

Information Dependency in Quantificational Subordination

Linton Wang, Eric McCready, and Nicholas Asher
University of Texas at Austin

Forthcoming in *Where Semantics Meets Pragmatics*, Elsevier Press

Abstract. The purpose of this paper is to (a) show that the received view of the problem of quantificational subordination (QS) is incorrect, and that, consequently, existing solutions do not succeed in explaining the facts, and (b) provide a new account of QS. On the received view of QS within dynamic semantic frameworks, determiners treated as universal quantifiers (henceforth *universal determiners*) such as *all*, *every*, and *each* behave as barriers to inter-sentential anaphora yet allow anaphoric accessibility in a number of situations. We argue that universal determiners are not intrinsic anaphora barriers and that anaphoric accessibility under them is enabled factors including lexicon information and discourse effects of universal determiners. In support of this viewpoint, we first provide a data survey on the phenomena of QS and its interactions with plurals, rhetorical relations, and adverbial quantification. The results of the survey show that judgments of (naive) native English speakers on the QS examples are quite different from what is claimed in the literature. We argue that the various solutions in the literature, which in general accept that universal determiners are intrinsic anaphora barriers, fail to account for the facts from the survey data. We then describe the approach we adopt, which denies that universal determiners are anaphora barriers and reconstructs their semantics so that information in their scope can be released for anaphora. The constraints on QS noted in the literature we model in Segmented Discourse Representation Theory (SDRT) as conditions on the discourse relations which can hold between subordinated constituents. We show that this approach accounts for the QS data.

keywords. Anaphora, Dynamic Semantics, Pronouns, Quantificational Subordination, Rhetorical Relations, SDRT, Telescoping, Universal Determiners

1 Introduction

The term *quantificational subordination* (QS) refers to instances of anaphoric dependence of pronouns on antecedents introduced under the scope of quantificational anaphora barriers (QAB). In formal semantic theories, and in dynamic semantics in particular, such barriers include the determiners *all*, *every*, and *each*. The problem of QS is that “robust” dynamic semantic theories, as presented in e.g. Kamp (1981), Heim (1982), and Groenendijk & Stokhof (1991), make any antecedent introduced under the scope of a QAB inaccessible for anaphoric binding; nonetheless, sometimes antecedents introduced in these positions are in fact anaphorically accessible. For example, consider the discourses (1), (2)¹ and (3). In the present section, we report judgments found in the literature without considering the results of our survey. The survey results with respect to these examples are presented in section 2.

- (1) a. Every hunter that saw *a deer*₁ shot *it*₁.
b. * *It*₁ was a female.
- (2) a. Every hunter that saw *a deer*₁ shot *it*₁.
b. *It*₁ died immediately.

¹(1) and (2) are simple variations of examples in Roberts (1996).

- (3) a. Every chess set comes with *a spare pawn*₁. (B. Partee, in Roberts 1987)
 b. *It*₁ is taped to the top of the box.

In these discourses, the universal determiner *every* is generally assumed to block anaphoric accessibility to the indefinite introduced within its scope. For instance, in (1), the pronoun *it* in (1b) cannot depend on the indefinite *a deer*, for it is introduced under the scope of the determiner *every* in (1a). This fact is predicted by standard dynamic semantics. However, the universal determiner *every* does not prevent the singular pronoun in (2b) and (3b) from accessing the indefinite under its scope in (2b) and (3b). This problem has been discussed by many authors (e.g. Sells 1985, Roberts 1987, 1989, 1996 Poesio & Zucchi 1992, Gawron 1996).²

1.1 Varieties of QS

It turns out that a number of phenomena also standardly considered within dynamic semantic theories interact in interesting ways with QS. In this section, we will consider the interactions of QS with plurals, rhetorical relations, and adverbial quantification. We will try to show that QS is not an isolated phenomenon, but shares many properties with other anaphoric phenomena, and that there are a number of distinct mechanisms that can facilitate QS.

QS and Plural Anaphora

Standard dynamic semantic theories have problems explaining the interaction of pronominal number with anaphoric accessibility. Standard dynamic accounts predict the unacceptability of (4a,b) and the unacceptability of (5a,b) as well.

- (4) a. *Every student*₁ went to school.
 b. **He*₁ brought lunch boxes.
- (5) a. *Every student*₁ went to school.
 b. *They*₁ brought lunch boxes.

The felicity of the discourse (5a,b) is generally considered to be data that a theory of plurals in dynamic semantics has to explain. However, just as in the problem of quantificational subordination, the assumption that the determiner *every* is a QAB plays a prominent role.³ Since the problem of plurals and the problem of QS are related to each other in dynamic semantics by virtue of this shared assumption, it seems desirable to solve both of them through the use of a single mechanism, if possible. To make this point even clearer, it is easy to find examples of discourses in which both problems appear simultaneously. An example is (6a,b).

²For many scholars, the problem of QS is closely related to the problem of modal subordination (MS), shown in (i) and (ii).

- (i) *A wolf*₁ might come in. *It*₁ would eat you first. (F. Landman, in Roberts 1987)
 (ii) *A wolf*₁ might come in. **It*₁ is hungry.

The difficulty here is how to explain the felicity of (i) and the infelicity of (ii). In the present paper, modal subordination is not our main focus; however, we believe that the approach described admits a straightforward extension to it. We will make some programmatic comments to this effect in the final section.

³See Kamp & Reyle (1993), van den Berg (1996), Krifka (1996), and Asher & Wang (2003) for different solutions to the problem.

- (6) a. *Every man*₁ loves *a woman*₂.
 b. *They*₁ send *them*₂/*her*₂ flowers.

In (6b), the pronoun *they* refers to the men introduced in (6a), an instance of the problem of plurals; in addition, the pronoun *them* in (6b) refers, on its most salient reading, to the women introduced under the scope of *every*, an instance of (plural) quantificational subordination. The challenge here is to keep dynamic semantics' good predictions for plural anaphora while rectifying its deficiencies for singular pronouns with antecedents within QAB.⁴

Plural anaphora also triggers another problem, exemplified by (7a,b),

- (7) a. *Every farmer*₁ who owns *a donkey*₂ beats *it*₂.
 b. *They*₁ treat *it*₂/*them*₂ cruelly.

Besides being a straightforward case of QS, (7b), like (6b), seems to be unselective with respect to pronominal number in the second argument. Approaches to donkey anaphora which take the pronoun *it* in (7a) to be number neutral, such as e.g. Neale (1990), argue that (at least) this type of pronoun lacks semantic number and requires only syntactic agreement. However, the unselective nature of the pronouns in (7b) and (6b) shows that this cannot be the whole story; given that both singular and plural pronouns are allowed, the effect of syntactic agreement is unclear here.⁵ Further, many people prefer the plural pronoun *them* in (7b) and (6b).

QS and Telescoping

As predicted by traditional dynamic semantics, the universal determiner *every* in (8a), as a QAB, makes the use of the singular pronoun *he* in (8b) unacceptable.

- (8) a. *Each student*₁ in the syntax class was accused of cheating on the exam.
 b. **He*₁ had a Ph.D. in astrophysics. (Roberts 1987) (Fodor and Sag check)

It often is noted in the literature that universally quantified NPs can also serve as antecedents for singular pronouns like *he* in (9b) and (10b). This phenomenon is known as *telescoping* (Roberts 1987 and Poesio & Zucchi 1992).

- (9) a. *Each student*₁ in the syntax class was accused of cheating on the exam.
 b. *He*₁ was reprimanded by the dean. (Fodor & Sag 1982)
- (10) a. *Each degree*₁ candidate walked to the stage.
 b. *He*₁ took his diploma from the dean and returned to his seat. (Sells 1985)

To explain telescoping, Roberts (1987) suggests that the acceptability of singular anaphora in (9a,b) and (10a,b) is due to the successful construction of rhetorical relations between (9a,b) and (10a,b). However, Roberts does not provide a mechanism to account for how rhetorical relations affect the acceptability of telescoping examples.

⁴There are also cases similar to plural quantificational subordination in modal subordination contexts. Discourses like "Some *mosquitos*₁ might come in the house. *They*₁ would bite you" or "A *mosquito*₁ might come in every room in this house. *They*₁ would bite someone in the room" seem to fall directly into this category. Just as in plural quantificational subordination, plural modal subordination turns out to cause trouble for previous approaches to modal subordination.

⁵See Kanazawa (2001) for more criticism of number neutral approaches.

QS and Adverbial Quantification

Although (11a,b) is predicted to be unacceptable by standard dynamic semantics, a number of similar examples in the literature such as (12a,b), which contains a quantificational adverb, are in fact felicitous. Standard dynamic semantics cannot explain the fact why adding a quantification adverb improves the acceptability of (11) substantially.

- (11) a. *A train₁* leaves every hour from Boston.
b. (?)*It₁* stops in New Haven.
- (12) a. *A train₁* leaves every hour from Boston.
b. *It₁* *always (sometimes)* stops in New Haven. (Sells 1985)

It has been suggested by e.g. Karttunen (1976) that the adverb in (12b) makes the singular pronoun *it* accessible to the antecedent under the QAB *every*, a notion implemented in the formal semantics of Gawron (1996). We will evaluate this claim after describing the results of our survey on examples like these in section 2.

1.2 Outline of the Paper

It is well known that the acceptability of QS examples is usually controversial. In an attempt to settle the issue of the correct data set to use, we conducted a data survey on a set of representative examples. We summarize the survey results in section 2. The results are surprising: judgments from linguistically unsophisticated native English speakers on QS examples differ from what is claimed in the literature. It turns out that many of the examples found in the literature are marginal for naive speakers. We used the examples judged acceptable by most speakers as a relatively uncontroversial set of data, which we then attempted to explain.

For a formal account, we must choose between two basic approaches. The first accepts that some logical operators, including universal determiners, are intrinsically barriers to anaphora. According to this view, a proper solution is to add extra machinery (e.g. antecedent accommodation in Roberts 1987, 1989, 1996) to extract information from under the operators in cases in which anaphora is possible. The second denies that these operators are in fact anaphora barriers, and reconstructs their semantics so that information in their scope can be released, as in Groenendijk & Stokhof 1989 and Kibble 1998. In section 3, we argue that the first approach in general over-generates in some cases, and, moreover, does not provide a powerful enough tool to deal with the problems of plural and plural quantificational subordination. In section 4, we further argue that quantificational subordination is not a completely unified phenomenon. In section 5, we adopt the second approach outlined above and in section 6 propose a formalism for determiners that can account for the release of information necessary for anaphora without making use of any mechanisms that are not independently required. In order to make distinctions between different types of quantificational subordination, a mechanism for *information attachment* is required; in our theory, this mechanism is based on accounts of rhetorical relations in SDRT (Segmented Discourse Representation Theory; cf. Asher 1993 and Asher & Lascarides 2003). We then use the formalism of universal determiners and SDRT to explain the possibility QS phenomena. We look at some examples in section 7. In section 8, we then tie all the strands together to show how contrasts in both felicity, and marginality, fall out of our approach. Finally, we close by discussing some implications of this approach for subordination involving modals, a line we will pursue in future work. The material we used for our data survey is presented in the appendix.

2 Results of a Data Survey

We distributed two surveys to approximately 80 students in a basic linguistics course taught by the second author. Students were offered extra credit points for their participation. Sixty surveys were returned for the first survey, and fifty nine for the second. The surveys consisted of a number of discourses: participants were asked to judge the felicity of the discourses on a four-point scale, with possible answers ‘totally out’, ‘very weird but still possible’, ‘a little weird’, and ‘fine’. We interpreted the data obtained in this manner as follows. Discourses were judged as acceptable if 2/3 or more of the responses were for ‘fine’ or ‘a little weird.’ Discourses were judged unacceptable if 2/3 or more of responses were ‘very weird but still possible’ or ‘totally out’. Discourses which showed no pattern in the responses were judged marginal. The results of the survey are shown in the appendix along with the examples presented.

Summarizing, three main results were obtained from the survey.

Result A: Standard QS Examples

For standard QS examples like (2) and (3a,b), we found that plural anaphors were generally preferred, when the antecedent was a universally quantified NP. We also found that rhetorical relations affect the acceptability of QS discourses. For example, according to SDRT (Asher 1993 and Asher & Lascarides 2003), the rhetorical relation between (3a) and (3b) is *Elaboration* and the rhetorical relation between (2a) and (2b) is *Result*. According to our survey data, QS discourses involving *Elaboration* are generally better than QS with other relations such as *Result*.

Result B: Telescoping

We first checked the possibility of anaphoric dependence on DPs which contains universal determiners like telescoping example (13).

- (13) a. *Each student*₁ went to school. *He*₁ walked.
b. *Each student*₁ went to school. *They*₁ walked.

First, our survey revealed that in telescoping examples plural pronouns are preferred over singular pronouns. Second, although the literature suggests that a difference between the universal determiners *each* and *every* is expected, no significant difference was observed in our survey. This result agrees with the experiment results reported in Carminati *et. al.* (2002).

Since every telescoping example with plural pronoun was judged acceptable, we checked whether different discourse relations result in different acceptability of telescoping examples with singular pronouns. For example, based on SDRT’s theory of rhetorical relations relations, the sentences in (9a,b), repeated as (14a,b), are linked by *Result* while those in (10a,b), repeated as (15a,b), are linked by *Narration*.

- (14) a. *Each student*₁ in the syntax class was accused of cheating on the exam.
b. *He*₁ was reprimanded by the dean. (Fodor & Sag 1982)
- (15) a. *Each degree*₁ candidate walked to the stage.
b. *He*₁ took his diploma from the dean and returned to his seat. (Sells 1985)

The survey results show that, even though no discourse relation provides an automatically acceptable discourse, the relation *Narration* facilitates telescoping better than other relations like *Background*, *Result*, and *Commentary*. As before, *Elaboration* discourses are also better than the others.

