Article

Language & Linguistics 語言暨語言學

Language and Linguistics
16(1) 3–42
© The Author(s) 2015
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1606822X14556594
lin.sagepub.com

SAGE

Global Approach to Scalar Implicatures in DRT*

Jae-Il Yeom

Hongik University

It has been disputed whether scalar implicatures (=SIs) arise globally or locally. Basically, SIs should be global because they are calculated by comparing the strengths of whole alternative statements arising from a scalar set. On the other hand, localists claim that local SIs are sometimes favored and they are obtained by applying an operator to part of a statement. In this paper, I show that seemingly local SIs are effects that follow from anaphoric bindings of some variables in global SIs or from the contributiveness constraint. Moreover, some SIs do not correspond to any clause-type syntactic structures. In this paper, I propose an analysis of SIs within the discourse representation theory. The maxim of quantity provides the basic SI, which is the weakest one. The basic SI undergoes some processes in which some variables can be bound or some conditions are eliminated. In these processes, global SIs get the effects of local SIs, and some SIs arise from non-sentential constituents. Then among the resulting meanings, we get the strongest possible meanings, restricted by contributiveness conditions and likelihood.

Key words: contributiveness, dynamic binding, global scalar implicature, likelihood, local scalar implicature, strongest meaning, weak negation

1. Introduction

It has been one of the hottest issues whether a scalar implicature (=SI) is calculated in the process of compositional interpretation of a sentence or as an inference from an utterance as a whole. Landman (1998), Chierchia (2002), Fox (2007), and Chierchia et al. (2012) took the former position. In that position, a scalar term introduces a set of alternatives and an operator is introduced locally to get an SI as part of the meaning of a sentence. To them, SIs are just a grammatical phenomenon, and they can be called **localists**. Neo-Griceans, on the other hand, assumed that SIs are obtained by applying the maxim of quantity by Grice (1975) to a whole utterance. They can be called **globalists**. Spector (2003), van Rooij & Schulz (2004), Sauerland (2004), and Chemla (2008) can be considered globalists. This issue is still in debate, and I will try to support the global approach to SIs by discussing new evidence and proposing a new analysis of SIs.

The maxim that is relevant to the calculation of SIs is the first submaxim in the maxim of quantity by Grice (1975):

(1) The maxim of quantity:
   a. Make your contribution as informative as is required (for the current purposes of the exchange).
   b. Do not make your contribution more informative than is required.

* This is an extended version of the paper presented at the 27th Pacific Asia Conference on Language, Information, and Computation. I thank the audience for the questions and comments. I benefited very much from the comments of two anonymous reviewers in clarifying some points I made and making the paper more organized and streamlined. All remaining errors are mine.

When a weaker statement is made and a stronger statement is still relevant in the context, it is implicated that the stronger statement is not true in the speaker's information state. Assuming that the speaker is well informed and that he knows the stronger alternative is false, the hearer accepts the negation of the stronger statement as an SI.

(2) ϕ_1 and ϕ_2 are relevant in the context;
$\phi_1 > \phi_2$;
ϕ_2 uttered

$\rightsquigarrow \neg\phi_1$ ('$\rightsquigarrow$': 'implicate(s)')

The (Neo)-Gricean idea was made more concrete by Horn's (1972, 1989) notion of **scalar set**; Gazdar's (1979) introduction of the epistemic operator; and Sauerland's (2004) stepwise derivation of an SI, which will be discussed in §3.2.[1] This predicts the following implicatures:

(3) Some students who drank a bottle of beer **or** a glass of wine were allowed to drive.
$\rightsquigarrow$ NOT[Some students who drank a bottle of beer **and** a glass of wine were allowed to drive.]
= **No** students who drank a bottle of beer **and** a glass of wine were allowed to drive.

In (3), the weaker expression *or* is used in the relative clause, but we get an SI by negating a stronger alternative of the whole statement, which we get by replacing *or* with *and*.

There are a lot of examples which favor a globalistic approach. In (3), the meaning with a local SI would mean that some students who drank a bottle of beer or a glass of wine, *but not both*, were allowed to drive. This leaves the possibility that some other students who drank both were allowed to drive. But globalists predict the implicature that no students who drank both were allowed to drive. In the given context, the global approach predicts a more plausible inference. In (4), a linguist at MIT is likely to have read at least one of the two books, and the global SI is more plausible.

(4) Every linguistics student at MIT has read LGB **or** Syntactic Structures. (Modified from Sauerland 2004, (58))
a. **NOT**[Every linguistics student at MIT has read LGB **and** Syntactic Structures.]
b. Every linguistics student at MIT **NOT**[has read LGB **and** Syntactic Structures.]
(= No linguistics students at MIT have read LGB and Syntactic Structures.)

However, localists claim that implicatures are included in the strengthened meaning of a statement when a CHUNK of a statement, as a scope site of a scalar expression, is interpreted.

[1] Horn (1972, 1989) suggested that an SI arises by comparing a set of alternative statements that arise by replacing a scalar term in the original statement with a stronger scalar alternative expression in the language system. But a scalar set can be formed pragmatically, as Hirschberg (1991) pointed out. Sauerland (2004) assumes that SIs have an epistemic status, following Gazdar (1979), but further assumes that SIs are derived in two steps: he derives primary implicates first, and then secondary implicates from them.

A scalar term generates a set of alternatives, and an operator takes that set of alternatives, together with the original chunk, and yields an SI, following Rooth's (1985) or Krifka's (1995) analyses of focus phenomena, in which a focus-sensitive operator takes a set of alternatives for its interpretation. The strengthened meaning of the sentence chunk α is the conjunction of the literal meaning $[\![\alpha]\!]$ and the implicature $\neg S(\alpha^{ALT})$, where $S(\alpha^{ALT})$ is the weakest alternative of α that entails α: $[\![\alpha]\!] \wedge \neg S(\alpha^{ALT})$. Chierchia (2004) introduces the negation operator to get a stronger meaning, but Fox (2007) and Chierchia et al. (2012) introduce the exhaustivity (*exh*, hereafter) operator.

There are cases where a local SI seems to be required. In example (5), you are forced to reinterpret the first sentence, to avoid contradiction when you take both phonology and semantics. Such implicatures are called **intrusive implicatures**. One thing in common is that the two disjuncts are not independent of each other. The discourse is not inconsistent, because even without reinterpretation we would not reach an absurd state but an information state in which you attend meeting A. However, the discourse is incoherent. A speaker who can truly assert the second sentence would not utter the first sentence without adjusting the meaning of the first sentence.[2] The meaning of the first sentence has to be adjusted so that it means that if you take phonology or semantics, but not both, you attend meeting A. In (6), *some* and *at least warm* are forced to to be interpreted as *some*, *but not all* and *at least warm*, *but not hot*, respectively.

(5) If you take phonology or semantics, you attend meeting A. But if you take both, you attend meeting B.

(6) a. John broke all or **some** glasses.
b. John wanted hot or at least **warm** water.

Even if local SIs are not required like this, there are cases where local SIs are preferred. In (7), it is more likely that a student who watched TV and played games failed math. For this reason, the global SI that no students who watched TV and played games failed math is not acceptable. Similarly, in (8), if either of the two requirements is sufficient to get a grade, it is more plausible to assume that no students satisfied both options. This corresponds to the local SI. The global SI that not every student did both is not likely to be accepted.

(7) Some students who watched TV **or** played games failed math.
$=\rightsquigarrow$ Some students who watched TV or played games, but **not both**, failed math.
('$=\rightsquigarrow$': 'convey(s)')
$\not\rightsquigarrow$ **NOT**[Some students who watched TV **and** played games failed math.]

(8) Every student wrote a paper **or** made a classroom presentation.
$=\rightsquigarrow$ Every student wrote a paper or made a classroom presentation but **did** not do **both**.
$\not\rightsquigarrow$ **NOT**[Every student wrote a paper **and** made a classroom presentation.]

[2] A simple way of distinguishing coherence and consistency is that a discourse is coherent if it is asserted by the same speaker, but a discourse is consistent if it does not lead to the absurd information state if uttered by different speakers. In the example at hand, if the two sentences are uttered by two different speakers, it does not lead to the absurd information state. But they cannot be uttered by the same speaker felicitously.

Localists' approaches may look more systematic, and more constrained than globalists', because they are based on syntactic structures and calculation of SIs is precisely incorporated into compositional interpretation. However, there are problems with their approaches, which are extensively discussed in Geurts (2010:151–160). One theoretical problem is that, as Horn (1989) pointed out, SIs do not arise within downward-entailing or non-monotonic contexts and SIs are based on strengths of statements as a whole.

(9) a. If John drank a bottle of beer **or** a glass of wine, he was not allowed to drive.
= ⇝ If John drank a bottle of beer or a glass of wine but **not both**, he was not allowed to drive.
⇝ It is **not** the case that if John drank both a bottle of beer **and** a glass of wine, he did not drive.

b. Every student who read **some** of the book passed the test.
= ⇝ Every student who read some, but **not all**, of the book passed the test.
⇝ **Not** every student who read **all** of the book passed the test.

The antecedent clause of a conditional introduces a non-monotonic context, and the use of the disjunction structure does not yield a local or global SI. The restrictor of the universal quantifier *every* also introduces a downward-entailing context, and the use of *some* in it does not yield an SI. Even if localists calculate SIs locally, they have to check whether an alternative involved in the calculation makes the whole statement stronger to see if it really yields a valid SI. In this respect, SIs are inherently global.[3]

Chemla & Spector (2011) and Chierchia et al. (2012) claim that there are cases where an expression in a non-monotonic context can yield an SI.

(10) Exactly three students wrote a paper or made a presentation.

They claim that the sentence can mean that exactly three students wrote a paper or made a presentation *but did not do both*. This seems different from a case of intrusive implicature. But this is not against the idea that SIs arise by comparing the strengths of alternative statements. The alternative statement with the stronger alternative does not entail the original statement. Any globalistic approach must be able to account for these cases.

In this paper, I accept the idea that SIs are inherently global. The question is *why and how we get the effects of local SIs from global SIs*. There are three factors involved. First, there is more than one kind of scale. SIs are generally assumed to arise from the use of a semantically and pragmatically weaker term. But there is another kind of scale: likelihood. The two scales may be in opposite directions. When they clash, we are forced to find an alternative SI. Behind this

[3] Geurts (2010) also points out an empirical problem with the localistic approach. In a propositional attitude context, it makes the wrong prediction:

(i) I hope you read some of the books. ⇝ I hope you did not read all of them.

The use of *some* does not implicate 'not all' in the desire context.

reasoning lies the assumption that there are multiple candidates for SIs that have different degrees of likelihood. Different SIs can be obtained from some of the truth conditions of a stronger alternative statement, ignoring others. We need a process of selecting a more likely one. Second, in some cases, the distinction of local and global SIs is not a matter of structure but a matter of information. Some seemingly local SIs are actually global SIs that involve anaphoric binding. To deal with such SIs, I will use discourse representation structures (DRSs). Third, there are cases where SIs are really local. To account for such cases, we need a separate condition called the contributiveness constraint. Some global SIs might make some expressions non-contributive. In such cases, we need to revise the SIs so that all expressions can be contributive. This is not part of an SI, and contributiveness is defined with respect to the main context. In this sense, it is not a local process.

The paper is organized in the following order. In §2, I discuss two factors that affect global SIs so that we would get the effects of local SIs: likelihood, and anaphoric binding in dynamic semantics. In §3, I propose a concrete analysis of SIs within the framework of discourse representation theory (DRT). Then I discuss cases where some global SIs are not accepted due to their unlikelihood. And there are cases where some variables in an SI can be bound by the corresponding variables in the main DRS and the conditions on them are eliminated in the SI DRS. I also deal with a case where an SI can arise due to the different degree of unlikelihood. Unlikelihood (or unexpectedness) is a factor that gives rise to an SI. Something unexpected is more informative than something trivial. I also discuss cases where a scalar term is embedded in the scope of a quantifier over individuals or a belief context. These cases can also be analyzed as cases where some variables in an SI are bound by corresponding variables in the main DRS. This process puts the SI in the scope of the quantifier, even though it happens globally. In §4, I discuss cases where a scalar term is embedded in a disjunction structure or a conditional. In these cases, simple negation of a stronger alternative statement makes some expressions non-contributive. To avoid this, I propose to use the weak negation operator, instead of the standard strong negation operator. To a weak SI with the weak negation operator is added contributiveness conditions for expressions in the sentence. Then the weak SI is strengthened and the strengthened SI applies to the domain in which it does not conflict with the contributiveness conditions. This is an additional way a global SI gets the effect of a local SI. In §5, I make some concluding summarizations and state some implications.

2. Why and how we seek local SIs

An SI is inherently global, but there are cases where we try to find a local SI. What is involved here is the ordering of likelihood. When a global SI of a statement is unlikely, we try to find a way to get a different SI, removing some part of the statement. The removed part is what is simply ignored as irrelevant or dynamically bound by the corresponding part of the original statement. These processes are discussed in this section.

