relation: "Infer P from the assertion of S0, and Q from the assertion of S1, where normally $P \rightarrow Q$." (Kehler, 2000, p. 541). The sort of implication that Kehler has in mind is that Q "plausibly follows from" P. Additional types of Cause-effect relations include Explanation (because), Violated expectation (but), and Denial of preventer (even though). The core of the Cause-effect relation is that the hearer is establishing a path of implication and the arguments of the relation are propositions. When ellipsis occurs in a clause participating in a Cause-effect relation it is a semantic relation that is reconstructed (apparently permitting the familiar constraints on 'surface' vs. 'deep' anaphora to be overcome; cf. Hankamer & Sag, 1976; Tanenhaus & Carlson, 1990; see also Simner, Garnham & Pickering, 2002, for discussion of did it as a VP anaphor vs. as the main verb do plus a pronoun). For example, in (2) with a cause-effect relation, VP ellipsis is permitted despite the absence of a syntactically appropriate antecedent.

(2) This problem was to have been looked into, but obviously nobody did. [look into the problem] (Vincent Della Pietra, in conversation, cited by Kehler, 2002).

Kehler (2002) classifies (2) as a cause-effect relation because it can be viewed as an implicational relation, and he contrasts (2) with the resemblance relation in (3).

(3) # This problem was looked into by John, and Bob did too. [look into the problem]

In a discussion of VP-ellipsis, Kehler (1995) proposed that a VP can elide only when all subsequently needed information is recoverable. That includes the semantics of the VP and the arguments to the discourse coherence relation. Cause-effect relations are established using only the propositions denoted by clauses, not by subclausal constituents, so no syntactic reconstruction is necessary to recover the arguments of the discourse coherence relation. By contrast, resemblance relations require that corresponding sub-clausal semantic arguments be identified and aligned. As Kehler notes, finding corresponding arguments will be facilitated when they occur as syntactically parallel arguments. Eliding a VP generally indicates that these corresponding arguments are shared and recoverable by reconstruction.

The idea that the type of discourse relation involved influences both ellipsis and anaphora resolution is appealing, and Kehler presents various intuitions in support of it. These include not only the acceptability of syntactically inappropriate antecedents for an elided constituent, but also intuitions about binding theory violations, e.g., the general requirement that a pronoun must be free in its local domain. For example, Kehler contrasts the cause-effect example in (4), where coreference between him and John is claimed to be acceptable, with the coreference in the similar resemblance relation example in (5), which is claimed to be unacceptable.

(4) John's_i lawyer defended him_i because he_i wouldn't.

(5) #John's_i lawyer defended him_i, and he_i did too.

Recent investigations of discourse coherence and pronoun resolution in sentences without VP ellipsis provide some experimental support for Kehler's ideas. Wolf, Gibson and Desmet (2004) examined parallel and nonparallel reference in sentences with resemblance or cause-effect relations, using sentences like those in (6). In a self-paced reading study, they found increased reading times for the pronoun when there was a nonparallel antecedent ONLY in the resemblance relation sentence.

(6)

a. Resemblance, Parallel Reference

Fiona complimented Craig and similarly James congratulated him after the match but nobody took any notice.

b. Resemblance, Nonparallel Reference

Fiona complimented Craig and similarly James congratulated her after the match but nobody took any notice.

c. Cause-Effect, Parallel Reference

Fiona defeated Craig and so James congratulated him after the match but nobody took any notice.

d. Cause-Effect, Nonparallel Reference

Fiona defeated Craig and so James congratulated her after the match but nobody took any notice.

Smyth (1994) and Chambers and Smyth (1998) also investigated “parallel function” effects in processing pronouns. We will take up those studies in the General Discussion, where they are most relevant.

Kehler’s theory embodies distinct claims. One claim is that discourse coherence relations provide the right dimension for distinguishing syntactically sensitive versus syntactically insensitive VP ellipsis. An even stronger claim is that the appropriate account of syntactically sensitive ellipsis involves syntactic reconstruction, where (some morphological details aside) the syntactic representation of an antecedent is copied into the elided VP. These claims may be assessed separately. However, what makes his theory elegant and explanatory is the connection between the operations needed to establish the appropriate discourse coherence relation and the applicability/non-applicability of syntactic conditions. On his view, the correlation between a particular discourse coherence relation and whether ellipsis is syntactically sensitive is not arbitrary. When only propositions (no sub-clausal constituents or entities) are required to establish the coherence relation, as in Cause-effect relations, syntactic conditions do not govern the ellipsis. When sub-clausal constituents/entities must be recognized in order to establish the similarities or contrasts needed for a (Resemblance) discourse coherence relation, then syntax becomes relevant and syntactic conditions on the elided material must be obeyed. In short, Kehler’s theory should probably be evaluated in terms of the separate claims involved but at the same time we should not lose sight of the fact that the explanatory core of his theory rests on the nonarbitrary connection between the information required for the establishment of a particular discourse coherence relation and the particular grammatical conditions that come into play.

Kehler’s theory is intriguing and theoretically important. If there really is a categorical distinction between ‘syntactic’ ellipsis and ‘semantic’ ellipsis that correlates with the distinction between kinds of coherence relations, this would solve many outstanding problems with theories of ellipsis, coordination and binding. The simplest form of theories of ellipsis, coordination and binding have numerous counterexamples, many of them summarized by Kehler (2002). Kehler proposes limiting the scope of the restrictive form of the theories (the form containing the most stringent syntactic constraints) to only a proper subset of discourse coherence relations. If limiting the scope of the theories would in fact eliminate the counterexamples, this approach could both simplify linguistic theory and capture an important generalization about linguistic structures. However, our own judgments are often less crisp than those presented by Kehler and we have doubts about the prediction that ellipsis in cause-

effect relations is syntactically insensitive and doubts about the role parallelism is attributed in his theory. While corpus data are surely valuable, relying on corpus data alone (as Kehler has largely done) has its limitations, we think, because the occurrence of minimal pairs will be rare. Linguists' intuitions about minimal pairs can be used to supplement the corpus data. But, especially when intuitions are subtle, intuition (or the assessment of it) may vary unintentionally depending on a person's beliefs. A formal comprehension experiment on nonlinguists allows judgments to be gathered without worrying about the unintended influence of current linguistic theories, though of course the experimental data will have their own limitations (typically underestimating our abilities due to the inclusion of some unmotivated or careless subjects).

Below we report several comprehension experiments that provide a test of the predictions of Kehler's theory. Experiments 1 and 2 test the prediction that a syntactically appropriate antecedent is required only with resemblance relations, not cause-effect relations. These experiments evaluate the claim that the relevant dimension for sorting ellipsis into syntactically sensitive vs. syntactically insensitive types is the discourse coherence relation involved. Experiment 3 turns to the prediction that rules of binding theory, in particular Principle A requiring a reflexive to be bound in its local domain, must be obeyed in ellipsis involving resemblance relations but not in ellipsis involving cause-effect relations. Experiments 4 and 5 look at the effect of syntactic parallelism between the antecedent clause and the ellipsis clause on the preference for strict or sloppy interpretations of pronominals (or reflexives) in examples where the discourse coherence relation does not vary. These experiments bear on the stronger claim that syntactically sensitive instances of ellipsis involve reconstruction of the syntax of the elided clause.

