Adverbs in VP ellipsis: an experimental investigation of antecedent selection

KEIR MOULTON
University of Toronto
keir.moulton@utoronto.ca

Abstract

This paper presents a case study of verb phrase ellipses with adverbially modified antecedents. It is shown experimentally that there is a preference for resolving ellipses in certain embedded clauses with unmodified VPs. The effect is hypothesized to reflect a general requirement to minimize the complexity of accommodated content. Four experiments support this hypothesis over plausible candidate hypotheses, including syntactic approaches to the effect (Matsuo 2001; Sailor 2014).

Keywords: VP ellipsis, adverbs, experimental syntax

Résumé

Cet article présente une étude de cas d'ellipses de syntagmes verbaux avec des antécédents modifiés de façon adverbiale. Il est démontré expérimentalement qu'il existe une préférence pour la résolution des ellipses dans certaines propositions enchâssées avec des SV non modifiés. On propose que cet effet provient d'une exigence générale visant à minimiser la complexité du contenu accommodé. Quatre expériences soutiennent cette hypothèse par rapport a d'autres hypothéses plausibles, y compris des analyses syntaxiques du phénomène (Matsuo 2001; Sailor 2014).

Mots-clés: ellipses de VP, adverbes, syntaxe expérimentale

1. INTRODUCTION

Quite often theoretical debates can turn on whether a phenomenon deserves a syntactic, semantic, pragmatic, or processing treatment.[1] Sometimes such issues are not trivial to resolve. This is particularly true in the case of ellipsis, where a diverse set of grammatical and extra-grammatical factors are at play, including subtle syntactic and semantic matching requirements (Sag 1976; Merchant 2001; Chung 2013), the discourse relations between antecedent and elided clause (Kehler 2000, 2002; Frazier and Clifton 2006), and the role of focus and de-accenting in governing what material is elided and what is retained (Rooth 1992; Merchant 2001). The very fact that elided material is silent makes the job even harder: we have only indirect access to what needs to be represented in the syntax and semantics. One prominent topic in the study of ellipsis is the 'size' of the elided material – what verbal projections in the clause are targeted by ellipsis (Akmajian and Wasow 1975; Fox and Takahashi 2005; Merchant 2008; Kim et al. 2011; Sailor 2014; Aelbrecht and Harwood 2015). This article examines the factors that determine the size of antecedents chosen in verb phrase ellipsis (VPE) and, using experimental methods, teases apart syntactic from discourse-pragmatic factors in ellipsis resolution. The main contribution is a case study which finds that readers resolve ellipses with 'smaller' antecedents in various embedded clauses compared to unembedded clauses. Based on these results, I will advance a hypothesis wherein pressures on presupposition accommodation drive the selection of smaller antecedents.

The case study involves instances of VPE where, in principle, either an adverbially modified or unmodified antecedent is available. Matsuo (2001) claimed that in antecedent-contained deletion structures (ACD, see Bouton 1970) such as (1) an unmodified antecedent is required. (Elided material is struck out.)

(1) John will quickly learn French from the same book his brother did ~~learn French/?~~ ~~quickly learn French~~.   (after Matsuo 2001: 110(30))

I call these 'small antecedents'. In non-ACD constructions, a modified or 'large' antecedent is more naturally chosen, as in (2), due to Potsdam (1998).

(2) Helga easily won her race and Sophie will ~~easily win the race/*win the race~~ too.
                                                                (Potsdam 1998: 402(19a))

Matsuo (2001) and later Sailor (2014) claim that the small-antecedent effect is due to the syntactic properties of ACD constructions. In an ACD, the ellipsis site is contained within the antecedent verb phrase. In the generative tradition, it is widely agreed that such contained ellipses are fundamentally illegitimate and that a syntactic operation re-arranges the structure of an ACD so that the ellipsis site is not contained in the VP that serves as its antecedent.[2] Matsuo and Sailor provide an explanation for

---

[1]Abbreviations used: ACD: antecedent-contained deletion; ACH: Antecedent Containment Hypothesis; CE: clause containing an ellipsis; MV: main verb; Op: operator; PH: Parsimony Hypothesis; QR: quantifier raising; RR: restrictive relative; VPE: verb phrase ellipsis.

[2]See Jacobson (2008) for an alternative.

the small-antecedent effect that rests on this operation, the details of which I present more fully in section 2.

Now Matsuo (2001) reported that the small antecedent is the only possible resolution in ACD structures like (1). However, a modified antecedent is available for the ACD in (3), where the context forces it.

(3) Fred and Maria are competing in a music competition in which they both sing the same repertoire. Both Fred and Maria sing all their songs well, but they each only sing one song flawlessly.

    a. Speaker A: Which song did Fred flawlessly sing?

    b. Speaker B: Fred flawlessly sang the song Maria did ~~flawlessly sing~~.

If the unmodified antecedent were chosen in [3b], the definite would fail to denote a unique song. The response in [3b] is felicitous, so a large antecedent is available as long as the context overwhelmingly rules out a small antecedent. Whatever the small-antecedent effect is, it is not a categorical judgment. This is where experimental methods become useful, as a way to probe gradient acceptability. I will show experimentally that in impoverished contexts, there is indeed a preference for resolving ellipses with a small antecedent when they appear in certain embedded clauses. Experiments 1a and 1b establish the small-antecedent effect in ACD constructions. Experiments 2–4 then seek to determine the source of the small-antecedent effect. Experiments 2 and 3 rule out an account of the small-antecedent effect that reduces the problem to that of the height at which the clause containing the ellipsis attaches, as in the proposals in Matsuo (2001) and Sailor (2014). Experiment 3 additionally shows that the small-antecedent effect extends to temporal adverbial subordinate clauses, which in turn helps establish that the small-antecedent effect is not the result of parallelism constraints on coordinated clauses. Experiment 4 tests antecedent choice in restrictive versus non-restrictive relatives and finds that restrictive relatives are more likely to exhibit the small-antecedent effect than non-restrictive relatives.

Putting these results together, I suggest that the small-antecedent effect is a reflection of the Principle of Parsimony (Crain and Steedman 1985; Altmann and Steedman 1988), which constrains what hearers and readers assume about a context. The key observation is that the small-antecedent effect arises in restrictors of presupposition triggers. Sentences containing presupposition triggers are only felicitous if the *common ground*—the shared knowledge of speaker and hearer—supports that presupposition (Stalnaker 2002). The English definite article *the*, for instance, presupposes a unique (or familiar) entity that meets its description. A presupposition failure arises if no such entity exists: *The king of France is bald* exhibits a presupposition failure because there is no such person that meets this description, let alone a unique one. In some cases, the hearer may *accommodate* the referent even if it is not part of the common ground (Karttunen 1974; Stalnaker 1974; Lewis 1979; see von Fintel 2004 for a recent discussion). For instance, if you tell me *I need to feed my cat*, and I didn't know that you have a cat, I may be likely to update my beliefs and accommodate that you have a cat. I will propose that, at least as far as processing ellipsis is concerned, the amount of content that is

accommodated can vary and that there is a preference to accommodate less. The idea is intuitive: in (1), the choice of a large antecedent requires the reader to accommodate that there exists a book and that the manner of the event described by the embedded VP was quick. However, in a null context, this level of specificity is not motivated. The choice of a small antecedent involves accommodating a "less informative" VP, in a way that I make concrete in Section 6. The gist is this: specifying the manner in which the book was read is not motivated in these contexts (compare to (3)) and so, by parsimony, is not targeted as the antecedent for ellipsis. I call this the Parsimony Hypothesis for the small-antecedent effect. In clauses that do not restrict presuppositional determiners, but rather assert as in (2) and in non-restrictive relatives, there is nothing that militates against adverbial modification, because there is no accommodation required.[3]

2.   POSSIBLE EXPLANATIONS FOR THE SMALL-ANTECEDENT EFFECT

This section outlines alternative analyses of the small-antecedent effect, both those that have appeared in the literature (the Antecedent Containment Hypothesis, or ACH, Matsuo 2001; Sailor 2014) and another viable candidate hypothesis related to parallelism, a factor known to influence antecedent selection.

2.1.   Antecedent Containment

Matsuo (2001) argued that the small-antecedent effect arises in (1) because this is an Antecedent-Contained Deletion (ACD) structure (Bouton 1970). In ACD, the clause that houses the ellipsis, C_E, is contained within the antecedent of ellipsis. In (4) the C_E is a relative clause that modifies the matrix object. (The gap in the relative is represented as a trace left by operator (Op) movement.)

(4)   John [$_{VP}$ sang [the song [$_{C_E}$ Op$_i$ that Fred did ~~sing t$_i$~~]]]

Elided VPs require matching antecedents. Matching is not possible in structures like (4) because no phrase can match a phrase it is properly contained in:[4] the elided [$_{VP}$ sing t$_i$] does not match the antecedent [$_{VP}$ sang the song [Op$_i$ that Fred did ~~sing t$_i$~~]]. To account for the grammaticality of ACD, it is generally assumed that some form of syntactic reorganization occurs so that the ellipsis can find an antecedent that does not contain it. The most widely adopted analysis involves quantifier raising (QR), which covertly moves the DP that contains the relative outside of the VP at the level of Logical Form (Sag 1976; May 1985).[5] The most recent account, due to Fox

---

[3]As explained in Section 6, while the most direct test of the Parsimony Hypothesis would involve restrictive relatives headed by non-presuppositional determiners, this is not possible since ACD independently requires presuppositional determiners (Diesing 1992).

[4]In earlier literature, this was often described as a problem of infinite regress. This conception does not fit with modern approaches to ellipsis, which emphasize the identity conditions on ellipsis.

[5]Baltin (1987) rejects a QR approach and simply extraposes the relative clause outside the VP projection that serves as the antecedent.

(2002), also invokes QR, but introduces the C_E via late-merger to the QR'd DP. Late-merger is a type of counter-cyclic operation in a which a phrase can adjoin to a non-root node (Lebeaux 1990). Fox takes late-merger one step further and proposes counter-cyclic adjunction to the higher, silent copy of the DP that undergoes QR. The corollary for this view is that QR is rightward and can be interleaved with overt operations like relative clause adjunction (Fox and Nissenbaum 1999).[6] Definite descriptions that host ACD can also undergo QR (Harley 2002). I will assume that QR adjoins to VP. The derivation is shown in (5). In (5a) the object undergoes rightward QR and adjoins to the VP. Shaded copies indicate QR'd material. In (5b), the relative clause C_E late-merges to the QR'd (and silent) head noun.

(5) John sang the song that Fred did.

    a. *Quantifier Raising (QR)*
       John [$_{VP}$ [$_{VP}$ sang t] [the song]]

    b. *Later-merger of C_E*
       John [$_{VP}$ [$_{VP}$ sang t] [the song [$_{C_E}$ that Fred did ~~sing t~~]]]

Now the elided VP can find a matching antecedent in the first clause, [$_{VP}$ sing t].

    Returning to the small-antecedent effect, Matsuo argues that QR targets a position below that of the adverb ("low QR") (6a), in which case the only option for ellipsis that doesn't violate containment is the unmodified antecedent. (I will assume, following Ernst 2001, that manner adverbs are VP-adjoined.) For Matsuo, the representation that would allow for a large antecedent, with QR adjoining to the higher VP that contains the adverb (6b), is not available.