2.1 Ordering of likelihood

Horn (1972, 1989) proposed the notion of scalar set (=a set of scalar terms) based on semantic strength. The scalar terms in the set mechanically generate a set of alternative statements. However, there is another scale that is not based on scalar terms or any explicit expressions. It is the scale

of likelihood, which determines the likelihood of propositions, based on background knowledge. It ultimately determines what SI we get from a stronger alternative sentence.

Scalarity of likelihood is not based on some explicit expressions, but on propositions.[4] Likelihood is defined in terms of the ordering of possible worlds. The ordering relation between possible worlds based on likelihood can be defined as follows, following Kratzer (1991), except that it is defined with respect to an information state:[5]

(11) Modal base *W*: a set of accessible worlds, with regard to likelihood.
Ordering source: if $w_1 \leq_w w_2$, then w_2 is more likely than w_1 in w.[6]
ϕ_1 is more likely than ϕ_2 in an information state s ($\phi_1 \sqsupseteq_s \phi_2$) iff
for all $w \in s$,
i. if ϕ_2 is true in w', then there is $w'' \geq_w w'$ such that ϕ_1 is true in w'';
ii. there is $w' \in s$ such that ϕ_1 is true in w' and for all $w'' \geq_w w'$, ϕ_2 is false in w''.

I assume that the ordering of likelihood is given in the model. A proposition ϕ_1 is more likely than ϕ_2 in an information state, which is represented as '$\phi_1 \sqsupseteq_s \phi_2$', iff for every possible world w in an information state s, (i) if ϕ_2 is true in w', then there is a ϕ_1-world w'' which is at least as likely as w' in w, and (ii) there is a ϕ_1-world w' such that no ϕ_2-world w'' which is as likely as, or more likely than, w' in w.

In (12), the scalar set relevant is <love, like>. If you like to study math, you are likely to do well in math and unlikely to fail in math. Then it is more likely to be true that no one who likes to study math does well in math than that no one who loves to study math does well in math. It is assumed that the speaker follows the maxim of quantity and the mentioned scalar term is understood as a limit. Thus *like* is an upper limit for the sentence to be true, and it leads to the meaning of *at most likes to study math*. On the other hand, it is more likely to be true that no one who loves to study math fails in math than that no one who likes to study math fails in math. Thus *like* is a lower limit and it is interpreted as meaning *at least likes to study math*. Similarly, it is more likely to be true that people who do well in math love to study math than that people who do well in math like to study math, and it is more likely to be true that people who fail in math like to study math than that people who fail in math love to study math. This leads to two different interpretations of *like to study math*: one with 'at most' and the other with 'at least'.

---

[4] Lewis (1973) and Kratzer (1991, 2012) capture the likelihood in terms of the ordering of possible worlds. Fine (1973) and Lassiter (2011) propose the ordering of probability. Which position we may take to capture likelihood, the scalarity of likelihood can be opposite to that of semantic scalarity.

[5] One reason I define it with respect to an information state is that, as information accumulates, some propositions that were more likely at an earlier stage can be ignored. For example, we know that birds fly but penguins do not. When we only know that an animal named Petra is a bird, it is likely that it flies. But when we know that it is a penguin, we come to the conclusion that it does not.

[6] If the truth values of generic sentences are given with respect to each possible world, the ordering of possible worlds will be defined in terms of generic sentences that are true in that possible world. Therefore the ordering of possible worlds is defined with respect to a possible world in the model.

(12) a. No one who likes to study math does well in math.
b. No one who likes to study math fails in math.

(13) a. People who do well in math like to study math.
b. People who fail in math like to study math.

A set of scalar alternatives introduces a set of propositions, which can be ordered with respect to their likelihood. And the ordering of likelihood might be opposite to the semantic ordering in the scalar set. In such cases, no global SIs arise from a stronger alternative. Take (3) and (7) for example. In (3), when a student who drank a bottle of beer or a glass of wine was allowed to drive, a student who drank both a bottle of beer and a glass of wine is not likely to have been allowed to drive.

(3) Some students who drank a bottle of beer or a glass of wine were allowed to drive.
$\rightsquigarrow$ No students who drank a bottle of beer and a glass of wine were allowed to drive.

(14) Likelihood Scale in (3): (ii) $\sqsupset_s$ (i)
i. Some students who drank a bottle of beer **and** a glass of wine were allowed to drive.
ii. Some students who drank a bottle of beer **or** a glass of wine were allowed to drive.

Drinking either a bottle of beer or a glass of wine is the **upper limit** for being allowed to drive. Thus (14ii) is more likely than (14i). For every possible world in which a student who did both was allowed to drive, there is a possible world at least as likely in which a student who did only one of the two was allowed to drive, but not the other way around. Then the SI that we get by negating the less likely proposition becomes more likely. Therefore it is compatible with this background knowledge and it is accepted.

In (7), on the other hand, when a student who watched TV or played games failed math, a student who did both is more likely to have failed math.

(7) Some students who watched TV or played games failed math.
$\not\rightsquigarrow$ No students who watched TV and played games failed math.

(15) Likelihood Scale in (7): (i) $\sqsupset_s$ (ii)
i. Some students who watched TV **and** played games failed math.
ii. Some students who watched TV **or** played games failed math.

Semantically (15i) is stronger than (15ii). However, watching TV or playing games is understood as the **lower limit** for failing math. Thus (15i) is more likely than (15ii): for every possible world in which a student who watched TV or played games failed math, there is a possible world at least as likely in which a student who watched TV and played games failed math, but not the other way around. Thus if there is a student who failed math, he is more likely to be one who did both than one who did only one of the two things. This blocks an SI from arising from the use of the weaker alternative, because the negation of the more likely proposition is less likely. This makes the SI unacceptable.

In this case, it is more likely to get the meaning that some students who watched TV or played games, but didn't do both, failed math. A local SI is a way of avoiding calculating an SI against the likelihood. However, in order not to be against the ordering of likelihood, we are ignoring the expression *failed math* in calculating an SI, because the likelihood we have talked about is based on the relation between doing one or both of the two things and failing math and not doing both is not likely to lead to failing math. If the expression is ignored, there is no ordering of likelihood in doing only one of the two things or both alone. If we ignore the expression *failed math*, the effect of the local SI can be derived from a global SI. This will be discussed below.

Consider (4) and (8). In both cases, the VPs with *or* are considered to be the upper limits that every (linguistics) student is expected to have done, and a global SI is expected. In (4), the stronger alternative with *and* can be globally negated. In (8), however, the global negation of the stronger alternative with *and* is trivial. No student was expected to both write a paper and make a class presentation. When a stronger alternative is considered to be out of the question, its negation leads to too trivial an SI. In such a case, we try to get a more informative SI, which is a local SI. This indicates that we only consider alternatives that can be considered plausible and informative enough.

The discussion of likelihood shows that a less plausible or trivial global SI leads us to find a more plausible or more informative local SI. This indicates that the factor of likelihood tells us when we get a global SI and when we need a local SI with some expression ignored. The question is how to get the effect of a local SI when we assume a globalistic approach. In the next subsection, I will show that getting an SI is not a matter of structure. It will give us a clue to the answer.

2.2 SIs corresponding to no syntactic constituents

In a syntactic analysis, an operator applies to a constituent that denotes a proposition. In Chierchia (2002), the negation operator applies to a scope site of a scalar expression. A scope site has a semantic type of a proposition, and structurally it is a clause. Fox (2007) used the *exh* operator instead. The two operators are assumed to take as arguments both the meaning of the constituent as a proposition and a set of alternative propositions to that.

But I am going to claim that there are cases where we can get an SI from a non-clausal expression or from an expression that intuitively does not correspond to (part of) the original statement. In (16), from the use of *some of the books*, a global approach predicts the SI that no boys who read all of the books passed the test, which is unlikely, if not impossible, if there are boys who read all of the books. A more plausible inference from this is that no boys read all of the books. Otherwise, it is likely that some or all boys who read all of the books passed the test.

This is more likely to be related to a local SI. A local approach predicts that the sentence conveys the meaning that some boys who read some, but not all, of the books passed the test.[7] This

[7] An anonymous reviewer suggested the meaning that only some boys who read only some of the books passed the test. It implies that there are some boys who read only some of the books and did not pass the test, and this is not the implicature from the use of *some* in the relative clause, which I am more concerned with. And it does not allow the possibility that there are boys who read all of the books and passed the test. This is not an impossible interpretation, but it is the least plausible one. I am claiming that for the most plausible interpretation, we need a different approach.

does not exclude the possibility that some or all boys who read all of the books passed the test. Considering the possibility, we can suppose that the speaker ignores the possibility and simply talks about boys who read some of the books. In this context, by the use of *some* in the relative clause the speaker can only implicate that the boys who passed the test did not read all of them, ignoring (and hence, not mentioning) boys who read all of the books or assuming that no boys read all of the books. The global and local approaches seem to yield the following two SIs.

(16) Some boys who read some of the books passed the test.
  a. No boys read all of the books.
  b. The boys who (read some of the books and) passed the test did not read all of the books.

The two possible inferences from the use of *some of the books* correspond to the following structures:

(17) a. Some boys who read some of the books.
  b. Some boys who passed the test read some of the books. (≠ (16))

The first inference comes from the NP *some boys who read some of the books* with the expression *passed the test* ignored. I also mentioned in (7) the possibility that the expression *failed math* is ignored in calculating an SI. And (16b) can be calculated from a quite different sentence than the original, if we follow the strict procedures of calculating an SI in either of the two syntactic theories. Moreover, in (16b) the subject should anaphorically refer to the boys who (read some of the books and) passed the test, which is introduced by the original statement. The SI itself does not have a corresponding constituent in the original statement. What I am suggesting is that what SI we get is not a matter of structure, but a matter of information.

This analysis can apply to cases of intrusive implicatures. In (18), when you read the first sentence, you do not get any SI. But when you read the second sentence, you might have to reinterpret the first as meaning that every boy who read some, but not all, of the books passed the test, if you assume a localistic approach. But implicatures are supposed to add more information to the context, and this should be done after the statement is interpreted. We can do this by assuming that the information that the boys who (read some of the books and) passed the test did not read all of the books is added, instead of reinterpreting the first sentence. This is not a revision but an addition of information, and the added information does not have a corresponding constituent in the original sentence.

(18) Every boy who read some of the books passed the test. Every boy who read all of the books is exempted from the test.
  ⇝ The boys who (read some of the books and) passed the test did not read all of the books.

One way of calculating SIs based on information, not on structures, and making the calculation manageable is to resort to *semantic representations* in dynamic semantics:

(19) Mary met a doctor. He treated some of the patients.
≃ Mary met a doctor who treated some of the patients.
He = the doctor who Mary met

Even though the pronoun *he* is used in a new separate sentence, the whole text seems to have the effect of embedding the second sentence in the first as a relative clause. Then the SI from the second sentence has the effect of a local SI.

The opposite of a similar effect is observed in (16) and (18). The sentences can be followed by a sentence like 'they did not read all of the books', where 'they' refers to the boys who read some of the books and passed the test. The result effect is the same as the local SIs.

(20) a. Some boys who read some of the books passed the test. ($\rightsquigarrow$) **They** did not read all of the books.
b. Some boys who read some, but not all, of the books passed the test.
c. i. Every boy who read some of the books passed the test. **They** did not read all of the books.
ii. Every boy who read some, but not all, of the books passed the test.

It has been pointed out in Kadmon (1990) that in (20) there is a subtle difference between the discourse with a pronoun and the one with a local SI. When a pronoun refers to the boys who read some of the books and passed the test, it is generally assumed that there is a maximal set of boys in the context that satisfies the conditions. This implies that there might have been no other boys who read all of the books and passed the test. However, when the same content is realized within the relative clause, it does not exclude the possibility that some other boys who read all of the books passed the test. If the speaker intends to mean this, it is a case where the scalar term *some* does not yield an SI: it could mean 'at least some'. Therefore, if the scalar term yields an SI at all, the former expresses the intended meaning better. Note again that this effect is normally predicted in dynamic semantics.

3. SIs in DRT

Geurts (2010) claimed that SIs are not a sentence-based phenomenon but a discourse-based phenomenon, and discussed in a DRT a case where an SI can involve anaphoric binding when an indefinite is used in a statement. But he did not discuss the possibility that some local readings can be analyzed as cases where global SIs have the effects of local SIs. When a scalar term is embedded in a syntactic island which does not introduce a semantic operator, the SI is not really an embedded one. This can be easily explained in a DRT. This is what I am going to discuss in this section.

3.1 Basics of DRT

In this paper, I employ a DRT in representing the meaning of a sentence for various reasons. I assume that SIs are calculated from the meaning of a sentence or a discourse. But the meaning of a sentence or a discourse should be manageable to calculate an SI from it. For this reason, we need

a representation theory of meaning, and I adopt the DRT. It allows us to manage more fine-grained and flexible chunks of meanings than the classical predicate logic or a syntactic theory. DRSs are less structured than sentence structures or classical predicate logic and, more importantly, no island effect is observed in them. In DRSs there are no scopes of existential quantifiers, unlike formulas in the classical predicate logic. And anaphoric relationships are easily captured across sentences. These properties of DRSs are necessary for us to account for some SIs that arise from an expression that is not a syntactic constituent and some global SIs that have the effects of local SIs.