2. EXPERIMENT 1

The first experiment was an attempt to determine if it is true that listeners would reject resemblance sentences that don't have a parallel antecedent, whereas they would accept such sentences if they involve a cause-effect relation. Listeners heard sentences like those in (7) and were asked to indicate whether they understood the sentence to their satisfaction or not. The frequency with which they indicated comprehension difficulty and their reaction time were recorded. The straightforward prediction from Kehler's theory is that the lack of a syntactically appropriate antecedent, in examples with an initial passive clause and a final active clause, will result in higher acceptability and more frequent and possibly faster understanding when the two clauses are conjoined with because, which induces a causal relation, than when they were conjoined with a non-causal 'resemblance' connective such as just like.

2.1. Method

2.1.1. Materials—Sixteen sentences like (7) were constructed with four versions of each. All sentences had a passive first clause followed by an active second clause containing VP ellipsis (did). Two (7a.,b) contained the causal connective because. Two (7c,d) involved a resemblance relation, just like for eight sentences and and...did too for eight sentences (see Appendix 1). The b and d versions contained an adverb after the agent simply to keep the agent out of clause-final position which may be particularly prominent. We did not have any specific predictions about the effect of the adverb. However, one could imagine that focusing the agent of the first clause might matter when it is the constituent that contrasts with the unelided constituent of the second clause.

(7)

- a. The problem was looked into by Kim because Lee did. (Because, no adverb)
- b. The problem was looked into by Kim last time because Lee did. (Because, adverb)

- c. The problem was looked into by Kim just like Lee did. (Resemblance, no adverb)
- d. The problem was looked into by Kim last time just like Lee did. (Resemblance, adverb)

Kehler's theory leads us to expect that (7c,d) should be very difficult indeed, because listeners will expect VP ellipsis in resemblance relations only with a syntactically appropriate antecedent. This should lead to low acceptance rates and possibly also long judgment times relative to the causal sentences, where a syntactically appropriate antecedent is not required.

The resulting 64 sentences were included in four counterbalanced lists, each list containing 4 sentences in each of the four forms illustrated in (7). The complete list of sentences appears in Appendix 1. The 16 experimental sentences in each counterbalanced list were included in a total of 90 sentences of a variety of forms, including sentences with possible pragmatic anomalies (A few students often failed the final exam) and ambiguities (Rene realized the committee debated the bill, not Lenora). These sentences included intonational phrase breaks at points of theoretical interest, and some of the remaining filler sentences, highly variable in form, had pragmatically questionable pitch accents, but the present 16 experimental sentences simply contained an intonational phrase boundary before the connective and each content word or phrase received an H* pitch accent. A practice list of seven sentences (none of them of the form illustrated in (7)) was also constructed.

2.1.2. Participants and procedures—Forty-eight University of Massachusetts students were tested in individual half-hour sessions. They were instructed that they would hear a series of sentences and should indicate whether or not they understood the sentence to their satisfaction by pulling one of two response triggers. They should pull the right-hand trigger if they “got” the sentence and the left-hand trigger if they did not “get it.” They were told to trust their intuitions, and to pull the left-hand trigger if they found the sentence confusing or they didn't really understand it, and were told to pull a trigger as quickly as possible after the end of the sentence. If they pulled the right-hand trigger, indicating understanding, they were visually presented with a simple question about the sentence (some true/false, some two-choice wh-questions) to ensure that understanding was satisfactory.

The experiment took place in a sound-attenuated chamber. Participants heard digitized sentences played by a computer over external speakers, at a comfortable listening level. The experimental session began with the seven-item practice list, and continued after a short break with the one of the four lists of 64 sentences including the 16 experimental sentences, all presented in individually-randomized order. Twelve participants heard each list.

2.2. Results and Discussion

Table 1 presents the mean percentages of times that sentences received a “got it” (understand) response and the mean reaction times for all responses.¹ No RT effects approached significance (all $F < 1.0$), and no effect of presence of adverb approached significance. However, causal sentences were accepted less often than their resemblance counterparts ($F(1,4) = 5.92$, $MSe = 0.05$, $p < .02$; $F(1,15) = 4.24$, $MSe = 0.02$, $p = .06$).² This is exactly the opposite of the pattern predicted by Kehler's theory. However, although the lack of a preference for the causal sentences is informative, their actual dispreference is not. It is possible that the causal sentences were not acceptable for some other reason, not the ellipsis per se. Experiment 2 investigated similar sentences including both those with syntactically appropriate and syntactically

¹Because the nearly 50-50 division of responses resulted in too many missing observations, it was not possible to analyze RTs to the two types of responses separately.

inappropriate antecedents to determine whether the difficulty of the Experiment 1 causal sentences was specific to the inappropriate syntactic form of the antecedent.

3. EXPERIMENT 2

Sixteen sentences like those in (8) were tested in a written acceptability judgment experiment. Experiment 2 had several purposes. One was to generalize the lack of a penalty for nonparallel resemblance sentences from auditory to written presentation. A second was to compare the effects of syntactic parallelism vs. nonparallelism on sentences with causal vs. resemblance relations. A third was to eliminate the possibility that the causal sentences in Experiment 1 were frequently rejected because of reasons unrelated to the syntactic mismatch between antecedent and elided constituent by ensuring that causal and resemblance sentences are equally acceptable in their syntactically matching forms. To this end, sentences were modified from those used in Experiment 1 by modifying the least-often accepted sentences. In most cases this was done by replacing the causal connective because with even though to eliminate unlikely causal relations.

(8)

- a. The problem was looked into by Kim even though Lee did.
- b. The problem was looked into by Kim just like Lee did.
- c. Kim looked into the problem even though Lee did.
- d. Kim looked into the problem just like Lee did.

3.1 Method

3.1.1. Materials—Sixteen experimental sentences like those in (7) were constructed (all appear in Appendix 2). There were four forms of each sentence, defined by the factorial combination of passive vs. active first clause and causal vs. resemblance connective. The second clause was always active, so a passive first clause results in a mismatching construction. The resulting 64 sentences were divided into four counterbalanced lists as in Experiment 1 and combined with 75 other items of a wide variety of constructions in a written questionnaire.