(6) a. *Low QR*
       John [$_{VP}$ quietly [$_{VP}$ [$_{VP}$ sang t] [the song [$_{C_E}$ that Fred did ~~sing t~~]]]]

    b. *High QR*
       John [$_{VP}$ [$_{VP}$ quietly sang t] [the song [$_{C_E}$ that Fred did ~~quietly sing t~~]]]

Matsuo adopts an economy approach to QR (Fox 2000) in order to rule out (6b): the object only needs to undergo low QR as in (6a) to resolve the containment problem; any longer application of QR, as in (6b), is thereby blocked and as a result a large antecedent is not available.

    In a sustained exploration of the VP projections targeted by ellipsis, Sailor (2014) claims that the small-antecedent effect with adverbs is just one instance of a larger pattern whereby ellipses in subordinate clauses target lower projections than those in coordinated clauses. One instance of this pattern is voice mismatch, which Sailor suggests is tolerated by VPEs in subordinate clauses more than in coordinated clauses (see Kehler 2002). Following Merchant (2013), Sailor argues that voice mismatches are possible when ellipsis targets a projection below the Voice head. Other instances of such low VPE, according to Sailor, are ellipses that allow

---

[6]This, of course, requires a copy theory of movement, but for expository convenience I illustrate the derivations with traces.

strict readings for reflexives (Hestvik 1995) and certain instances of non-recovered auxiliaries. (To demonstrate these effects, Sailor (2014) focuses on adverbial subordinate clauses, including *before* and *after* clauses. Moulton (2008) showed that the small-antecedent effect extends to such adverbial clauses, and Experiment 3 in this article confirms this.)

Like Matsuo, Sailor connects the various small-antecedent effects he identifies to a containment problem: subordinated C_Es are adjoined to a position lower than coordinated C_Es. In the case of voice mismatches, this position is lower than the material that enforces voice match (the voice head), therefore allowing mismatch. In the case of adverbs, Sailor says, the C_E is adjoined below the site of the matrix adverb, which I have sketched in (7a). (For clarity of exposition, I have omitted the articulated VP structure spelled out in Sailor's analysis. This does not affect the main point.) Without a movement operation to produce a structure like (7b), the choice of a large antecedent is ruled out by containment. Sailor (2014, p. 63) says that "high VPE in subordination structures [i.e. (7a)] is marked because recovery of such a large ellipsis site […] would yield an antecedent-contained deletion configuration." Sailor then goes on to propose an economy condition which militates against the representation in (7b) that allows for a large antecedent.

(7) a. *Low C_E attachment*
         John [$_{VP}$ quietly [$_{VP}$ [VP sang the song] [$_{C_E}$ before Fred did ~~sing the song~~]]]

    b. *High C_E attachment*
         John [$_{VP}$ [$_{VP}$ quietly sang the song] [$_{C_E}$ before Fred did ~~sing the song~~]]

Like Matsuo (2001), Sailor defines this economy condition in terms of QR: (7b) is less economical than (7a), which requires no applications of QR to be interpreted (i.e. to avoid a containment violation).[7]

Now both Matsuo (2001) and Sailor (2014) took the small-antecedent effect to be categorical, in that subordinate C_Es do not allow 'large' antecedents at all. We saw that this was not true in (3b), where a large antecedent is possible if the context is set up to promote it. We can, however, reinterpret Matsuo's and Sailor's proposals as a processing preference for short QR. An application of longer QR is less economical than a shorter application. This is a plausible hypothesis: in important work on ACD resolution, Hackl et al. (2012) found that there is a preference for resolving ACD with an embedded antecedent rather than a matrix one, as the latter involves a more complex application of long QR.[8] On this view, the small-antecedent effect is the result of the general impossibility of contained deletions and an economy condition on QR that favours the structure in (6a),

---

[7]Sailor does not in fact commit to QR as the operation that attaches the subordinate clause high (Sailor 2014: 73, ft. 27). Fox and Nissenbaum (2003), however, do explicitly argue for a QR analysis of adverbial ACD. Moreover, Sailor does not include canonical ACD configurations involving relative clauses under his proposal, since both high and low attachments of the C_E require QR and are "equally costly". But canonical ACD configurations *do* show a small-antecedent effect, so we do want an analysis that covers both these and adverbial subordinate C_Es.

[8]See Jacobson and Gibson (2014) for a critical appraisal of Hackl et al.'s conclusion.

which only allows an unmodified antecedent, over (6b). I call this the Antecedent Containment Hypothesis (ACH). Experiments 2 and 3 test the ACH with stimuli in which there is no containment problem and shows that the small-antecedent effect persists as long as the C_E is subordinated.

2.2. Parallelism

While the ACH is the main rival to the Parsimony Hypothesis, another possible explanation of the pattern of antecedent preferences involves parallelism. There are two types of parallelism to distinguish. Identity conditions on non-focused material force parallelism between the antecedent VP and the elided (or de-accented) material (Rooth 1992; Tancredi 1992). There are also parallelism requirements between the clause that houses the antecedent and the clause that houses the ellipsis site, particularly in coordinate clauses (Frazier et al. 2000; Frazier and Clifton 2001; Carlson 2003).[9] The preference for a large antecedent in coordination as opposed to relative clauses could very well be a reflection of parallelism requirements in the former but not the latter. Indeed non-elided versions of coordinated sentences where one VP is modified and the other is not are marked (8a), in comparison to non-elided versions of the relative clause (8b).

(8)   a.   John quickly learned French #and Bill learned French too.

      b.   John quickly learned French from the same book Bill learned French from.

If parallelism can boost the preference for larger antecedents, then using coordination as a "baseline" for the small-antecedent effect may not be appropriate.

Another source of parallelism is additive particles such as *too*, as in (8a), which presuppose identity of non-focused material. Given that Matsuo's original ACD example does not contain *too*, whereas coordinated clauses with ellipses often do, this could also introduce a confound. Both Experiments 2 and 3 address the parallelism confounds by controlling for the effects of coordination and additive particles like *too*.

3.   EXPERIMENTS 1A AND 1B

The first two experiments establish the basic effect: that in ACD, but not ellipses in a coordinated clause, a small antecedent is more likely to be chosen. The only difference between the two experiments is the determiner heading the DP containing the ACD: a definite or a universal. The experiments were conducted separately so they are reported separately here.

---

[9]An additional source of parallelism may be coherence relations. Kehler (2000) and Kehler (2002) have argued extensively that coherence relations influence VPE. The central coherence relations are resemblance, cause-effect, and contiguity. Kehler argues that resemblance relations strongly require syntactically matching antecedents for VPE (e.g., in not tolerating voice mismatch).

3.1. Experiment 1a

Methods

Two types of clauses containing ellipsis (C_Es) were compared: a coordinated clause and a relative clause modifying a definite description. The clause containing the antecedent (C_A) contained a pre-verbal manner adverb.

(9) a. Eddy quietly sang the song and Fred did too. C_E = coordinated.

    b. Eddy quietly sang the song that Fred did. C_E = relative clause (ACD)

Two versions of nine sentences were created following the pattern in (9). Materials appear in Appendix A. These items were included among 61 other one- or two-sentence discourses from an unrelated experiment. Each of the sentences was followed by a forced-choice question like (10). (The order of answer options was counterbalanced.)

(10) What did Fred do?

    a. quietly sing a song

    b. sing a song

Participants were instructed to choose the answer that corresponded to their "understanding of what the sentence said". They were instructed to read at a normal speed and answer with their first impression. In principle, a response of the unmodified verb phrase is compatible with participants resolving the ellipsis with a modified VP. That is, because of the entailment relation, (10b) will be a true albeit under-informative response even if a participant resolves the ellipsis with (10a). The experiment thus relies on the assumption that even if participants are reporting under-informative responses, they would do so equally across embedded and coordinated conditions. Twenty-four native English speakers completed the questionnaire via Amazon Mechanical Turk using Turktools (Erlewine and Kotek 2016) and were paid $1.50 for their participation. Two counterbalanced lists were created, so that no participant saw any one item in more than one condition. Participants were instructed to read each sentence or two-sentence discourse and click on the radio button next to their answer.

Results

The results are presented in Figure 1, which shows the mean proportion of large (i.e. modified) antecedent responses for each condition (71% vs. 17%). A logistic mixed-effects model (Baayen et al. 2008; Jaeger 2008) was fit using the lme4 package (Bates 2005) in R (R Development Core Team 2012). The model included C_E type as a fixed factor, and participant and item as random factors. The models also included random intercepts and random slopes for both these factors and their interaction (a maximal random effects structure, as recommended by Barr et al. 2013). There was a significant effect of C_E type (Est.=−5.1210, SE = 1.1609, z = −4.411, p < 0.001).
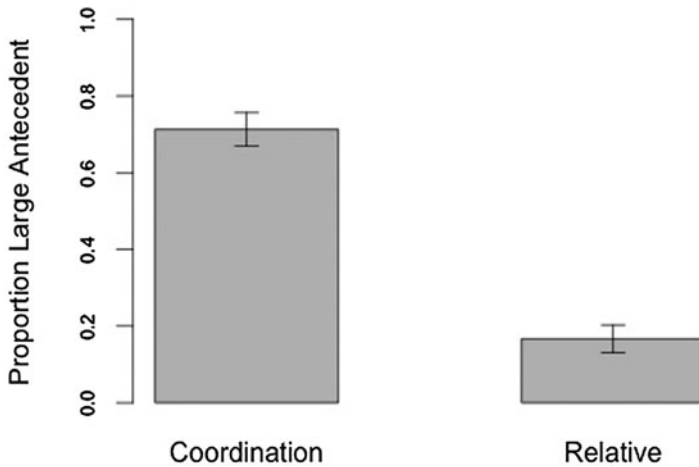
Figure 1: Mean proportion of large-antecedent responses in Experiment 1a

3.2. Experiment 1b

Methods

Experiment 1b tested whether the small-antecedent effect extends to ellipses within restrictors of universally quantified nouns. Looking ahead to the Parsimony Hypothesis, note that quantifiers such as *every* and *all* presuppose that the set forming their restrictors is non-empty (Lappin and Reinhart 1988; Diesing 1992). This is what makes the sentences in (11) pragmatically anomalous, since they presuppose the existence of American kings and unicorns.

(11) a. All/every American king(s) lived in New York.

    b. All unicorns have accounts at the Chase Manhattan Bank.

Experiment 1b tested the same materials as Experiment 1a, but with a universal *every* instead of a definite heading the relevant noun. All other aspects of the experiment were the same, except that participants (N = 9) were undergraduate students who completed a pen-and-paper written questionnaire for course credit. All were native English speakers.

Results

The results are presented in Figure 2, which shows the mean proportion of large (i.e. modified) antecedent responses for each condition (84% vs. 6%). A logistic mixed-effects model was fit with C_E type as a fixed factor and participant and item as random factors.[10] The analysis showed a significant effect of C_E type, such that coordinated C_Es differed from relative clause C_Es (Est. = −8.560, SE = 2.086, z = −4.104, p < 0.001).