Result C: QS and Adverbs

We tested the effect of quantificational adverbs on QS. The participants of our study did not judge discourses including adverbs as more felicitous than similar discourses without them, i.e. (12a,b) is not obviously better than (11a,b). We conclude that quantificational adverbs do not contribute a significant effect to QS phenomena.

Other results of our survey will become clear from the rest of the paper.

3 Problems for Extraction Approaches

The majority of approaches to quantificational and modal subordination use a variety of special mechanisms to override constraints on anaphoric accessibility, such as antecedent accommodation in Roberts (1987), (1989), (1996) and Poesio & Zucchi (1992), quantification domain introduction in Gawron (1996), presupposition accommodation in Geurts (1999), and contextual dependency in Frank (1997).⁶ These approaches, in addition to leaving the issue of the marginality of many QS examples unexplained, in general exhibit three problems (in addition to the power of the machinery they introduce): (a) they over-generate in certain cases, meaning that their predictions are too liberal, (b) they are not powerful enough to explain some very simple cases of anaphora, and (c) the information dependency issues related to rhetorical relations, which create much of the complexity of the QS problem, remain entirely untouched.

The explanation of quantificational (and modal) subordination in these approaches relies totally on the special mechanisms used to make otherwise inaccessible information available. We call these approaches *extraction approaches*. The felicity of discourse subordination between two informational constituents is tied to the availability of a mechanism which extracts information from the first constituent and interprets the second with respect to that information. Such a mechanism is assumed to operate in cases like (2) and (3), so that anaphora becomes possible despite assumptions about the anaphora blocking nature of the universal determiners; it is further assumed that this mechanism is fully general and applies equally to instances of quantificational, modal, and conditional subordination.

To see that these approaches are too liberal in their predictions, consider the example (2), repeated as (16), and (17).

(16) a. Every hunter that saw a *deer*₁ shot *it*₁.

b. *It*₁ died immediately.

(17) a. *Every hunter*₁ that saw a *deer*₂ shot *it*₂.

b. * *He*₁ intended to kill *it*₂.

According to Roberts (1987), (1989), (1996) and Poesio & Zucchi (1992), (16a,b) receives an interpretation consistent with the conditional paraphrase *if a hunter saw a deer then he shot it and it died immediately* or simply *if a hunter shot a deer then it died immediately*. The mechanism inducing this interpretation, however, also allows the following reading of (17a,b): *if a hunter saw a deer then he intended to kill it*. According to the survey result, (16a,b) is marginal and (17a,b) is even worse than (16a,b). These accounts incorrectly predict the discourses (16a,b) and (17a,b) to be

⁶Except for the work of Gawron and Poesio & Zucchi, who focus on QS, the other authors discussed here direct their work primarily to modal subordination rather than quantificational subordination. Roberts and Geurts claim that their approaches can be generalized to deal with QS. Frank (1997) does not say anything explicit about quantificational subordination, but her approach can be extended in a similar way.

felicitous.⁷ The Roberts-style theories, then, over-generate in cases like these. This problem also applies to the approach outlined in Geurts (1999) and to possible extensions of Frank (1997), since these theories, while developed differently, produce essentially similar logical representations. In order to handle this difficulty, these approaches must provide clear and detailed constraints on the mechanism of information extraction; however, despite many attempts, definitively spelling out such constraints has proved to be a non-trivial task and still remains to be done.

For Gawron (1996), (16a,b), (17a,b), and even (3a,b) are predicted to be unacceptable since no quantificational element is present in (16b) and (17b). But he predicts that (18a,b) should be acceptable, and our survey shows the opposite.

- (18) a. *Every farmer*₁ who owns *a donkey*₂ beats *it*₂.
 b. *?He*₁ always treats *it*₂ badly.

Gawron's theory not only does not allow for good QS examples like (3a,b) but also wrongly predicts (given our survey) that quantificational adverbs can substantially facilitate QS.

Next, to see that the extra machinery approaches are not powerful enough to explain some very simple cases, consider the following example from our survey.

- (19) a. Every hunter who saw *a deer*₁ shot *it*₁.
 b. *They*₁ died immediately.

According to our survey, example (19a,b) is acceptable. But Roberts-style approaches cannot handle examples involving plurals, even such simple examples as this one.

Consider now more complex examples like the following.

- (20) a. *Every man*₁ loves *a woman*₂.
 b. *They*₁ send *them*₂/*her*₂ flowers.

In order to deal with QS in examples (20a,b), the semantic machinery should be able to explain, at a minimum, the following facts: if (20a) is interpreted on the $\forall\exists$ reading relevant for QS, antecedents introduced under the QAB *every* are accessible to both singular and plural pronouns. It is not obvious to us how the extra machinery theories can handle this issue. One possibility is to assume accommodation of the antecedent or presupposition by use of *plural discourse referents* as established in Kamp & Reyle (1993), but this approach faces two problems. First, as Krifka (1996) has argued, the mechanism for plural discourse referents in Kamp & Reyle (1993) is comparatively *ad hoc* and is not powerful enough to deal with some simple examples of plural anaphora. Second, there will be difficulties with discourses like (21)

- (21) *Each student*₁ wrote *two papers*₂. *They*₁ sent *them*₂/**it*₂ to L&P.

Examples like (21) raise the question under what circumstances a plural discourse referent should be accommodated, as opposed to a singular discourse referent. This problem becomes

⁷Even though Roberts (1996) and Poesio & Zucchi (1992) provide some constraints on antecedent accommodation, their constraints do not make (17a,b) infelicitous. The reason, perhaps, lies in the nature of their constraints, which are tied relatively closely to the particular examples which they consider.

especially acute in examples like (21), in which plural pronouns and singular pronouns are not interchangeable.⁸

The issues of QS involve another level of complexity related to rhetorical relations and information degradation. The first issue involves the relation between rhetorical relations and felicity in the QS examples. It has already been noted in Roberts (1987) that an explanation of the felicity of telescoping examples like (10), repeated as (22), should be based on the successful construction of rhetorical relations between (22a) and (22b).⁹

- (22) a. *Each degree*₁ candidate walked to the stage.
b. *He*₁ took his diploma from the dean and returned to his seat. (Sells 1985)

Intuitively, the infelicity of (8a,b), repeated as (23a,b), also seems to be due to the failure to construct a proper rhetorical relation between (23a) and (23b)—that is, a relation which is capable of supporting the necessary information for QS.

- (23) a. *Each student*₁ in the syntax class was accused of cheating on the exam.
b. **He*₁ had a Ph.D. in astrophysics. (Roberts 1987)

We agree with Roberts' intuition about the importance of rhetorical relations in the explanation of QS examples. However, the final story should be more complex. First, not all rhetorical relations can facilitate QS. For example, while the sentences in (9a,b), repeated as (24a,b), when analyzed in a theory like SDRT, are certainly connected by the rhetorical relation *Result*, the felicity of this example is controversial; QS researchers such as Fodor & Sag (1982) and Roberts (1987) judge this example good, but the naive informants in our survey participants find it unacceptable.

- (24) a. *Each student*₁ in the syntax class was accused of cheating on the exam.
b. **He*₁ was reprimanded by the dean. (Fodor & Sag 1982)

Also, according to our survey, the discourse (3a,b), linked by the *Elaboration* relation, is much more acceptable than discourse (2a,b), linked by the *Result* relation. Accounting for this difference requires a well-developed theory of rhetorical relations, such as that found in SDRT.

Second, a naive theory of rhetorical relation cannot help us to explain the felicity of (25a,b).

- (25) a. *Each student*₁ in the syntax class was accused of cheating on the exam.
b. *They*₁ had Ph.D.s in astrophysics.

⁸A similar issue also arises in modal subordination examples involving plurals, such as the following examples. In order to deal with the following example (i),

- (i) *A mosquito*₁ might come in every room in this house. *They*₁/*It*₁ would bite someone in the room.
(ii) *Two mosquitoes*₁ might come in every room in this house. *They*₁/**It*₁ would bite someone in the room.

whatever extraction machinery is assumed may accommodate the antecedent or presupposition with either a plural or singular discourse referent; either choice is felicitous. However, the accommodation of a singular discourse referent for (ii) is impossible.

⁹According to our survey, none of the telescoping examples are really acceptable. Example (10) is one of the most acceptable examples, judged significantly much better than others.

Both (23a,b) and (25a,b) are linked by the *Background* relation. The only difference between (23b) and (25b) is on the plurality and singularity of pronouns. If the infelicity of (23a,b) is based on the failure of inferring a rhetorical relation between (23a) and (23b), how to explain the success of this operation in the case of (25)?

Another phenomena related to the complexity of the QS issue is the way that information degrades in discourse. According to our survey, while (26a) is acceptable, (26b) is less good as (26a), presumably because of the way the discourse is continued.

- (26) a. *Every student₁ wrote a paper₂. They₁ sent it₂ to a journal.*
b. *Every student₁ wrote a paper₂. They₁ sent it₂ to a journal. They₁ had worked very hard on ?it₂.*

As seen here, the availability of an antecedent to a pronoun in discourse can degrade; but it can also be rescued to some degree. The discourse (28a,b) is better than (27a,b) simply because of the presence of a floating quantifier *each* in (28b).

- (27) a. *Three students₁ each wrote a paper₂.*
b. *They₁ sent it₂ to a journal.*
(28) a. *Three students₁ each wrote a paper₂.*
b. *They₁ each sent it₂ to a journal.*

It seems to be the case that the floating quantifier *each* in (28b) maintains the availability of the anaphoric antecedent, improving the felicity of the discourse.

These issues about rhetorical relations and availability degradation as they apply to the QS examples have been completely ignored up to the present in the literature. We will provide an attempt to deal with the complication in section 8. It will turn out that the complexity issue also plays a role in another poorly understood issue, the marginality of QS examples, that we discuss in section 8.3.

4 QS De-Generalized

We now describe our reanalysis of the essential problem of QS. Proponents of the first sort of approach, that discussed in the previous section, make one crucial assumption that turns out to be at the root of many of the observed problems: that universal determiners are anaphora barriers and a special information extraction mechanism is required to account for good QS examples, although the information extraction mechanism may be activated or guided by different factors. In fact, as we show below, this assumption is not correct. Examining why this is so gives substantial insight into the nature of QS and is the first step on the road to a true solution.

Let us begin by considering a case of conditional subordination.

- (29) a. *If a person₁ goes to school, he₁ will learn a lot of new things.*
b. **He₁'s a pretty knowledgeable guy.*

We are interested in the reading of the conditional in (29a) on which the pronoun *he* does not refer to a particular individual, but rather is linked to the nonspecifically interpreted indefinite introduced in the antecedent. We use the term *local information* for information subsumed within

the scope of a semantic operator in this manner, to bring out the idea that such information is restricted to the operator domain. As shown by the infelicity of continuing the discourse in (29), the standard dynamic formulation of conditionals as inducing this local quality on their content seems to be correct.

Now, however, let us reconsider the quantificational subordination cases. In both DPL and DRT, the interpretation of universal determiners exploits the semantics of conditionals; thus, information introduced under universal determiners turns out to be local information which cannot escape the universal domain.¹⁰ As a result, universal determiners serve as anaphora barriers in DRT and DPL. Because of this analysis, the problem of quantificational subordination seems to be the same as the problem of making nonlocal use of local (proximate) information; the accommodation technique used in approaches of the first type functions extract this local information from the domain of the QAB, allowing anaphoric links to it.¹¹

However, a wider range of natural language data shows that the claim that sentences like *every student went to school* provides only proximate information about students is false, despite their conditional-based analysis within dynamic semantic frameworks. Anaphoric dependence is available relatively unrestrictedly to objects introduced by universal determiners, as in the example (20a), repeated as (30a).

- (30) a. *Every man₁ loves a woman₂.*
 b. *They₁ send them₂/her₂ flowers.*

From this perspective, the reason that example (10) is significantly better than other telescoping examples, is not because of the manner in which the pronouns in (10b) access the ‘proximate’ information in (10a), but rather because of how the information introduced by (10a) can be updated by continuing the discourse with (10b).

¹⁰In dynamic semantics, the conditional implementation of universal determiners is generally called internally dynamic but externally static. Groenendijk & Stokhof (1989) provide an externally dynamic notion of universal determiners to account for the QS cases. Some problems may be worth mentioning about their approach. First, as they themselves note, it is not clear when a universal determiner should be treated as externally static or externally dynamic. They suggest that the issue may be related to the discourse relation between sentences, but provide no further details. Poesio & Zucchi (1992) and Roberts (1995) take this to be an essential problem with the solution in Groenendijk & Stokhof (1989) to the QS examples. Second, even though we may be able to determine in what situations externally static and externally dynamic interpretations should be used, we still need to deal with the semantic and discourse-level differences among different universal determiners discussed above. Third, moreover, universal determiners are always externally dynamic to plural anaphora. The internal-external dynamic distinction does not really apply for universal determiners. Fourth, universal determiners and conditionals have different discourse effects. While

- (i) Every students went to school. They brought lunch boxes.
 (ii) If x is a student, then he went to school. They brought lunch boxes.

(i) is fine but (ii) is unacceptable.

¹¹Quantificational subordination also occurs in examples which do not contain universal determiners. If we take the determiner *most* to be a QAB, the discourse (i)–(ii), similar to an example in Heim (1990), is a straightforward example of QS.

- (i) Most books contain a *table of content*₁.
 (ii) *It*₁/*They*₁ is/are at the end.

It is not clear in the literature whether the determiner *most* should be taken as a QAB. Our account of universal determiners in the dynamic semantic tradition as exploiting conditional semantics may not be able to be generalized to the cases containing determiners like *most*. This issue should not affect the main points in this paper.

On our account, the problem does not lie in the nature of the extraction mechanism that accounts for anaphora into quantificational contexts, but in the kind of semantic objects that are introduced by such contexts, and in the kinds of objects can then depend on them for their meaning. The infelicity of (31), for us, is not directly caused by the QAB *each*.