3.1.2. Participants and procedures—Forty-eight University of Massachusetts students completed a written questionnaire in individual half-hour sessions. Twelve completed each counterbalanced form of the questionnaire. They read instructions indicating that they were to rate the acceptability of some sentences on the questionnaire on a five-point scale (they were told that other sentences were to be evaluated for their meaning and that they were to indicate which of two alternatives best expressed the meaning of the sentence). They were told to indicate that the sentence is a “1” (unacceptable) if it is one that they would not say except by mistake or one that they would not expect to hear a native speaker of English saying. They were told to circle an intermediate number (e.g., 3) if they think that the sentence is one you’d probably not say and one that you’d be surprised or slightly confused if you heard an English speaker saying. They were told to indicate that the sentence is a “5” (acceptable) if it is one

²Additional analyses were conducted dividing the Experiment 1 sentences into the eight just like and the eight and....did too resemblance relations. The advantage of resemblance over cause-effect sentences was equally present for both the first eight sentences (60.5 vs 54.1% accepted) and the second eight sentences (59.6 vs. 52% accepted). We call attention to Experiment 4, in which we explicitly manipulated the presence of did too in an attempt to manipulate parallelism, but note that did too did not induce structural parallelism in Experiment 1 because an active second clause was conjoined to a passive first clause in this experiment. In addition, we analyzed the first vs second half of each participant’s data separately, to see if the observed difference between causal and resemblance sentences held true for both halves. If anything, the difference was somewhat larger in the first half of the experiment, where it averaged 11%, diminishing to 8% over the course of the entire experiment.

they could easily imagine themselves or another native speaker of English saying or hearing and not noticing anything odd about it.

3.2. Results and Discussion

The mean ratings appear in Table 2. The goal of equating the causal and resemblance sentences in the matching (active-active) forms was clearly met. Ratings for both causal and resemblance sentences were significantly reduced in the mismatching (passive-active) form ($F(1,47) = 134.9, p < .001$; $F(1,15) = 70.2, p < .001$). The size of the reductions did not differ between causal and resemblance sentences (interaction $F_s \ll 1.0$). Post-hoc analyses were conducted separately on the four items with an even though causal relation and the 12 items with a because relation. The same pattern was observed for each set of items. The effect of matching antecedent may have been somewhat smaller for the even though than the because items, with means of 3.47 vs 4.28 for mismatching vs matching even though sentences and means of 2.92 vs 4.37 for because sentences (a difference that cannot be treated too seriously because of the small number of even though items). Most importantly, the effects of syntactic matching in each set of items was very closely similar for causal and resemblance sentences, never differing by more than 0.09 points on the rating scale.

The results of Experiment 2 show a clear advantage for sentences with a syntactically appropriate antecedent for both the causal and the resemblance discourse coherence relations. The causal relation did not influence acceptability nor did it interact with syntactic form.

3.3. Discussion: Experiments 1 and 2

The results of Experiments 1 and 2 do not overall provide much comfort for the discourse coherence approach to ellipsis. Examples like (2) and (3) provided some of the central evidence for Kehler's theory. Yet, when we manipulate the syntactic appropriateness of the antecedent in a systematic way, without changing focus structure across examples or the presence of presuppositional elements (e.g., too), we find that intuitions do not vary in the predicted manner. Instead, what we find is that syntactic matching is preferred for resemblance relations AND for cause-effect relations. We will take up the importance of this observation, and possible limitations, below after reporting three experiments on the processing and interpretation of pronominals and reflexives in VP ellipsis sentences.

4. EXPERIMENT 3

In Experiment 3 we test the claim that a reflexive may have a strict reading only with a causal discourse coherence relation, not with a resemblance coherence relation. Prior psycholinguistic research on processing VP ellipsis has exploited a priming technique to show that the strict interpretation ("the fireman defended the policeman") is computed, at least temporarily, in sentences like (9), where the strict interpretation is clearly unpreferred (Shapiro & Hestvik, 1995). This has been reported to occur even in sentences like (10) where the strict reading is impossible due to the nature of the verb (Shapiro, Hestvik, Lesan & Garcia, 2003). These results suggest that the strict reading is computed even for sentences with a resemblance coherence relation and even for sentences which ultimately may not receive a strict interpretation.

(9) The policeman defended himself and the fireman did too, according to someone who was there.

(10) The policeman perjured himself and the fireman did too.

The basic finding is that words semantically related to the first clause subject, the antecedent of the reflexive on the strict reading, are activated following *did*. The idea is that in copying the VP from the first clause, the index on the reflexive also gets copied, resulting in activation

of the phrase which binds the index (policeman). (See also Frazier & Clifton, 2000, for studies of bound variable interpretations of pronominals, including but not limited to examples in VP ellipsis contexts.)

Note that the strict interpretation of (9) and (10) requires that the elided reflexive is not bound in its local domain, violating Principle A (Chomsky, 1981). Kehler's theory claims that Principle A must be obeyed in ellipsis involving resemblance relations. However, if 'semantic' ellipsis is possible when cause-effect relations are involved, Principle A may be ignored. Thus, Kehler's proposal predicts that a strict interpretation of a reflexive in an elided VP is available only when a cause-effect, not a resemblance, relation is involved. This prediction was tested in a written questionnaire study using sentences like those in (11), including examples in which the resemblance relation was conveyed by a simple conjunction (see Appendix 3).

(11)

- a. Doug blamed himself for the band's collapse because everyone else did.
- b. Doug blamed himself for the band's collapse just like everyone else did.

4.1 Method

4.1.1. Materials—Sixteen sentences like those in (11) were constructed, with a because and a just like or and version of each. The examples were intuitively open to both a strict and a sloppy interpretation. They were included in a written questionnaire containing 74 other sentences from different experiments, all of which were unambiguous and were to be rated for acceptability. Multiple counterbalancing forms of the questionnaire were constructed so that half the forms had the causal version of each sentence and half had the resemblance version. Each form of the questionnaire presented all the sentences in a different random order.

4.1.2. Participants and Procedure—Forty eight University of Massachusetts undergraduates completed one form of the questionnaire each, working individually. They were instructed to choose one of two possible interpretations of each Experiment 3 sentence, as indicated by paraphrases (e.g., Everyone else blamed Doug/Everyone else blamed himself). They were also instructed to rate the remaining 74 sentences (treated as fillers from the perspective of Experiment 3) on a five-point scale for acceptability. However, the Experiment 3 sentences were not rated for acceptability.

4.2. Results and Discussion

The causal relation sentences (11a) received an average of 53% strict interpretations ("Everyone else blamed Doug"), while the resemblance sentences (11b) received an average of 48% strict interpretations. These values did not differ significantly ($t(47) = 1.07$).³ There was no substantial difference in the probability of assigning a strict interpretation to a reflexive in VP ellipsis in cause-effect relations vs. in resemblance relations. According to Kehler's hypothesis, a strict interpretation should have been available only in the cause-effect relations (11a). Thus the results of Experiment 3 again disconfirm the hypothesis.