---

[10]The model with the maximal random effects structure (Barr et al. 2013) did not converge. The model reported included intercepts and slopes for participants but not random slopes for items.
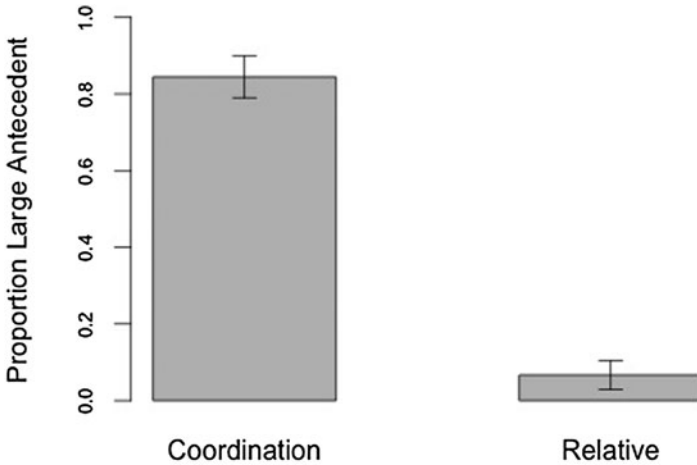
Figure 2: Mean proportion of large-antecedent responses in Experiment 1b

Discussion

In out-of-the-blue contexts like those presented, participants chose a small antecedent in relative clause conditions (ACD) but not in coordinated conditions, both under definites (Experiment 1a) and universals (Experiment 1b).

These experiments establish the basic small-antecedent effect. I now turn to the predictions of the proposed hypotheses, the Parsimony Hypothesis (PH) and the antecedent-containment hypothesis (ACH).

The accounts in Matsuo (2001) and Sailor (2014) – which I have grouped together as the non-categorical ACH – attribute the small-antecedent effect to the syntactic height at which the C_E is attached: economy conditions on QR preferentially place the site of C_E below the adverb, and this precludes a large antecedent as that would be an instance of antecedent containment. In syntactic configurations where there is no containment problem precluding a large antecedent, even for an embedded C_E, the ACH does not predict the kind of small-antecedent preference found for the C_Es that must invoke QR to allow a large antecedent. The competing Parsimony Hypothesis (PH), on the other hand, makes a positive prediction: even if the overt syntax makes available a large antecedent, a small antecedent would be chosen according to the PH if the C_E is presupposed as part of a definite description or strong quantifier.

These two approaches can be tested with (12a), where a post-verbal adverb is followed by an extraposed relative clause C_E. The landing site of QR can be diagnosed by the position of the extraposed relative clause (Williams 1974; Fox and Nissenbaum 1999). Because the relative clause is outside the VP constituent containing the adverb, the processor *must* give (12a) the parse in (12b) where QR has applied; in this case, economy considerations on QR do not militate against a high attachment, for the simple reason that the only possible parse for (12a) is one in which the containment problem is avoided.

(12) a. John sang the song quietly that Fred did.
      b. John [$_{VP1}$ [$_{VP2}$ sang t quietly] [the song [$_{CE}$ that Fred did]]]

The PH and the ACH make divergent predictions about sentences like (12a) with post-verbal adverbs, and those tested in Experiments 1a and 1b with pre-verbal adverbs as in (13).[11] These predictions are tested in Experiment 2.

(13) John quietly sang the song that Fred did.

While the ACH predicts a preference for small antecedents in (13) (against a coordinated baseline) it does not make the same prediction about the extraposed C_E in (12a), since economy constraints cannot rule out a small antecedent. The ACH predicts either that the choice of antecedent will be at chance, or perhaps that it will pattern with coordinated cases.[12] Either way, according to the ACH, we have no reason to expect that (13) and (12a) would pattern alike in terms of antecedent selection.The Parsimony Hypothesis (PH) on the other hand does make the strong, positive prediction that both (12a) and (13) will pattern alike and exhibit a small-antecedent effect, and to the same extent. That is because in both cases, the ellipsis is part of the restrictor of a definite description and hearers, according to the PH, are less willing to accommodate the specification of the manner of the event.

4. EXPERIMENT 2

Methods

Experiment 2 manipulated two factors, each with two levels. C_E TYPE: *main* (coordinated) or *embedded* (relative) and ADVERB POSITION: *pre-verbal* or *post-verbal*. These two factors were crossed to create the four conditions shown in (14).

(14) a. Edna quietly entered every room that David did.   *embedded/pre*

      b. Edna entered every room quietly that David did.   *embedded/post*

      c. Edna quietly entered the room and then David did.   *main/pre*

      d. Edna entered the room quietly and then David did.   *main/post*

The PH predicts that the extraposed C_E in (14b) will pattern with the non-extraposed C_E in (14a) in preferring a small antecedent compared to the main clause conditions in

---

[11]Sailor (2014) does discuss post-verbal adverbs in footnote 21 (p.23) noting that Cinque (1999) argues that these have a complex derivation, possibly involving predicate VP fronting. Sailor does not spell out how such derivations for post-verbal adverbs would interact with ACD – and whether they too would predict a small-antecedent effect. I believe the null hypothesis is that the parser will not treat the structures in (12a) as involving containment, and the economy principle will not predict a small antecedent.

[12]The difference between these two possibilities hinges on whether, when containment is not an issue, the largest antecedent must be chosen, or whether a smaller target is available. See Sailor (2014) for discussion.

(14c) and (14d). The ACH predicts an interaction between C_E TYPE and ADVERB POSITION such that (14a) will show a small-antecedent effect compared to (14b), which is predicted to either allow large antecedents at the rate of coordination or at least at a higher rate than (14a).[13]

Sixty-four sentences were created in the four conditions, following the pattern in (14). Ten items used *every* as the determiner; six used *the*. In Experiments 1a and 1b, relative clauses were compared to highly parallel coordinated clauses that contained the particle *too*. This baseline might overestimate the expectation for large antecedents for coordinated clauses, because such particles demand parallelism of de-accented material (and *a fortiori* elided material, see section 2.2). In an attempt to control for this confound, half of the items in Experiment 2 did not use such a particle in the coordinated case and instead used *then* as in (14). The other half of the items included *also* in both the coordinated and embedded ellipses as in (15).[14]

(15) a. Larry energetically played every sport that Fred also did.   *embedded/pre*

b. Larry played every sport energetically that Fred also did.   *embedded/post*

c. Larry energetically played every sport and Fred also did.   *main/pre*

d. Larry played every sport energetically and Fred also did.   *main/post*

If differences in antecedent selection are simply due to the presence or absence of particles such as *too* or *also*, we expect to find no small-antecedent effect in Experiment 2.

Four counterbalanced lists were created in which each item appeared once in only one condition. Each item was followed by a forced-choice question as in Experiments 1a,b. The placement of the adverb in the answers was counterbalanced. The 16 items were included among six unrelated experiments, all followed by a comprehension question or a naturalness rating. Materials are provided in Appendix B. Forty-eight native English speaking undergraduates participated in a computerized questionnaire. The items werepresented in random order with the items from the other experiments.

Results

Figure 3 shows the mean proportion of large-antecedent responses for each condition for 47 participants.[15] *Main* clause (coordinated) conditions showed agreater proportion of large-antecedent responses than the *embedded* (relative) conditions. In both cases, the large antecedent was chosen more often when the adverb was post-verbal.

---

[13]The post-verbal condition (14b) may be less natural than the other three conditions, in part because the extraposed relative is not appreciably heavier than the adverb it moves past. This potential confound is overcome in Experiment 3. Moreover, even though the extraposed conditions may be less natural than their coordination counterparts, I do not know of a reason why this would lead to a preference for small antecedents in just those cases.

[14]*Also* was chosen over *too* because speakers preferred it in ACD in an informal questionnaire.

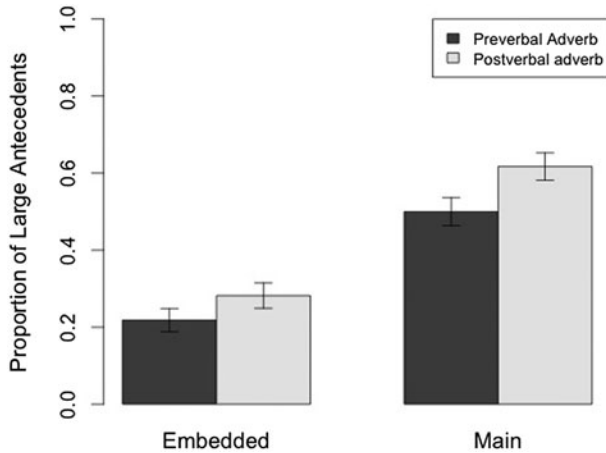[15]The file containing one participant's data was corrupted.

Figure 3: Mean proportion of large-antecedent responses in Experiment 2

The results were analyzed with a logistic mixed-effects model. The model reported here included as fixed factors C_E Type and Adverb Position, and participant and item as random factors. The models also included random intercepts and random slopes for both these factors and their interaction. There was a significant effect of C_E Type (Est. = 1.8514, SE = 0.3364, z = 5.504, p < 0.001) and a significant effect of Adverb Position (Est. = −0.6889, SE = 0.2372, z = −2.905, p < 0.01). There was no interaction.

Table 1 shows the means broken down by the presence of *also*. Items with *also* showed a slight numerical increase in large-antecedent responses in the post-verbal condition, for both the embedded and coordinated ellipses. There was no significant effect of the presence of *also* (Est. = 0.04668, SE = 0.32706, z = .143, p > .5).[16]

Discussion

The results of Experiment 2 were not those predicted by the ACH. Regardless of the position of the adverb, *embedded* C_Es were more likely to take small antecedents than *main* C_Es. If the small-antecedent preference were due to the low attachment of the C_E when the adverb is pre-verbal, then we should have observed an interaction. The results are consistent with the Parsimony Hypothesis, which predicts that it is the nature of the C_E that is at stake: more semantically specified (but contextually unmotivated) material is less preferred in presupposed clauses. An asserted C_E imposes no such requirements and a large antecedent is significantly more available. It is important to understand that while the rates of large-antecedent responses are higher for the extraposed embedded C_E than the non-extraposed embedded

---

[16]There was no significant difference between the use of *the* vs. *every* (p > 0.19). The numerical trend was for a greater proportion of large-antecedent responses with *every* than *the*. This was true for both *embedded* and *main* C_Es.

| ADVERB POSITION | Without *also* | | With *also* | |
	embedded	main	embedded	main
pre-verbal	.24 (.37)	.50 (.38)	.19 (.30)	.51 (.41)
post-verbal	.23 (.31)	.58 (.39)	.33 (.04)	.64 (.37)

Table 1: Experiment 2, mean proportion of large antecedents (standard deviations)

C_E – which might give the impression that the two conditions do not pattern alike, as the ACH predicts – this difference is a part of a second *main* effect of ADVERB POSITION that holds across C_E-type. This means that the difference between the two embedded conditions can be factored out: adverb position is an additional factor that promotes small antecedents across clause types, which I will now address.[17]

The main effect of ADVERB POSITION was not predicted by either the ACH or PH. Nonetheless, I would like to briefly outline a possible explanation for this effect. It is well known that there are semantic differences between pre-verbal and post-verbal adverbs (Jackendoff 1972; Peterson 1997; Shaer 2000). Morzycki (2008) argues that pre-verbal adverbs may be interpreted non-restrictively whereas post-verbal adverbs must always restrict the VP. If Morzycki makes the bet in (16a) and it turns out that Floyd performed the nose job but not easily, he loses. In (16b), where the adverb is pre-verbal, it is not clear whether the bet is won or lost.