Before I start explaining how SIs are calculated, I will briefly explain the basics of the DRT I assume in this paper. In the standard DRT, the meaning of a DRT is defined as the truth conditions with respect to an assignment in a single model. I need to deviate from the standard DRT. I assume that an assignment is defined as a function that assigns a semantic entity to each of the natural numbers, and that a context consists of a set of information states, each of which is a set of pairs of a possible world and an assignment.[8] For this reason, the rules (21ci) and (21cii) are added to the standard theory of a DRT. I assume that a DRS updates a context by updating each information state in it. An information state is updated with a DRS by eliminating a pair of a possible world and an assignment which does not verify the conditions in the DRS. In this process, an information state is converted to a context with one member, which is updated with a DRS and is then converted back to an information state with the help of ↓. In this process we extend the domain *dom*(*g*) of an assignment g to g, if necessary, as given in (21ciii). If a DRS has more than one condition, a pair <*w*, *g*> must verify each of them. A non-DRS condition *P*(*i*) is verified by <*w*, *g*> if *g*(*i*) is a member of the meaning of *P*, which is given by the model. The operators like negation, disjunction, a conditional and a quantifier introduce conditions of the DRS form. They all involve some temporary extension of an assignment. In (21cix), where a quantifier is interpreted, the DRS for the restrictor is verified again when the DRS for the nuclear scope is interpreted. This is due to the conservativity property of quantifiers.

(21) a. For a given model <*W*, *D*, *F*, *G*>, where *W* is a set of possible worlds, *D* a set of individuals, *F* an interpretation function for constants, and *G* a given set of assignments, each of which is a function from a set of natural numbers *N* to some semantic entities.

b. A context *c* is a subset of a power set of *W* × *G*, that is, a set of information states; an information state *s* is a set of pairs of a possible world w and an assignment *g*;

c. A context *c* is updated with a DRS <*V*, *C*>, where *V* is a set of variables and *C* a set of conditions:

[8] A similar assumption was made in Beaver (1993), where an information state is assumed to be a set of epistemic alternatives, each of which is a set of possible worlds. Davis (2011) proposes a set of possible update relations to choose an optimal one from, assuming a single information state. Multiple interpretations lead to multiple information states. The question is whether multiplicity is just a matter of meaning or a matter of context. I suppose that multiplicity is a matter of contextual information we need to resolve through a discourse.

i. $c + \langle V, C\rangle = \{\downarrow (\{s\} + \langle V, C\rangle): s \in c\} (\downarrow (\{s\}) = s)$;

ii. $\{s\} + \langle V, C\rangle = \{\{\langle w, g'\rangle: \langle w, g\rangle \in s, g \subset_V g',$
$\langle w, g, g'\rangle \models \langle V, C\rangle\}\}$
($g \subset_V g'$ means $dom(g) \cup V = dom(g')$;
$\models$ means 'verifies'; $dom(g) = \{i \in N: \exists x[g(i) = x]\}$.)

iii. $\langle w, g, g'\rangle \models \langle V, C\rangle$ iff $g \subset_V g'$, and $\langle w, g'\rangle \models C$;

iv. $\langle w, g\rangle \models \phi; \psi$ iff $\langle w, g\rangle \models \phi$ and $\langle w, g\rangle \models \psi$.

v. $\langle w, g\rangle \models P(i)$ iff $g(i) \in F(P)(w)$;

vi. $\langle w, g\rangle \models \neg\langle V, C\rangle$ iff there is no g' s.t. $\langle w, g, g'\rangle \models \langle V, C\rangle$;

vii. $\langle w, g\rangle \models \langle V_i, C_i\rangle \vee \langle V_{ii}, C_{ii}\rangle$ iff
there is a g' such that $\langle w, g, g'\rangle \models \langle V_i, C_i\rangle$ or
there is a g'' such that $\langle w, g, g''\rangle \models \langle V_{ii}, C_{ii}\rangle$;

viii. $\langle w, g\rangle \models \langle V_i, C_i\rangle \Rightarrow \langle V_{ii}, C_{ii}\rangle$ iff
for all g' s.t. $\langle w, g, g'\rangle \models \langle V_i, C_i\rangle$,
there is a g'' s.t. $\langle w, g', g''\rangle \models \langle V_{ii}, C_{ii}\rangle$;

ix. $\langle w, g\rangle \models \langle V_i, C_i\rangle \forall_i \langle V_{ii}, C_{ii}\rangle$ iff
$\{x \in D$: there is a g' s.t. $\langle w, g, g'\rangle \models \langle V_i, C_i\rangle, g'(j) = x\} \subseteq$
$\{x \in D$: there are g' and g'' s.t. $\langle w, g, g'\rangle \models \langle V_i, C_i\rangle,$
$\langle w, g', g''\rangle \models \langle V_{ii}, C_{ii}\rangle, g''(j) = x\}$

Accessibility relations for anaphora phenomena are reflected in the interpretation, and they can be generalized as follows:

(22) In one of the following configurations, a variable in (*i*) is accessible for a variable in (*ii*), but a variable in (*ii*) is not accessible for a variable in (*iii*), and not vice versa:
$\langle V_i, \ldots \langle V_{ii}, \ldots\rangle \ldots\rangle$
$\langle V_i, \langle V_{ii}, \ldots\rangle \vee \langle V_{iii}, \ldots\rangle\rangle$
$\langle V_i, \ldots\rangle \Rightarrow$ (or $\forall$) $\langle V_{ii}, \ldots\rangle$

A variable newly introduced in a DRS is accessible for any variable in the conditions in it. In a disjunction structure, a variable in one disjunct is not accessible to a variable in another disjunct. In a conditional, a variable introduced in the antecedent clause is accessible for a variable in the consequent clause, but not vice versa. In a quantificational DRS, a variable in the restrictor DRS is accessible to a variable in the nuclear scope DRS.

3.2 SIs from non-clauses

Sauerland (2004) proposed a nice theory for SIs. He is a neo-Gricean and assumes a scalar set. The use of a scalar item triggers a set of scalar alternatives. Stronger alternatives yield weaker implicatures, which he calls primary implicatures. They are characterized as the wide scope of the negation over the epistemic operator (¬K 'the speaker does not know (that)'). Each weak SI is

tentatively strengthened to a stronger one (K¬ 'the speaker knows it is not the case (that)') to see if the stronger one can be a valid inference. If the stronger one can be compatible with the assertion plus other established implicatures, it becomes a secondary implicature, which is what we call an SI. This idea is slightly improved by the notion of innocent excludability in Fox (2007), in which a group of stronger alternatives is tentatively strengthened, instead of each alternative.

Their analyses of SIs are problematic in that they make some unmotivated assumptions, which will be discussed below, but their proposal suggests a couple of directions for the analysis of SIs. One is that a weaker implicature with '¬K' is strengthened to a stronger implicature with 'K¬'. I will take a similar strategy in the next section. I propose a weaker negation, which is somehow strengthened to the stronger negation later. The other is that an SI can be derived from part of a stronger alternative, not from all of the stronger alternatives. When part of a stronger alternative yields an SI, if it corresponds to a matrix clause, it is called a global SI, and if it corresponds to an embedded clause, it is called a local SI, whether it may be projected or not.

When some part of a stronger alternative is ignored in calculating an SI, it can be simply left out because it is not quite relevant. But it is also possible that it is not involved in calculating an SI because the variables in it are bound by some variables in the main DRS already established. Once a statement is interpreted, the information in it is supposed to be presupposition for a DRS that is added later, in the sense that it is taken for granted in interpreting the coming DRS, following Stalnaker (1978, 2002). Therefore if an SI is added to the main DRS after the original sentence is interpreted, part of the stronger alternative DRS can be bound by the DRS of the original sentence, and thus it does not have to be part of the SI we would get.

Take (16) for example. If the sentence is interpreted into the DRS in (23i), then we can get the negation of the DRS from a stronger alternative as in (23ii). It is added to the main DRS and we get the main DRS of (16) with the stronger alternative DRS in it:

(23)

(i)

X Y (Z) (k)
boys(X)
books(Z)
Y⊆Z
read(X,Y)
test(k)
passed(X,k)

(ii) ¬

X′ Y′
boys(X′)
books(Z)
l | l ∈ Z ⇒ | read(X′,l)
test(k)
passed(X′,k)

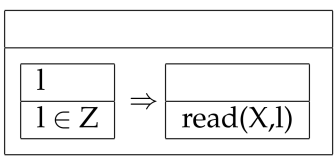
(iii)

X Y (Z) (k)
boys(X)
books(Z)
Y⊆Z
read(X,Y)
test(k)
passed(X,k)
¬ [X′ | boys(X′) | l | l ∈ Z ⇒ | read(X′,l) | passed(X′,k)]

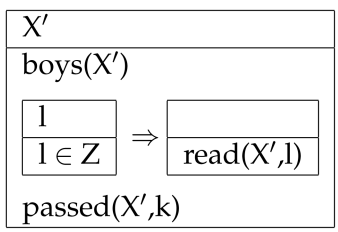
But in the negation of the stronger alternative, the variables *Z* and *K* are definite. Thus they are bound by the same variables in the main DRS, and the conditions 'book(*Z*)' and 'test(*k*)' are eliminated, which yields the DRS (23iii).[9]

We can consider a possibility that some other variables and conditions in the alternative DRS can be taken to be part of the presuppositions. The variable *X′* (i) may, or (ii) may not, be bound by the variable *X* in the main DRS:

(24) (i) ¬



(ii) ¬



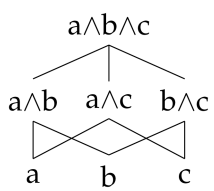
If *X′* is bound by *X*, the alternative DRS we get is like (24i), where the conditions 'boys(*X*)' and 'passed(*X*, *k*)' are transferred to the scope site of the antecedent variables. In this case, they are eliminated from the alternative DRS because they are already introduced by the main DRS. The resulting DRS can be paraphrased in various ways, because the conditions are not structured:

(25) a. Some boys who read some of the books passed the test. They (=*X* 'the boys') did not read all of them (=*Z* 'the books').
b. Some boys who passed the test read some, but not all, of the books.
c. Some boys who read some, but not all, of the books passed the test.

In the main DRS, the SI DRS is not embedded in another DRS. Thus it can be paraphrased as a separate sentence, as in (25a). It does not have to be embedded in a relative clause, as in (25b, c). It makes no sense to say whether an SI is global or local. The paraphrase (25c) can be taken to be a local SI, in localists' analysis. But in my analysis, it is one of the paraphrases of the result of a global SI with the variable *X′* bound by the discourse referent introduced by the original sentence.

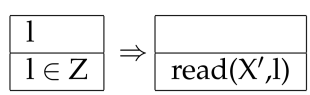
If *X′* is not bound as in (24ii), the SI DRS consists of three conditions, part of the conditions can be ignored in calculating an SI. From the three conditions, we can get up to seven possible combinations:

(26)



a: boys(X′)

b:

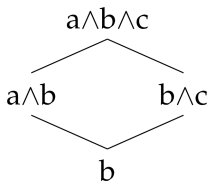


c: passed(x′,k)

[9] A similar process is proposed in van der Sandt (1992) and Yeom (1998) to account for presupposition projection.

Since we are trying to calculate an SI from the use of *some*, we cannot ignore the condition (b) from the stronger alternative of *some*. From this condition we can get the following four combinations:

(27)

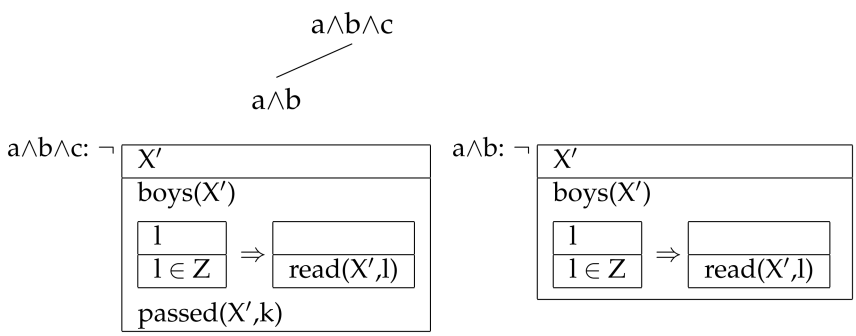


From these, we can get the following SIs:

(28) a. No boys read all of the books and passed the test.
b. No boys read all of the books.
c. No one read all of the books and passed the test.
d. No one read all of the books.

Among these, the last two SIs are not relevant, because they are about people in general. We are talking about X', which is introduced by an NP *some boys who read some of the books*. Talking about people in general is too general. Hence the condition 'boys(X')' should be involved in calculating an SI. This leaves the following two combinations:

(29)



The first one is adopted in a context where there were some boys who read all of the books but did not pass the test. If there were no boys who read all of the books, the condition 'passed(X', k)' is not relevant and the second combination is chosen. Of the two combinations, which one is preferred?