- (31) a. *Each degree*₁ candidate walked to the stage.
 b. **He*₁ was working in a high-tech company.

The infelicity is simply that the discourse does not properly support the anaphoric accessibility, because the discourse relation that holds between the two segments (*Background*) does not easily combine with the dependency relation necessary for use of a singular pronoun. This point will be elaborated in the next section.

Note that, however, the issue of making use proximate information seems to reappear in certain situations, for instance in the discourse (32).

- (32) a. If *a person*₁ goes to school, *he*₁ will learn a lot of new things.
 b. *He*₁ will become a pretty knowledgeable guy.¹²

Here, the information introduced by *a person* in (32a) truly seems to be proximate; it has no specific reading. As we can see, the only difference between (29b) and (32b) is presence of the modal *will*. Nonetheless, the antecedent information is available to the pronoun in (32b) despite its proxy status; but such is not the case for (29b). Some means of making use of proxy information seems to be needed in (32b), as well as some way of restricting access to it (for 29b).

5 Informal Elaboration of Our Approach

In the previous section we stated that, in our view, there are two crucial elements needed for a true account of quantificational subordination: an explanation of why anaphoric relations are possible in the good cases of subordination, and an explanation of why many discourses involving quantification subordination are *marginal*, or even infelicitous for many speakers. To answer the first question, we redefine the received formalism of universal determiners in order to induce the release of the correct sort of information from them.¹³ Doing so, of course, means that we massively

¹²An anonymous referee points out to us that the discourse in (i) and (ii) seems to be more acceptable than (32a,b).

- (i) If someone goes to school they will learn a lot of new things.
 (ii) They will become a pretty knowledgeable person.

The referee indicated that the plural pronoun *they* in both (i) and (ii) is used to avoid gender bias, in a context where it receives a singular interpretation. The referee also indicated the possibility that many of the survey examples in which the plural pronouns are preferred to singular pronouns may be explained by this singular usage of plural pronouns; the plural pronouns in the survey may not always be meant to be plural.

There might be some truth to this notion. The authors are unsure how far this idea can be taken. A further survey may be required to check this out. But the authors also notice that this explanation will not work in all cases. Consider the examples (iii) and (iv).

- (iii) Every Italian loves their mothers.
 (iv) *Each Italian loves their mothers.

It has been noted that (iii) is fine for the reason that the plural *their* is not really plural but used to avoid gender bias. However, this kind of gender bias avoiding usage does not work out in the example (iv) in which *each* is used.

¹³This can be viewed as a kind of generalization of the external dynamics of universal determiners in Groenendijk & Stokhof (1989).

over-generate “good” discourses; we appeal to the notion of discourse coherence in use in SDRT to solve this problem. The fine-grained distinctions between rhetorical relations made in SDRT provide us with a way to identify the rhetorical relations which facilitate QS and those which block QS. The apparent marginality of QS cases we explain by utilizing a dynamic formalism which allows expression of notions of discourse processing and optionality. We will delay our discussion of the issue of marginality until section 8.3, here presenting only the first part of the story.

5.1 Universal Determiners and Modes of Processing

Our formal semantics for universal determiners follows the formalism for universal determiners in Asher & Wang (2003). To model the truth conditional and discourse effects of universal determiners, they appealed to *counting predicates* and *modes of processing* (dynamic transitions, or jumps), within the framework of dynamic semantics. Counting predicates will play an important role when we define the meaning of determiners in section 6. Basically, a determiner is defined as a complex composed of a quantificational force and a counting predicate, which is used to control the proper witness output related to the meaning of the determiner. The modes of processing can be seen as programs for controlling the processing of witnesses. We model the information dependency relations relevant to anaphora and truth-conditional semantics by using these two special features.

Let us take the DP *all students* as an example. In our formalism, this DP is represented as $\exists x; all_x(student(x))$.¹⁴ In this formula, $\exists x$ is a quantificational force which is modelled by resetting in a way similar to that employed by Groenendijk & Stokhof (1991). The dynamic conjunction (relational composition) defined in Groenendijk & Stokhof (1991) is written as “;”, and $all_x(student(x))$ plays the role of controlling the proper witness output through use of the counting predicate *all*. The formula $\exists x; all_x(student(x))$ is roughly interpreted as the following: for information states S_F and S'_F which are sets of structured assignments, if $\langle S_F, S'_F \rangle \in [\exists x; all_x(student(x))]_M$, then the assignment functions in S'_F assign students to the variable x and the set of objects assigned by the assignment functions in S'_F consists of exactly the set of students, call them witnesses, defined in the model M . In other words, the quantificational force of the determiner introduces a discourse variable, and the counting predicate tells us how many objects should be related to this variable. Thus, the formula $all_x(\varphi)$, which makes use of the counting predicate *all*, returns the entire set of witnesses which satisfy $\varphi(x)$. For similar reasons, the formula $some_x(\varphi)$, where *some* is a counting predicate, returns some of the witnesses which satisfy $\varphi(x)$; and the formula $a_x(\varphi)$, which includes the predicate *a*, returns one of the witnesses which satisfy $\varphi(x)$.

The semantics we provide for universal determiners is quite different from traditional dynamic semantics. For example, in DPL (Groenendijk & Stokhof 1991), universal determiners, interpreted as universal quantifiers, are universal tests, i.e. if the test succeeds then the output is the same as input, and otherwise nothing is output. A special feature of a universal test is that the output may contain no information about the objects quantified over by the universal quantifier. On the contrary, in our semantics, the output of a translated DP is always related to the meaning of the DP by witnesses. There is no output that does not contain relevant information about the DP.

There is no difference between the witness sets introduced by the three universal determiners *all*, *every*, and *each*, for they each make use of the counting predicate *all*. The source of the differences among the three determiners is the manner in which they employ the witnesses they introduce, or, in other words, in how these witnesses are processed. We explicate the difference by introducing different modes of processing. In this paper, we employ three processing modes:

¹⁴Concerning the compositionality matter, it can be written as $\lambda Q.[\exists x; all_x(student(x)); Q(x)]$ in the clause level.

a simple distributive mode $Dis(x)$, a dependent mode $Dep(x, y)$, and a ‘strong’ dependent mode $S_Dep(x)$.¹⁵

For example, in (33), every determiner has a simple distributive reading. We symbolize the distributive reading through use of the distributive mode of processing $Dis(x)$. The logical forms of (33a,b, and c), when processed with the simple distributive mode, can be represented in our formalism by the formulas in (33’):

- (33) a. All students went to school.
 b. Every student went to school.
 c. Each student went to school.

(33’) $\exists x; all_x(student(x)); Dis(x); went_to_school(x)$.

In our formal semantics, we assume this simple distributive reading to be the default reading for universal determiners in our initial input states.

The dependent mode of processing occurs in examples like those in (34).

- (34) a. All students like a teacher.
 b. Every student likes a teacher.
 c. Each student likes a teacher.

The $\forall\exists$ reading is more obvious in (34b) and (34c) than in (34a). To model this reading, which we call the dependent reading, we allow use of the dependency mode $Dep(x, y)$ when processing the determiners *every* and *each*. We can think of the dependency mode $Dep(x, y)$ as a two place predicate which claims existence of a dependency relation between the objects assigned to the variables x and y . For example, we can represent the $\forall\exists$ reading of (34b) and (34c) as follows by exploiting existential closure.

(34’) $\exists x; \exists y; all_x(student(x)); Dep(x, y); a_y(teacher(y)); like(x, y)$.

These formulas admit the interpretation that for every different student there is a teacher he likes, which is the dependent reading.¹⁶ These mechanisms will provide a way to account for anaphoric accessibility in examples like (20a,b).¹⁷

The final mode of processing we use here, the strong dependent mode, models the difference in the discourse effects of the determiners *every/each* and *all*. An example of this difference is given in (35).

- (35) a. *All men₁ love a different woman₂. They₁/*He₁ send(s) her₂ flowers.*
 b. *Each/Every man₁ loves a different woman₂. They₁?He₁ send(s) her₂ flowers.*

¹⁵The semantics in this paper is a natural extension of the semantics of plurals in Asher & Wang (2003). In Asher & Wang (2003), modes of processing like collective mode and cumulative modes are also provided to account for collective readings and cumulative readings.

¹⁶This mechanism will provide us a way to account for scoping *in situ*. See Asher & Wang (2003) for more detail.

¹⁷In fact, there already is an antecedent for this dependency interpretation of different readings for (34b) and (34c) in the literature by quantifier elimination using Skolem functions. For example, the logical formula $\forall x\exists xFxy$ can be Skolemized by the formula $\forall xFxf(x)$ for some function f .

The use of plural pronouns in (35a,b) is acceptable for many speakers. However, according to our survey, use of a singular pronoun is more appropriate in (35b) than in (35a). In our formal semantics, the acceptability of the plural pronouns in (35a,b) can be modelled by use of the dependency mode, but to capture the difference in the acceptability of singular pronouns in (35a,b), we make use of the strong dependent mode, $S_Dep(x, y)$. This mode is able to apply in the case of *each* and *every*, but is not usually available with *all*.

The following QS examples exhibit a distinction which we take to be similar.

- (36) a. All_1 degree candidates walked to the stage. $*He_1$ took his diploma from the dean and returned to his seat.
- b. $Each/Every$ degree candidate₁ walked to the stage. He_1 took his diploma from the dean and returned to his seat.

According to our survey, (36b) is only a little better than marginal; but (36a) is even worse. Just as with the dependent mode, the strong dependent mode maintains a dependency relation; but it also creates a way to produce a witness which is accessible to singular pronouns. The strong dependent mode $S_Dep(x, y)$ is able to create a dependency relation not only for y on x , but also for x on x itself. In cases like (36b), the strong dependent mode produces a *reflexive* dependency on x , $S_Dep(x, x)$, since there is no other object introduced on which a dependency relation can be produced. Note that we do not allow the dependent mode to be reflexive, i.e. $Dep(x, x)$ is ill-formed. It should also be noted that the semantic difference between the dependent and strong dependent modes is not truth-conditional, but is simply a difference in discourse effects.¹⁸

Different modes of processing make the information released from universal determiners sensitive to the usage of singular or plural pronouns, and by doing so, produce different environments for information dependency in subsequent discourse. We will explicate the formal details of how this sensitivity is possible in section 6. For clarity, we here summarize in the table below the modes of processing or dependency relations each determiner makes available.

Determiner	Dependency Relations		
	<i>Dis</i>	<i>Dep</i>	<i>S_Dep</i>
all	✓	(?✓)	
every	✓	✓	✓
each	✓	✓	✓

Note that though the mode *Dep* is marginal for the determiner *all*, it is still available. The option chosen will depend on the preferences and tendencies of a given speaker. In this sense, the choice modes of processing in real sentence interpretation is pragmatic in nature.¹⁹

We are assuming that different modes of processing do create ambiguity rather than underspecification for universal determiners.²⁰ In section 6 we will also introduce the predicates *singular* and *plural* into our logic. They will be used to check the number agreement of singular and plural pronouns with their antecedents.

¹⁸In section 6, we will see that the strong dependent mode is definable by the dependent mode. But from a linguistic viewpoint, it is still important to separate them.

¹⁹Originally, we, like many others, believed that there must be some substantial difference between *every* and *each*. What we suspected is that the strong dependent mode is in general available to *each* but not usually available to *every*. However, according to the survey results, there is no such difference. This is an unexpected result.

²⁰See Asher & Wang (2003) for a more detailed argument.

5.1.1 Linguistic Evidence for Modes of Processing

In certain languages, there is direct linguistic evidence for the modes of processing we postulate. In this paper, we present some facts from Chinese, Japanese, and Hungarian which support the dependent and strong dependent mode of processing. Consider first the Chinese example (37).

- (37) a. *mei-yi-ge xue-sheng xi-huan yi-ge lao-shi.*
 every-one-classifier student like one-classifier teacher
 ‘Every student likes a teacher.’
- b. *mei-yi-ge xue-sheng ge xi-huan yi-ge lao-shi.*
 every-one-classifier student each like one-classifier teacher
 ‘Each student likes a teacher.’

Both the $\forall\exists$ reading and the $\exists\forall$ reading are possible for (37a) but only the $\forall\exists$ reading is possible for (37b). The word *ge* in example (37b) forces the dependent reading of the ‘indefinite’. We can simply interpret *ge* as a lexical expression which induces obligatorily the dependent mode of processing. Note that without *ge* the dependent mode of processing is still a possible choice in processing the sentence (37a). The role of *ge* in Chinese can be thought of as similar to the role of floating *each* in English.²¹

In English, the sentence “Every student likes a different teacher” has to be interpreted on the $\forall\exists$ reading (aka the dependent reading) in order to make sense of it. In Chinese, without the presence of *ge* to force the dependent mode, (38a) is marginal. With *ge*, however, as in (38b), it becomes fine.

- (38) a. *?mei-yi-ge xue-sheng xi-huan yi-ge bu-tong-di lao-shi.*
 every student like one-classifier different teacher
 ‘Every student likes a different teacher.’
- b. *mei-yi-ge xue-sheng ge xi-huan yi-ge bu-tong-di lao-shi.*
 every student each like one-classifier different teacher
 ‘Every student likes a different teacher.’

The Chinese *ge* also sometimes introduces the strong dependent mode. In example (39), *ge*, used with in combination with the *explicit* reflexive pronoun *zi*, represents the strong dependent mode structurally. We can represent this in our framework as $S_Dep(x, x)$.

- (39) a. *mei-yi-ge xue-sheng ge-zi shui-le.*
 every student each-self sleep-aspect
 ‘Each student slept.’

The Chinese data about *ge* also shows interesting phenomena relating to information degradation rescue like (28a,b). Consider (40).

- (40) a. *mei-yi-ge xue-sheng ge xi-huan yi-ge lao-shi.*
 every-one-classifier student each like one-classifier teacher
 ‘Each student likes a teacher.’