³Note that half of the experimental items required a change in number of the reflexive for the sloppy reading (himself -> themselves), while half did not. Separate analyses of these two halves indicated a numerically greater tendency to assign a strict reading to the former (number change) than the latter (no number change) items, 61 vs 39%, but this difference was not significant by items ($F(2(1,14)) = 2.40$, $p = .14$). There was some hint of a greater preference for the strict reading in the no number change items for cause-effect than for resemblance sentences, a difference of 8% as opposed to a difference of 1% for number change items, but the interaction between number change and type of relation did not approach significance ($F(2(1,14)) = 0.31$).

5. EXPERIMENTS 4A AND 4B

The question now is why intuitions do go along with Kehler's hypothesis in the examples he cites. We suspect this is not due to the particular type of discourse coherence relation, but due to parallelism itself (correspondence between the syntax of the antecedent clause and the syntax of the ellipsis clause) along with the effects of presuppositional items like too which reinforce parallelism. In many of Kehler's examples of VP-ellipsis with resemblance relations too is included (virtually always in sentences involving the connector and). With cause-effect connectives, too is seldom present. Kehler (2002) reports a particularly relevant example (his example 102), presented here as (12).

(12)

- a. This problem was looked into by John, even though Bob already had.
- b. #This problem was looked into by John, even though Bob already had too.

In this example, it is only the inclusion of too that makes the example completely unacceptable. To test this hypothesis, Experiment 4 examined the interpretation of ambiguous VP ellipsis sentences in highly parallel sentences (with did and too) vs. sentences that were less parallel (e.g., with an auxiliary like would, and without too), using only sentences with a resemblance relation. Its goal was to explore the possibility that sloppy interpretations are simply more preferred when the conjoined clauses are made more highly parallel by adding a final too. If this proves to be the case, then it is possible that Kehler's observation that strict interpretations were relatively more acceptable with cause-effect relations and sloppy interpretations with resemblance relations could be attributed not to the difference in type of relation but instead to the frequent use of a final too in the resemblance relation sentences he examined.

A written questionnaire study (Experiment 4A) and an online comprehension study (Experiment 4B) investigated the interpretation of sentences like those in (13) and (14).

(13)

- a. John saw a snake near him and Bill did too.
- b. John saw a snake near his backpack and Bill did too.
- c. John saw a snake near him and Bill would soon.
- d. John saw a snake near his backpack and Bill would soon.

(14)

- a. Fred kicked a cockroach away from him and Henry did too.
- b. Fred kicked a cockroach away from his sleeping bag and Henry did too.
- c. Fred kicked a cockroach away from him or Henry did.
- d. Fred kicked a cockroach away from his sleeping bag or Henry did.

Two forms of these VP ellipsis sentences (the a and c forms in (13) and (14)) contained the personal pronoun him or her in the first clause. The remaining two forms (b and d) contained the possessive pronoun his or her. This manipulation was included simply to explore whether possessive pronouns, like reflexives, might promote the sloppy interpretation (independent of the degree of parallelism manipulation that was the main focus of the experiments). In the elided VP, the pronoun could be interpreted as bound by the second clause subject (the sloppy reading) or it could be coreferential with the first clause subject (the strict reading). Crucially, the a and b forms of the sentences were highly parallel: the two clauses were conjoined with and and the second clause contained the word too. By contrast, the c and d forms were less

parallel and never contained the word *too*. In half the items, the c and d form contained an auxiliary (only) in the second clause, as in (13). In the other half, the c and d form contained *or* as the clausal connective. However, none of the sentences, in any form, involved a cause-effect relation and all of them at least arguably involved a resemblance relation.

The question is whether the parallel forms will result in fewer strict interpretations and more sloppy interpretations than the less parallel forms, even though the discourse coherence relation remains a resemblance relation. Kehler's theory does not have any basis for predicting a difference, but such a difference may be at the root of many of the example sentences he has presented in support of his theory.

5.1. Method

5.1.1. Experiment 4A: Questionnaire—Forty-eight University of Massachusetts students completed a written questionnaire. Eight sentences like (13), and eight like (14) appeared in the questionnaire, in one of the four versions illustrated in (13) and (14). Each sentence was followed by an appropriate two-choice question, e.g. Where was the snake that Bill saw? Near John ____ Near Bill ____, or Where would the snake that Bill saw be? Near John's backpack ____ Near Bill's backpack. The question was designed to determine whether the participant gave the preceding sentence a strict or a sloppy interpretation. The strict reading alternative was always presented as the first of the two answers.

The resulting 64 sentences (16 items, each in four versions) were assigned to four counterbalanced forms of the questionnaire, with each questionnaire form containing four items in each version. The 16 items in one questionnaire form were randomly intermixed with 56 other sentences, all unrelated to the Experiment 4A items, and each followed by a two-choice or a YES-NO question. Each participant completed the written questionnaire individually, working at his or her own pace.

5.1.2. Experiment 4B: On-line comprehension—Forty-eight University of Massachusetts students were tested in individual half-hour sessions. A self-paced phrase-by-phrase reading procedure was used. Participants read the 16 sentences described in Experiment 4A in four counterbalanced lists (for instances of each sentence form in each list) as a part of a list of 124 sentences in all, including sentences from other unrelated experiments and filler sentences.. Order of presentation was individually randomized, and the experimental list was preceded by a 7-item practice list. A trial began with the video screen displaying an underscore where each character of the words of the sentence would appear. When the participant pressed a thumb trigger, the first presentation region replaced the corresponding underscores; each successive trigger press replaced the current presentation region with underscores and presented the next region. Each Experiment 4B sentence was presented as two separate regions, clause by clause. Participants were urged to read as quickly as they comfortably could, and the time taken to read each region was measured (but proved not to be relevant to the current experimental topic). Following the end of the sentence, a question with two possible answers, similar to those used in Experiment 4A, appeared on the screen. One answer was congruent with the sloppy interpretation, and one with the strict. The participant pulled a response trigger under the answer that fit his/her interpretation of the sentence.

5.2 Results and Discussion

5.2.1. Experiment 4A—The left-hand data column of Table 3 presents the percentages of sloppy interpretations of sentences in the questionnaire. Analyses of variance indicated a significantly higher frequency of sloppy interpretations for the parallel (13 and 14a,b) than the nonparallel (13 and 14c,d) versions of the sentences ($F(1,47) = 9.88, p < .01$; $F(1,15) = 4.56, p < .05$). Additional analyses contrasting the eight sentences like (13) and the eight like (14)

indicated that there was no significant effect or interaction of the different ways of manipulating parallelism, nor was there any effect of whether the pronoun was a personal pronoun or a possessive.

5.2.2. Experiment 4B—The right-hand column of Table 3 presents the percentages of sloppy interpretations of sentences that were read in an on-line, speeded fashion. Clearly, the parallel sentences received more sloppy interpretations than the nonparallel ones ($F(1, 47) = 26.66$, $MSe = .04$, $p < .001$; $F(1, 15) = 8.99$, $MSe = .04$, $p < .01$.) Whether the pronoun was a personal pronoun or a possessive had no significant impact on interpretations (maximum $F = 1.54$).