In my analyses, there are multiple SIs possible, and we can assume that the possible meanings constitute a preorder with respect to their strengths. There should be a guideline to tell us which SI to choose. In this selection the preference of the strongest meaning in terms of likelihood is involved:

(30) Preference of the strongest meaning:
If a sentence has multiple meanings ordered, the strongest meaning is preferred unless other principles disfavor it.

Based on the constraint, if $\neg(\phi \wedge \psi)$ is simply compared with $\neg\phi$, the latter is preferred.[10]

The reasoning is that, by using a scalar term, the speaker intends to convey a certain meaning and the hearer is supposed to notice it. When there are multiple meanings available that are semantically ordered, if the strongest meaning possible is not adopted, the hearer will not notice what meaning the speaker intends. To make sure communication succeeds, the speaker and the hearer follow a principle that allows them to pick out a unique SI. The preference of the strongest meaning is such a principle.

We are dealing with cases where conditions in an SI DRS are **conjunctive**. The negation of an alternative with more conjuncts leads to a weaker SI. Thus we tend to eliminate conditions to get a stronger SI. Then when is a weaker candidate adopted? If a weaker one is to be selected, there should be some motivation. The motivation is that the stronger is not acceptable. When $\neg\phi$ and $\neg(\phi \wedge \psi)$ are candidates, the latter is adopted when ϕ is, or is likely to be, true.[11] Then the fact that the stronger is not adopted leads to the conclusion that ψ is false. This can be generalized as follows:

(31) For any two candidates $\neg(\phi_1 \wedge \ldots \wedge \phi_m)$ and $\neg(\phi_1 \wedge \ldots \wedge \phi_m \wedge \phi_{m+1} \wedge \ldots \wedge \phi_n)$, the latter is adopted only if $(\phi_1 \wedge \ldots \wedge \phi_m) \wedge \neg(\phi_{m+1} \wedge \ldots \wedge \phi_n))$ is more likely.

In (28), (b) is preferred, and when (a) is adopted, we come to the conclusion that there are some boys who read all of the books, but they didn't pass the test.

So far I have shown that in the DRT, SIs can be dealt with independently of syntactic structures. This allows us to get some SIs that do not correspond to an expression in the original sentence. In this respect, my analysis is better than traditional approaches. Possible candidates for an SI are ordered and the strongest meaning is preferred.

3.3 An SI from a weaker but less likely alternative

When we talk about the strength of a statement, we think it means semantic strength. However, this is not necessarily the case. I have claimed that the ordering of likelihood plays an important role in determining an SI, especially when the direction of semantic strength is opposite to the direction of likelihood.

Suppose that a set of boys must steer a ball through blocks, and that a boy who touches fewer blocks wins. And if a boy plays well, he will go to the final. In this context, consider the following statement:

---

[10] The tendency of taking the strongest meaning is stronger in presupposition projection, because presuppositions are what are taken for granted when some expression is used. When a presupposition is accommodated, it is accommodated to a scope site which yields the strongest meaning possible. On the other hand, a stronger alternative to a scalar term or its negation is not what is taken for granted, but a pragmatic inference the hearer makes, in addition to the asserted meaning. Even in this case, the hearer tends to get a stronger inference unless there is evidence against it.

[11] Here ϕ is necessarily included because it includes a scalar term.

(32) Some boys who touch many blocks cannot go to the final.

(33) ¬

x
a:boys(x)
b: [y | blocks(y)] ALL_y [| touch(y)(x)]
c:¬ can(go_to_final(x))

¬ (a + b + c) = ??No boys who touch all blocks cannot go to the final.
(=Some boys who touch all blocks can go to the final.)
¬ (a + b) = ?No boys touch all blocks.

First, the variable *x* in the SI DRS can be bound by the variable introduced by the original utterance. This leads to the meaning that some boys who touch many, but not all, of the blocks cannot go to the final. We are not concerned with this interpretation. Suppose that the variable *x* is not bound. According to the ordering of likelihood, replacing *many* with *all* makes the statement more likely. Then the negation of it makes the proposition that some boys who touch all the blocks can go to the final, but the SI is not likely. Hence it is not acceptable. Then a next candidate is the negation of (a + b), which makes the proposition that *no boys touch all the blocks*. This is not quite plausible, but it is the most plausible SI we can get from the semantically stronger alternative. In this case, touching blocks is not related to going to the final, and the likelihood from the relation does not tell us whether it is plausible.

As I claimed in §2.1, likelihood is more important than semantic strength in some cases. In the example at hand, a semantically weaker *some* yields a more plausible SI.

(34) ¬

x, y
a:boys(x)
d:blocks(y)
e:touch(y)(x)
c:¬ can(go_to_final(x))

¬ (a + d + e + c) = No boys who touch any blocks cannot go to the final.
(=Some boys who touch some blocks can go to the final.)
¬ (a + d + e) = ??No boys touch any blocks.

Replacing *many* with *some* will make the statement weaker, but the statement is less likely. So the negation of (a + d + e + c) makes the proposition that some boys who touch some blocks can go to the final. The inference is quite an informative SI and very likely, which makes it acceptable. In this case, we do not have to try to find an alternative SI from the negation of (a + d + e). But the possibility is not excluded and we can get the SI that no boys touch any blocks. This SI is not acceptable.

It is generally assumed that an SI is derived from a semantically (or pragmatically) stronger alternative to an original statement. The stronger alternative is what we get by replacing an

expression in the original statement with a semantically (or pragmatically) stronger expression.[12] But we have seen a case where an SI obtains from a semantically weaker alternative. An SI can be understood as an effort to get more information from a statement. Informativeness can generally be measured by comparing semantic (or pragmatic) strength. But unexpectedness can be related to informativeness. In some sense, a less likely statement can be considered a more informative statement. Therefore we can revise the pattern in (2) as follows:

(2′) ϕ_1 and ϕ_2 are relevant in the context;
ϕ_1 is more informative than ϕ_2;
ϕ_2 uttered
$\rightsquigarrow \neg\phi_1$

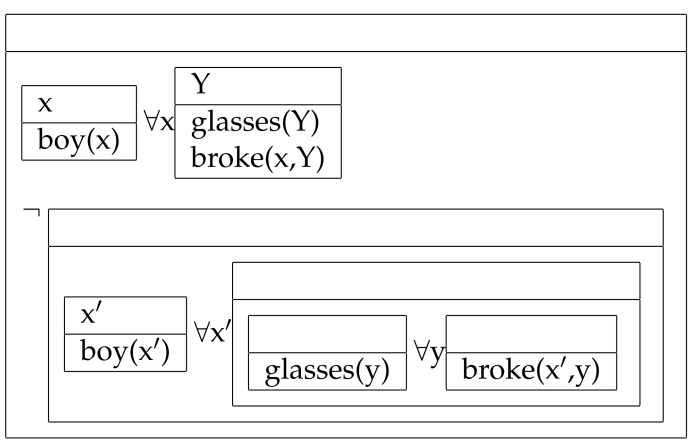
(35) ϕ_1 is more informative than ϕ_2 iff
i. ϕ_1 is semantically stronger than ϕ_2,
ii. ϕ_1 is pragmatically stronger than ϕ_2, or
iii. ϕ_1 is **less likely** than ϕ_2.

If ϕ_1 is less likely and true, then it should be mentioned. Not doing so implicates that ϕ_1 is not true. Since ϕ_1 is less likely, the negation is more likely and accepted as an SI.

3.4 Quantifiers and local SIs

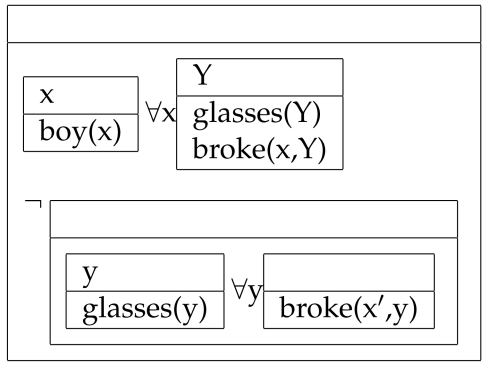
There are cases where a quantifier is involved, and in such cases, we really observe the effect of a local SI. When a quantifier is involved, the projection of an SI might violate the condition that a bound variable cannot become free:

(36) a. Every boy broke some glasses.
b.



[12] Hirschberg (1991) showed that there are scalar sets which are not semantically ordered.

c.



In (36b), if the universal quantificational DRS in the nuclear scope of the global SI DRS is tentatively projected as in (36c), then the variable *x′* will become free, and it is not a well-formed DRS. Thus in this case, the SI cannot be projected.

Then how can (8) be explained? (It is repeated here for convenience.)

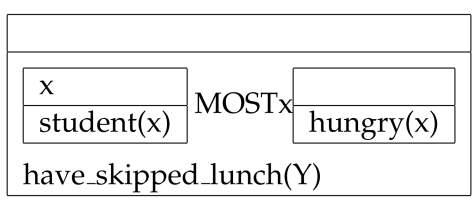
(8) Every student wrote a paper or made a classroom presentation.
⇝ Every student wrote a paper or made a classroom presentation, but did not do both.

This cannot be explained based on the analysis I have just given. We need to come up with a different explanation.

One phenomenon that gives us a clue to a solution is that a quantifier introduces another variable:

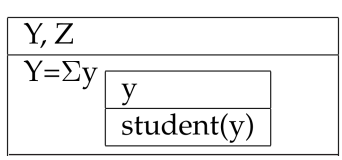
(37) a. Most students are hungry. *They* have skipped lunch.
(≃ Most students have skipped lunch and are hungry.)

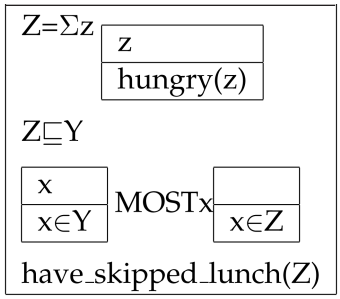
b.



We could assume that the antecedent of *they* (or, the variable *Y* in the DRS) should be something like *the students who are hungry*. Then how is the antecedent variable for *Y* introduced? I assume that a quantifier introduces three discourse referents, and modify the DRS minimally:

(38) [most$_x$ [students]$_y$]$_z$ [are hungry]

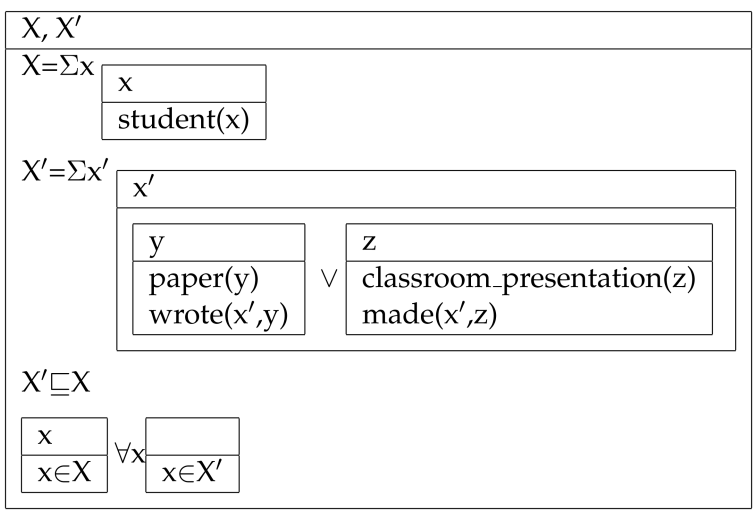




The NP *students* introduces a variable *Y* that takes the maximum number of students given in the context. This is represented as '*Y* = Σ*y* such that students(*y*)'. Similarly, the whole DP *most students* introduces another variable *Z* which takes as its value a maximum subset of the students that consists of hungry students. And the determiner *most* introduces a variable *x* that ranges over individual students.

Then (8) can be analyzed similarly. The sentence can be interpreted into the following DRS:

(39)

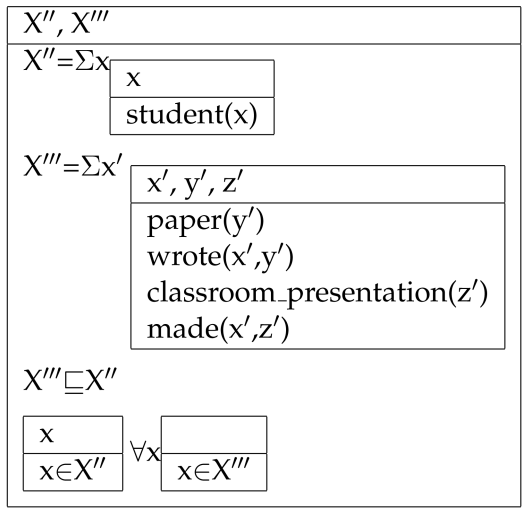


The restrictor and the nuclear scope are introduced as *X* and *X*′, respectively. And the meaning of quantification applies to the members of both sum individuals.