(16) a. I'll bet you $80 that Floyd could perform a successful nose job in a moving taxi easily.

   b. I'll bet you $80 that Floyd could easily perform a successful nose job in a moving taxi.   (Morzycki 2008: 105–106(17,18))

This is because pre-verbal adverbs have both a restrictive and non-restrictive reading and a non-restrictive modifier can 'escape' the propositional content of the

---

[17]An anonymous reviewer contends that the results are consistent with the ACH, though they are not "fully predicted by them" because the ACH makes no predictions about which antecedent is chosen when the embedded clause is extraposed. This is true, but what matters for evaluating these competing hypotheses is that the ACH *fails* to predict the results we found whereas the PCH *does* predict those results—namely, that both the extraposed and non-extraposed conditions show a small-antecedent effect compared to their respective baselines, and the difference with respect to those baselines is not significantly different. The reviewer suggests that the results "could be seen as evidence that something else in addition" to the ACH is needed to explain the preference for small antecedents even in the extraposed conditions. But this is a misinterpretation of the difference between the two embedded conditions, which is attributable to the independent main effect of ADVERB POSITION. The null hypothesis is that the main effect of CLAUSE TYPE has one source, not, as the reviewer suggests, two separate sources that are indistinguishable (one for extraposed embedded C_Es and another for non-extraposed embedded C_Es). The ACH does not predict such a single source for small antecedents across embedded C_Es, whereas the PH does.

complement clause. (See Potts 2005 on how this dimension of 'not-at-issue' content is integrated.) Relevant to the present discussion is the fact that there is existing experimental evidence to support the idea that the search for antecedents for VPE need not target non-restrictive material. Frazier and Clifton (2011) show that, rather unexpectedly, the subject of the second clause in (17a) can be interpreted as though modified by the restrictive relative in the subject of the first clause. However, the same effect was not found with non-restrictive relatives (17b).

(17) a. A man who had climbed Mt. Greylock called and a woman did too.

b. A man, who had climbed Mt. Greylock, called and a woman did too.

They argue that parallelism requirements are responsible for the enrichment to the meaning of the subject of the second clause in (17a). They suggest that not-at-issue material like the non-restrictive relative can escape the kinds of parallelism requirements imposed on (17a). It is possible, then, that a somewhat similar explanation applies to the main effect of ADVERB POSITION in Experiment 2: a non-restrictive pre-verbal adverb similarly escapes parallelism requirements and need not be part of the ellipsis. Post-verbal adverbs, which are only restrictive, necessarily form part of the main assertion of the antecedent clause. As established by Hardt and Romero (2004) and Frazier and Clifton (2005), content from a main assertion is more likely to be resolved as part of an ellipsis. While this proposal for the effect of ADVERB POSITION requires further testing, it bears repeating that in Experiment 2, this effect held regardless of C_E type: it will not by itself explain the small-antecedent effect when C_E is part of a presupposed restrictor clause.

Another informative result of Experiment 2 concerns the role of *also*, a parallelism-promoting item. Numerically, the presence of *also* does appear to increase large-antecedent responses in the post-verbal conditions, but the effect was not significant. So while such items may boost large-antecedent responses, their presence cannot explain the contrast between main and embedded C_Es.[18]

One potential drawback is that in Experiments 1a, 1b, and 2, coordinated clauses are used as a baseline. Even without discourse particles like *too* and *also*, coordinated clauses may impose parallelism constraints that demand not just matching antecedent and ellipsis, but overall parallelism between the C_A and C_E as discussed in section 2 (Frazier et al. 2000; Frazier and Clifton 2001; Carlson 2003). To ensure that the small-antecedent effect is not a by-product of comparing embedded C_Es to parallelism-seeking coordinated clauses, Experiment 3 was designed to hold the relation

---

[18]There was, however, a large drop in the proportion of large-antecedent responses for the coordinated ellipses compared to Experiment 1a/1b from 71%/84% to 56%. One possibility is that the absence of *also* in half the items was responsible. But this would not entirely explain the difference between the experiments, since the coordinated cases with *also* (most comparable to Experiment 1 *and*-conditions) still showed only 64% large-antecedent responses. Perhaps the items that did not include *also* primed participants to entertain a small antecedent in the coordination cases. At the outset, I assumed that main clauses would take a large antecedent as the default, but the factors at play in main clauses may themselves be more complex. In addition, the raw numbers should not be given too much importance.

between C_E and C_A constant. To do this, VP ellipsis was tested in temporal modifiers headed by *after, when* and *before* as in (18a).

(18) a. Sally quickly read the book after/when/before Fred did.

b. After/when/before Sally quickly read the book, Fred did.

These were compared to minimally different sentences in which the ellipsis was in a main clause (18b). Neither involves coordination, and so we do not expect overall parallelism constraints to enforce maximal similarity between the two clauses. The only difference is that the ellipsis is in a matrix clause in (18b), but in a temporal embedded clause in (18a).

These stimuli offer an additional way to test the PH against the ACH. Sailor (2014) concentrates almost exclusively on adverbial subordinate C_Es like these. Investigating these provides a further test of the family of approaches the ACH encapsulates. Furthermore, the PH makes predictions about ellipsis in temporal adjuncts because they are presuppositional, containing definite descriptions of times. A number of other authors argue that the prepositions *after* and *before* embed expressions that refer to unique times (Beaver and Condoravdi 2003; von Stechow 2002, 2009; Condoravdi 2010; Sharvit 2014). Heim (1987), for instance, gives a paraphrase with a definite description:

(19) Mary left before/after John arrived.
    → Mary left before/after **the time** at which John arrived.

As von Stechow (2009) and others point out, this paraphrase incorrectly presupposes that John arrived just once. Beaver and Condoravdi (2003) propose that the *before*-clause contains an operator equivalent to *earliest*, which does not require one event but does still presuppose there is a unique earliest time:[19]

(20) "A before B" = 1 iff $(\exists t \in A)\ t < earliest(B)$
    "A after B" = 1 iff $(\exists t \in A)\ t > earliest(B)$

(21) Sally quickly read the book before Fred did.
    → Sally quickly read the book before "the earliest time" that Fred read the book.

When-clauses have likewise been argued to semantically contain a definite description of times. Caponigro (2004), Caponigro and Pearl (2009) and Hall and Caponigro (2010) argue that *when*-clauses are free relatives, and they may contain a maximality operator (in the sense of Link 1983): the *when*-clause denotes a set of time intervals and the maximality operator shifts this set to the maximal element in that set. In (22) the maximal element is the largest time interval in which Sally read the book.[20]

---

[19]von Stechow (2009) suggests that the operator in *after*-clauses might be *latest* rather than *earliest*.

[20]There are non-maximal interpretations for when-clauses, as in (i):

(i) Call me when you have a moment.   (Hall and Caponigro 2010: 559(76))

The experimental materials were designed to rule out non-maximal interpretations for *when*-clauses by describing past episodic events in both clauses.

(22) John quickly read the book when Sally did.
      → John quickly read the book at/during **the** time interval in which Sally read it.

Since (22) has this "hidden" definite description, we expect its restrictor (the set of times denoted by the C_E) to be subject to constraints on parsimony, just as relative clauses that form the restrictor of the overt definite article are.[21] Resolving to a large antecedent requires hearers to accommodate in (22) not just the particular time at which Sally read the book, but also the manner of that event. The Parsimony Hypothesis predicts readers will choose the unmodified antecedent more often in the embedded clause ((18a), (22)) than the main clause (18b).

5.   EXPERIMENT 3

Methods

Experiment 3 tested two factors each with two levels. C_E TYPE: *main* vs. *embedded* (using *after*, *before*, or *when*); and ADVERB POSITION: *pre-verbal* vs. tpost-verbal. The two factors were crossed to create the four conditions in (23).

(23) a. Peter quietly walked into the house after Lesley did.   *embed/pre*
      b. Peter walked into the house quietly after Lesley did.   *embed/post*
      c. After Peter quietly walked into the house, Lesley did.   *main/pre*
      d. After Peter walked into the house quietly, Lesley did.   *main/post*

Twenty items were created in four conditions following the pattern in (23), using the three kinds of temporal subordinators (*when, after, before*). These items were included among twenty other items, which consisted of twelve three- or four-sentence discourses testing an unrelated phenomenon. Eight filler items were included. Each sentence was followed by a forced-choice question with the modified and unmodified antecedents as options. Four counterbalanced lists were created, so that each participant saw each item once in only one condition. Materials appear in Appendix C. Forty native English speaking undergraduates participated in a written comprehension questionnaire. As with Experiment 2, the ACH predicts an interaction between C_E TYPE and ADVERB POSITION whereas the PH predicts a main effect of C_E TYPE.

Results

Figure 4 shows the mean proportion of large-antecedent responses. When the ellipsis was presented in a main clause, participants chose the large antecedent more often

---

[21]A reviewer points out that not all adverbial ACDs may be amenable to a definite description or quantificational treatment (e.g., clauses introduced by *even though* or *although*, but see Schlenker (2004) for such an analysis of conditionals). I predict that these will not show a small-antecedent effect. This requires further testing.
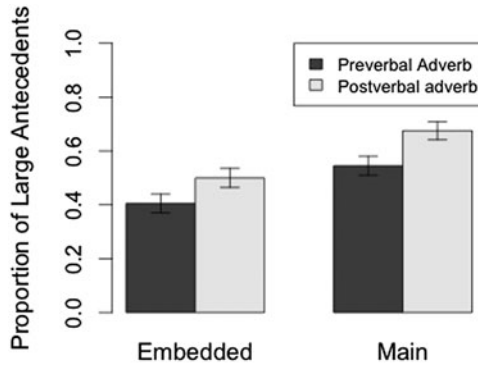
Figure 4: Mean proportion of large-antecedent responses in Experiment 3

than if the ellipsis was found in a temporal embedded clause. In both cases, more large-antecedent responses were observed if the adverb was post-verbal in the antecedent.

A mixed model was fit, which had as fixed factors C_E TYPE and ADVERB POSITION, and participant and item as random factors with a maximal random effects structure. The analysis showed a significant effect of C_E TYPE (Est. = 1.1405, SE = 0.3703, z = 4.135, p < .001) and a significant effect of ADVERB POSITION (Est.=−0.7754, SE = 0.2383, z= −3.254 , p < .001). There was no interaction.

Discussion

The results of Experiment 3 are not compatible with the hypothesis that the small-antecedent effect is due only to a countervailing preference for large antecedents in coordination due to parallelism. The embedded clauses differed from non-coordi-nated main clauses in the likelihood of large-antecedent selection. Additionally, the experiment confirmed that the small-antecedent effect is not due to a containment problem, as the ACH predicts: for the condition where the C_E is not contained in the modified VP (the *post-verbal embedded* condition), readers were as likely to chose an unmodified antecedent compared to main clauses as they were for the non-contained *pre-verbal embedded* conditions. The separate main effect of adverb position repli-cates that found in Experiment 2 and bears out Frazier and Clifton's (2011) proposal that non-restrictive material is less likely to be targeted as part of an antecedent.