The results of Experiment 4 suggest that manipulating parallelism *per se* has an effect on the interpretation that is preferred for a pronominal in a VP ellipsis context. Manipulating the discourse coherence relation is not necessary, suggesting that it is possible that some of the judgments on which Kehler based his theoretical claims may have actually reflected differences in parallelism, not differences in the type of relation between the clauses.

6. EXPERIMENT 5

Experiment 4 tested claims about the ungrammaticality of the strict identity interpretation of pronominals in sentences with a resemblance discourse coherence relation. Experiment 5 tested similar sentences with a reflexive instead of a pronominal as in (15) and (16) in a written questionnaire like that used in Experiment 4A. As discussed in the introduction to Experiment 3, according to Kehler's theory, we would expect only sloppy ('bound') responses for the sentences with reflexives (15b,d) and (16b,d).

(15)

- a. John kicked a snake away from him and Bill did too.
- b. John kicked a snake away from himself and Bill did too.
- c. John kicked a snake away from him and Bill would soon.
- d. John kicked a snake away from himself and Bill would soon.

(16)

- a. Brad detonated a bomb 20 yards away from him and Willy did too.
- b. Brad detonated a bomb 20 yards away from himself and Willy did too.
- c. Brad detonated a bomb 20 yards away from him or Willy did.
- d. Brad detonated a bomb 20 yards away from himself or Willy did

6.1 Method

Four versions of 16 sentences like those in (15) and (16) were included in the questionnaire described in Experiment 2 (see Appendix 4). All procedures described for Experiment 2 hold for Experiment 5. Note, though, that the Experiment 5 items were evaluated for meaning as in Experiment 4A rather than for acceptability, as in Experiment 2. A pair of alternative answers followed each sentence, as in Experiment 4A.

6.2 Results and discussion

Table 4 shows that bound ('sloppy') interpretations were significantly more frequent for parallel than for nonparallel sentences ($F(1, 47) = 52.58$, $p < .001$; $F(1, 15) = 19.55$, $p < .001$) and significantly more frequent for reflexives than for simple pronouns ($F(1, 47) = 26.09$, p

< .001; $F_2(1,15) = 18.78$). The size of the effect of parallelism was closely comparable for reflexives and pronouns (interaction $F_1(1,47) = 1.60$, $p > .20$; $F_2(1,15) = .1.40$, $p > .20$).

Both parallelism and the nature of the pronominal, reflexive versus pronoun, increased the number of bound ('sloppy') interpretations, but the two factors did not interact. This shows that the results of Experiment 4 were not peculiar to the type of pronominal tested: Parallelism increases the number of sloppy interpretations comparably for simple pronouns, possessive pronouns, and reflexives. We suspect that the presence of too in the parallel cases plays a critical role in our examples, as well as in many of Kehler's. Note also that although Kehler's analysis claims that only the sloppy reading would be available for resemblance sentences with reflexives, the frequency of sloppy readings did not approach 100% and in fact exceeded 50% only for the parallel sentences.

7. GENERAL DISCUSSION

At the level of empirical generalizations, Experiments 1 and 2 provide evidence that even in VP ellipsis sentences without syntactically appropriate antecedents, a resemblance relation (just like) may be considered at least as acceptable as a corresponding sentence with a cause-effect relation. Further, sentences which contain a syntactically appropriate antecedent for the elided constituent are rated more acceptable than ones that don't even if they involve a cause-effect relation. In short, lack of a syntactically appropriate antecedent leads to lesser acceptability in VP ellipsis sentences in general, not just with resemblance relations. Turning to strict and sloppy identity in VP ellipsis sentences, native speakers apparently allow a strict interpretation even in sentences with a reflexive (see discussion in Fiengo & May's, 1994), as shown by Experiments 3 and 5. The results of Experiment 3 also showed that strict interpretations were no more likely in cause-effect sentences than in resemblance sentences. Experiments 4 and 5 showed that holding the type of coherence relation constant, increasing parallelism by using the and... did too construction increased the number of sloppy interpretations, for ordinary pronouns, possessive pronouns, and reflexives.

To this point we have evaluated Kehler's theory as if only a single discourse coherence relation could apply to a given pair of clauses. But Kehler (2000, p. 552) assumes "for some examples there may be more than one operative coherence relation. In particular, merely using a connective indicating a Cause-effect relationship is not necessarily enough to avert the recognition of a Parallel relation also..." (Note: A "Parallel relation" in Kehler's terms is a Resemblance discourse coherence relation of the sub-type where arguments are related by similarity, as opposed to contrast.) The possibility that more than one type of discourse coherence relation may be simultaneously active complicates Kehler's theory. For present purposes, it suggests that the theory is best evaluated by looking at the acceptability of Cause-effect relations lacking syntactic parallelism between the antecedent clause and the ellipsis clause. In such cases, only the Cause-effect relation should be inferred. Consequently, it shouldn't matter in the (pure) Cause-effect relation whether a syntactically appropriate antecedent for the ellipsis is available whereas it should matter for a Resemblance relation. Experiment 2 results are perhaps most relevant here. The mismatching causal sentences, the ones without a syntactically appropriate antecedent, were predicted to be grammatical whereas the mismatching resemblance sentences were predicted to be ungrammatical. Yet the mismatching causal sentences were rated no more acceptable than the predicted ungrammatical mismatching resemblance sentences.

In Experiment 3, sloppy interpretations were assigned roughly 50% of the time regardless of whether the ellipsis clause was because everyone else did or and everyone else did. If only the because clause supported a Cause-effect relation, this finding would be sharply inconsistent with Kehler's claims. However, if the theory permits a second (Cause-effect or Contiguity)

coherence relation to be assigned to the and examples, it could permit a strict reading to be available in them (though not if the requirements of each coherence relation must be met). We don't find this attempt to reconcile the data with the theory to be particularly enlightening. The lack of any difference between the clear Cause-effect clauses and the intended Resemblance clauses requires explanation. One could assume that comprehension is a maximal inference sort of process and that readers assign all possible additional coherence relations essentially without evidence even when they already have inferred one discourse coherence relation that is supported by the connective. But we think the elegance and explanatoriness of Kehler's theory would be severely restricted by making such assumptions.

Perhaps the point where the availability of a second coherence relation becomes most relevant is in the discussion of the too examples in Experiments 4 and 5. It is reasonable to assume that processing of too leads comprehenders to compute a Resemblance relation even in sentences that may also be assigned another discourse coherence relation. Consequently, in Experiments 4 and 5 perhaps readers assigned a Resemblance relation more readily to the too (parallel) examples than to the would soon (nonparallel) examples. In this case, the results of Experiments 4 and 5 don't really crucially distinguish among the predictions of a syntactic account of ellipsis versus Kehler's account.