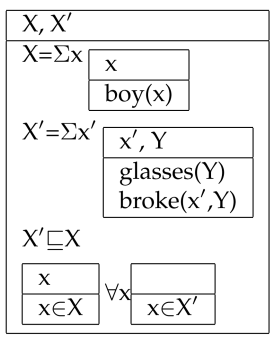
The alternative SI DRS for the sentence is given in (40a), which is embedded in the main DRS (39). In the SI DRS, *X*″ and *X*‴ are bound by *X* and *X*′, respectively. Then the meaning of quantification is also eliminated because it is redundant. And the negation applies only to the DRS for the conjunction alternative. Thus we get the sum *X*′ with the SI DRS included in it, as given in (40b). This gives the effect of a local SI and the resulting meaning is as in (40c):[13]

[13] Under the new interpretation of a quantifier, (36a) can be explained slightly differently:

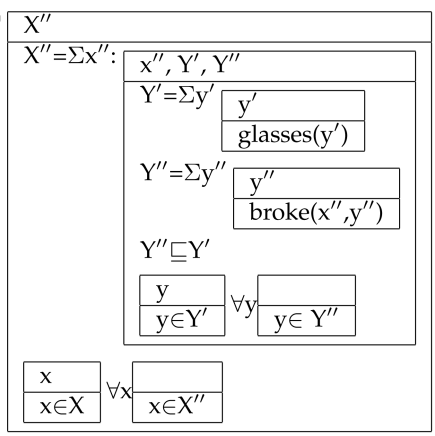
(40) a. ¬



i. a.

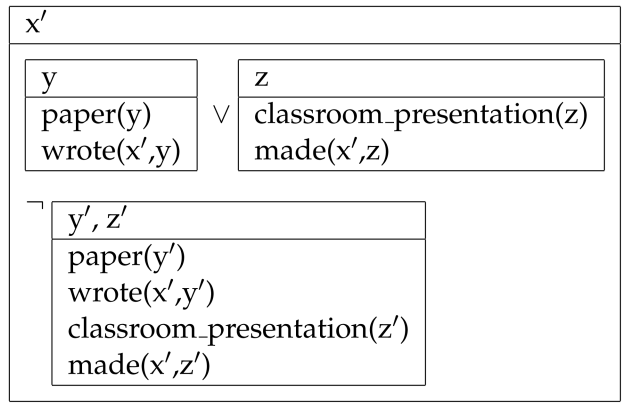


b. ¬



The sentence is interpreted as (ia) and the SI DRS can be represented as in (ib). In this DRS, the stronger alternative to *some*, *all*, introduces two variables Y' and Y'' in the nuclear scope of the quantifier *every*, and the change in the nuclear scope introduces a new variable X''. Thus no pronoun can be introduced to refer to the original nuclear scope variable X'. A more important thing is that X in the SI DRS is bound by the same variable in the main DRS and it leads to the effect of a local SI. So the DRS has the meaning that every boy broke some, but not all, of the glasses.

b. $X' = \Sigma x'$



c. Every student wrote a paper or made a classroom presentation.
They did not do both.
≃ Every student wrote a paper or made a classroom presentation, but did not do both.

This analysis allows us to account for cases where an SI arises in a non-monotonic context. The quantifier *exactly three students* introduces a non-monotonic context, in which a scalar term *or* is used:

(10) Exactly three students wrote a paper or made a presentation.

Chemla & Spector (2011) and Chierchia et al. (2012) observed that a local reading is possible. This can be easily explained in the analysis proposed in this paper as a case where binding is involved between the assertion and the SI. The students who wrote a paper or made a presentation can be referred back to by a pronoun.

3.5 Belief contexts

So far I have shown how the effects of local or global SIs can be obtained when a scalar term is embedded in a disjunction or conditional structure. Even if we get the effects of local SIs, the calculations are still globalistic because all conditions are based on whole statements. Such an analysis can account for the effects of local SIs in belief contexts.

In Chierchia (2002), it is claimed that a scalar expression in a propositional attitude context yields a local SI:

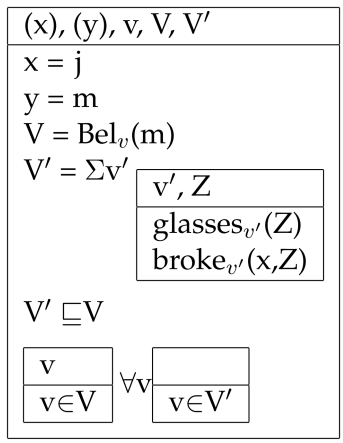
(41) Mary believes that John broke some glasses.
⇝ Mary believes that John did not break many/all glasses.
⇝ NOT(Mary believes that John broke many/all glasses.)

All we can get from a globalistic approach is that it is not the case that Mary believes that John broke many/all glasses. Thus it is a serious problem with a globalistic approach.[14]

[14] Spector (2006) explains such an example based on the process of a belief formation. A belief statement is generally based on the agent's actual statement. The belief statement is uttered because the speaker heard Mary say that John broke some glasses. When the speaker heard the statement, he got the SI that he did not break all glasses.

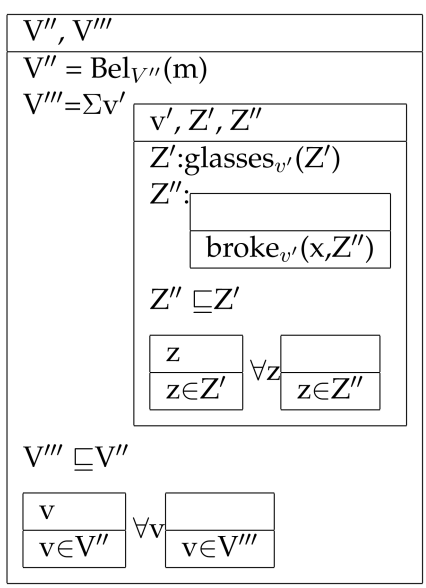
One clue to the solution is that the verb *believe* is a universal quantifier over doxastic alternatives: Mary's belief worlds constitute the domain of quantification, which I represent as 'Bel_v(m)' and the complement clause provides the nuclear scope of quantification. So far I have ignored a variable for possible worlds, but here I add variables for possible worlds to get accessible doxastic alternatives and express quantification over possible worlds. In the example at hand, the variable for the quantification domain, that is, the belief context in the SI, can be bound by the same variable in the main DRS. To show how, I assume that (41) is interpreted into the following DRS:

(42)



In this DRS, the variables V and V' are accessible for anaphora from the main DRS. Now by replacing *some* with *all*, we can get the SI DRS as follows:

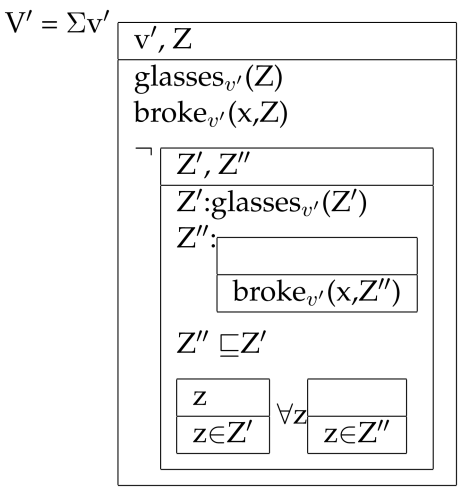
(43)



This DRS is embedded in the main DRS and the variables V'' and V''' can be bound externally by V and V' in the main DRS, deleting the conditions on them (including '$V''' \sqsubseteq V''$' and the universal

quantification between V'' and V'''). This results in the following strengthened SI DRS embedded in the main DRS:

(44)



Since the variables V and V' are bound by Mary's belief context and possible worlds in which the complement clause of the original statement is true, the strengthened SI DRS holds in the possible worlds in V'. In some sense, the two variables act as if they are referred to anaphorically, by equating the meaning of 'Mary believes' in the original statement and the same meaning in the stronger alternative statement. This generates the effect that the universal quantification over doxastic alternatives has wide scope over the negation operator, and hence the effect of a local SI.

This is one of the possible SIs we can get, which is derived by using the mechanism of binding. As we have observed in other cases, a universal quantifier over epistemic alternatives and a universal quantifier over deontic alternatives yield weaker SIs:

(45) John must have broken **some** glasses. (epistemic)
$\rightsquigarrow$ ¬John must have broken **all** glasses.
(= John may not have broken **all** glasses.)
$\not\rightsquigarrow$ John must ¬[have broken **all** glasses].
(= John cannot have broken **all** glasses.)

(46) John must read **some** of the books. (deontic)
$\rightsquigarrow$ ¬John must read **all** of the books.
(= John does **not have to** read **all** of the books.)
$\not\rightsquigarrow$ John must **not** read **all** of the books.

Here the use of *some* is supposed to yield an SI. If the statement with the epistemic necessity operator *must* is negated, the SI we get is an epistemic possibility proposition, which is a global SI. A similar result is obtained when a scalar term is used in an obligation statement. This is what we get if no binding is involved in calculating an SI.

Considering these observations, belief statements are exceptional in that they always yield SIs which have the effect of a local SI. We can think of a couple of reasons for this. When we talk about Mary's belief, we can think of Mary's statements the speaker heard, and they are likely to yield SIs, which are also taken to be part of Mary's belief even in the speaker's report. Another reason can be that a belief operator generally does not show scope interactions with the negation operator. The negation of a belief statement is generally understood as a statement of negative belief. This may be related to the lack of a salient existential counterpart of the universal quantifier *believe* over epistemic alternatives.

4. Contributiveness constraint in disjunction structures and conditionals

4.1 Weak negation

I have assumed that a context consists of a set of information states. I will explain why a DRS is interpreted with respect to a set of information states. I will start with the original idea of calculating an SI. When a weaker statement is made, we do not jump to the conclusion that the negation of a stronger statement is true. Since the speaker follows the maxim of quantity, we can suppose that a stronger statement is not supported in the speaker's information state. But if we use the negation operator to get an SI, as we have been doing so far, we run into problems with cases where a scalar term is embedded in a disjunction structure or a conditional. The reason is that a disjunction structure is weaker than each disjunct and a conditional is weaker than the antecedent or the consequent clause. In a globalistic analysis, when a disjunct or one clause of a condition contains a scalar term, the SI would be the negation of a disjunction structure or a conditional as a whole. Then the SI would become too strong. To avoid such problems, we need to start with a weaker negation than the ordinary negation, and it is motivated by the basic idea of how an SI is derived by the maxim of quantity. Therefore we need to go back to the basic idea of SIs.

The basic reasoning when we calculate an SI is this. We first make the weakest inference as to what the speaker's information state is like, considering (i) the pragmatic principles the speaker follows and (ii) what proposition could be asserted in such an information state. Therefore an SI should be primarily conditions on information states. On the other hand, we do not know what the speaker's information state is like. For this reason, we need to deal with a set of information states as candidates for the speaker's information state. Thus when we get an SI, the context is assumed to only include information states that satisfy the conditions, excluding the ones that do not satisfy the conditions. Suppose that ϕ is uttered in the current context, instead of a stronger one ψ. It results in a context that consists of information states that support ϕ, excluding those that support even ψ:

(47) a. $c + \phi = \{s \in c: \{s\} + \phi = \{s\}\} = c'$
b. $\{s \in c': \{s\} + \psi \neq \{s\}\} = \{s \in c: \{s\} + \phi = \{s\}, \{s\} + \psi \neq \{s\}\}$
$= (c + \phi) \backslash (c + \phi + \psi)$

The effect of eliminating information states that support ψ is different from the meaning of $\neg\psi$. If an information state does not support ψ, it means there is at least one possible world, but not

necessarily every possible world, in which ψ is false. But if a context is updated with $\neg\psi$, ψ is false in every possible world in an information state. The former gives us a weaker SI, like a primary implicature in Sauerland's analysis, while the latter gives a stronger SI, like a secondary implicature in Sauerland's terms.[15]

To get the weaker SI, we can define a new negation operator '$\sim$', the meaning of which is defined with respect to information states, not to possible worlds:

(48) $c + \sim\phi = c\backslash(c + \phi)$

I will call '$\sim$' the weak negation operator and $\neg$ the strong negation operator. The weak negation operator represents the operation of set difference between contexts. In a context, '$\sim\phi$' removes information states that support ϕ from a set of information states in the context. Thus the meaning of it can be paraphrased as 'it is not supported that ϕ', which is weaker than 'it is supported that $\neg\phi$'. The weak negation operator yields the minimal conditions on information states. The result is the weakest SI. This has an upside and a downside.