(only $\forall\exists$)

²¹In English, the sentence “Every student each likes a teacher” is ungrammatical. This might indicate that the dependent mode is already present in the determiner *every*, making *each* unnecessary.

- b. *ta men ge song ta/*ta-men yi-fen li-wu.*
 they each send her/*them one-classifier present
 ‘They each send her a present.’

If *ge* appears in (40b), then the discourse (40a,b) is acceptable if the singular pronoun *ta* is used but unacceptable if the plural pronoun *ta-men* is used. However, in discourse (41a,b), without using *ge* in (41b), neither the singular pronoun nor the plural pronoun is very good.

- (41) a. *mei-yi-ge xue-sheng ge xi-huan yi-ge lao-shi.*
 every-one-classifier student each like one-classifier teacher
 ‘Each student likes a teacher.’

(only $\forall\exists$)

- b. *ta men song ?ta/*ta-men yi-fen li-wu.*
 they send ?her/*them one-classifier present
 ‘They send her a present.’

These data suggest that a continuation with *ge* is required to halt the information degradation and then sustain the anaphoric relation.

In Hungarian, as argued in Farkas (1997), the dependent reading is marked by determiner reduplication, *D-reduplication* for short. For example, (42a) unambiguously has only the $\forall\exists$ reading but (42b) is ambiguous.

- (42) a. *Minden gyerek olvasott egy-egy könyvet.* (Farkas 1997)
 every child read a-a book-ACC
 ‘Every child read a book.’

- b. *Minden gyerek olvasott egy könyvet.* (Farkas 1997)
 every child read a book-ACC
 ‘Every child read a book.’

The second feature of Hungarian D-reduplication is that there must be something to fill in the dependency relation. In other words, the dependency relation must be relational. As a result, the following sentence is infelicitous, as it lacks anything capable of carrying out this function.

- (43) **Het-het gyerek szalad.* (Farkas 1997)
 seven-seven child runs
 ‘Seven children are running.’

The third feature is that *ki-ki* (‘who-who’) in Hungarian must get a dependent reading, and is incompatible with a simple distributive reading.

- (44) a. *Ki-ki leult egy-egy szekre.* (Farkas 1997)
 who-who sat-down a-a chair
 ‘Everybody sat down on a chair.’

- b. **Ki-ki leult.* (Farkas 1997)
 who-who sat-down
 ‘Everybody sat down.’

Evidence for modes of processing can also be found in Japanese.²² Japanese has a number of ways to express universal determiners: the one we discuss here is *kaku*, written with the same character as Chinese *ge* above. In example (45a), as with the Chinese example (38), when the sentence is produced without the use of the floating quantifier *sorezore*, the sentence is marginal.²³

- (45) a. *?kaku gakusei-ga chigau sensei-ga suki da.*
 each student-NOM different teacher-NOM like COP
 ‘Each student likes a different teacher.’
- b. *kaku gakusei-ga sorezore chigau sensei-ga suki da* ‘Each student likes
 each student-NOM that-that different teacher-NOM like COP
 a different teacher.’

The evidence from these three languages clearly shows the existence of the dependent mode of processing.²⁴ Note that while all of these languages explicitly mark the dependent and strong dependent mode, they also have expressions that exhibit an ambiguity between the dependent mode and the distributive mode, just as English determiners do. It is interesting to speculate on whether there are languages which require a choice between modes in all instances, that is, whether certain languages lack ambiguous determiners of this sort entirely. Our guess is that there are not, but this is a matter that requires further empirical work.²⁵

5.2 The Necessity of Integrating A Rich Notion of Discourse Structure

Rhetorical relations play two essential roles in our explanation of QS. First, it is necessary to check whether any rhetorical relation can be inferred between two discourse constituents at all; if not, QS will be infelicitous in the sense that the discourse would be infelicitous regardless of the facts about anaphora. More importantly for our account, it is possible to make a distinction between rhetorical relations that are capable of supporting the kind of information dependency needed for QS and relations that are not. Our survey indicates that the only relation with the former property is *Elaboration*; although *Narration* discourses are better than others, they are still rather marginal. Essentially, then, our account involves the following three parts: (i) check whether the determiner allows release of the proper witness set for the following pronoun (discussed in section 5.1), (ii) check whether any rhetorical relation can be inferred, and, crucially, (iii) check whether the inferred relation is capable of supporting QS. Failure at any of these steps causes the discourse to become infelicitous.²⁶

In SDRT terms, an incoherent discourse is understood as a discourse in which there is a discourse constituent or speech act (which here we can just consider as an instance of a proposition in a discourse) that does not seem to be connected in any meaningful way to the rest of the discourse (paraphrased from Asher & Lascarides 2003, p. 4). The following is another example of an incoherent discourse.

²²Thanks to Norihiro Ogata for help with the Japanese evidence.

²³An interesting complication here is that when the DP *kaku gakusei* is used with the topic-marking particle *-wa* rather than nominative *-ga*, the sentence also becomes fine. It may be that *-wa* has the ability to induce a dependent reading itself. We will not discuss this issue further in the present paper.

²⁴For more linguistic evidence readers may check German distributivity marker *je* (Link 1998) and the Korean distributivity marker *siik* (Cho 1988).

²⁵Another question for future research is whether the presence of overt markers of dependency relations varies systematically with other elements of the semantic system of the language; this question, too, we must put aside for the moment.

²⁶We present the analysis algorithmically for expository purposes. In fact, it is fully declarative.

- (46) a. John arrived in Edinburgh by train. (Asher & Lascarides 2003)
 b. ??Max's hair was black.

The infelicity of (48a,b) can be simply explained by discourse incoherence; it's not obvious how the two speech acts conveyed by these two sentences are meant to be connected, and so the discourse makes no sense as a whole.

We will not go into much detail here about how rhetorical relations can be computed and inferred, but instead refer our readers to Asher (1993) and Asher & Lascarides (2003) for details. The essence of the idea, however, is that a variety of information sources, including lexical semantics and domain knowledge, determines what relations, if any, can be computed between discourse-level propositions; given the right world knowledge and (relevant parts of) the semantic information contained in the speech acts at issue, speakers are able to infer how these speech acts are connected. In SDRT, the particular connections inferred to hold between speech acts also affect the semantic interpretation of the discourse; for instance, if the relation *narration*(π_1, π_2), is inferred, where π marks a speech act, then it will be a logical consequence of the discourse that the event described by π_1 takes place before the event described by π_2 .

In these terms, a discourse failing at condition (ii) will be one in which the two discourse constituents have no discernible connection at all. This situation will ordinarily not arise in cases of QS, because in the general case the presence of an anaphoric pronoun is enough to signal that a connection is intended with another constituent in the discourse. The following variation of the discourse above is as close as we can come:

- (47) a. *Every male student*₁ arrived in Edinburgh by train.
 b. ??! *His*₁ hair was black.

Nevertheless, we should note that an example like (47) immediately becomes acceptable when we substitute the plural pronoun in for the singular one in (47b). Thus, it appears that the problem is that some discourse connections cannot be established when a dependent mode of processing is required. In fact, this case may not be very different from the cases that clause (iii) is designed to deal with, a clause to which we now turn.

Clause (iii) accounts for examples like the following, the infelicity of which we stated in section 4 to be related to rhetorical relations.

- (48) a. *Each degree*₁ candidate walked to the stage.
 b. **He*₁ was working in a high-tech company.

On our account, the universal determiner *each* is not a QAB, and does license a witness for the singular pronoun in (48b) through the use of the strong dependent mode. We propose that the reason for the infelicity of (48) should not be directly related to the failure of the anaphoric link, but should be accounted for by the inability of the discourse relation between the two constituents, here *Background*, to license the continuation properly in a sense to be clarified below.

Another instance is the following. We stated above that, according to our survey, the example (10), repeated as (49), is more acceptable than other telescoping examples.

- (49) a. *Each degree*₁ candidate walked to the stage. (Sells 1985)
 b. ?*He*₁ took his diploma from the dean and returned to his seat.

We showed above that the universal determiner does not block anaphoric dependence of the singular pronoun in (49b) on the content of that DP. It is also possible to infer the *narration* rhetorical relation from (49a,b), given world knowledge about how graduation ceremonies proceed. With these facts, the logical form for the discourse (49) will be roughly like the following:

(49') $\pi_1 : each_x(degree_candidate(x); walked(x)); \pi_2 : took_and_returned(x); Narration(\pi_1, \pi_2)$

Note that when we say that universal determiners do not block anaphoric links to pronouns, we do not intend to say that the witnesses introduced under universal determiners are always accessible to pronouns. Whether the witnesses are accessible depends not only on whether proper witnesses are released but also on (a) whether a meaningful rhetorical relation can be deduced and (b) whether the deduced rhetorical relation can facilitate QS. Later in the paper (in section 7) we will see how this works out in more detail. And a more detailed elaboration of the construction of rhetorical relations is required to account for the complexity of QS mentioned in section 3. We will explain more about the issue in section 8.

6 Formal Semantics

The semantics we provide here is based on augmented dynamic predicate logic (DPL⁺) proposed in Asher & Wang (2003), which is an extension of DPL.²⁷ In DPL⁺, the language of DPL is expanded to include information about different sorts of counting predicates, which provide witnesses for certain situations, and modes of processing, which function to process witnesses in particular ways.

Definition 1 Language of DPL⁺

1. Logical Symbols

The usual DPL logical symbols (including variables, connectives ($\neg, ;$) and quantifier \exists).

2. Non-Logical Symbols

(a) The Usual Predicate Symbols

(b) Transition Predicates: j^{Dep} and $j^{S.Dep}$ (of arity 2), and j^{Dis} (of arity 1)

(c) Counting Predicates: *all, some, a, one, two, three, ..., many, most, ...*²⁸

Definition 2 Syntax of DPL⁺

1. $P(t_1, \dots, t_n)$ is an atomic formula, in which t_i is a term and P an n-ary predicate.

2. the usual DPL clauses for formulas involving $;$, \neg , and \exists .

3. For any counting predicates D , such as *all, n, the, a, and some*, $D_{t_i}(\varphi)$ is a formula if φ is a formula and t_i is an variable.

²⁷DPL⁺ has considerably more expressive power than DPL or first order logic, although it does not have the expressive power of full second order logic since we don't quantify directly over arbitrary sets of objects in the domain.

²⁸This set of counting predicates is richer than we strictly speaking need for this paper.

The syntax of DPL^+ is similar to DPL , excluding formulas which represent information about modes of processing and the counting predicate formulas.

A model for the language of DPL^+ is defined by a pair, $M = \langle D, I \rangle$, where D is a non-empty set of objects and I is an interpretation function. For a constant c_j , $I(c_j) \in D$. For a n -place predicate P^n , $I(P^n) \subseteq D^n$. For an assignment function g , $g : (V \cup C) \rightarrow D$, where V is the set of variables, C is the set of constants, and $g(c_i) = I(c_i)$. An information state $S_F = \{ \langle g, f_g \rangle \mid g \in S \text{ and } S \subseteq \$ \}$, in which $\$$ is the set of assignment functions, and $f_g : V \rightarrow (\$ \cup \emptyset(\$)) \times D$. The function f_g , defined such that if $f_g(t_i) = (A, B)$, then $g \in A$, is a function which records dependency and assignment information on assignments. We will make use of this function in defining the semantics of transition operations. The projection function $\mu^{1,2}$ yields either the first or the second element of the pairs that make up $f_g(t_i)$. $\mu^1(f_g(t_i))$ partitions the set of assignments according to how many objects are assigned to x by assignments in a given information state, while $\mu^2(f_g(t_i))$ is the actual assignment to x by the pair $\langle g, f_g \rangle$ in the information state.

The collection function δ collects the values assigned to a variable t_i by an element in the partition on assignments given by $\mu^1(f_{g_j}(t_i))$ and is defined as follows: $\delta(\mu^1(f_{g_j}(t_i))) = \{g_h(t_i) \mid g_h \in \mu^1(f_{g_j}(t_i))\}$. The interpretation of formulas $[\cdot]_M$ is a function that $[\varphi]_M \subseteq \$_F \times \$_F$, where $\$_F$ is the set of information states.²⁹ For any information state S_F , $S = \{g \mid \langle g, f_g \rangle \in S_F\}$ and $F = \{f_g \mid \langle g, f_g \rangle \in S_F\}$. The initial input information state S_F must satisfy the following two constraints: (i) for any g_j and t_i , $\mu^1(f_{g_j}(t_i)) = S$, and (ii) for any g_j and t_i , $\mu^2(f_{g_j}(t_i)) = g_j(t_i)$. These two constraints cause the distributive mode to be the default processing mode. In definition 3, we give the semantics for the transition or jump formulas.

Below we detail the semantics of the atomic formulas, which for us include not only standard DPL formulas but also those transition formulas (modes of processing) that have an effect on how assignments are structured and what variables in fact get assigned in information states.

Definition 3 Semantics for Atomic Formulas of DPL^+

1. $S_F[P(t_1, t_2, \dots, t_n)]_M S'_F := S_F = S'_F$, and for any $g_j \in S'$, $\langle \mu^2(f'_{g_j}(t_1)), \dots, \mu^2(f'_{g_j}(t_n)) \rangle \in P^M$.
2. $S_F[j^{Dep}(t_i, t_h)]_M S'_{F'} :=$
 - (a) $S = S'$,
 - (b) for any $g_j \in S'$ and t_k , $\mu^2(f'_{g_j}(t_k)) = \mu^2(f_{g_j}(t_k))$, and
 - (c) for any $g_j \in S'$, $\mu^1(f'_{g_j}(t_h)) = \{g_m \mid g_m(t_i) = g_j(t_i) \text{ and } g_m \in \mu^1(f_{g_j}(t_h))\}$, and for any t_k such that $t_k \neq t_h$, $\mu^1(f'_{g_j}(t_k)) = \mu^1(f_{g_j}(t_k))$.
3. $S_F[j^{S-Dep}(t_i, t_j)]_M S'_F :=$
 - (a) $S = S'$,
 - (b) for any $g_j \in S'$ and t_k , $\mu^2(f'_{g_j}(t_k)) = \mu^2(f_{g_j}(t_k))$, and
 - (c) for any $g_j \in S'$ and $t_k = t_i$ or t_j , $\mu^1(f'_{g_j}(t_k)) = \{g_h \mid g_h(t_i) = g_j(t_i) \text{ and } g_h \in \mu^1(f_{g_j}(t_k))\}$, and for $t_k \neq t_i$ or t_j , $\mu^1(f'_{g_j}(t_k)) = \mu^1(f_{g_j}(t_k))$.