We are still left with the question then whether syntactic parallelism between the antecedent clause and the ellipsis clause comes into play with all discourse coherence relations, as we interpret our data as suggesting, or it only comes into play with Resemblance coherence relations. The latter suggestion does assume that syntactic parallelism in discourses with types of coherence relations other than Resemblance may lead to the assignment of a second, Resemblance, coherence relation (and we note that Kehler makes this assumption). The literature on processing pronouns may be relevant here. (We thank Julie Sedivy for pointing this out.) Smyth (1994) studied "parallel function" assignments of antecedents to pronouns. He began with sentences originally studied by Crawley et al. (1990) where a subject-antecedent preference had been established even for pronouns appearing in object position. Smyth adapted Crawley et al.'s materials to make the clause containing the pronoun syntactically parallel to the clause containing the antecedent. He then found 88% object-antecedent assignments in a study where people read the test sentences and then indicated their interpretation. The connective was always and then and the sentences "unambiguously described a sequence of events" (Smyth, 1994, p. 211). In other words, in sentences with a Contiguity relation like those in (17) and (18), syntactic parallelism between the clauses led to overwhelming parallel function assignments.

(17) John pushed Sammy and then Evelyn kicked him.

(18) Sarah visited Cathy at home and then Charles phoned her at work.

Although it would be possible to claim that syntactic parallelism per se leads to the assignment of a second, Resemblance coherence, relation, this seems forced for examples like (17) and possibly (18). In (17), a Contiguity relation and a Cause-effect relation seem more salient to us. To claim a Resemblance relation is also assigned would seem to make the theory rather empty.

An alternative approach is to account for syntactic parallelism effects directly, without resort to an intermediate step of having syntactic parallelism lead to assignment of a Resemblance discourse coherence relation. Here we follow the lead of Carlson (2002). In her book, devoted to the effects of parallelism especially in processing ellipsis, she proposed the Parallelism hypothesis in (19) (Carlson, 2002, p. 6).

(19) Carlson's Parallelism Hypothesis: The processor favors analyses in which DPs that share internal properties (have similar syntactic, prosodic, and semantic features) share external properties (appear in similar structural positions within their respective clauses or phrases), and vice versa.

Carlson's insight is that once structural similarities between clauses are recognized, analyses entailing further similarities are favored. This approach seems right to us. Throughout the literature on parallelism, once similarities across clauses exist, further similarities lead to faster processing times and tend to favor analyses that allow further similarities (Carlson 2002; Frazier, Munn and Clifton, 2000; Frazier, Taft, Roeper, Clifton, & Ehrlich, 1984, Henstra, 1996).

Looking at Kehler's (2002) examples (e.g., (2)–(5) above but also many others), what often differs is not just the hypothesized discourse coherence relation, but also the presence of too in the resemblance relation sentences but not in the cause-effect sentences. Given the nature of too, native speakers may be happiest if satisfying its presupposition allows the property predicated of the second clause subject to be the same as the property predicated of the first clause subject at all levels of linguistic analysis (see Schwarz, 2005, Carlson et al., in progress, for evidence). Even a change from a Proper Name to a definite description seems slightly worse in sentences containing too, as in (20a) than in sentences lacking too (at least, when pitch accents are placed on the UPPERCASE words).

(20)

- a. John perjured himself and the dentist did TOO.
- b. John perjured himself AND the dentist did.

From this perspective, it is not surprising that sentences containing too should tend to have a stronger syntactic parallelism requirement than sentences lacking it, and should tend to favor sloppy reading of VP ellipsis sentences containing reflexives, since on the sloppy reading both the first and second clause contain a reflexive predicate (as in The policeman defended himself and the fireman did too.).

Hendriks (2003) discussed Kehler's proposal, suggesting that discourse coherence relations are not quite the right notion. She suggested that what differs between resemblance relations and cause-effect relations is how the topic is established. Basically she argued that resemblance relations involve contrastive topics, whereas cause-effect relations involve non-contrastive topics. In effect, in a cause-effect sentence, the clauses make independent contributions to discourse. The cause clause and the effect clause are analyzed as if they were independent sentences. This approach raises many interesting issues some of which are connected to the issues discussed here and some of which go far beyond the scope of the present paper. For present purposes, we want to note that in parallel clauses with contrasted objects (John kissed MARY and he hugged SUE.), the notion of contrastive topic does not seem appropriate for capturing the effects of parallelism. Also, in at least some cause-effect sentences (Ana left BECAUSE Tim arrived.), it is primarily the connection between the two clauses which is asserted, so it is not clear how these clauses could really be analyzed as if they were independent sentences. Thus, although in a subset of syntactically parallel sentences contrastive topics may be involved, we doubt that a contrastive topic approach could provide a fully general account of the phenomena discussed here.

In conclusion, we acknowledge that Kehler's theory is an elegant and provocative apparent solution to the important problem of under- vs. over-generation of VP ellipses. However, we must raise the cautionary note that the examples he has presented in support of his theory may not have quite the force he ascribes to them. Our own approach to the overgeneration problem

is to embed a syntactic account of ellipsis in a theory of processing in which the processor patches up an antecedent at LF if it is not of the required form (see Arregui, et al, under revision, for evidence). This, coupled with Carlson's approach to parallelism, could account for the graded acceptability of ellipsis without a syntactically appropriate antecedent and simultaneously account for the anaphora facts reviewed here (the availability of strict identity even with 'Resemblance' coherence relations, and the preference for sloppy identity and for "parallel function" assignment of antecedents in anaphora).

We acknowledge that it is possible that refinements to Kehler's theoretical position may improve its ability to account for our data. Kehler (2000) discussed the possibility that manipulating the connective alone might not be sufficient to alter the type of discourse coherence relation at stake. However, if discourse coherence is what's involved, surely it should. That is, using because as a connective, one has no choice but to analyze the relation between two clauses as a cause-effect relation. While it certainly may be true that more elaborate discourses might contain more information dictating the type of coherence relation involved, we do not see how sentences that differ only in the connective can fail to show some effect in the direction predicted by the discourse coherence theory account if indeed coherence relations are playing the causal role that Kehler proposed.

The data we have presented in this paper indicate that parallelism and the presence of presuppositional words like too can directly influence the interpretation and acceptability of ellipsis sentences. Given Kehler's selection of supporting examples, these factors could appear to result in support for his discourse coherence theory because the presence of parallelism is a large part of what is used to identify an example as being an instance of a resemblance relation. Thus, while we find Kehler's theory very elegant and attractive, we find that there is good reason to be cautious in evaluating the empirical evidence for the theory.