The upside is that the use of the weak negation operator can save the global approach to disjunction structures and conditionals. When neo-Griceans try to get an SI by applying the strong negation operator to a disjunction structure with one disjunct containing a scalar term, it leads to an unwanted result. Take (6) for example, assuming that the first disjunct makes ϕ_1, and the second ϕ_2, together with the rest of the statement, and that the stronger alternative to ϕ_2 is ϕ_2'. Then the meaning we get from the negation of the stronger alternative $\neg(\phi_1 \vee \phi_2')$ is that for every possible world in an information state, ϕ_1 and ϕ_2' are false:

(49) a. John spilt wine or broke some glasses.
ϕ = John spilt wine;
ψ = John broke some glasses;
ψ' = John broke all/many glasses

b. $(\phi_1 \vee \phi_2) \wedge\neg (\phi_1 \vee \phi_2')$
$= (\phi_1 \vee \phi_2) \wedge (\neg\phi_1 \wedge \neg\phi_2'))$
$= ((\phi_1 \wedge (\neg\phi_1\wedge\neg\phi_2')) \vee (\phi_2 \wedge (\neg\phi_1\wedge\neg\phi_2')))$
$= (\phi_2 \wedge (\neg\phi_1\wedge\neg\phi_2'))$
$= (\neg\phi_1\wedge\phi_2\wedge\neg\phi_2')$
$(\{s\in c$: for every $w\in s$, ϕ_2 is true but ϕ_1 and ϕ_2' are false$\})$

The problem with this result is that the proposition from the first disjunct is false in every possible world. The SI with the strong negation operator is too strong. This is not empirically supported. Sentence (49a) does not entail or implicate that John did not spill wine. It only implicates that John did not break all or many of the glasses.

[15] Geurts (2010) calls these a weak implicature and a strong implicature. He also needs a process of strengthening a weak implicature into a strong one. Grice and Sauerland assume that a speaker somehow knows/believes that a stronger alternative is true or false. From this assumption, a weaker implicature is strengthened into a stronger one.

If we apply the weak negation operator instead, we can avoid this bad result. Removing information states that support $(\phi_1 \vee \phi_2')$ does not have the effect of negating the proposition ϕ_1 from the first disjunct. This is clear when an information state is considered that supports $(\phi_1 \vee \phi_2)$, but not $(\phi_1 \vee \phi_2')$. Such an information state can include some ϕ_1-worlds: '$\sim(\phi_1 \vee \phi_2')$' does not necessarily eliminate an information state that includes ϕ_1-worlds. This is different from the result in (49b), where '$\neg(\phi_1 \vee \phi_2')$' eliminates information states with ϕ_1-worlds completely. This shows that the weak negation is the one we need in calculating an SI, when SIs are calculated globally.

(50) $c + (\phi_1 \vee \phi_2) + \sim(\phi_1 \vee \phi_2')$
$= (c + (\phi_1 \vee \phi_2))\backslash(c + (\phi_1 \vee \phi_2) + ((\phi_1 \vee \phi_2')))$
$= \{s \in c\colon \{s\} + (\phi_1 \vee \phi_2) = \{s\},\ \{s\} + ((\phi_1 \vee \phi_2') \neq \{s\}\}$

Let's consider conditionals. Assuming that $(\phi_1 \rightarrow \phi_2) = (\neg\phi_1 \vee \phi_2)$, the standard way of calculating a global SI leads to a context that consists of information states that support that $(\phi_1 \wedge \phi_2 \wedge \neg\phi_2')$.

(51) $(\phi_1 \rightarrow \phi_2) + \neg\ (\phi_1 \rightarrow \phi_2')$, where ϕ_2' is stronger than ϕ_2.
$= (\neg\phi_1 \vee \phi_2) \wedge \neg(\neg\phi_1 \vee \phi_2')$
$= (\neg\phi_1 \vee \phi_2) \wedge (\phi_1 \wedge \neg\phi_2')$
$= ((\neg\phi_1 \wedge (\phi_1 \wedge \neg\phi_2')) \vee (\phi_2 \wedge (\phi_1 \wedge \neg\phi_2')))$
$= (\phi_1 \wedge \phi_2 \wedge \neg\phi_2')$

This is not the right result, because in such information states the speaker would simply assert the conjunction of the antecedent and consequent clauses of the conditional. She would not need to use a conditional from the beginning. A sentence like *If John is bored*, *he will read some of the books on the desk* implies that John will not read all of the books, but not that John is bored. The SI with the strong negation operator is too strong.

If we apply the weak negation operator to a stronger alternative, we can avoid the wrong result of the standard way of calculating a global SI:

(52) $c + (\phi_1 \rightarrow \phi_2) + \sim(\phi_1 \rightarrow \phi_2')$, where ϕ_2' is stronger than ϕ_2.
$= (c + (\phi_1 \rightarrow \phi_2))\backslash(c + (\phi_1 \rightarrow \phi_2')))$
$= \{s \in c\colon \{s\} + (\neg\phi_1 \vee \phi_2) = \{s\},\ \{s\} + (\neg\phi_1 \vee \phi_2')) \neq \{s\}\}$

This results in a context that consists of information states which can include $\neg\phi_1$-worlds. This again shows that we need to use the weak negation operator '$\sim$' to calculate SIs.

The global SI with the weak negation operator allows us to avoid the problems with the standard globalistic approach of SIs. However, the SI with the weak negation operator is too weak to be an informative enough inference. This is a downside of using the weak negation operator. To get a significant inference, we need to strengthen the weak negation. But simply strengthening it to the strong negation $\neg$ leads to the problems discussed in (49) and (51). Therefore the weak negation operator can be strengthened to the strong negation only when it does not cause such problems as in (49) and (51).

But if strengthening the weak negation to the strong negation leads to problems discussed in (49) and (51), we need a new process that will solve the remaining problems in (50) and (52) before strengthening it. In (50), ϕ_1-worlds are not completely eliminated, but they are not ensured to stay in every information state in the context. If they do not remain in every information state, there is no reason to mention ϕ_1. In (52), $\neg\phi_1$-worlds are not completely excluded, but they may not stay in every information state in the context. If they do not, we do not need to use a conditional. When we try to strengthen the weak negation to the strong negation, it is part of the strengthening process to make sure those possible worlds remain in every information state in the context. This leads to the discussion of the constraint of contributiveness.

The contributiveness constraint is that in a sentence, every expression has to be contributive to the meaning of the sentence, except some expressions that are required only for grammatical reasons.[16] If an SI with strong negation makes some expression semantically non-contributive, it has to be excluded. As we will see, the contributiveness conditions are more robust than SIs. Therefore the SI applies only to some subsets of an information state so that the effect is not incompatible with the contributiveness conditions. This is what I am going to show in this section.

4.2 Disjunction structures

Even neo-Griceans do not take a globalistic approach in dealing with disjunction structures. Sauerland (2004) assumes that the scalar alternatives of *or* are {and, L, R, or}. If one disjunct includes a weaker scalar term, an SI from the scalar term becomes an SI of the whole sentence. Assuming that *all* is a stronger alternative to *some*, we can get an SI from a disjunct as the SI of the whole statement:

(49a) John spilt wine or broke some glasses. (= p)

(53) a. calAlt(p) = { John spilt wine and broke all glasses,
John spilt wine and broke some glasses,
John spilt wine, John broke all glasses,
John broke some glasses,
p}

b. K¬(John spilt wine and broke all glasses)
K¬(John spilt wine and broke some glasses)
K¬(John broke all glasses)

Among the three SIs in (53b), the last one is what we would get if the second disjunct were not embedded in the disjunction structure. The assumption of the operators *L* and *R* is not well

[16] In the sentences below, the expressions *of*, *be*, and *a* can be taken to be semantically empty, even if they can be theoretically interpreted otherwise:

(i) John can **be** proud **of** the work.
(ii) John is **a** doctor.

They are used to make the sentences grammatical, but they can be assumed to be semantically null.

motivated but it has the effect of projecting a local implicature from each disjunct into the main context.

We might think that the disjunction structure is transparent for the projection of an SI from a disjunct. However, it is not necessarily the case. We have already seen in (6) cases where a stronger alternative to a disjunct yields only a local SI. In this case, Sauerland simply says that no SI arises. But as I pointed out, there has to be some modification to the meaning of the second disjunct so that the two disjuncts do not overlap in their meanings. This has the effect of a local SI.

Moreover, Sauerland (2004) explains these observations based on the assumption that the operators *L* and *R* are alternatives to *and* and *or*. However, the assumption is not motivated in various respects. First, they are not witnessed as overt expressions in natural language. Second, scalar alternatives should have the same complexity, but expressions conjoined by *and* or *or* are more complex than each of the conjuncts/disjuncts. Therefore they are not plausible alternatives. Third, if 'A' or 'B' is a weaker alternative to 'A and B', then it would be possible to claim that *pretty L young* is an alternative to *pretty* (*and*) *young* and *pretty L smart* is an alternative to *pretty* (*and*) *smart*, etc. But the two conjunction structures with *L* are the same as *pretty*. Then is *pretty* an alternative to all conjunction structures with *pretty* as the left conjunct? If it were, then the use of *pretty* would yield myriad SIs, which is very strange.[17]

After having considered Sauerland's analysis of disjunction structures, we need to solve a couple of problems. First, we are assuming the weak negation operator in calculating SIs, but somehow we need to convert it into the strong negation to make an informative inference. Second, we do not assume the unmotivated operators *L* and *R*. In Sauerland's analysis, with the help of the two operators, when a scalar term is embedded in a disjunct of a disjunction structure, an SI from it behaves like the SI of the whole sentence. The question is how an SI in a embedded structure can be projected to the main context without the help of the operators *L* and *R*. The answer lies in the constraint of contributiveness.

The basic idea of the contributiveness constraint is that every expression has to have a meaning contribution to the meaning of the whole sentence.[18] Meaning contributions are captured by semantic entities, and they should be defined pragmatically with respect to an information state.

(54) a. Contributiveness constraint:
Every expression in a sentence ϕ contributes to the meaning of ϕ.

b. The meaning contribution of an expression α in a sentence ϕ in an information state s =
(i) $\downarrow (\{s\} + \phi) \setminus \downarrow (\{s\} + \phi[\emptyset_{\tau(\alpha)} / \alpha])$ or
(ii) $\downarrow (\{s\} + \phi[\emptyset_{\tau(\alpha)} / \alpha]) \setminus \downarrow (\{s\} + \phi)$
(Here $\phi[\emptyset_{\tau(\alpha)}$ is a null element corresponding to the type of α.)

---

[17] Alonso-Ovalle (2008) proposed using the Alternative Semantics to dispense with the two operators, without giving up the idea by Fox (2007).

[18] Chierchia et al. (2012) use Hurford's constraint, which says that in a disjunction structure, one disjunct cannot entail another. The constraint is not exactly formulated in that there are examples that do not satisfy the constraint. Moreover, it does not explain why it is necessary. And it only applies to disjunction structures. The contributiveness condition is more general and it explains why it is necessary.

c. An expression α in a sentence ϕ is **contributive** in an information state s iff the meaning contribution of α in s is not the empty set.

In an information state, the way one expression in a sentence contributes to the meaning of the whole sentence has to be defined with the effect of eliminating or adding possible worlds with respect to an information state.[19] The contribution of an expression α in a sentence ϕ can be calculated by comparing the original sentence with α with an alternative sentence $\phi[\emptyset_{\tau(\alpha)} / \alpha]$ we get by replacing α with its corresponding null element in ϕ.

This is intuitively clear, but the null element $\emptyset_{\tau(\alpha)}$ must be defined more precisely. In a disjunction structure, the null element of type <e, t> is a function from a set of individuals to 0 (false). In a conjunction structure, on the other hand, the null element for a conjunct of type <e, t> is a function from a set of individuals to 1 (true). One crucial thing is that a null element is something that makes the rest of the sentence determine the meaning of the whole sentence. Roughly, the null elements can be defined as follows:

(55)

semantic types	disjunction structure	conjunction structure
t	0	1
$\langle e, t\rangle$	λx . 0	λx . 1
$\langle e, \langle e, t\rangle\rangle$	$\lambda x \lambda y$. 0	$\lambda x \lambda y$. 1
$\langle\langle e, t\rangle, t\rangle$	λP . 0	λP . 1
...		

When an expression is contributive by eliminating some possible worlds, they are not witnessed in the resulting information states, because they are already eliminated from the information states. They can be separately calculated as $\downarrow (\{s\} + \phi[\emptyset_{\tau(\alpha)} / \alpha])\backslash \downarrow (\{s\} + \phi)$. On the other hand, when an expression is contributive by adding some possible worlds, they can be witnessed as $\downarrow (\{s\} + \phi)\backslash \downarrow (\{s\} + \phi[\emptyset_{\tau(\alpha)} / \alpha])$ in the resulting information states.