²⁹For a specific formula, its interpretation a relation rather than a function since the existential quantifier is defined as non-deterministic rather than deterministic.

4. $S_F[j^{Dis}(t_i)]_M S'_F :=$
 - (a) $S = S'$,
 - (b) for any $g_j \in S'$, $\mu^2(f'_{g_j}(t_i)) = g_j(t_i)$, and for any $t_k \neq t_i$, $\mu^2(f'_{g_j}(t_k)) = \mu^2(f_{g_j}(t_k))$,
and
 - (c) for any $g_j \in S'$, $\mu^1(f'_{g_j}(t_i)) = S'$, and for any $t_k \neq t_i$, $\mu^1(f'_{g_j}(t_k)) = \mu^1(f_{g_j}(t_k))$.
5. $S_F[\exists t_i]_M S'_F := S =_{t_i} S'$, i.e. for any $g_j \in S$, there is an $g'_k \in S'$ that $g_j =_{t_i} g'_k$ and for any $g'_k \in S'$, there is $g_j \in S$ that $g_j =_{t_i} g'_k$. And for any t_j and $g \in S'$, $\mu^1(f_g(t_j)) = S'$ and $\mu^2(f_g(t_j)) = g(t_j)$.
6. $S_F[t_i]_M S'_F := S_F = S'_F$.

Note that transition formulas, though atomic, are *not* simple tests, but transform the structure of the output assignments.

We now tackle the recursive rules for complex formulas. We provide the details of the semantics of the counting formulas, which occur either in the presupposed or in the asserted part of a determiner's lexical semantics, as these are special to our approach and affect anaphora.

Definition 4 Semantics for Complex Formulas of DPL⁺

1. $S_F[\varphi; \psi]_M S'_F := \exists S_{F''}, S_F[\varphi]_M S''_F$ and $S''_F[\psi]_M S'_F$.
2. $S_F[\neg\varphi]_M S'_F := S = S_{F'}$, $\neg\exists S_{F''}, S_F[\varphi]_M S''_F$
3. $S_F[some_{t_i}(\varphi)]_M S'_F := S_F[\varphi]_M S'_F$, and for any $g_j \in S'$, $|\delta(\mu^1(f'_{g_j}(t_i)))| \geq 1$.
4. $S_F[at_{t_i}(\varphi)]_M S'_F := S_F[\varphi]_M S'_F$, and for any $g_j \in S'$, $|\delta(\mu^1(f'_{g_j}(t_i)))| = 1$.
5. $S_F[n_{t_i}(\varphi)]_M S'_F := S_F[\varphi]_M S'_F$, and for any $g_j \in S'$, $|\delta(\mu^1(f'_{g_j}(t_i)))| = n$.
6. $S_F[all_{t_i}(\varphi)]_M S'_F := S_F[\varphi]_M S'_F$, and for any $g_j \in S'$, $\delta(\mu^1(f'_{g_j}(t_i))) = \varphi^M(t_i)$, in which $\varphi^M(t_i) = \{g_i(t_i) | g_i \in S \text{ and } S_F \in \{S'_F | \exists S_F, S_F[\varphi]_M S'_F\}\}$.
7. $S_F[most_{t_i}(\varphi)]_M S'_F := S_F[\varphi]_M S'_F$, and for any $g_j \in S'$, $|\delta(\mu^1(f'_{g_j}(t_i)))| \geq |\varphi^M(t_i) - \delta(\mu^1(f'_{g_j}(t_i)))|$.
8. $S_F[more_than_n_{t_i}(\varphi)]_M S'_F := S_F[\varphi]_M S'_F$, and for any $g_j \in S'$, $|\delta(\mu^1(f'_{g_j}(t_i)))| > n$.
9. $S_F[at_least_n_{t_i}(\varphi)]_M S'_F := S_F[\varphi]_M S'_F$, and for any $g_j \in S'$, $|\delta(\mu^1(f'_{g_j}(t_i)))| \geq n$.
10. $S_F[the_{t_i}(\varphi)]_M S'_F := S_F[all_{t_i}(\varphi)]_M S'_F$, and for any $g_j \in S'$, $|\delta(\mu^1(f'_{g_j}(t_i)))| = 1$.

The definitions of satisfaction, truth, consequence, and validity follow the definitions in DPL in Groenendijk & Stokhof (1991).

Definition 5 Satisfaction. φ is satisfied with respect to M, S_F written by $M, S_F \models \varphi$ iff $\exists S'_F$ such that $S_F[\varphi]_M S'_F$.

Definition 6 Truth. φ is true with respect to M written by $M \models \varphi$ iff for any $S_F M, S_F \models \varphi$.

Definition 7 Consequence. $\varphi \models \psi := \forall M, S_F, S'_F$, if $S_F[\varphi]_M S'_F$ then $\exists S''_F$, that $S'_F[\psi]_M S''_F$.

Definition 8 Validity. $\models \varphi := \forall M, S_F, \exists S'_F, S_F[\varphi]_M S'_F$.

Definition 9 provides a naive semantic definition for number features of pronoun (*pronoun numbers* in shot). This definition should be compared to the definitions in Krifka (1996), van den Berg (1996), and Asher & Wang (2003).

Definition 9 $S_F[\textit{singular}(t_i)]_M S'_F := S_F[\textit{one}_{t_i}(t_i)]_M S'_F$.³⁰

For number agreement with plural pronouns, a possible naive semantic definition is given in definition 10, which was suggested by Krifka (1996), or in definition 11 suggested by Asher & Wang (2003).

Definition 10 $S_F[\textit{plural}(t_i)]_M S'_F := S_F[\textit{more_than_one}_{t_i}(t_i)]_M S'_F$.

Definition 11 $S_F[\textit{plural}(t_i)]_M S'_F := S_F[\textit{at_least_one}_{t_i}(t_i)]_M S'_F$.

Neither the combination of definition 9 and 10 nor the combination of definition 9 and 11 provides a correct account of pronoun number. The examples (50a) and (50b) are direct counterexamples to definition 9, 10, and 11.

- (50) a. *Mary*₁ wrote *one or two articles*₂. *She*₁ sent *them*₂/**it*₂ to L&P. (Krifka 1996)
 b. *Mary*₁ wrote *an article*₂. *She*₁ sent **them*₂ to L&P. (Krifka 1996)

Any model in which Mary wrote only one article will make the plural pronoun in (50a) unacceptable on definition 10 and make the singular pronoun in (50a) acceptable by definition 9. Both of these two consequences are unwelcome. Adopting definition 11 will also cause problems in making (50b) acceptable.

These problems made Krifka give up a semantic account of pronoun numbers, using instead a syntactic account of number agreement for pronoun numbers. However, we take a different direction. We think that the problem with the definitions 9, 10, and 11 is that they take semantic number agreement of pronouns to be a purely extensional phenomenon. We instead analyze the semantic number of pronouns as an (epistemic) intensional concept. To model the intensional notion of pronoun numbers, we introduce an intensional version of DPL⁺ (IDPL⁺).

Definition 12 IDPL⁺

1. *Syntax and Model:* IDPL⁺ has the same syntax as DPL⁺ and builds an intensional semantics on top of the extensional model theory for DPL⁺. A model M for IDPL⁺ is a first-order intensional model without the accessibility relations, i.e. $M = \langle W, D, I \rangle$ in which W is a set of possible worlds, D is a set of objects, and I is an interpretation function which assignment interpretation for predicates.
2. *Quasi-IDPL⁺ Interpretation:* $[\cdot]_M^q$ is an interpretation function that $[\varphi]_M^q \subseteq (W \times \mathcal{S}_F) \times (W \times \mathcal{S}_F)$ in which W is the set of possible worlds and \mathcal{S}_F is the set of DPL⁺ like information states. For any information state (w, S_F) and (w^1, S_F^1) ,

³⁰For any variable t_i , $S_F[t_i]_M S'_F$ for any information state S_F .

- $(w, S_F)[\varphi]_M^q(w^1, S_F^1) := w = w^1$ and $S_F[\varphi]_M^q S_F^1$ in DPL⁺ fashion, e.g.
 $(w, S_F)[P(t_1, t_2, \dots, t_n)]_M^q(w_1, S_F^1) := w = w_1, S_F = S_F^1$, and for any $g_j \in S^1$,
 $\langle \mu^2(f_{g_j}^1(t_1)), \dots, \mu^2(f_{g_j}^1(t_n)) \rangle \in I(w_1, P)$.

3. *IDPL⁺ Interpretation*: An information state σ in IDPL⁺ is a subset of $W \times \$_F$ in which W is the set of possible world and $\$_F$ is the set of DPL⁺ like information state. An interpretation of IDPL⁺ formula φ is a function $[\cdot]_M$ that $[\varphi]_M \subseteq \wp(W \times \$_F) \times \wp(W \times \$_F)$. For any information state σ and σ_1 , $\sigma[\varphi]_M \sigma_1 :=$

- (a) For any $\kappa \in \sigma$, $\exists \kappa_1 \in \sigma_1$ that $\kappa[\varphi]_M^q \kappa_1$, and
- (b) For any $\kappa_1 \in \sigma_1$, $\exists \kappa \in \sigma$ that $\kappa[\varphi]_M^q \kappa_1$.³¹

The definition of satisfaction, truth, consequence, and validity in IDPL⁺ is similar to DPL⁺. The semantics of pronoun numbers are defined in definition 13.

Definition 13 Intensional Semantics for Pronoun Numbers.

1. $\sigma[\textit{singular}(t_i)]_M \sigma_1 := \sigma = \sigma_1$ and $\forall \kappa \in \sigma$, $\kappa[\textit{one}_{t_i}(t_i)]_M^q \kappa$, i.e. $\sigma[\textit{one}_{t_i}(t_i)]_M \sigma_1$.
2. $\sigma[\textit{plural}(t_i)]_M \sigma_1 := \sigma = \sigma_1$ and $\exists \kappa \in \sigma$, $\kappa[\textit{more_than_one}_{t_i}(t_i)]_M^q \kappa$.³²

Using definitions 13, the problem from (50) can be easily solved. The following facts justify that our semantics are adequate to account for (50).

Fact 1 Facts to Account for (50)

1. For any M , if $\sigma[\exists x; a_x(\textit{article}(x)); \textit{wrote}(\textit{mary}, x)]_M \sigma_1$ then $\sigma_1[\textit{plural}_{t_i}(t_i)]_M = \emptyset$.
2. Assume $\textit{one_or_two}_x(\varphi)$ is defined as $\textit{one}_x(\varphi) \vee \textit{two}_x(\varphi)$. In order to make sense of the accessibility of plural pronouns in (50a), we introduce the idea of intensional presupposition for the interpretation of $\textit{one_or_two}_x(\varphi)$: the disjunction *or* introduces genuine (epistemic) alternative possibilities. According to this intensional presupposition, for any M , if $\sigma[\exists x; \textit{one_or_two}_x(\textit{article}(x)); \textit{wrote}(\textit{mary}, x)]_M \sigma_1$, then

- $\exists \kappa \in \sigma \& \exists \kappa_1 \in \sigma_1$ that $\kappa[\textit{one}_{t_i}(\varphi)]_M^q \kappa_1$ and $\exists \kappa \in \sigma \& \exists \kappa_1 \in \sigma_1$ that $\kappa[\textit{two}_{t_i}(\varphi)]_M^q \kappa_1$

so $\sigma_1[\textit{plural}_{t_i}(t_i)]_M \sigma_1$ and $\sigma_1[\textit{singular}_{t_i}(t_i)]_M = \emptyset$.

The failure of satisfying intensional presupposition does not necessary fail the semantic evaluation, but will fail predicates which requires the intensional presupposition. For example, in the world w that Mary wrote exactly one article, the sentence *Mary wrote exactly one or two articles* is true in w , but the plural pronoun in (50a) cannot be satisfied. Here we see that the intensional presupposition of pronominal number does not directly affect truth evaluation, but rather in a sense introduces a requirement that the number be appropriate in terms of the intensional model.

The following facts follow directly from definition 13.

³¹This definition applies to all formulas except mentioned specifically in the following.

³²The plural predicate can be understood as the epistemic model formula $\Diamond \textit{more_than_one}_{t_i}(t_i)$ defined in the sense of Veltman (1996).

Fact 2 Some Facts about Pronoun Numbers

1. $\forall \sigma$ and σ' , if $\sigma[one_x(x)]_M\sigma'$, then $\sigma'[singular(x)]_M\sigma'$.³³
2. $\forall \sigma$ and σ' , if $\sigma[one_x(x)]_M\sigma'$, then $\sigma'[plural(x)]_M = \emptyset$.
3. $\forall \sigma$ and σ' , if $\sigma[two_x(x)]_M\sigma'$, then $\sigma'[singular(x)]_M = \emptyset$.
4. $\forall \sigma$ and σ' , if $\sigma[two_x(x)]_M\sigma'$, then $\sigma'[plural(x)]_M\sigma'$.