Summarizing, Experiment 3 shows once again that the small-antecedent effect arises in embedded clauses generally, and it cannot be attributed to a containment problem alone as per the ACH. Instead, the predictions of the PH are borne out. Like ACD constructions headed by *the* and *every* which showed the small-antecedent effect, temporal adverbial clauses, which are likewise presuppositional, exhibit a small-antecedent effect compared to asserted main clauses. In the next section, I provide a more explicit formulation of the PH and show how it is part of a more general constraint on accommodation in processing as argued by Crain and Steedman (1985) and Altmann and Steedman (1988).

6. PARSIMONY AND ACCOMMODATION IN PROCESSING

Crain and Steedman (1985) claim that the famous garden path in is a result of prag-
matic considerations, not parsing constraints (see Frazier 1978).

(24) The horse raced past the barn fell. (Bever 1970)

The target, but not initially pursued, restrictive relative (RR) parse requires the listener
to not only accommodate a (unique) horse that was raced past the barn, but also a con-
trast set of horses which do not have the property ascribed by the relative.[22] The more-
readily pursued main verb (MV) analysis only requires the accommodation of the
horse, not a contrast set. In this sense, the MV parse requires a more parsimonious
or simpler discourse model. Crain and Steedman (1985) propose a Principle of
Parsimony: the reading that carries fewer unsupported presuppositions[23] will be
favored over one that carries more. However this is implemented formally, the intuition
is clear that the RR parse requires the reader to imagine a more elaborate discourse
model. Grodner et al. (2005) found that even without ambiguous input, readers find
restriction difficult to process in a context that does not motivate restriction. They com-
pared (unambiguous) restrictive and (unambiguous) non-restrictive relatives and found
that in null contexts, the restrictive relatives were read more slowly. (When supporting
contexts were provided, the effect reversed direction.) They argue that the results
support the Principle of Parsimony, in that the restrictive relative requires a more elab-
orate discourse model (evoking a contrast set) which the non-restrictive does not.[24]
Additionally, a number of studies have shown that individuals prefer unmodified
NPs in a number of ambiguous constructions and that the Principle of Parsimony is
responsible (Ni 1996; Sedivy 2002; Spivey-Knowlton and Sedivy 1995).

The small-antecedent effect has a similar, albeit not identical flavour. In the case
of ellipses in presupposed clauses, choosing a large, modified antecedent, as in (25b),
leads to a richer discourse model than the small antecedent (25a): the large resolution
ascribes not only the property of being read by Mary to the book in question, but the
manner as well.

(25) a. John quickly read the book that Mary did ~~read~~.

b. John quickly read the book that Mary did ~~quickly read~~.

If the ellipsis is part of the restrictor of a presuppositional element (a definite descrip-
tion, a universal), it requires accommodation in a null context (Karttunen 1974;

---

[22]Modifiers do not always implicate contrast – see below – but the claim here is that readers
are more likely to pursue a contrastive interpretation for the relative clause (Sedivy 2003).

[23]Crain and Steedman (1985) used the word *presupposition* to describe the contrast set
component, but Portner (1989) and Sedivy (2003) show that the contrast set evoked by a
restrictive relative is a conversational/Gricean implicature.

[24]The principle of parsimony relies, in turn, on a more general pressure following from the
Gricean Maxim of Quantity: Speakers should say as much as needed to be informative, without
saying more than is necessary (Grice 1975). That is, since the simple discourse model is pre-
ferred, any additional modification in the description is redundant.

Stalnaker 1974; Lewis 1979; see von Fintel 2004 for a recent overview). When readers are given a choice between parses that require accommodation, they choose the parse that requires the 'least' accommodation, in the sense that the relative clause with an unmodified (small) VP denotes a superset of the entities picked out by a relative with a modified (large) VP.

(26) $\underbrace{\{x : \text{Mary quickly read x}\}}_{\text{large antecedent}} \subseteq \underbrace{\{x : \text{Mary read x}\}}_{\text{small antecedent}}$

Put another way, the unmodified relative is less informationally rich and, in the null context, more parsimonious.[25]

In this way, Parsimony not only militates against introducing more individuals into a discourse model (as in Crain and Steedman 1985) but it evaluates between different "amounts" of accommodation, where the subset relation ranks the information content of the restrictors.

(27)   *Parsimonious Accommodation*
       All else being equal, in a null context, given two parses [Det A] and [Det B], where Det is a presuppositional determiner and A and B are properties such that $A \subset B$, choose parse [Det B].

Parsimony only regulates restrictor material that requires accommodation. When the elided phrase is part of asserted content, the pressures of parsimony are called off. This is consistent with Experiments 1–3, where main clauses did not show a small-antecedent preference. In the last section, we saw that temporal adjunct clauses could be profitably analyzed as involving definite descriptions of times. If their content is subject to accommodation, then here too a less informationally rich (i.e. unmodified) ellipsis is chosen.

It is important to stress that Parsimonious Accommodation is intended here only as a processing heuristic. There is certainly no barrier to accommodating "more" than is strictly necessary in, for instance, bridging scenarios:

(28)   John read a book about Schubert and wrote to the author.   (Heim 1982: 371)

As Heim (1982, p. 371ff.) points out, the author we accommodate is not just any author (that would certainly be more parsimonious) but the author of the book mentioned in the first clause. In my terms, this amounts to accommodating a richer set of properties of the individual. But here the pressures of discourse coherence motivate the bridging.[26] My claim is that Parsimonious Accommodation governs parsing in

---

[25]This proposal resembles Schlenker's (2005) *Minimize Restrictors!*, which militates against the inclusion of redundant modifiers. But that is a constraint on how much one says in a given context. The present claim is a little different: readers do not know what is redundant or not, since the context is so impoverished. The Principle of Parsimony is intended to govern the richness of the discourse model a hearer constructs.

[26]The reader might wonder how Parsimonious Accommodation relates to the principle of *Maximize Presupposition* (Heim 1991), which requires the speaker to choose a presuppositional competitor (e.g., *the*) over a non-presuppositional one (e.g., *a*) when the presupposition

impoverished contexts: when faced with a syntactic ambiguity, readers appeal to it. This, of course, was also the intended application of the Principle of Parsimony in Crain and Steedman (1985).

A reviewer asks how the processor enforces (27): does the processor compare the available options in parallel and evaluate which one requires fewer commitments? Another option, as pointed out by the reviewer, is that the reader might pursue the small antecedent first without entertaining the large antecedent. We leave these interesting questions about time-course to future research.

The natural prediction is that non-presuppositional determiners will lift the small-antecedent effect. The simplest direct test of this prediction would be to choose a relative clause C_E that is headed by a non-presuppositional determiner. This is not so straightforward, however, since only strong, presuppositional determiners license ACD (Carlson 1977; Diesing 1992). For instance, the weak cardinal reading of *few* in (29a) does not license ACD, whereas the partitive version (which is always presuppositional) does (29b):[27]

(29) a. ??I read few books that you did.

b. I read few of the books that you did.   (Diesing 1992: 71)

Since non-presuppositional determiners and quantifiers cannot head ACD, the most direct test of the Parsimony Hypothesis is not possible. Instead, I test non-restrictive relatives, in which the content of the relative is asserted and therefore not subject to Parsimonious Accommodation.

7. EXPERIMENT 4: RESTRICTIVE VS. NON-RESTRICTIVE

Unlike restrictive relatives, the content of non-restrictive relatives does not serve as the restrictor of the head noun, and so is not subject to accommodation. Instead, non-restrictive relatives form a type of secondary assertion (see Potts 2005).[28]

(30) a. John read the book that is on the table.
    assertion: John read the book
    presupposition: there is a (unique) book on the table

---

is satisfied – in Heim's words "Make your contribution presuppose as much as possible!" Maximize Presupposition might be characterized as a speaker-based principle, while Parsimonious Accommodation is a hearer-based heuristic for impoverished contexts, and so I do not think the two principles are in conflict.

[27]Fiengo and May (1994) demonstrate the effect with the ambiguity of *many*:

(i) Dulles suspected many spies that Angleton did.
    (Fiengo and May 1994: 242(12c))

This is only grammatical with a strong presuppositional interpretation of *many*. Fiengo and May (1994) argue that it is actually not ACD that forces a presuppositional quantifier, but the presence of a restrictive relative.

[28]It has been argued (e.g., Emonds 1979; McCawley 1981, among others) that non-restrictive relatives are in a high syntactic position at LF. The recent literature suggests that this is not the case (Potts 2002).

    b. John read Moby Dick, which is on the table.
       assertion: John read Moby Dick
       secondary assertion: Moby Dick is on the table

Parsimonious Accommodation is not expected to constrain ellipsis choice in non-restrictive relatives because their content is not presupposed. Testing non-restrictives, however, is not without problems, as non-restrictive relatives often resist ACD:

(31) *Dulles suspected Philby, who Angleton did.   (May 1985: 12(23))

Vanden Wyngærd and Zwart (1991) report that non-restrictive ACD is ameliorated with *as well* or *not*. (See also Pesetsky 2000; Lasnik 1999.)

(32) ?Dulles suspected Philby, who Angleton did not/as well.

                            (Vanden Wyngærd and Zwart 1991: 154(19))

To the extent that non-restrictive ACDs are acceptable, my judgement is that a large antecedent is available.

(33) a. Dulles wrongly suspected Philby, who Angleton did ~~wrongly suspected~~ as well.

    b. Maria quickly spoke to Regina, who Bill did not ~~quickly speak to~~.

However, we saw in Experiment 2 that additive modifiers like *as well* could boost the proportion of large antecedents (due to the parallelism constraints it imposes). To counteract this effect, Experiment 4 tested restrictive and non-restrictive relatives both with *as well*. This may increase the overall number of large antecedents, but I assume this would apply equally across C_E-TYPE.

(34) a. Rob secretly suspected Mary, who Peter did as well.   *non-restrictive*

    b. Rob secretly suspected the woman that Peter did as well.   *restrictive*
  What did Peter do?
      Secretly suspect Mary
      Suspect Mary

The restrictive relative forms the descriptive content of the definite determiner, and so requires accommodation. I have argued that accommodation in processing is subject to parsimony, and readers will choose the less informative NP meaning (the unmodified VP). The non-restrictive relative is not presupposed, and instead serves as a type of secondary assertion (Potts 2005). If the small-antecedent effect is due to pressures on accommodation, as hypothesized, we predict a difference between the two relatives in antecedent choice such that small antecedents will be chosen more often in restrictive relatives.

Methods

Twelve item sets like (34) were created with the two-level factor C_E-TYPE: *non-restrictive* and *restrictive*. The order of answer options (modified, unmodified) was counterbalanced. Twenty participants completed one of two lists using Amazon Mechanical Turk and were paid $1.25 for their participation. All participants identified themselves as native English speakers.

Results

The mean proportion of large antecedents is reported in Figure 5. Non-restrictive relatives received 39.1% large-antecedent responses; restrictives, 27.5%. A mixed model was fit with C_E type as a fixed factor, and participant and item as random factors. The model also included random intercepts and random slopes for participants and items. There was a significant effect (Est.$=-0.8087$, SE $= 0.3995$, $z = -2.025$, p < 0.05).