Adding a disjunct in a disjunction structure makes a statement weaker, and it leaves some possible worlds in an information state as the evidence of contributiveness:

(56) In a disjunction structure $(\phi \vee \psi)$, the meaning contribution of a disjunct ϕ in an information state $s = \downarrow (\{s\} + \phi)\backslash \downarrow (\{s\} + \psi)$
$= \{w \in s$: ϕ is true, but ψ is false, in $w\}$

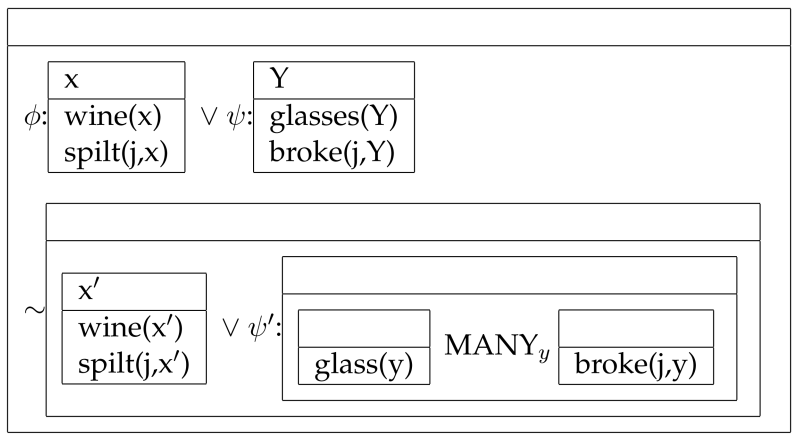
The meaning contribution of a disjunct ϕ in a disjunction statement $(\phi \vee \psi)$ is a set of possible worlds in s in which only ϕ is true. We can say the same thing about ψ.[20]

[19] This can be defined precisely in some cases but not in others. To calculate the contributiveness of an expression, we need to define a null element, which has the same semantic type as the expression but has no meaning contribution at all, and compare information states that support the original statement and the statement with that expression replaced with its corresponding null element.

[20] This is similar to Hurford's (1974) constraint that each disjunct cannot entail the other disjunct. But the contributiveness constraint is more general. It can apply to any sentences. If we say 'if John drinks, he drinks', the second *drinks* should have a different meaning from the first; otherwise it is not contributive at all. We will see how the constraint can apply to conditionals.

This will make sure that when a context is updated with a disjunction sentence ($\phi \vee \psi$), the resulting information states include ϕ-worlds and ψ-worlds. First, consider (49a). The composition of possible worlds in the result information state s' should satisfy the following DRS:

(57)



As I said in (50), the DRS with the weak negation opens the possibility that there are some ϕ-worlds in the context, but it does not ensure that every information state contains ϕ-worlds.

To ensure it, we need the contributiveness constraint, as given in (4.2). In a structure ($\phi \vee \psi$), the meaning contribution of ϕ is the following:

(58) Contributiveness conditions for ϕ and ψ in an result information state s:
  a. contributiveness of ϕ: $\downarrow (\{s\} + (\phi \vee \psi)) \setminus \downarrow (\{s\} + \psi)$
  $= \downarrow (\{s\} + \phi) \setminus \downarrow (\{s\} + \psi)$
  $= \downarrow (\{s\} + (\phi \wedge \neg\psi)) \neq \emptyset$
  b. (It is possible John spilt wine and broke no glasses.)
  contributiveness of ψ: $\downarrow (\{s\} + (\psi \wedge \neg\phi)) \neq \emptyset$
  (It is possible John broke some glasses but did not spill wine.)

The meaning contribution of ϕ in an information state is that there should be some ϕ-worlds in which ψ is false, in every information state in the context. Thus $\downarrow (\{s\} + (\phi \wedge \neg\psi)) \neq \emptyset$. The contributiveness constraint also requires ψ to be contributive. That is, $\downarrow (\{s\} + (\psi \wedge \neg\phi)) \neq \emptyset$.

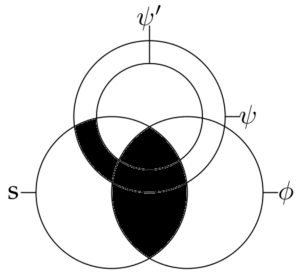
On the other hand, an information state in the context is not supposed to support ($\phi \vee \psi'$), which is required by the minimal SI. Together with the contributiveness conditions of ϕ and ψ, the effect of ψ' does not apply to a set of possible worlds for the contributiveness conditions for ϕ, but to a set of possible worlds for the contributiveness conditions for ψ. The weak SI can be strengthened to the strong one in the possible worlds for the contributiveness conditions for ψ. This leads to the condition that ψ' is false in possible worlds in which ϕ is false and ψ is true. That is, the SI from ψ' with the strong negation operator applies only to the possible worlds in which ϕ is false and ψ is true.

The contributiveness of ϕ is witnessed in $\neg\psi$-worlds, and in those worlds ψ' is false too. The contributiveness of ψ is witnessed in $\neg\phi$-worlds and in these worlds, ψ' is not supposed to be

supported. One thing which is unclear is whether ψ' is true or false in possible worlds in which ϕ and ψ are true. Such a result information state looks like one of the two diagrams:

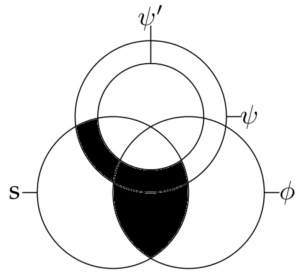
(59)

(i)



$(\phi \vee (\neg\psi' \wedge \psi))$

(ii)

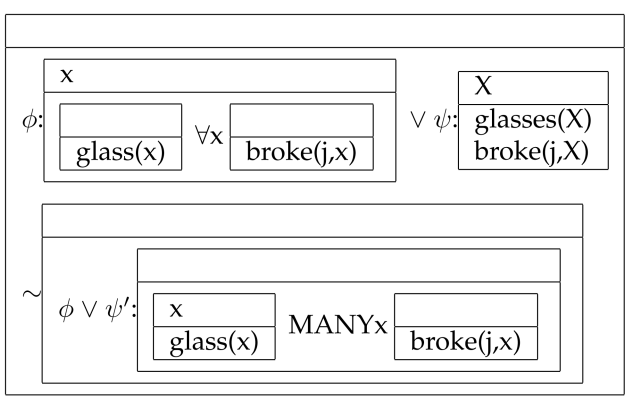


$((\phi \vee \psi) \wedge \neg\psi')$

In this case, the contributiveness condition does not play any role, and it is indeterminate whether the actual SI is local or not. To decide which should be the actual information, we need a separate strengthening process. In (59), (ii) is stronger than (i). The general strategy in making a pragmatic inference is to select the strongest possible reading, as given in (30). Therefore, if there is no other restriction, (ii) is preferred. Therefore we can conclude that $((\phi \wedge \psi) \wedge \neg\psi'$, if it is compatible with the rest of the assertion plus other inferences.

Now consider (6a), where one disjunct is stronger than the other. From the fact that $(\phi \vee \psi)$ is uttered in an information state s, instead of $(\phi \vee \psi')$, where ψ' is stronger than ψ, normally the contributiveness condition of ϕ in a result information state s' cannot be satisfied: the possible worlds in which ϕ is true but ψ is false is the empty set, because ϕ is stronger. In this context, the only way ϕ satisfies the contributiveness condition is to interpret ψ as $(\neg\phi \wedge \psi)$. This is the way the weaker disjunct is strengthened. Note that this is not the effect of SI, but the result of the contributiveness constraint.

(60)

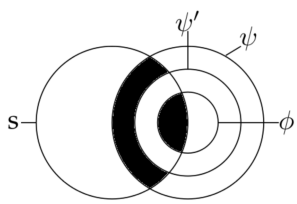


i. $\phi \Rightarrow \psi' \Rightarrow \psi$ (ordering in strengths of expressions)
ii. $\{s'\}+ (\phi \vee \psi) = \{s'\}$ (original statement)
iii. $\downarrow (\{s'\} + \phi) \setminus \downarrow (\{s'\} + \psi) \neq \emptyset$ (contributiveness of ϕ)

iv. $\downarrow (\{s'\} + \psi) \backslash \downarrow (\{s\} + \phi) \neq \emptyset$ (contributiveness of ψ)
v. ψ should be strengthened to $(\neg\phi \wedge \psi)$ (from (i), (iii))
vi. $\downarrow (\{s'\} + \phi \vee (\neg\phi \wedge \psi))) \backslash \downarrow (\{s'\} + (\neg\phi \wedge \psi))$
$= \downarrow (\{s'\} + \phi) \backslash \downarrow (\{s'\} + (\neg\phi \wedge \psi))$
$= \downarrow (\{s'\} + \phi) \neq \emptyset$ (contributiveness of ϕ with ψ revised)
vii. $(\{s'\} + (\phi \vee (\neg\phi \wedge \psi'))$
$= \{s'\} + \psi' \neq \{s'\}$ (weakest implicature with ψ' revised)
viii. $(\{s'\} + \neg\phi) + \psi' \,/\, (\{s'\} + \neg\phi)$

Assuming the revised interpretation of ψ, the meaning contribution of ϕ is $\downarrow (\{s\} + \phi)$. This ensures that ϕ is true in a non-empty set of possible worlds in which $(\neg\phi \wedge \psi)$ is false. Together with this condition, the weakest SI $\sim (\phi \vee \psi')$ has the effect that ψ' is not supported in possible worlds in which ϕ is false. The result information state should look like the following:

(61)



This means that John broke all or some but not all/many of the glasses.

4.3 Conditionals

The same procedures we applied to disjunction structures can also apply to conditionals. A conditional $(\phi \rightarrow \psi)$ is assumed to be the disjunction of the negation of the antecedent clause and the consequent clause. In conditionals in general, however, the antecedent clause and the consequent clause are closely related in a given context. And a conditional asserts that if the antecedent clause is satisfied, the consequent clause will hold. In this case, the negation of a stronger alternative of the consequent clause is likely to be part of the result of the situation in which the antecedent clause holds, as in (62). However, there is still a possibility that the negation of a stronger consequent clause holds, regardless of whether the antecedent clause holds or not, as in (63):

(62) If John throws a party, Sue or Mary will come.
$\not\rightsquigarrow$ Not both will come.
$\rightsquigarrow$ If John throws a party, not both will come.

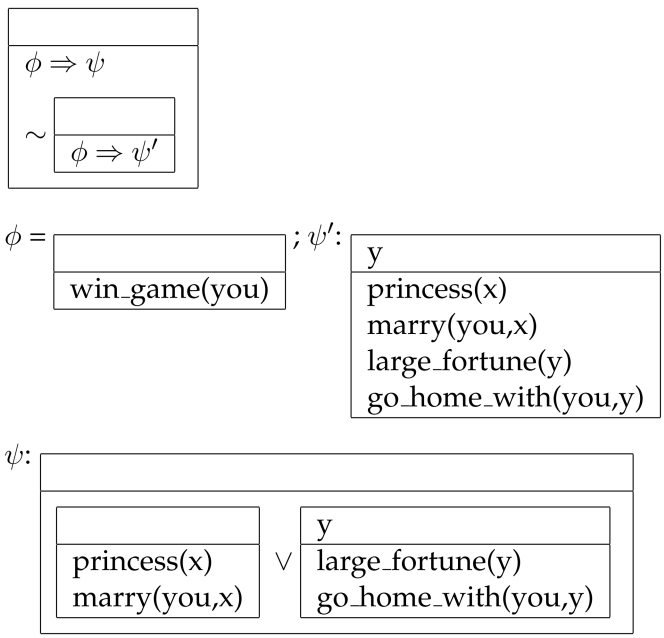
(63) If you win the game, you will marry the princess or go home with a large fortune.
$\rightsquigarrow$ You will not both marry the princess and go home with a large fortune.
$\not\rightsquigarrow$ If you win the game, you will not both marry the princess and go home with a large fortune.

In (63), the condition described by the antecedent clause is relevant to each of the two options described by the disjunction structure in the consequent clause. But it might not be related to the

stronger alternative to the consequent clause: that is, the proposition from the stronger alternative might not hold, regardless of whether the antecedent clause holds or not.

This can be explained in my DRT analysis easily. The disjunction DRS is embedded in the consequent clause of the conditional DRS. If the weakest SI DRS is embedded in the main DRS, we can get the following DRS:

(64)



In a conditional, an expression in the antecedent clause does not yield an SI because it introduces a non-monotonic context. Suppose that ($\phi \rightarrow \psi$) is uttered in an information state s, instead of ($\phi \rightarrow \psi'$), where ψ' is stronger than ψ. The meaning contribution of ϕ is the set of possible worlds added to s due to the addition of ϕ: a set of possible worlds in which only $\neg\phi$ is true (i.e. a set of possible worlds in which neither ϕ nor ψ is true).[21] The meaning contribution of ψ is a set of possible worlds added to s due to the addition of ψ, instead of its corresponding semantically

[21] The semantics of a conditional can be defined more precisely, reflecting the accessibility for anaphora and the property of being affected by the context. For anaphoric accessibility, a conditional sentence has to be interpreted as follows, as in Heim (1982):

(i) s + "if ϕ, then ψ" = $s\backslash((s + \phi)\backslash(s + \phi + \psi)) = s'$

That is, a pronoun in the antecedent clause can refer to something in the main context, and a pronoun in the consequent clause can refer to something in the main context or the antecedent clause. We can assume this rule, but it would lead to a more complex calculation. For the sake of convenience, I assume a propositional logic in which a conditional is equivalent to the disjunction of the negation of the antecedent clause and the consequent clause.

null element. If ψ is null, s would consist of $\neg\phi$-worlds. If the ψ is non-null, possible worlds in which ϕ and ψ are true are added.