In order to make universal determiner phrases like *every student* accessible to plural pronouns like *they*, again, we assume an *intensional presupposition* for universal determiners: universal determiners presuppose the *possibility* that there are at least 2 objects denoted by the restrictor in the intended information state.³⁴

We translate natural language determiners into DPL⁺ formulas by making use of the DPL⁺ counting predicates. For example, we interpret (in an approximate way) the following determiners in subject position by the combination of a quantifier meaning and a determiner meaning: *a student* as $\exists x; a_x(student(x))$; *some students* as $\exists x; some_x(student(x))$; *all students* as $\exists x; all_x(student(x))$; *every/each student* as $\exists x; all_x(student(x)), \exists x; all_x(student(x)); j^{Dep}(x, ?)$, or $\exists x; all_x(student(x)); j^{S-Dep}(x, ?)$. The symbol “?” is underspecified and will be instantiated to variables and filled in during the compositional process.

The following provides an example to introduce the formal semantics. Consider (51a), which has the logical form (51b). The semantic processing of this example can be represented as (51c).

- (51) a. *Every man₁ loves a woman₂. They₁ send her₂ flowers.*
- b. $\exists x; \exists y; all_x(man(x)); j^{Dep}(x, y); a_y(woman(y)); love(x, y); plural(x); singular(y); send(x, y)$
- c. $\{(w_1, S_F^1)\} \xrightarrow{\exists x} \{(w_2, S_F^2)\} \xrightarrow{\exists y} \{(w_3, S_F^3)\} \xrightarrow{all_x(man(x))} \{(w_4, S_F^4)\} \xrightarrow{j^{Dep}(x, y)} \{(w_5, S_F^5)\}$
 $\xrightarrow{a_y(woman(y))} \{(w_6, S_F^6)\} \xrightarrow{love(x, y)} \{(w_7, S_F^7)\} \xrightarrow{plural(x)} \{(w_8, S_F^8)\} \xrightarrow{singular(y)} \{(w_9, S_F^9)\}$
 $\xrightarrow{snd(x, y)} \{(w_{10}, S_F^{10})\}$

Assume that in the input information (w, S_F^1) , $S^1 = \{g_1^1, g_2^1, g_3^1\}$.³⁵ Since distributive mode is the default, so the following is true for the input information state. For any variable t_i ,

- $\mu^1(f_{g_1^1}(t_i)) = \mu^1(f_{g_2^1}(t_i)) = \mu^1(f_{g_3^1}(t_i)) = \{g_1^1, g_2^1, g_3^1\}$
- $\mu^2(f_{g_1^1}(t_i)) = g_1^1(t_i)$, $\mu^2(f_{g_2^1}(t_i)) = g_2^1(t_i)$, and $\mu^2(f_{g_3^1}(t_i)) = g_3^1(t_i)$

The semantic evaluation of (51c) is detailed as follows. For simplicity, the information in the information state not directly affected by interpretation of the following example is considered unchanged in the process of semantic interpretation.

1. $(w_1, S_F^1) \xrightarrow{\exists x} \{(w_2, S_F^2)\}$

³³This formula does not imply that singular pronouns have an “exactly one” presupposition.

³⁴In case that we have a singleton information state which assigns only one object for N of *every N*, the intensional presupposition fails. For more detail about the intensional semantics for pronoun numbers and intensional presuppositions for determiners, see Wang & Asher (2004).

³⁵In the following elaboration, we drop the set notation when doing so will not cause confusion.

- (a) $w_1 = w_2$,
- (b) $S^2 = \{g_1^2, g_2^2, g_3^2\}$, that $g_1^2 =_x g_1^1$, $g_2^2 =_x g_2^1$, and $g_3^2 =_x g_3^1$, and
- (c) The simple distributive mode still holds, i.e. (a) $\mu^1(f_{g_1^2}(x)) = \mu^1(f_{g_2^2}(x)) = \mu^1(f_{g_3^2}(x)) = \{g_1^2, g_2^2, g_3^2\}$, and (b) $\mu^2(f_{g_1^2}(x)) = g_1^2(x)$, $\mu^2(f_{g_2^2}(x)) = g_2^2(x)$, and $\mu^2(f_{g_3^2}(x)) = g_3^2(x)$.
2. $(w_2, S_F^2) \xrightarrow{\exists y} (w_3, S_F^3)$: This step is similar to step 1.
- (a) $w_2 = w_3$,
- (b) $S^3 = \{g_1^3, g_2^3, g_3^3\}$, that $g_1^3 =_y g_1^2$, $g_2^3 =_y g_2^2$, and $g_3^3 =_y g_3^2$, and
- (c) The simple distributive mode still holds.
3. $(w_3, S_F^3) \xrightarrow{all_x(man(x))} (w_4, S_F^4)$: For expository purposes, we assume that $I(w_3, man) = \{m_1, m_2\}$ and $g_1^3(x) = m_1$, $g_2^3(x) = m_2$, and $g_3^3(x) = m_1$. According to the IDPL⁺ semantics, we first check whether every object assigned to the variable x is a man. By assumption, it turns out yes. Then we check whether the set of object assigned to the variable x is the set of men relative to the world w_3 . By assumption, it turns out yes, i.e. $\delta(\mu^1(f_{g_1^3}(x))) = \delta(\mu^1(f_{g_2^3}(x))) = \delta(\mu^1(f_{g_3^3}(x))) = \{m_1, m_2\}$. So:
- (a) $w_3 = w_4$,
- (b) $S_4 = \{g_1^4, g_2^4, g_3^4\} = S_3$, and
- (c) The simple distributive mode still holds.
4. $(w_4, S_F^4) \xrightarrow{j^{Dep(x,y)}} (w_5, S_F^5)$
- (a) $w_4 = w_5$,
- (b) $S_5 = \{g_1^5, g_2^5, g_3^5\} = S_4$, and
- (c) The simple distributive on variable y has been switched to dependent mode, i.e. (a) $\mu^1(f_{g_1^5}(y)) = \mu^1(f_{g_3^5}(y)) = \{g_1^5, g_3^5\}$ and $\mu^1(f_{g_2^5}(y)) = \{g_2^5\}$, and (b) $\mu^2(f_{g_1^5}(y)) = g_1^5(y)$, $\mu^2(f_{g_2^5}(y)) = g_2^5(y)$, and $\mu^2(f_{g_3^5}(y)) = g_3^5(y)$.
5. $(w_5, S_F^5) \xrightarrow{a_y(woman(y))} (w_6, S_F^6)$: Again, for simplicity of exposition, we assume that $g_1^5(y) = a_1$, $g_2^5(y) = a_2$, $g_3^5(y) = a_1$, and a_1 and $a_2 \in I(w_5, woman)$. According to the IDPL⁺ semantics, we first check whether every object assigned to the variable y is a woman. By assumption, it turns out yes. Then we check whether the set of objects assigned to the variable y and corresponding to the first argument of the functional structure has cardinality 1, it turns out yes, i.e. $|\delta(\mu^1(f_{g_1^5}(y)))| = |\delta(\mu^1(f_{g_3^5}(y)))| = |\{a_1\}| = 1$ and $|\delta(\mu^1(f_{g_2^5}(y)))| = |\{a_2\}| = 1$. So:
- (a) $w_5 = w_6$,
- (b) $S_6 = \{g_1^6, g_2^6, g_3^6\} = S_5$, and
- (c) The dependency mode still holds.

6. $(w_6, S_F^6) \xrightarrow{\text{love}(x,y)} (w_7, S_F^7)$: Assume that $\langle m_1, a_1 \rangle$ and $\langle m_2, a_2 \rangle \in I(w_6, \text{love})$. According to the IDPL⁺ semantics, we check whether objects assigned to x and y are pairs of *love*. By assumption, it turns out yes. So:
- (a) $w_7 = w_6$,
 - (b) $S_7 = \{g_1^7, g_2^7, g_3^7\} = S_6$, and
 - (c) The dependency mode still holds.
7. $(w_7, S_F^7) \xrightarrow{\text{plural}(x)} (w_8, S_F^8)$: Since $(w_7, S_F^7)[\text{more_than_one}_x(x)](w_7, S_F^7)$, i.e. $|\delta(\mu^1(f_{g_1^7}(x)))| = |\delta(\mu^1(f_{g_2^7}(x)))| = |\delta(\mu^1(f_{g_3^7}(x)))| = |\{m_1, m_2\}| \geq 1$. So:
- (a) $w_8 = w_7$,
 - (b) $S_8 = \{g_1^8, g_2^8, g_3^8\} = S_7$, and
 - (c) The dependency mode still holds.
8. $(w_8, S_F^8) \xrightarrow{\text{singular}(y)} (w_9, S_F^9)$: Since $(w_8, S_F^8)[\text{one}_x(x)](w_8, S_F^8)$, i.e. $|\delta(\mu^1(f_{g_1^8}(y)))| = |\delta(\mu^1(f_{g_2^8}(y)))| = |\delta(\mu^1(f_{g_3^8}(y)))| = |\{a_2\}| = 1$. So:
- (a) $w_9 = w_8$,
 - (b) $S_9 = \{g_1^9, g_2^9, g_3^9\} = S_8$, and
 - (c) The dependency mode still holds.
9. $(w_9, S_F^9) \xrightarrow{\text{snd}(x,y)} (w_{10}, S_F^{10})$: Assume that $\langle m_1, a_1 \rangle$ and $\langle m_2, a_2 \rangle \in I(w_9, \text{send})$. According to IDPL⁺ semantics, we check whether objects assigned to x and y are pairs of *send*. By assumption, it turns out yes. So:
- (a) $w_{10} = w_9$,
 - (b) $S_{10} = \{g_1^{10}, g_2^{10}, g_3^{10}\} = S_9$, and
 - (c) The dependency mode still holds.

We hope that this detailed example has helped to clarify how the formalism works.

7 Information and Coherence in Quantificational Subordination: Some Cases Explained

Our analysis of quantificational subordination is based on three factors discussed in previous sections: (a) compatibility conditions between the output of various universal determiners and pronouns, (b) the availability of the inference of rhetorical relations between discourse segments, and (c) whether an inferred discourse relation can facilitate QS. The first part of the analysis corresponds to the formalism presented in section 6; the second part and the third was informally addressed in section 5.2. We now make the discussion there more concrete.

The first part of the analysis can be summarized as follows. The universal determiners *all*, *every*, and *each* are not intrinsically barriers to anaphora; however, the (dynamic) procedure by

which they are processed outputs objects that are intrinsically singular or plural, meaning that only pronouns of the correct type are able to pick them up as antecedents. For instance, in the examples (6a,b), repeated by (52a) and shown paired with their corresponding logical forms, *every* output plural objects from the restrictor position, which plural pronouns may pick up; however, the presence of the ‘jump’ operator j^{Dep} in (52b) makes the object information accessible to both singular and plural pronouns.

- (52) a. *Every man₁ loves a woman₂. They₁ send them₂/her₂ flowers.*
 b. $\exists x; \exists y; all_x(man(x)); j^{Dep}(x, y); a_y(woman(y)); love(x, y); plural(x); singular(y); send(x, y)$
 c. $\exists x; \exists y; all_x(man(x)); j^{Dep}(x, y); a_y(woman(y)); love(x, y); plural(x); j^{Dis}(y); plural(y); send(x, y)$

Note that in order to get the pronoun *them* in the second sentence of (52a), we have to put $j^{Dis}(y)$ in order to switch the dependency of y on x back to simple distributive information on y . According to our survey, the usage of a plural pronoun in the second sentence of (52b) is better than usage of a singular pronoun. Remember that we defined the simple distributive mode as default in input contexts. The survey participants’ preference shows that the tendency to switch back from the dependent mode to the simple distributive mode is natural for naive speakers. This phenomenon will play an important role in our understanding of information degradation, which we return to in section 8.1.

From the perspective of transitional information processing, then, the universal determiners *all*, *every* and *each* are ambiguous, and create different situations for anaphoric dependence on each reading.

Similar facts hold in the case of telescoping. In example (10), repeated as (53), even though using the strong dependent mode makes the witness output from (53a) available for the singular pronoun in (53b), the inferred discourse relation, *Narration*, does not support QS well, so the discourse turns out little better than marginal.

- (53) a. *Each degree₁ candidate walked to the stage.* (Sells 1985)
 $\exists x; all_x; candidate(x); j^{S-Dep}(x, x); walk(x)$.
 b. *?He₁ took his diploma from the dean and returned to his seat.*
 $singular(x); take(x); return(x)$.

However, in our survey, we found that if the singular pronoun in (53b) is changed to a plural pronoun, as in (54b), then the discourse is acceptable.

- (54) a. *Each degree₁ candidate walked to the stage.* (Sells 1985)
 $\exists x; all_x; candidate(x); j^{S-Dep}(x, x); walk(x)$.
 b. *They₁ took their diplomas from the dean and returned to their seats.*
 $j^{Dis}(x); plural(x); take(x); return(x)$.

To explain the difference between the acceptability of (53a,b) and (54a,b), we refer again to the default reversion to distributive mode. Since we assume that it is a natural tendency to revert to the simple distributive mode, it is obvious that (54a,b) should be more acceptable than (53a,b). Another reason for (54a,b) to be more acceptable than (53a,b) is that simple distributive and dependent

modes are also choices for processing (54a) and (53a) which support (release witnesses for) the plural pronoun in (54b) rather than the singular pronoun in (53b).

Similarly, the infelicitous instances in (55) fall out of the inability to infer a rhetorical relation which supports the strong dependency relation. As a result, the the right sort of witnesses to serve as antecedent to a singular pronoun will not be obtained.

- (55) a. ?? *Every dog*₁ came in. *It*₁ lay down under the sofa. (Poesio & Zucchi 1992)
b. *If *every cat*₁ purrs, *it*₁ is happy. (Poesio & Zucchi 1992)
c. *John likes *every dog*₁ and Sam feeds *it*₁. (Hornstein 1984)

The infelicity of example (23), repeated as (56), receives an explanation by the failure to compute a discourse relation of the right type between (56a) and (56b).

- (56) a. *Each student*₁ in the syntax class was accused of cheating on the exam.
b. **He*₁ had a Ph.D. in astrophysics.