References

- Carlson, K., C. Clifton and Frazier, L.: In progress, 'VP-ellipsis parallelism and focus structure'.
- Chambers CG, Smyth R. 'Structural Parallelism and Discourse Coherence: A Test of Centering Theory'. *Journal of Memory and Language* 1998;39:593–608.
- Chomsky, N.: 1981, *Lectures on Government and Binding* Dordrecht: Foris.
- Crawley RA, Stevenson RJ, Kleinman D. 'The use of heuristic strategies in the interpretation of pronouns'. *Journal of Psycholinguistic Research* 1990;19:245–264. [PubMed: 2231480]
- Dalrymple M, Shieber S, Pereira F. 'Ellipsis and Higher-Order Unification'. *Linguistics and Philosophy* 1991;14:399–452.
- Fiengo, R. and May, R.: 1994, *Indices and Identity* Cambridge, MA: MIT Press.
- Frazier L, Clifton C. 'On bound variable interpretations: The LF-only hypothesis'. *Journal of Psycholinguistic Research* 2000;29:125–139. [PubMed: 10709179]
- Frazier L, Munn A, Clifton C Jr. 'Processing co-ordinate structures'. *Journal of Psycholinguistic Research* 2000;29:343–370. [PubMed: 10953824]
- Frazier L, Taft L, Roeper T, Clifton C Jr, Ehrlich K. 'Parallel structure: A source of facilitation in sentence comprehension'. *Memory & Cognition* 1984;12:421–430.
- Gordon PC, Grosz BJ, Gilliom LA. 'Pronouns, names and the centering of attention in discourse'. *Cognitive Science* 1993;17:311–347.
- Hankamer J, Sag I. 'Deep and surface anaphora'. *Linguistic Inquiry* 1976;7:391–428.
- Hardt, D.: 1993, 'Verb phrase ellipsis: form, meaning and processing', Ph.D. dissertation, University of Pennsylvania..
- Hardt D. 'Dynamic Interpretation of Verb Phrase Ellipsis'. *Linguistics and Philosophy* 1999;22:187–221.
- Hendricks, P.: 2003, 'Ambiguous And: Coreference and contrast in coordination', University of Groningen manuscript.

- Henstra, J.A.: 1996, *On the parsing of syntactically ambiguous sentences: Coordination and relative clause attachment* Unpublished PhD Thesis, University of Sussex, Sussex, England.
- Hestvik A. 'Reflexives and Ellipsis'. *Natural Language Semantics* 1995;3:211–237.
- Kehler, A.: 1993, 'The effects of establishing coherence in ellipsis and anaphora resolution', *Proceedings of the 31st Annual Meeting of the Association of Computational Linguistics*. Columbus: Ohio State University.
- Kehler, A.: 1995, *Interpreting Cohesive Forms in the Context of Discourse Inference*, Ph.D. dissertation, Harvard University.
- Kehler A. 'Coherence and the Resolution of Ellipsis'. *Linguistics and Philosophy* 2000;23(6):533–575.
- Kehler, A.: 2002, *Coherence, Reference and the Theory of Grammar*, Stanford:CSLI Publications.
- Sag, I.: 1976, *Deletion and Logical Form* Ph.D. dissertation, MIT.
- Schwarz, F.: 2005, "Presuppositions in processing: A case study of German 'auch'", Presented at Sinn und Bedeutung, October 2005.
- Shapiro LP, Hestvik A. 'On-line comprehension of VP-ellipsis: Syntactic reconstruction and the semantic influence'. *Journal of Psycholinguistic Research* 1995;24:517–532. [PubMed: 8531171]
- Shapiro LP, Hestvik A, Lesan L, AR Garcia. 'Charting the time-course of VP-ellipsis sentence comprehension: Evidence for initial and independent structural analysis'. *Journal of Memory and Language* 2003;49:1–19.
- Simner, J., Pickering, M. and Garnham, A.: 2003, 'Discourse cues to ambiguity resolution: Evidence from "Do it" Comprehension', *Discourse Processes* 36,
- Smyth R. 'Grammatical determinants of ambiguous pronoun resolution'. *Journal of Psycholinguistic Research* 1994;23:197–230.
- Tanenhaus M, Carlson G. 'Comprehension of deep and surface verb phrase anaphors'. *Language and Cognitive Processes* 1990;5:257–280.
- Williams E. 'Discourse and Logical Form'. *Linguistic Inquiry* 1977;8:101–139.
- Wolf F, Gibson E, T Desmet. 'Discourse coherence and pronoun resolution'. *Language and Cognitive Processes* 2004;19:665–675.

Appendix 1: Experimental materials, Experiment 1. Optional adverb indicated in (); alternatives indicated by |

1. The problem was looked into by Kim (last time) because|just like Lee did.
2. The magazine was accused of plagiarism by the Academy (last year) because|just like the Historical Association did.
3. The artist was followed by the reporter this morning because|just like the fans did.
4. The airplane was checked again by the airlines (often) because|just like Security did.
5. Classes were cancelled by the teachers (last time) because|just like the administration did.
6. The lawyers were excluded by the chemists (typically) because|just like the biologists did.
7. The child was reprimanded by the cleaning lady (yesterday) because|just like the babysitter did.
8. The radio was turned up for the news by the teenager (often) because|just like the parents did.
9. The cause of the accident was investigated by the police (most times) because the insurance company did.|and the insurance company did too.
10. The assignment was rewritten by Angela (last time) because Fran did| and Fran did too.

11. The box was taken apart completely by Frank (yesterday) because Robert did| and Robert did too.
12. The applicant was questioned by the supervisor (usually) because the owner did.|and the owner did too.
13. The coffee was made with cinnamon (this morning) by Anita because Marica did.| and Marica did too.
14. The incident was followed up by the principal (usually) because the teacher did.|and the teacher did too.
15. The snow was shoveled by the resident (yesterday) because the landlord did.|and the landlord did too.
16. The plants were watered daily by the caretaker (usually) because| the gardener did.| and the gardener did too.

Appendix 2: Experimental materials, Experiment 2. Alternatives indicated by |. Rating scale indicated for item 1a

1a. The problem was looked into by Kim even though|just like Lee did.

Unacceptable 1 2 3 4 5 Acceptable

1b. Kim looked into the problem even though|just like Lee did.

2a. The magazine was accused of plagiarism by the Academy because|just like the Historical Association did.

2b. The Academy accused the magazine of plagiarism because|just like the Historical Association did.

3a. The artist was followed by the reporter because|just like the fans did.

3b. The reporter followed the artist because|just like the fans did.

4a. The airplane was checked again by the airlines because|just like Security did.

4b. The airlines checked the airplane again because|just like Security did.

5a. Classes were cancelled by the teachers because|just like the administration did.

5b. The teachers cancelled classes because|just like the administration did.

6a. The lawyers were excluded by the chemists because|just like the biologists did.

7a. The chemists excluded the lawyers because|just like the biologists did.

8a. The child was reprimanded by the cleaning lady even though|just like the babysitter did.

8b. The cleaning lady reprimanded the child even though|just like the babysitter did.

8a. The radio was turned up for the news by the teenager because|just like the parents usually did.

8b. The teenager turned up the radio for the news because|just like the parents usually did.