(65) In a conditional $(\phi \rightarrow \psi)$ $(= (\neg\phi \vee \psi))$, $\{s\} + (\neg\phi \vee \psi) = \{s\}$.
i. the meaning contribution of ϕ in an information state s:
$\downarrow (\{s'\} + (\neg\phi \vee \psi))\backslash \downarrow (\{s'\} + \psi)$
$= \downarrow (\{s'\} + \neg\phi)\backslash \downarrow (\{s'\} + \psi)$
$= \downarrow (\{s'\} + (\neg\phi \wedge \neg\psi)) \neq \emptyset$
(It is possible that you don't win game and you don't marry the princess or go home with a large fortune.)
ii. the meaning contribution of ψ in an information state s:
$\downarrow (\{s'\} + (\neg\phi \vee \psi))\backslash \downarrow (\{s'\} + \neg\phi)$
$= \downarrow (\{s'\} + (\phi \wedge \psi)) \neq \emptyset$
(It is possible that you win the game and you marry the princess or go home with a large fortune.)

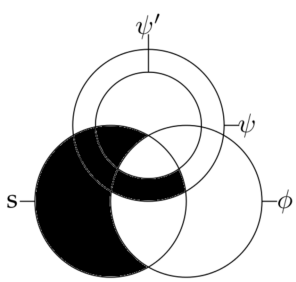
As I mentioned in the previous section, the SI DRS itself does not ensure that the possible worlds for the antecedent clause remain in the information states in the resulting context. But with the contributiveness conditions, every information state in the context contains some worlds in which ϕ and ψ are false and some worlds in which both are true.

The context is expected not to support $(\phi \rightarrow \psi')$. But considering the contributiveness of ϕ, the condition that ψ' is not supported does not apply to the possible worlds for the contributiveness conditions of ϕ, but in possible worlds for the contributiveness conditions for ψ, namely possible worlds in which ϕ and ψ are true, and in those possible worlds the weak SI can be strengthened. Thus ψ' is false in possible worlds in which ϕ and ψ are true. This leads to the meaning that it is possible that you win the game and you marry the princess or go home with a large fortune, but not both.

The contributiveness of ϕ is witnessed in possible worlds in which ψ is false, and in those possible worlds ψ' is false too. The contributiveness of ψ is witnessed in possible worlds in which ϕ is true, and in these possible worlds ψ' is not supported. One thing unclear is whether that condition from the SI applies to a set of possible worlds in which ϕ is false and ψ is true. This leads to one of the two possible result information states here:

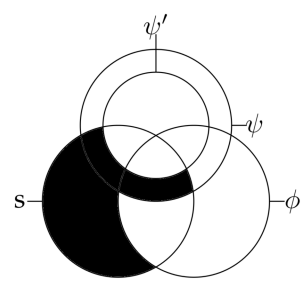
(66)

(i)



$(\phi \rightarrow (\neg\psi' \wedge \psi))$

(ii)



$((\phi \rightarrow \psi) \wedge \neg\psi')$

If the stronger alternative is still related to the antecedent clause, the resulting information state is like (i). But if it is not, then we get an information state like (ii). This can be explained with the help of the preference of the strongest meaning, which is given in (30). With no other restriction, (ii) is preferred to (i) because it is stronger. Thus we primarily get the meaning that $((\phi \rightarrow \psi) \wedge \neg\psi')$, unless it is not acceptable and the other meaning is required. This is why an SI in the consequent clause of a conditional is preferably projected.

I have claimed that an SI is primarily derived with the weak negation operator applying globally to a stronger alternative of the whole statement. One possible problem is that the weaker SI itself might make a segment DRS from an expression containing no scalar term non-contributive. The contributiveness constraint excludes such a case. It also determines whether the effect of a stronger alternative applies overall or only to a limited set of possible worlds in a sentence. The discussions show that disjunction structures and conditionals can be dealt with in the same way. Sauerland (2004) and other linguists like Fox (2007) and Chierchia et al. (2012) assume unmotivated operators *L* and *R*, but then they would have to assume similar operators for conditionals. In my analysis, no special operators are assumed, but the right results are predicted based on principles that are independently motivated.

Before closing this section, I need to mention that contributiveness conditions are stronger than SIs. We have seen cases where SIs are restricted by contributiveness conditions, but not vice versa. There is no case where an inference from the contributiveness condition is canceled by an SI. The reason is clear. Contributiveness conditions come from explicit expressions. If the speaker cancels them, he or she has no reason to use the expressions. On the other hand, SIs are inferences from what the speaker does not explicitly utter. If the speaker cancels them, it is not actually a cancellation but just a clarification. This is why SIs are easily canceled.

5. Conclusion

In this paper, the main claim is that SIs are always global. There are many cases where it looks as if we get local SIs. However, I claim that all SIs, even seemingly local SIs, have to be calculated as global procedures. I explain how the effects of local SIs obtain in the framework of DRT, which dispenses with local SIs altogether.

Calculating SIs is extremely complex but it starts with the use of the weak negation operator. It basically reflects the idea behind the maxim of quantity. It generates the weakest SI. The rest of the process is to strengthen the basic SI. The motivations of this strengthening process are the principle of the preference for the strongest meaning and the ordering of likelihood. The first one is necessary for successful communication because it tells the hearer which SI should be considered, and the second one tells the hearer whether he or she needs to find another SI.

Behind this analysis is the assumption that a scalar term can generate a set of SIs which are related to each other. When a stronger SI is rejected for some reason and a weaker SI is adopted as the actual SI, it has the effect of negating the stronger ones. This is quite natural because a weaker one is adopted when the stronger one is not accepted. This has the effect of strengthening the weaker one. All these processes work together to yield the strongest possible SI.

Getting a set of SIs has to be systematic. We discussed the following devices in the framework of DRT. This starts with the weakest SI with the weak negation operator.

- Strengthening an SI by removing some of the conditions in an SI
  - some variables in the weakest SI being bound by variables in the main DRS;
  - stripping some conditions in weakest implicatures when they are not relevant.
- Imposing the contributiveness conditions on the disjunctive conditions in an SI.
- When the weak negation operator is strengthened to the strong negation operator, the resulting SI applies only to a smaller set of possible worlds in which it does not conflict with the contributiveness conditions.

Including the weakest SI, a set of SIs we get by these processes constitutes a partial order. When a variable in an SI is bound by a variable in the main DRS, it enhances the relatedness of the SI to the context. In this respect, this is also a strengthening process. Stripping some of the conjunctive conditions may make the resulting SI stronger because the negation of a smaller number of conditions leads to a stronger SI. But when a scalar term is embedded in a semantic operator as in a disjunction structure or a conditional, removing a condition would make the SI weaker. Thus instead of removing some conditions, the contributiveness conditions are imposed, and the SI applies only to some domain in which it does not conflict with the contributiveness conditions. This allows the weak negation to be strengthened to the strong negation operator in the limited set of possible worlds.

The contributiveness condition can yield an inference different from an SI: it is stronger than an SI. The reason is that it comes from an overt expression. If an expression that is used in a sentence does not make any semantic contribution, there is no reason to use that expression. Thus an inference from the contributiveness constraint is not cancelable. On the other hand, an SI is an inference from what is not overtly expressed. Therefore an SI can be canceled easily.

References

Alonso-Ovalle, Luis. 2008. Innocent exclusion in an Alternative Semantics. *Natural Language Semantics* 16.2:115–128.

Beaver, David I. 1993. The kinematics of presupposition. Manuscript. Amsterdam: University of Amsterdam.

Chemla, Emmanuel, & Benjamin Spector. 2011. Experimental evidence for embedded scalar implicatures. *Journal of Semantics* 28.3:359–400.

Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. *Structures and Beyond*, ed. by Adriana Belletti, 39–103. Oxford & New York: Oxford University Press.

Chierchia, Gennaro, Danny Fox, & Benjamin Spector. 2012. Hurford's constraint and the theory of scalar implicatures: evidence for embedded implicatures. Manuscript.

Davis, Christopher. 2011. *Constraining Interpretation: Sentence Final Particles in Japanese.* Amherst: University of Massachusetts at Amherst dissertation.

Fine, Terrence L. 1973. *Theories of Probability: An Examination of Foundations*. New York: Academic Press.

Fox, Danny. 2007. Free choice disjunction and the theory of scalar implicatures. *Presupposition and Implicature in Compositional Semantics*, ed. by Uli Sauerland & Penka Stateva, 71–120. Basingstoke & New York: Palgrave Macmillan.

Gazdar, Gerald. 1979. *Pragmatics: Implicature*, *Presupposition*, *and Logical Form*. New York: Academic Press.

Geurts, Bart. 2010. *Quantity Implicatures*. Cambridge & New York: Cambridge University Press.

Grice, Paul. 1975. Logic and conversation. *Speech Acts*, ed. by Peter Cole & Jerry L. Morgan, 41–58. New York: Academic Press.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Amherst: University of Massachusetts at Amherst dissertation.

Hirschberg, Julia B. 1991. *A Theory of Scalar Implicature*. New York: Garland.

Horn, Laurence R. 1972. *On the Semantic Properties of Logical Operators in English*. Los Angeles: University of California at Los Angeles dissertation.

Horn, Laurence R. 1989. *A Natural History of Negation*. Chicago: University of Chicago Press.

Hurford, James R. 1974. Exclusive or inclusive disjunction. *Foundation of Language* 11.3:409–411.

Kadmon, Nirit. 1990. Uniqueness. *Linguistics and Philosophy* 13.3:273–324.

Kamp, Hans. 1981. A theory of truth and semantic representation. *Formal Methods in the Study of Language*, ed. by J. A. G. Groenendijk, T. M. V. Janssen & M. B. J. Stokhof, 277–322. Amsterdam: Mathematisch Centrum.

Kratzer, Angelika. 1991. Modality. *Semantics: An International Handbook of Contemporary Research*, ed. by Arnim von Stechow & Dieter Wunderlich, 639–650. Berlin: Walter de Gruyter.

Kratzer, Angelika. 2012. *Modals and Conditionals*. Oxford & New York: Oxford University Press.

Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25: 209–257.

Landman, Fred. 1998. Plurals and maximalization. *Events and Grammar*, ed. by Susan Rothstein, 237–271. Dordrecht: Kluwer.

Lassiter, Daniel. 2011. *Measurement and Modality: The Scalar Basis of Modal Semantics*. New York: New York University dissertation.

Lewis, David K. 1973. *Counterfactuals*. Oxford: Basil Blackwell.

van Rooij, Robert, & Katrin Schulz. 2004. Exhaustive interpretation of complex sentences. *Journal of Logic*, *Language*, *and Information* 13.4:491–519.

Rooth, Mats Edward. 1985. *Association with Focus*. Amherst: University of Massachusetts at Amherst dissertation.

Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27.3: 367–391.

Spector, Benjamin. 2006. *Aspects de la pragmatique des opérateurs logiques*. Paris: Université Paris 7 dissertation.

Spector, Benjamin. 2003. Scalar implicatures: exhaustivity and Gricean reasoning. *Proceedings of the European Summer School in Logic*, *Language and Information Student Session 8*, ed. by Balder ten Cate, 277–288. Vienna, Austria. http://users.soe.ucsc.edu/~btencate/esslli03/stus2003proc.pdf.

Stalnaker, Robert C. 1978. Assertion. *Syntax and Semantics*, Vol. 9: *Pragmatics*, ed. by Peter Cole, 315–332. New York: Academic Press.

Stalnaker, Robert C. 2002. Common ground. *Linguistics and Philosophy* 25.5–6:701–721.
van der Sandt, Rob A. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9.4:333–377.
Yeom, Jae-Il. 1998. *A Presuppositional Analysis of Specific Indefinites: Common Grounds as Structured Information State*. New York: Garland.

[Received 14 February 2014; revised 25 June 2014; accepted 23 July 2014]

English Language and Literature
Hongik University
94 Wausan-ro, Sangsu-dong, Mapo-gu
Seoul 121-791, Korea
jiyeom@hongik.ac.kr

篇章表徵理論之等級含意分析：整體的觀點

廉在一
弘益大學

過去文獻對於等級含意究竟來自部分或整體一直是爭議的議題，本文認為表面上看起來是來自部分結構的含意可由整體觀點之代詞約束或者貢獻約制取得，因此認為等級含意並未與句法結構相應。本文採用篇章表徵理論論證等級含意之整體的觀點，根據語用學量的原則，等級含意得到最小的意涵，而在言談過程中，有些變數受到約制，有些條件被刪除，等級含意之整體觀點涵蓋部分觀點，或者來自句子以外之意涵，最後得到的最大意涵既來自於貢獻約制或者可能性。

關鍵詞：語意隱涵，部分等級含意，整體等級含意，貢獻約制，語意可能性，弱否定