On our account, the infelicity of (56) is not directly related to the problem of anaphoric accessibility. Rather, the rhetorical relation here is just not of the right type to support the needed processing mode. This constraint, we argue, accounts for what appears to be a failure of anaphoric accessibility. This last piece of the story is presented in section 8.2.

On our account, then, the infelicity of the impossible cases of telescoping follow directly from independently necessary constraints on the coherence of discourse, not from any *ad hoc* constraints that must be defined on the availability of special update mechanisms. This account is related to, but not identical with, the script-based account of Poesio & Zucchi (1992); unlike their theory, however, our analysis is not specific to quantificational subordination, but falls out of general pragmatic constraints on possible discourses. Our analysis also incorporates a distinction between the information released by various universal quantifiers, which is overlooked in their work. In addition, the explicit invocation of knowledge bases which we make allows us to give an explanation for the marginality of the QS construction in general.

8 Complexity in QS

In this section, we are going to deal with the issues we put aside in the previous sections: the issues of information degradation, rhetorical relations, and the marginality of QS. The facts here are all very subtle, and are completely untouched by previous theories. The ideas we present here represent a first attempt at providing an explanation for these issues.

8.1 QS and Information Degradation

In previous sections, we argued that QS examples can be easily explained in terms of information dependence between discourse constituents, using information content relevant to pronominal interpretation which is released from determiners, and the possibility of inferring discourse relations. An interesting phenomenon observed in Asher & Wang (2003) is that dependency relations between objects degrade rapidly in discourse. In example (57), while (57a) is fine, increasing the number of sentences quickly causes problems; the discourses in (57b and c) have already become degraded.

- (57) a. *Each student*₁ wrote a *paper*₂. *They*₁ sent *it*₂ to L&P.
 b. *Each student*₁ wrote a *paper*₂. *They*₁ sent *it*₂ to L&P. *They*₁ had worked very hard on *it*₂.
 c. *Each student*₁ wrote a *paper*₂. *They*₁ sent *it*₂ to L&P. *It*₂ was a good paper.

In fact, degradation also occurs in shorter discourses, although here it is dependent on the presence of particular quantificational elements.

- (58) a. *Three students*₁ each wrote a *paper*₂. *They*₁ sent *it*₂ to L&P.
 b. *Three students*₁ each wrote a *paper*₂. *They*₁ each sent *it*₂ to L&P.

For many speakers, (58b) is obviously better than (58a). Even though the required information dependency for the proper interpretation of pronouns in (58a and b) is already created by the floating *each*, as we previously showed, the repetition of this element nevertheless increases the acceptability of the discourse in (58b) by preventing the information released by the determiner from degrading.

In order to explain the degradation effects, we appeal to two ideas: (a) information degradation in QS is caused by the natural tendency of speakers to switch back to the simple distributive mode; (b) determiners and certain rhetorical relations can ‘rescue’ information from degradation by blocking the tendency to revert to the simple distributive mode. As stated above, we assume that the simple distributive mode is the default, and that the other modes tend to revert back to it over time unless supported. Our formal semantics set the *simple distributive* mode as the *default* in the input information state. Given this, the decay of information dependency can be explained as a *tendency* to switch back to the simple distributive mode from the dependent and strong dependent modes. In regard to issue (b), we mentioned in sections 3 and 5 that some rhetorical relations facilitate QS better than others. We will provide details in the next section. The use of determiners also prevents information degradation by making the dependent or strong dependent mode able to persist through discourse. In example (59), (59a) and (59c) are better than (59b).

- (59) a. *Each man*₁ loves a *woman*₂. *They*₁ send *them*₂ flowers.
 b. *Each man*₁ loves a *woman*₂. *They*₁ send *her*₂ flowers.
 c. *Each man*₁ loves a *woman*₂. *They*₁ each send *her*₂ flowers.

In (59a), the degradation of dependent information from the first sentence naturally leads to a preference for the plural pronoun *them* rather than the singular pronoun *her* in (59b). In (59c), the determiner *each* in the second sentence activates the dependent mode to prevent information degradation, i.e. we interpret the floating *each* as a kind of mode of processing, either dependent or strong dependent.

8.2 Rhetorical Relations and Information Dependency

We explain the infelicity of (56a,b) by appealing to the failure to construct the right sort of rhetorical relations between (56a) and (56b). Given this account, it may be puzzling that while (56a,b) is infelicitous, its plural variation (25a,b), repeated as (60a,b), is felicitous.

- (60) a. *Each student*₁ in the syntax class was accused of cheating on the exam.

- b. *They*₁ had a Ph.D. in astrophysics.

The answer to this objection lies in our analysis of modes of processing. As we can see from the release of the proper information for the plural pronoun in (60b), no dependent or strong dependent mode of processing is used. The felicity of (60a,b) shows that a rhetorical relation is constructed between (60a) and (60b). The rhetorical relation is in fact the *Background* relation. However, in order to release the proper information for singular pronouns in (56b), the strong dependent mode is required. We attribute the infelicity of (56a,b) to the nature of the discourse relation *Background*, which is not *strong* enough in some sense to facilitate the passing of the strong dependent mode to (56b). To explain how rhetorical relations differ in their licensing the persistence of the strong dependent mode, we appeal to the concept of a *hierarchy of information connectivity* in rhetorical relations.

The notion of strength of information connectivity in rhetorical relations is based on the strength of the connection between the information constituents that the relations connect. The stronger the connection, the higher the relation is placed in the hierarchy. We roughly distinguish information connectivity in rhetorical relations by the following three level hierarchy.

- Weak Connectivity: *Background, Explanation, Result, Commentary*
- Intermediate Connectivity: *Narration*.
- Strong Connectivity: *Elaboration*.

The notion of connectivity can be intuitively understood as the extent to which the second argument of the rhetorical relation in question can stand independently in the discourse. For instance, the second argument of *Background* is understood simply as presenting additional information which may be relevant to the first argument. However, the second argument of *Elaboration* has no independent ‘life’; without the first argument, it is irrelevant or meaningless, for it serves only to provide details of the event described by the first argument. Even though the hierarchy is rough and a more detailed theory of information connectivity still needs to be worked out, the hierarchy intuitively seems right.³⁶

The hierarchy outlined here relates to the QS cases in the following way. The *Background* relation provides only weak information connectivity. However, the strong dependent mode requires that the connection be strong, since that mode of processing is required to provide a suitable antecedent for the singular pronoun. We can also predict that a rhetorical relation which has stronger information connectivity has more power to prevent information degradation. This prediction is confirmed by examples in our survey, e.g. (3a,b) which is connected by *Elaboration* is much more acceptable than (2a,b) which is connected by *Result*.

8.3 Explaining Marginality

We stated above that an explanation of the marginality of QS constructions is important, if not necessary, for a full account of the phenomenon. In this section, we show how our account can provide at least a route toward an explanation, although we are not yet prepared to give a complete one.

³⁶Note that the difference between information connectivity among rhetorical relations does not correspond to the difference between subordinating and coordinating rhetorical relations.

The three components of our account—the modes of processing that are optionally bundled with universal determiners, the inference of discourse relations for discourse coherence, and the effect of information degradation—are all partly pragmatic in nature, in that the processing modes are optional, the inference of discourse coherence is dependent on various information sources beyond simple compositional semantics, and the degradation effect represents the tendency and preference in mode of processing. Because they are pragmatic, they can vary in availability, as follows. The possibility of use of a particular mode of processing is related to whether a particular language user has a preference for using that mode.³⁷ Similarly, inference of discourse relations in the strict context of QS and telescoping discourses is related to information connectivity provided by rhetorical relations and particular world knowledge that a given speaker may or may not have. The information degradation represents a tendency in the process of processing dependency information. However, pragmatic tendencies of this sort may not be equally shared by every language user. Consequently, these components can fail quite easily given the right context or the wrong speaker. From our perspective, the notorious variation in acceptability from speaker to speaker in examples like these is attributable to the nature of the conditions that allow anaphora.

Understanding the differences of universal determiners provides us a way to explain the marginality of QS examples. For example, use of a singular pronoun is not totally impossible in (35a). This may be explained by the idea that although some speakers may have a preference for processing the determiner *all* using the distributive mode and dependent mode, certain people are able to process *all* using the strong dependent mode. Conversely, for some people, the singular pronoun in (35b) is not totally acceptable. Given that the modes of processing are optional but admit preferences and dis-preferences, this fact can be explained by stating that the strong dependent mode is not preferred by those speakers for whom (35b) is marginal.

The following examples from Asher (2001) will make clearer our explanation of marginality due to the availability of processing modes.

- (61) a. *All students*₁ wrote *a paper*₂. **He*₁ submitted *it*₂ to L & P.
 b. *Every student*₁ wrote *a paper*₂. ?*He*₁ submitted *it*₂ to L & P.

For most speakers of English, (61b) is better than (61a). Nonetheless, some people still find (61b) to be unacceptable. According to our semantics for universal determiners, to use a singular pronoun felicitously in the second sentences of (61a and b) it is necessary to make use of the strong dependent mode of processing. However, the strong dependent mode is not available equally to every determiner: it is almost completely unavailable to *all*, but in general available to *every*. This availability of modes of processing explains the hierarchy in felicity of the examples in (61).

Another parameter in the felicity judgments has to do with how discourse relations interact with the modes of processing. For some speakers, a discourse relation like *narration* may suffice to support a strongly dependent processing mode, even though for many this appears not to be the case. This variation, we argue, is another cause of the marginality of telescoping constructions, though more research is needed to investigate the exact interaction between different discourse relations and modes of processing.

³⁷There are some other advantages to making use of modes of processing. For example, it provides a consistent and compositional way to make sense of the different readings of plurals. It also makes available a way to understand different quantifier readings without utilizing scope. See Asher & Wang (2003) for details.

9 Concluding Remarks

We summarize what we have done in this paper. We began by showing the result of our survey which shows different judgment on QS examples from what is claimed in the literature. And then, we discuss some previous theories of quantificational subordination that used special mechanisms to extract information from otherwise inaccessible semantic components, arguing that they were inadequate for a general solution of the problem. We showed that such approaches either over-generate or are unable to handle certain types of discourses that, in general, are perfectly felicitous, and argued that the problem of QS is generated from the conditional semantics used by dynamic theories to handle universal quantification. We solved this part of the problem by proposing a semantic formalism which releases information from universal determiners directly. We then moved to block the over-generation made possible by this formalism with the mechanisms for inferring discourse coherence utilized in SDRT. By using these two components, we showed, a unified theory of QS can be developed which is located within a broader landscape of discourse anaphora and discourse structure, rather than standing by itself as a special case. We also showed that our approach provides a means of modelling the marginality of many QS examples.

We close with a pointer to future research. We believe that the account presented here can be generalized to other kinds of discourse subordination. In quantificational subordination, licensing depends on compatibility between the object introduced by a universal quantifier and the pronoun which depends on it. Abstracting away from the quantificational case, subordination comes with a compatibility requirement between a dependent object and its antecedent. In the modal domain, this requirement corresponds to a need for the domains of modal operators to ‘fit’ one another. In the literature, it has been noted that certain types of modality work together, and others do not:

- (62) a. *A wolf*₁ might come in. *It*₁ would eat you first. (epistemic; subjunctive)
b. *John*₁ should buy *a car*₂. *He*₁ would drive *it*₂ (deontic; subjunctive)
- (63) a. *A wolf*₁ will come in. **It*₁ would eat you first. (indicative; subjunctive)
b. *John*₁ might buy *a bottle of wine*₂. **He*₁ should drink *it*₂. (epistemic; deontic)

In our terms, the felicity/infelicity of the above examples corresponds to the compatibility of the modal operators involved.

This observation itself is not new; scholars of modal subordination (Roberts 1996, Frank 1997, Geurts 1999) have all noted it in one form or another. However, we think that our theory provides a new perspective on the facts by claiming that this need for compatibility is not limited to the modal case, but rather holds for all sorts of subordinating contexts. We also take the perspective that, parallel to the quantificational cases which are our main focus here, (a) modal operators are not anaphoric barriers, and (b) anaphoric accessibility is also controlled by whether a discourse is coherent.

Acknowledgements

The authors would like to thank the participants in the conference (*Where Semantics Meets Pragmatics*, Michigan, U.S.A., 2003) and an anonymous referee for useful comments and suggestions. We would also like to thank Sheila Asher and Neil Sinhababu for checking our data.

A Appendix: Survey Materials

We use the following schema to indicate the judgements of our informants on the examples in the surveys. For a full explanation of the terms used, see section 2. We also provide the rhetorical relations that hold between the two elements in each discourse, where it is relevant for QS in our discussion.

- A: Acceptable.
- >A: Better than acceptable
- UA: Unacceptable.
- <UA: Worse than Unacceptable.
- M: Marginal.
- <M: Better than unacceptable but worse than marginal.
- >M: Better than marginal but worse than acceptable.