9a. The cause of the accident was investigated by the police even though|just like the insurance company already did.

9b. The police investigated the cause of the accident even though|just like the insurance company already did.

10a. The assignment was rewritten by Angela because|just like Fran did.

10b. Angela rewrote the assignment because|just like Fran did.

11a. The box was kicked by Billy because|just like his older brother did.

11b. Billy kicked the box because|just like his older brother did.

12a. The applicant was questioned by the supervisor because|just like the owner did.

12b. The supervisor questioned the applicant because|just like the owner did.

13a. The coffee was made with cinnamon by Anita because|just like Marica always did.

13b. Anita made the coffee with cinnamon because|just like Marica always did.

14a. The incident was followed up by the principal because|just like the teacher did.

14b. The principal followed up the incident because|just like the teacher did.

15a. The snow was shoveled by the husband on weekends because|just like the wife did during the week.

15b. The husband shoveled the snow on weekends because|just like the wife did during the week.

16a. The plants were watered by the caretaker today even though|just like the gardener already did|did yesterday.

16b. The caretaker watered the plants today even though|just like the gardener already did|did yesterday

Appendix 3: Materials for Experiment 3. Alternatives indicated by |. Question indicated for item 1

1. Doug blamed himself for the band's collapse because|and everyone else did. What did everyone else do?
 blame Doug
 blame themselves
2. Fernando nominated himself for the position because|and everyone else did.
3. Carl congratulated himself for the success of the plan because|and everyone else did.
4. Peter expected himself to solve the problem because|and everyone else did.
5. Antonio incriminated himself because|just like everyone else did.
6. Sandy expected himself to work on weekends because|just like everyone else did.
7. Mr. Simeon assigned himself overtime because|just like everyone else did.

8. Robert voted for himself because|just like everyone else did.
9. Patrick defended himself at the meeting because|and Bill did.
10. Shawn described himself on the internet because|and David did.
11. Ian ridicules himself in front of others because|and his father does.
12. Jules expected himself to solve the problem because|and Gregory did.
13. Gordon thought better of himself after the confession because|just like Josh did.
14. Jeff puffed himself up for the interview because|just like Russell did.
15. Jackson made fun of himself at the party because|just like Randolph did.
16. Bill restrained himself during the argument because|just like Tim did.

Appendix 4: Materials from Experiment 4. Alternatives are indicated by |; presentation regions by ^. Question alternatives are illustrated for the first item

1. John saw a snake near him|near his backpack^and Bill did too|would soon.
Where was the snake that Bill saw?; Near John Near Bill
Where was the snake that Bill saw?; Near John's backpack Near Bill's backpack
Where would the snake that Bill saw be?; Near John Near Bill
Where would the snake that Bill saw be?; Near John's backpack Near Bill's backpack
2. Ernie heard a dog near him|his house^and Bob did too|would too.
3. Max pushed a gun away from him|his foot^and Larry did too|have
4. Ernest told a nurse about him|his history^and Sam did too|could have.
5. Mr. Johnson spilled a beer on him|his lap^and Craig did too|on him.
6. Lou noticed a deer near him|his shed^and Tom did too|or Tom did.
7. Greg told a pretty reporter about him|his success^and Larry did too|or Larry did.
8. Fred kicked a cockroach away from him|his sleeping bag^and Henry did too|or Henry did.
9. Karl dribbled some spaghetti on him|his tie^and George did too|would soon.
10. Thomas detected a tick on him|his shirt^and Roy did too|or Roy did.
11. Jason shooed some bees away from him|his face^and Tim did too|though Tim could have.
12. Jim photographed some people near him|his cabin^and Michael did too|could have.
13. Jake withheld some secrets about him|his family^and Brian did too|should have.
14. Kendall pushed the dessert away from him|his plate^and Billy did too|or Billy did.
15. Gary sent the doctor away from him|his room^and Glen did too|or Glen did.
16. Brad pointed at some people near him|his motorcycle^and Willy did too|or Willy did.

**Appendix 5: Materials for Experiment 5. Alternatives are indicated by |.
Question alternatives are indicated for the first item**

1. John kicked a snake away from him|himself and Bill did too|would soon.
Who did Bill kick the snake away from? From John___ From Bill___
Who did Bill kick the snake away from? From John___ From Bill___
Who would Bill kick the snake away from? From John___ From Bill___
Who would Bill kick the snake away from? From John___ From Bill___
2. Ernie threw some trash far from him|himself and Bob did too|would soon.
3. Max pushed a gun away from him|himself and Larry did too|could have.
4. Ernest told a nurse about him|himself and Sam did too|could have.
5. Mr. Johnson spilled a beer on him|himself and Craig did too|would soon.
6. Karl dribbled some spaghetti on him|himself and George did too|would soon.
7. Jason shooed some bees away from him|himself and Tim did too|could have.
8. Jim pushed some customers away from him|himself and Michael did too|could have.
9. Jake withheld some secrets about him|himself and Brian did too|or Brian did.
10. Lou threw a grenade far from him|himself and Tom did too|or Tom did.
11. Greg told a pretty reporter about him|himself and Larry did too|or Larry did.
12. Fred kicked a cockroach away from him|himself and Henry did too|or Henry did.
13. Thomas detected a tick on him|himself and Roy did too|or Roy did.
14. Kendall pushed the dessert away from him|himself and Billy did too|or Billy did.
15. Gary tossed a firecracker far away from him|himself and Glen did too|or Glen did.
16. Brad detonated a bomb 20 yards away from him|himself and Willy did too|or Willy did.

Acknowledgements

This research was supported in part by NSF Grant BCS 0090674 "Prosody in language comprehension" and NIH Grant HD19708 "Language comprehension" to the University of Massachusetts. We are grateful to Andrew Kehler, Julie Sedivy and three anonymous reviewers for helpful comments on this work.

Table 1
Percentage “got-it” responses and mean reaction times (ms), all responses, Experiment 1

Condition	Percentage “got-it”	Measure	RT
1. Causal, no adverb	53.1		1169
2. Causal, adverb	51.5		1078
3. Resemblance, no adverb	61.9		1064
4. Resemblance, adverb	58.8		1086

Table 2
Mean acceptability ratings (1 = unacceptable, 5= acceptable), Experiment 2

Condition	Rating
1. Mismatching, Causal	3.016
2. Mismatching, Resemblance	3.089
3. Matching, Causal	4.333
4. Matching, Resemblance	4.396

Table 3
Percentage of sloppy (bound) interpretations, Experiments 4A and 4B

Condition	Questionnaire (4A)	Experiment
		On-line Interpretation (4B)
Parallel, pronoun	42	38
Parallel, possessive	45	43
Nonparallel, pronoun	33	26
Nonparallel, possessive	27	23

Table 4
Percentage of sloppy (bound) interpretations, Experiment 5

Pronoun	Parallel	Nonparallel
Personal	54	32
Reflexive	78	46