Discussion

The results point in favour of the hypothesis that a pressure for small antecedents is conditioned by the presuppositional versus assertoric nature of the clause containing the ellipsis. Relative clauses that form the restrictor of a definite must be accommodated, and this accommodation is subject to Parsimony, giving rise to a greater preference for an unmodified VP resolution. Non-restrictive relatives do not form the restrictor of the presuppositional determiner, but are instead secondary assertions that do not require accommodation. Hence, there is not the same pressure to choose a small antecedent.

   The raw scores for large antecedents in non-restrictive relatives are not overwhelmingly high, as they are in main clauses in Experiments 1a, 1b or Experiment 2. In fact, they do not reach chance. A reviewer points out, however, that there may be several reasons for "small antecedents", Parsimonious Accommodation being only one of them. In this particular case, there may be several reasons why non-restrictives do not prefer large antecedents. First, the baseline grammaticality of ACD in non-restrictive relatives may not be high, and this could have the effect of creating more noise in the results if less acceptable sentences are harder to judge. Second, this experiment only employed pre-verbal adverbs, and we have already seen that pre-verbal adverbs in the antecedent are less likely to correlate with modified ellipsis interpretations. Finally, it is important to keep in mind that raw numbers are hard to interpret, especially without a baseline of coordination. The important result is that the non-restrictives are different from restrictive relatives, even when both include the parallelism-inducing *as well*, which supports the claim that Parsimonious
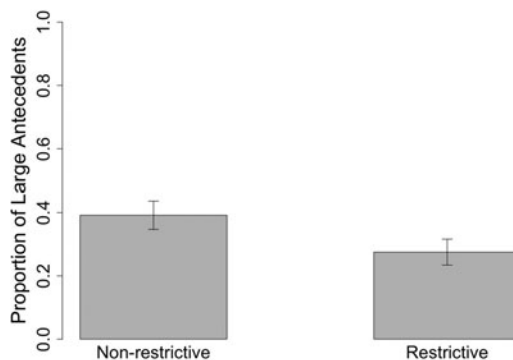

Figure 5: Mean proportion of large-antecedent responses in Experiment 4

Accommodation plays a role, although further research may find that there are additional factors that militate against large antecedents.

8. GENERAL DISCUSSION

I began with the intuition, reported by Matsuo (2001), that in ACD constructions manner adverbs are not targeted as part of the ellipsis. I showed that while that is not an absolute, grammatical, prohibition, it is a replicable trend in speakers' interpretation of elided sentences. I set out to determine the reason for the small-antecedent effect, ruling out along the way a very plausible syntactic explanation (the ACH). Ultimately, the explanation that fits the experimental results best is the one that attributes the small-antecedent effect to a very general pragmatic principle that militates against accommodating a richer context than needed (Crain and Steedman 1985). When readers are given a choice between parses that require accommodation, in a null context they choose the parse that requires the 'least' amount of modification in the sense of Parsimonious Accommodation. Evidence in favour of the Parsimony Hypothesis—and Parsimonious Accommodation in general—involved comparing restrictive to non-restrictive relatives; in the latter case the content of the relative does not need to be accommodated because it is a (secondary) assertion.

It becomes an interesting question whether adverbial modification is dis-preferred in embedded contexts generally, even outside of ellipsis, as in (35):

(35) John read the book that Fred quietly read.

Of course, we are forced in (35) to accommodate the adverb because it is overtly present. This is unavoidable. What the small-antecedent effect has shown us is that when faced with an ambiguity, readers will choose a syntactic parse that does not force "more" accommodation than necessary. This demonstrates that ellipsis resolution is readily influenced by pragmatic principles governing general expectations about information exchange and complexity in a discourse. This is generally recognized in the literature (Hardt and Romero 2004; Frazier and Clifton 2005) but we are only beginning to understand how to dissociate these pressures from syntacic-semantic principles governing the identity conditions on ellipsis resolution. I should point out, though, that Grodner et al. (2005) found that in null contexts, even unambiguous restrictive relative clauses were read more slowly than non-restrictives. They suggest that the Principle of Parsimony governs routine syntactic parsing, not just in cases of syntactic ambiguity. This raises the possibility that "excess" modification within a restrictive relative, perhaps as in (35), will pose processing costs. I leave it to future research to investigate this.

As noted, the simplest direct test of the PH would have used relative clause C_Es headed by a non-presuppositional determiner. This is not possible, however, as only strong, presuppositional determiners license ACD (Carlson 1977; Diesing 1992; Fiengo and May 1994). Weak determiners can, however, contain relative clauses—especially in presentational contexts like (36a) (*sm* is the reduced version of *some*, and it is unambiguously weak):

(36) a. Here are *sm* books that John bought.

    b. Here are some books that John bought.

With *sm*, (36a) does not *presuppose* that the restrictor is non-empty; it introduces to the common ground that there is a set of books John bought. The non-reduced *some* in (36b) is ambiguous. While weak determiners do not allow ACD, they can have elided VPs as long as the VP is not contained in the antecedent, for example if the antecedent is sentence-external as in (37):

(37) a. Speaker A: The candlesticks that were carefully cleaned aren't here.

    b. Speaker B: Here are *sm* that were ~~carefully cleaned~~.

A large antecedent is certainly possible in (37b), as the proposal predicts, but the fact that the manner-modified antecedent VP itself requires accommodation might prime the reader to further accommodate the same kind of VP event in the elided clause.

    Since Parsimonious Accommodation does not tie the small-antecedent effect to anything particular about ACD, it is predicted that *any* ellipsis in a presuppositional DP will be constrained by parsimony and show a small-antecedent effect. In (38a), the elision is part of a subject relative under a definite; in (38b), the elision is in a presentational (non-presupposed) relative. (I have put the antecedent VP in an asserted clause to avoid priming accommodation.)

(38) The man quietly sang the song and left.

    a. The woman that did ~~sing the song~~ stayed.

    b. There's a woman that did ~~quietly sing the song~~ and she stayed.

While testing is required, to my ears (38a) does not commit to the manner of the elided VP in the same way that (38b) does. Nonetheless, the effect may be attenuated because in the creation of such examples, the context becomes already quite a bit richer. I leave this for further research.

    Aside from pragmatic processing questions, the small-antecedent effect also bears on prominent issues in the syntax of ellipsis. One of these issues is the claim that ellipses target the largest constituent available (MAXELIDE). Another is the representation of manner adverbs in the VP. I turn to these in the concluding subsections.

8.1. Syntax of small antecedents and MAXELIDE

The small-antecedent effect might appear at first glance to be counter-evidence for the widely-advocated principle MAXELIDE (Merchant 2008). As it turns out, MAXELIDE is orthogonal to the issue, but it is worth seeing why. In its original form, MAXELIDE requires that ellipses target the largest possible constituent in environments involving trace binding like (39). This is why the VP ellipsis in (39a) sounds deviant compared to the sluiced TP in (39b).

(39) a. ?They said they heard about a Balkan language, but I don't know which they did ~~hear about t~~$_i$.

   b. They said they heard about a Balkan language, but I don't know which$_i$ ~~they heard about t~~$_i$.

MAXELIDE has been extended by other authors to regulate even elided material that does not involve trace binding (Fox and Takahashi 2005; Kim et al. 2011).[29] Kim et al. (2011) propose a generalized MAXELIDE, , and argue that it is a processing heuristic:

(40) MAXELIDE (Kim et al. 2011, p. 16)
      VP ellipsis preferentially targets configurationally higher rather than lower nodes.

Kim et al. use MAXELIDE to distinguish the highly acceptable case of voice-matching ellipsis in (41) from the degraded voice mismatched ellipsis in (42). Assuming active and passive voice are syntactically encoded by a voice head ($v_{active/passive}$) (Merchant 2013), the match cases involve elision of a larger constituent (vP) while the mismatch cases involve only elision of a VP (stranding a voice head), in violation of MAXELIDE.

(41) *Voice match*
      Jill [$_{vP}$ v_{act} [$_{VP}$ betrayed Abby]], and Matt did [$_{vP}$ ~~v_{act}~~ [$_{VP}$ ~~betrayed Abby~~]] too.

(42) *Voice mismatch*
      Jill [$_{vP}$ v_{act} [$_{VP}$ betrayed Abby]], and Matt was [$_{vP}$ v_{pass} [VP ~~betrayed by Jill~~]] too.

While it is tempting to see the small-antecedent effect as conflicting with MAXELIDE, it turns out to be consistent with a certain interpretation of (40). MAXELIDE makes no distinction between the large-antecedent choice (43) and the small-antecedent choice (44): both involve the elision of the largest verb phrase available in the C_E.

(43) *Large Antecedent*
      J. [$_{vP}$ v [$_{VP}$ quickly [$_{VP}$ read the book]]] after B. did ~~[$_{vP}$ v [$_{VP}$ quickly [$_{VP}$ read the book]]]~~

(44) *Small antecedent*
      J. [$_{vP}$ v [$_{VP}$ quickly [$_{VP}$ read the book]]] after B. did ~~[$_{vP}$ v [$_{VP}$ read the book]]~~

The difference between (42) and (44) is that in the former the C_E *must* contain the *v* head as a matter of basic grammaticality, and this is material that could *in principle* be elided but cannot be in this instance because it does not match with *v* in the antecedent. In the small-antecedent case in (44) the C_E does not contain an adverb and therefore involves elision of the largest candidate for VPE, consistent with MAXELIDE. The important point is that MAXELIDE does not govern the *choice* of ellipsis; rather, for any given grammatically available resolution of ellipsis, MAXELIDE demands elision of the largest verbal constituent available. The two principles are simply orthogonal.

---

[29]The picture is likely much more complicated once auxiliaries are considered; see Aelbrecht and Harwood (2015) on preservation of auxiliaries in ellipsis.
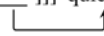
8.2. Position of post-verbal adverbs

The small-antecedent effect bears upon how manner adverbs are represented syntactically. If adverbs are introduced as specifiers of dedicated functional heads (Cinque 1999), and those heads may be omitted if there is no adverb, the predictions remain the same. However, the results are inconsistent with Larson's proposal that post-verbal adverbs are actually merged low as sisters to the verb (Larson 1988; Stroik 1990, 1996; Alexiadou 1997; Larson 2004). This is shown in (45), where it is assumed that the object is a specifier in VP and the verb moves to a v node.

(45) $[_{vP}\ v + \text{read}\ [_{VP}\ \text{the book}\ [_{V'}\ \_\_\_\ \text{quickly}\ ]]]$

If this were the representation of the antecedent vP, there would be no vP or VP constituent that contained the small antecedent to the exclusion of the adverb. Advocates of such a structure would require movement of the adverbitself – perhaps rightward[30] – to produce a small antecedent:

(46) $[_{vP_2}\ [_{vP_1}\ v + \text{read}\ [_{VP}\ \text{the book}\ [_{V'}\ \_\_\_\ \_\_\_\ ]]]\ \text{quickly}\ ]$

The string dominated by vP_1 is a constituent that excludes the adverb, at least as far as surface syntax is concerned. Such movements must be available on the V-sister analysis of adverbs because adverbs can be stranded by ellipsis:

(47) John read the book quietly and Bill did ~~read the book~~ loudly.