A.1 Survey 1 Material

- Every student went to school. He brought lunch boxes. (<UA)
 - Every student went to school. They brought lunch boxes. (>A)
- Each student went to school. He walked. (UA)
 - Each student went to school. They walked. (>A)
- All students went to school. He walked. (<UA)
 - All students went to school. They walked. (>A)
- All men love a woman. He sends her flowers. (UA)
 - All men love a woman. They send her flower. (A)
 - All men love a woman. He sends them flowers. (<UA)
 - All men love a woman. They send them flowers. (>A)
- Every man loves a woman. He sends her flowers. (A)
 - Every man loves a woman. He sends them flowers. (UA)
 - Every man loves a woman. They send her flowers. (A)
 - Every man loves a woman. They send them flowers. (>A)
- Each man loves a woman. He sends her flowers. (A)
 - Each man loves a woman. He sends them flowers. (UA)
 - Each man loves a woman. They send her flowers. (M)
 - Each man loves a woman. They send them flowers. (A)
- All students wrote a paper. He submitted it to a journal. (UA)
 - Every student wrote a paper. He submitted it to a journal (M)
 - Each student wrote a paper. He submitted it to a journal. (A)
- All students wrote a paper. They submitted it to a journal. (>A)
 - Every student wrote a paper. They submitted it to a journal (>A)

- c. Each student wrote a paper. They submitted it to a journal. (>A)
9. a. All students wrote a paper. He submitted them to a journal. (UA)
 b. Every student wrote a paper. He submitted them to a journal (UA)
 c. Each student wrote a paper. He submitted them to a journal. (M)
10. a. All students wrote a paper. They submitted them to a journal. (>A)
 b. Every student wrote a paper. They submitted them to a journal (>A)
 c. Each student wrote a paper. They submitted them to a journal. (>A)
11. a. Each student wrote a paper. They sent it to a journal. (>A)
 b. Each student wrote a paper. They sent it to a journal. They had worked very hard on it. (A)
 c. Each student wrote a paper. They sent it to a journal. It was a good paper. (M)
12. a. A train leaves every hour from Boston. It stops in New Haven. (>A)
 b. A train leaves every hour from Boston. It always stops in New Haven. (>A)
 c. A train leaves every hour from Boston. They stop in New Haven. (M)
 d. A train leaves every hour from Boston. They always stop in New Haven. (M)
13. a. Each student in the syntax class was accused of cheating on the exam. He was reprimanded by the dean. (Result, UA)
 b. Each student in the syntax class was accused of cheating on the exam. He had a Ph.D. in astrophysics. (Background, <UA)
 c. Each student in the syntax class was accused of cheating on the exam. They were reprimanded by the dean. (Result, >A)
 d. Each student in the syntax class was accused of cheating on the exam. They had Ph.D.s in astrophysics. (Result, A)
14. a. Every student in the syntax class was accused of cheating on the exam. He was reprimanded by the dean. (Result, UA)
 b. Every student in the syntax class was accused of cheating on the exam. He had a Ph.D. in astrophysics. (Background, <UA)
 c. Every student in the syntax class was accused of cheating on the exam. They were reprimanded by the dean. (Result, >A)
 d. Every student in the syntax class was accused of cheating on the exam. They had Ph.D.s in astrophysics. (Result, A)
15. a. Each degree candidate walked to the stage. He took his diploma from the dean and returned to his seat. (Narration, >M)
 b. Each degree candidate walked to the stage. He was working in a high-tech company. (Background, <UA)
 c. Each degree candidate walked to the stage. They took their diplomas from the dean and returned to their seats. (Narration, >A)
 d. Each degree candidate walked to the stage. They were working in a high-tech company. (Background, A)
16. a. All degree candidates walked to the stage. He took his diploma from the dean and returned to his seat. (Narration, UA)
 b. All degree candidates walked to the stage. He was working in a high-tech company. (Background, <UA)
 c. All degree candidates walked to the stage. They took their diplomas from the dean and returned to their seats. (Narration, >A)
 d. All degree candidates walked to the stage. They were working in a high-tech company. (Background, A)

17. a. Every chess set from that store came with a spare pawn. But then a boy lost it. (Narration, UA)
- b. Every chess set from that store came with a spare pawn. It is taped to the top of the box. (Elaboration, >A)
- c. Every chess set from that store came with a spare pawn. But then a boy lost them. (Narration, A)
- d. Every chess set from that store came with a spare pawn. They are taped to the top of the box. (Elaboration, >A)
18. a. Each chess set from that store came with a spare pawn. But then a boy lost it. (Narration, <M)
- b. Each chess set from that store came with a spare pawn. It is taped to the top of the box. (Elaboration, >A)
- c. Each chess set from that store came with a spare pawn. But then a boy lost them. (Narration, A)
- d. Each chess set from that store came with a spare pawn. They are taped to the top of the box. (Elaboration, >A)
19. a. All chess sets from that store came with a spare pawn. But then a boy lost it. (Narration, UA)
- b. All chess sets from that store came with a spare pawn. It is taped to the top of the box. (Elaboration, >A)
- c. All chess sets from that store came with a spare pawn. But then a boy lost them. (Narration, A)
- d. All chess sets from that store came with a spare pawn. They are taped to the top of the box. (Elaboration, >A)
20. a. Each student in the syntax class cheated cheating on the exam. He did a really bad thing. (Commentary, <UA)
- b. Each student in the syntax class cheated cheating on the exam. They did a really bad thing. (Commentary, A)

A.2 Survey 2 Material

1. Three students wrote a paper. He sent it to a journal. (Narration, UA)
2. Three students wrote a paper. They sent it to a journal. (Narration, >A)
3. Three students wrote a paper. He sent them to a journal. (Narration, <UA)
4. Three students wrote a paper. They sent them to a journal. (Narration, >A)
5. Every student in the chemistry class was caught cheating on the exam. He wrote the Periodic Table on his hand. (Explanation, <UA)
6. Every student in the chemistry class was caught cheating on the exam. They wrote the Periodic Table on their hands. (Explanation, >A)
7. Every student in the chemistry class was caught cheating on the exam. He was reprimanded by the dean. (Result, UA)
8. Every student in the chemistry class was caught cheating on the exam. They were reprimanded by the dean. (Result, >A)
9. Each student in the chemistry class was caught cheating on the exam. He wrote the Periodic Table on his hand. (Explanation, UA)
10. Each student in the chemistry class was caught cheating on the exam. They wrote the Periodic Table on their hands. (Explanation, A)
11. Each student in the chemistry class was caught cheating on the exam. He was reprimanded by the dean. (Result, UA)
12. Each student in the chemistry class was caught cheating on the exam. They were reprimanded by the dean. (Result, >A)

13. All students in the chemistry class were caught cheating on the exam. He wrote the Periodic Table on his hand. (Explanation, <UA)
14. All students in the chemistry class were caught cheating on the exam. They wrote the Periodic Table on their hands. (Explanation, >A)
15. All students in the chemistry class were caught cheating on the exam. He was reprimanded by the dean. (Result, UA)
16. All students in the chemistry class were caught cheating on the exam. They were reprimanded by the dean. (Result, >A)
17. Every hunter who saw a deer shot it. It died immediately. (Result, >M)
18. Every hunter who saw a deer shot it. It usually died immediately. (Result, A)
19. Every hunter who saw a deer shot it. They died immediately. (Result, A)
20. Every hunter who saw a deer shot it. They usually died immediately. (Result, A)
21. Every hunter who saw a deer shot it. He intended to kill it. (Background, M)
22. Every hunter who saw a deer shot it. He intended to kill them. (Background, <UA)
23. Every hunter who saw a deer shot it. They intended to kill it. (Background, A)
24. Every hunter who saw a deer shot it. They intended to kill them. (Background, >A)
25. Every farmer who owns a donkey beats it. He treats it badly. (Commentary, M)
26. Every farmer who owns a donkey beats it. He always treats it badly. (Commentary, M)
27. Every farmer who owns a donkey beats it. They treat it badly. (Commentary, M)
28. Every farmer who owns a donkey beats it. They always treat it badly. (Commentary, >M)
29. Every student in the chemistry class cheated on the exam. He wanted to get an A. (Explanation, UA)
30. Every student in the chemistry class cheated on the exam. They wanted to get an A. (Explanation, >A)
31. Each student in the chemistry class cheated on the exam. He wanted to get an A. (Explanation, M)
32. Each student in the chemistry class cheated on the exam. They wanted to get an A. (Explanation, >M)
33. All of the students in the chemistry class cheated on the exam. He wanted to get an A. (Explanation, UA)
34. All of the students in the chemistry class cheated on the exam. They wanted to get an A. (Explanation, >A)
35. John asks every girl he sees for a date. She usually says yes. (Result, >M)
36. John asks every girl he sees for a date. They usually say yes. (Result, >A)
37. John asks every girl he sees for a date. She says yes. (Result, >UA)
38. John asks every girl he sees for a date. They say yes. (Result, >A)
39. Every taxi that comes through here is occupied. It's usually a Yellow Cab. (Background, A)
40. Every taxi that comes through here is occupied. They're usually Yellow Cabs. (Background, >A)
41. Every taxi that comes through here is occupied. It's a Yellow Cab. (Background, UA)
42. Every taxi that comes through here is occupied. They're Yellow Cabs. (Background, >A)

References

- Asher, N. (1993), *Reference to Abstract Objects in Discourse*, Kluwer Academic Publishers.
- Asher, N. (2001), "Plurals in Discourse," manuscript, University of Texas at Austin.
- Asher, N. (2002), "Modal and Quantificational Subordination in Discourse," in *Proceedings of In the Mood*, Frankfurt, June 2002.
- Asher, N., & Lascarides, A. (2003), *Logics of Conversation*, Cambridge University Press.

- Asher, N. & Wang, L. (2003), "Ambiguity and Anaphora with Plurals in Discourse", to appear in *Semantic and Linguistic Theory XIII (SALT XIII)*.
- Carminati, M. N., Frazier, L., & Rayner, K. (2002), "Bound Variables and C-Command," *Journal of Semantics*, 19, pp. 1-34.
- Farkas, D. (1997), "Dependent Indefinites," in F. Corblin & D. Godard & J.-M. Marandin (1997), (eds.), *Empirical Issues in Formal Syntax and Semantics*, Peter Lang Publishers, pp. 243-268.
- von Stechow, K. (1994), *Restrictions on Quantifier Domains*, dissertation, University of Massachusetts at Amherst.
- Fodor, J., & Sag, I. (1982), "Referential and Quantificational Indefinites," *Linguistics and Philosophy*, 5, pp. 355-398.
- Frank, A. (1997) *Context Dependence in Modal Constructions*. Dissertation, University of Stuttgart.
- Frank, A., & Kamp, H. (1997), "On Context Dependence in Modal Constructions," *Proceedings of Semantic and Linguistic Theory (VII)*.
- Gawron, M. (1996), "Quantification, Quantificational Domains, and Dynamic Logic," in S. Lappin (1996), (ed.), *The Handbook of Contemporary Semantic Theory*, MA: Blackwell Publishers, pp. 247-267.
- Geurts, B. (1999), *Presuppositions and Pronouns*, Oxford: Elsevier.
- Groenendijk, J., & Stokhof, M. (1989), "Dynamic Montague Grammar," in Kálmán, L. et al. (1989), (eds.), *Proceedings of the Second Symposium on Logic and Language*, Budapest.
- Groenendijk, J., & Stokhof, M. (1991), "Dynamic Predicate Logic," *Linguistics and Philosophy*, pp. 39-100.
- Heim, I. (1990), "E-Type Pronouns and Donkey Anaphora," *Linguistics and Philosophy*, 13, pp. 137-177.
- Hornstein, N. (1984), *Logic as Grammar*, MA: MIT Press.
- Kamp, H. (1981), "A Theory of Truth and Semantic Representation," in J. Groenendijk, Janssen, T. & Stokhof, M. (1984), (eds.), *Truth, Representation, and Information*, Dordrecht: Foris, pp. 277-322.
- Kamp, H., & Reyle, U. (1993), *From Discourse to Logic: Introduction to Model Theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*, Dordrecht; Boston: Kluwer Academic.
- Kanazawa, M. (2001), "Singular Donkey Pronouns are Semantically Singular," *Linguistics and Philosophy*, 24, pp. 383-403.
- Karttunen, L. (1976), "Discourse Referents," *Syntax and Semantics Vol. 9*, pp. 363-385.
- Kibble, R. (1998), "Modal Subordination, Focus and Complement Anaphora," *The Tbilisi Symposium on Language, Logic and Computation: Selected Papers*, CSLI/SiLLI.
- Krifka, M. (1996), "Parametrized Sum Individuals for Plural Anaphora," *Linguistics and Philosophy*, pp. 555-598.
- Link, G. (1998), *Algebraic Semantics in Language and Philosophy*, Stanford, Calif.: CSLI Publications.
- Neale, S. (1990), *Descriptions*, Mass.: MIT Press.
- Poesio, M., & Zucchi, A. (1992), "On Telescoping," *Proceedings of the Conference on Semantics and Linguistic Theory*.
- Roberts, C. (1987), *Modal Subordination, Anaphora, and Distributivity*, Dissertation, UMass, Amherst.
- Roberts, C. (1989), "Modal Subordination and Pronominal Anaphora in Discourse," *Linguistics and Philosophy*, 12(6), pp. 683-721.

- Roberts, C. (1995), "Domain Restriction in Dynamic Interpretation," in E. Bach & E. Jelinek & A. K. & B. Partee (1995), (eds.), *Quantification in Natural Language*, Dordrecht: Kluwer.
- Roberts, C. (1996), "Anaphora in Intentional Contexts," in S. Lappin (1996), (ed.), *The Handbook of Contemporary Semantic Theory*, Massachusetts: Blackwell Publishers, pp. 215-247.
- Sells, P. (1985), *Restrictive and Non-Restrictive Modification*, (CSLI-85-28): Center for the Study of Language and Information.
- van den Berg, M. (1996), *Some Aspects of the Internal Structure of Discourse: The Dynamics of Nominal Anaphora*, Dissertation, University of Amsterdam.
- Veltman, F. (1996), "Defaults in Update Semantics," *Journal of Philosophical Logic*, 25, pp. 221-261.
- Wang, L. & Asher, N. (2004), "An Intensional Semantics for Pronoun Numbers", paper presented at the 32nd Society for Exact Philosophy, University of Maryland at College Park, Maryland.

Linton Wang
Department of Philosophy
1 University Station
University of Texas at Austin
Austin, Texas 78712, U.S.A.
kikiwang@mail.utexas.edu

Eric McCready
Department of Linguistics
1 University Station
The University of Texas at Austin
Austin, Texas 78712, U.S.A.
mccready@mail.utexas.edu

Nicholas Asher
Department of Philosophy
1 University Station
University of Texas at Austin
Austin, Texas 78712, U.S.A.
nasher@mail.utexas.edu