There are two problems that arise for this solution to the small antecedents discovered in the experiments. In the case of (47), it is clear that *both* adverbs bear focus. It is not clear to me that in the sentences that have small-antecedent interpretations, such as (48), the adverb in the antecedent clause is focused:

(48) John read the book quietly after Bill did ~~read the book~~.

If *quietly* does not bear focus in (48), and if it does not vacate the VP as in (47), then there would be no suitable small antecedent available, contrary to fact. The second problem arises from the fact that surface structure alone cannot provide an antecedent for VP ellipsis (Sag 1976). The antecedent vP in (46) is really one with a trace of an adverb; the elided vP in (47) matches its antecedent because it too will have a trace position for an adverb (the trace left by *loudly*). But in (48) – where a small antecedent is chosen – the ellipsis site does not contain a trace of an adverb so it would not be able to match the antecedent, even if the adverb vacated the vP. In sum, treating post-verbal adverbs as V-sisters is not consonant with the small-antecedent effect.[31]

However tentative these concluding remarks may be, and the explanation for the small-antecedent effect in terms of Parsimonious Accommodation, the experimental evidence securely demonstrates that such ellipses are possible and often preferred in embedded clauses. Moreover, the effect cannot be due to syntactic containment, as

---

[30]This could also be leftward movement of the adverb followed by remnant vP movement.

[31]Pesetsky (1995) suggests that right-ascending and right-descending structures are simultaneously available. In that case, the small antecedent would be consistent only with the right-ascending structure. Further tests would need to determine if this can be independently corroborated.

Matsuo ([2001](#)) and Sailor ([2014](#)) predict.[32] Any theory of VP ellipsis needs to account for this.

REFERENCES

Aelbrecht, Lobke, and William Harwood. 2015. To be or not to be elided: VP ellipsis revisited. *Lingua* 153: 66–97.

Akmajian, Adrian, and Thomas Wasow. 1975. The constituent structure of VP and AUX and the position of the verb *be*. *Linguistic Analysis* 1(3): 205–245.

Alexiadou, Artemis. 1997. *Adverb Placement: A case study in antisymmetric syntax.* Amsterdam: John Benjamins.

Altmann, Gerry, and Mark Steedman. 1988. Interaction with context during human sentence processing. *Cognition* 30(3): 191–238.

Baayen, Harald R., Douglas J. Davidson, and Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4): 390–412.

Baltin, Mark. 1987. Do antecedent-contained deletions exist? *Linguistic Inquiry* 18(4): 579–596.

Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3): 255–278.

Bates, Douglas M. 2005. Fitting linear mixed models in R: Using the lme4 package. *R News: The Newsletter of the R Project* 5: 27–30.

Beaver, David, and Cleo Condoravdi. 2003. A Uniform Analysis of *Before* and *After*. In *Proceedings of SALT XIII*, ed. Robert B. Young and Yuping Zhou, 37–54. Cornell University, Ithaca, NY: CLC Publications.

Bever, Thomas. 1970. The cognitive basis for linguistic structures. In *Cognition and the development of language*, ed. John R. Hayes, 279–362. New York: Wiley.

Bouton, L. F. 1970. Antecedent contained pro-forms. In *Papers from the sixth regional meeting of the Chicago Linguistics Society*, 154–167. Chicago: Chicago Linguistics Society.

Caponigro, Ivano. 2004. The semantic contribution of *wh*-words and type shifts: Evidence from free relatives crosslinguistically. In *Proceedings of SALT XIV*, ed. Robert B. Young, 38–55. Cornell University, Ithaca, NY: CLC Publications.

Caponigro, Ivano, Lisa Pearl. 2009. The nominal nature of *Where*, *When*, and *How*: Evidence from free relatives. *Linguistic Inquiry* 40(1): 155–164.

Carlson, Gregory N. 1977. Amount relatives. *Language* 53(3): 520–542.

Carlson, Katy. 2003. *Parallelism and prosody in the processing of ellipsis sentences.* New York: Routledge.

Chung, Sandra. 2013. Syntactic identity and sluicing: How much and why. *Linguistic Inquiry* 44(1): 1–44.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective.* New York: Oxford University Press.

---

[32]This does not mean that the other types of small-antecedent effects Sailor identifies, as described in section 2, are not the result of a containment problem, as Sailor suggests.

Condoravdi, Cleo. 2010. NPI-licensing in temporal clauses. *Natural Language and Linguistic Theory* 28(4): 877–910.

Crain, Stephen, and Mark Steedman. 1985. On not being led up the garden path: The use of context by the psychological parser. In *Natural language parsing*, ed. David Dowty, Lauri Karttunen and Arnold Zwicky, 94–128. Cambridge: Cambridge University Press.

Diesing, Molly. 1992. Bare plural subjects and the derivation of logical representations. *Linguistic Inquiry* 23(3): 353–380.

Emonds, Joseph. 1979. Appositive relatives have no properties. *Linguistic Inquiry* 10(2): 211–243.

Erlewine, Michael Yoshitaka, and Hadas Kotek. 2016. A streamlined approach to online linguistic surveys. *Natural Language and Linguistic Theory* 34(2): 481–495.

Ernst, Thomas. 2001. *The syntax of adjuncts*. Cambridge: Cambridge University Press.

Fiengo, Robert, and Robert May. 1994. *Indices and Identity*. Cambridge, MA: MIT Press.

von Fintel, Kai. 2004. Would you believe it? The king of France is back! In *Descriptions and beyond*, ed. Marga Reimer and Anne Bezuidenhout, 315–341. Oxford: Oxford University Press.

Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, MA: MIT Press.

Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33(1): 63–96.

Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and scope: A case for overt QR. In *Proceedings of the West Coast Conference on Formal Linguistics 18*, ed. Sonya Bird, Andrew Carnie, Jason D. Haugen and Peter Norquest, 132–144. Somerville, Massachusetts: Cascadilla Press.

Fox, Danny, and Jon Nissenbaum. 2003. VP ellipsis and the position of adverbs. *Snippets* 7: 7–8.

Fox, Danny, and Shoichi Takahashi. 2005. MaxElide and the re-binding problem. In *Proceedings of SALT 15*, ed. Effi Georgala and Jonathan Howell, 223–240. Cornell University, Ithaca, NY: CLC Publications.

Frazier, Lyn. 1978. On comprehending sentences: Syntactic parsing strategies. Doctoral dissertation, University of Connecticut, Storrs.

Frazier, Lyn, and Charles Clifton. 2001. Parsing coordinates and ellipsis: Copy *α*. *Syntax* 4(1): 1–22.

Frazier, Lyn, and Charles Clifton. 2005. The syntax-discourse divide: Processing ellipsis. *Syntax* 8(2): 121–174.

Frazier, Lyn, and Charles Clifton. 2006. Ellipsis and discourse coherence. *Linguistcs and Philosophy* 29(3): 415–346.

Frazier, Lyn, and Charles Clifton. 2011. Dynamic interpretation: Finding an antecedent for VPE. In *Processing linguistic structure: University of Massachusetts occasional papers in linguistics 38*, ed. Jesse Harris and Margaret Grant, 23–36. Amherst, MA: GLSA.

Frazier, Lyn, Charles Clifton, and Alan Munn. 2000. Parsing coordinate structures. *Journal of Psycholinguistic Research* 29(4): 343–370.

Grice, H. P. 1975. Logic and conversation. In *Syntax and Semantics, vol. 3: Speech acts*, ed. Peter Cole and Jerry L. Morgan, 41–58. New York: Academic Press.

Grodner, Daniel, Edward Gibson, and Duane Watson. 2005. The influence of contextual contrast on syntactic processing: Evidence for strong-interaction in sentence comprehension. *Cognition* 95(3): 275–296.

Hackl, Martin, Jorie Koster-Hale, and Jason Varvoutis. 2012. Quantification and ACD: Evidence from real-time sentence processing. *Journal of Semantics* 29(2): 145–206.

Hall, David P., and Ivano Caponigro. 2010. On the semantics of temporal *when*-clauses. In *Proceedings of SALT 20*, ed. Nan Li and David Lutz, 544–563. Cornell University, Ithaca, NY: CLC Publications.

Hardt, Daniel, and Maribel Romero. 2004. Ellipsis and the structure of discourse. *Journal of Semantics* 21(4): 375–414.

Harley, Heidi. 2002. WCO, ACD, and QR of DPs. *Linguistic Inquiry* 33(4): 659–664.

Heim, Irene. 1982. The semantics of definite and indefinite Noun Phrases. Doctoral dissertation, University of Massachusetts, Amherst.

Heim, Irene. 1987. Where does the definiteness restriction apply? Evidence from the definiteness of variables. In *The representation of (in)definiteness*, ed. Eric J. Reuland and Alice G. B. ter Meulen, 21–42. Cambridge, MA: MIT Press.

Heim, Irene. 1991. Artikel und Definitheit [Articles and definiteness]. In *Semantik: Ein internationales Handbuch der Zeitgenüssischen Forschung* [Semantics: An international handbook of contemporary research], ed. Arnim von Stechow and Dieter Wunderlich, 487–535. Berlin: de Gruyter.

Hestvik, Arild. 1995. Reflexives and ellipsis. *Natural Language Semantics* 3(2): 211–237.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jacobson, Pauline. 2008. Direct compositionality and variable-free semantics: The case of antecedent-contained deletion. In *Topics in ellipsis*, ed. Kyle Johnson, 30–68. Cambridge: Cambridge University Press.

Jacobson, Pauline, and Edward Gibson. 2014. Processing of ACD gives no evidence for QR. In *Semantics and Linguistic Theory*, Vol. 24, 156–176.

Jaeger, T. Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59(4): 434–446.

Karttunen, Lauri. 1974. Presuppositions and linguistic context. *Theoretical Linguistics* 1(1–3): 181–194.

Kehler, Andrew. 2000. Coherence and the resolution of ellipsis. *Linguistics and Philosophy* 23(6): 533–575.

Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford: CSLI Publications.

Kim, Christina S., Gregory M. Kobele, Jeffrey T. Runner, and John T. Hale. 2011. The acceptability cline in VP ellipsis. *Syntax* 14(4): 318–354.

Lappin, Shalom, and Tanya Reinhart. 1988. Presuppositional effects of strong determiners: A processing account. *Linguistics* 26(6): 1021–1038.

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19(3): 335–392.

Larson, Richard K. 2004. Sentence-final adverbs and "scope". In *Proceedings of the North East Linguistic Society 34*, ed. Keir Moulton and Matthew Wolf, Vol. 1, 23–44. Amherst, MA: GLSA.

Lasnik, Howard. 1999. *Minimalist analysis*. Oxford: Blackwell.

Lebeaux, David. 1990. Relative clauses, licensing, and the nature of the derivation. In *Proceedings of the North East Linguistic Society 20*, ed. Juli Carter, Rose-Marie Déchaine, Bill Philip, and Tim Sherer, 318–332. Amherst, MA: GLSA.

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8(1): 339–359.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, use and interpretation of language*, ed. R. Baeuerle, C. Schwarze, and Arnim von Stechow. Berlin: DeGruyter.

Matsuo, Ayumi. 2001. Asp-shell Structure in VP-ellipsis and ACD. In *Proceedings of WCCFL 20*, ed. Karine Megerdoomian and Leora Anne Bar-el, 20, 386–399. Somerville, MA: Cascadilla Press.

May, Robert. 1985. *Logical form: Its structure and derivation*. Cambridge, MA: MIT Press.

McCawley, James D. 1981. The syntax and semantics of English relative clauses. *Lingua* 53(2): 99–149.

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.

Merchant, Jason. 2008. Variable island repair under ellipsis. In *Topics in ellipsis*, ed. Kyle Johnson, 132–153. Cambridge: Cambridge University Press.

Merchant, Jason. 2013. Voice and ellipsis. *Linguistic Inquiry* 44(1): 77–108.

Morzycki, Marcin. 2008. Nonrestrictive modifiers in nonparenthetical positions. In *Adjectives and adverbs: Syntax, semantics, and discourse*, ed. Chris Kennedy and Louise McNally, 101–122. Oxford: Oxford University Press.

Moulton, Keir. 2008. Small antecedents: Syntax or pragmatics? In *Proceedings of the thirty-seventh annual meeting of the North East Linguistics Society*, ed. Emily Elfner and Martin Walkow, vol. 1, 45–58. Amherst, MA: GLSA.

Ni, Weijia. 1996. Sidestepping garden paths: Assessing the contributions of syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes* 11(3): 283–334.

Pesetsky, David. 1995. *Zero syntax*. Cambridge, MA: MIT Press.

Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge, MA: MIT Press.

Peterson, Philip. 1997. *Fact, Proposition, Event*. Dordrecht: Kluwer.

Portner, Paul. 1989. Processing indefinite noun phrases in quantified sentences. In *Five College Cognitive Science Paper*, vol. 89. Amherst, MA: Five College Cognitive Science Program.

Potsdam, Eric. 1998. A Syntax for Adverbs. In *Proceedings of the twenty-seventh Western Conference on Linguistics*, ed. Elly van Gelderen and Vida Samiian, 397–411. Fresno: Department of Linguistics, California State University, Fresno.

Potts, Christopher. 2002. The lexical semantics of parenthetical-*as* and appositive-*which*. *Syntax* 5(1): 55–88.

Potts, Christopher. 2005. *The logic of conventional implicature*. Oxford: Oxford University Press.

R Development Core Team. 2012. R: A language and environment for statistical computing. <http://www.R-project.org>.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1): 117–121.

Sag, Ivan. 1976. Deletion and logical form. Doctoral dissertation, Massachusetts Institute of Technology.

Sailor, Craig. 2014. The variables of VP ellipsis. Doctoral dissertation, UCLA.

Schlenker, Philippe. 2004. Conditionals as definite descriptions. *Research on Language and Computation* 2(3): 417–462.

Schlenker, Phillippe. 2005. Minimize Restrictors! Notes on definite descriptions, Condition C, and epithets. In *Proceedings of Sinn und Bedeutung 9*, ed. Emar Maier, Corien Bary, and Janneke Huitink, 385–416. Nijmegen: Nijmegen Centre of Semantics.

Sedivy, Julie C. 2002. Invoking discourse-based contrast sets and resolving syntactic ambiguities. *Journal of Memory and Language* 46(2): 341–370.

Sedivy, Julie C. 2003. Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research* 32(1): 3–32.

Shaer, Benjamin. 2000. Syntactic position and the readings of 'manner' adverbs. In *ZAS Papers in Linguistics 17*, ed. Catherine Fabricius-Hansen, Ewald Lang, and Claudia Maienborn, 265–286. Berlin: Zentrum Allgemeine Sprachwissenschaft.

Sharvit, Yael. 2014. On the universal principles of tense embedding: The lesson from 'before'. *Journal of Semantics* 31(2): 263–313.

Spivey-Knowlton, Michael, and Julie C Sedivy. 1995. Resolving attachment ambiguities with multiple constraints. *Cognition* 55(3): 227–267.

Stalnaker, Robert. 1974. Pragmatic presuppositions. In *Semantics and Philosophy*, ed. M. Munitz and P. Unger, 197–214. New York: New York University Press.

Stalnaker, Robert. 2002. Common ground. *Linguistics and Philosophy* 25(5): 701–721.

von Stechow, Arnim. 2002. Temporal prepositional phrases with quantifiers: Some additions to Pratt and Francez (2001). *Linguistics and Philosophy* 25(5–6): 755–800.

von Stechow, Arnim. 2009. Tenses in compositional semantics. In *The expression of time*, ed. Wolfgang Klein and P. Li, 129–166. Berlin: Mouton de Gruyter.

Stroik, Thomas. 1990. Adverbs as V-sisters. *Linguistic Inquiry* 21(1): 654–661.

Stroik, Thomas. 1996. *Minimalism, scope, and VP structure*. Thousand Oaks, CA: Sage Publications.

Tancredi, Christopher. 1992. Deletion, deaccenting and presupposition. Doctoral dissertation, Massachusetts Institute of Technology.

Vanden Wyngærd, G., and J.-W. Zwart. 1991. Reconstruction and Vehicle Change. *Linguistics in the Netherlands* 1: 151–160.

Williams, Edwin. 1974. Rule ordering in syntax. Doctoral dissertation, Massachusetts Institute of Technology.

APPENDIX A: MATERIALS USED IN EXPERIMENT 1A,1B

(49) John carefully read {every/the} book {and/that} Bill did (too).
     What did Bill do? (i) read a book.   (ii) carefully read a book.

Edna could easily finish {every/the} chore {and/that} Lisa could (too).
Larry energetically played {every/the} sport {and/that} Fred did (too).
John closely watched {every/the} movie {and/after/that} Bill did (too).
Lesley eagerly read {every/the} novel {and/that} Fred did (too).
Angela gradually gave up {every/the} bad habit {and/that} Foster did (too).
Nadia quickly found {every/the} solution {and/that} George did (too).
The Judge firmly rejected {every/the} claim {and/after/that} the lawyer did (too).
Eddy quietly sang {every/the} song {and/that} Fred did (too).

APPENDIX B: MATERIALS USED IN EXPERIMENT 2

The adverbs are shown in both positions, collapsing two distinct conditions in the experiment; the two connective signalling the type of CE are found in brackets.

(50) John (carefully) read every book (carefully) and then Bill did / John (carefully) read every book (carefully) that Bill did.
     What did Bill do? (i) read books (ii) (carefully) read books (carefully)

John (closely) watched every movie (closely) {and then/that} Bill did.

Angela (gradually) gave up every bad habit (gradually) {and then/that} Foster did.

The amateur singer (flawlessly) sang every song (flawlessly) {and then/that} the professional did.

Ed (violently) ripped out every page in the book (violently) {and then/that} Lesley did.

Fred (gently) petted the kitten (gently) {and then/that} Bill did.

Edna (quietly) entered the room (quietly) {and then/that} David did.

Tina (slowly) approached the gravestone (slowly) {and then/that} Joan did.

Edna will (easily) finish every chore (easily) {and/that} Lisa also will.

Larry (energetically) played every sport (easily) {and/that} Fred also did.

Lesley (eagerly) read every novel (eagerly) {and/that} Fred also did.

Heidi (calmly) spoke to the guests (calmly) {and/that} Nancy also did.

Nadia (quickly) found every solution (quickly) {and/that} George also did.

Larry (hesitantly) rejected every claim (hesitantly) {and/that} the lawyer also did.

Eddy (quietly) sang the song (quietly) {and/that} Fred also did.

Nina (loudly) played the song (loudly) {and/that} the band also did.

APPENDIX C: MATERIALS USED IN EXPERIMENT 3

(51)   Fred (calmly) spoke to the class (calmly) when Jill did./When Fred (calmly) spoke to the class (calmly), Jill did.
   What did Jill do? (i) speak to the class (ii) speak to the class calmly

Nina (loudly) played the national anthem (loudly) when the band did./ When Nina (loudly) played the national anthem (loudly), the band did.

Peter (quietly) walked into the house (quietly) after Lesley did./After Peter (quietly) walked into the house (quietly), Lesley did.

Edna (quietly) entered the room (quietly) after David did./ After Edna (quietly) entered the room (quietly), David did.

Tim (forcefully) rejected the argument (forcefully) after Maria did./ After Tim (forcefully) rejected the argument (forcefully), Maria did.

Fred (gently) petted the kitten (gently) when Bill did./ When Fred (gently) petted the kitten (gently), Bill did.

Nancy (gracefully) climbed the steps (gracefully) before Joan did./ Before Nancy (gracefully) climbed the steps (gracefully), Joan did.

Tina (slowly) approached the gravestone (slowly) when Fred did./ When Tina (slowly) approached the gravestone (slowly), Fred did.

The amateur singer (flawlessly) sang the song (flawlessly) after the professional did./ After the amateur singer (flawlessly) sang the song (flawlessly), the professional did.

Edna (easily) finished every chore (easily) after Lisa did./ After Edna (easily) finished every chore (easily), Lisa did.

Polly (secretly) spoke to the spy (secretly) after Fred did./ After Polly (secretly) spoke to the spy (secretly), Fred did.

John (carefully) studied every article (carefully) before Bill did./ Before John (carefully) studied every article (carefully), Bill did.

Larry (energetically) played the game (energetically) when Fred did./ When Larry (energetically) played the game (energetically), Fred did.

John (closely) watched every movie (closely) before Fred did./ Before John (closely) watched every movie (closely), Fred did.

John (carefully) read every book (carefully) when Bill did./ When John (carefully) read every book (carefully), Bill did.

Angela (gradually) gave up every bad habit (gradually) when Foster did./ When Angela (gradually) gave up every bad habit (gradually), Foster did.

Lesley (eagerly) read every novel (eagerly) when Fred did./ When Lesley (eagerly) read every novel (eagerly), Fred did.

Nadia (quickly) found the solution (quickly) when George did./ When Nadia (quickly) found the solution (quickly), George did.

Fred (hesitantly) rejected every claim (hesitantly) after Bill did./ When Fred (hesitantly) rejected every claim (hesitantly), Bill did.

Jeffrey (politely) spoke to the teacher (politely) when Fred did./ When Jeffrey (politely) spoke to the teacher (politely), Fred did.

APPENDIX D: MATERIALS USED IN EXPERIMENT 4

Rob secretly suspected {Mary, who Peter did as well / the woman that Peter did as well}.

What did Peter do? Secretly suspect Mary/Suspect Mary

Kim clearly saw {Damian, who Joyce did as well / the doctor that Joyce did as well}.

Joseph accidentally met {Mr. and Mrs. Wong, who Sawyer did as well / the couple that Sawyer did as well}.

Justine gently hugged {Kelly, who Max also did / the baby that Max also did}.

Nathaniel happily greeted {Tammy, who Juan also did / the guest that Juan also did}.

Cara eagerly listened to {Shel Silverstein, who Cory also did / the poet that Cory also did}.

Mark kindly taught {Matt, who Tasha also did / the boy that Tasha also did}.

Sam eventually paid {Mr. Baskin, who Craig also did / the contractor that Craig also did}.

Ellen heroically fought {Robin, who Henry also did / the intruder that Henry also did}.

Ethan skillfully bandaged {Ryan, who Jack also did / the patient that Jack also did}.

Jeff formally addressed {Dr. McKenzie, who Heather also did / the speaker that Heather also did}.

Aaron awkwardly danced with {Lisa, who Mike also did / the girl that Mike also did}.