Determiners and Movement*

Kyle Johnson
University of Massachusetts at Amherst

September 2009

A commonplace treatment of determiners has it that they name relations between predicates of individuals. In (1), for example, every says that the set of students is included in the set of things that ran.

(1) Every student ran.

∀x [x is a student → x ran]

TP

λQ. ∀x [x is a student → Q(x) = 1] λx. x ran

DP

TP

∀x[P(x) = 1 → Q(x) = 1]

D

| every

λP. λQ.

| student

νP

ν | VP

T

NP

There are two kinds of modifications to this view that are required. One is made necessary by the discovery that predicates are not simple descriptions of individuals, but

*My thanks to Satoshi Tomioka, Jan Anderssen, Shoichi Takahashi, Angelika Kratzer, Emily Elfrner, Margaret Grant, Jesse Harris, Misato Hiraga, Wendell Kimper, Pasha Siraj, Martin Walkow, Yoko Hattori, Yumi Kawamoto and the participants of the conference.

1 See Barwise and Cooper (1981). I will give TP and vP the same denotations in this paper. This seriously misrepresents things, but it also greatly simplifies matters in a way that allows us to concentrate on the quantification introduced by determiners. I will also systematically obscure the independent contributions of the lexical verbs and voice, represented here with ν.
instead describe events.² Instead of (1) we require something along the lines of (2).

(2) Every student ran.

\[ \lambda e'. \forall x [x \text{ is a student in } e' \rightarrow x \text{ ran in } e'] \]

\[ \text{TP} \]

\[ \lambda Q. \lambda e'. \forall x [x \text{ is a student in } e' \rightarrow Q(x)(e') = 1] \]

\[ \lambda x. \lambda e. x \text{ ran in } e \]

\[ \text{TP} \]

\[ \lambda P. \lambda Q. \lambda e'. \forall x [P(x)(e') = 1 \rightarrow Q(x)(e') = 1] \]

\[ \lambda x. \lambda e. x \text{ is a student in } e \]

\[ \text{NP} \]

\[ \text{T} \]

\[ \nu P \]

\[ \text{every} \]

\[ \text{student} \]

\[ \text{ran} \]

\[ \text{D} \]

\[ \text{NP} \]

\[ \text{vP} \]

\[ \text{VP} \]

The other comes from the study of reconstruction effects, in which a moved DP seems to be interpreted in two positions. (3) illustrates such an example. Under treatments of constituent questions that claim that the interrogative phrase is semantically interpreted in its surface position,³ the disjoint reference effect between John and he in (3) indicates that the moved phrase is interpreted in the position it moved from as well.

(3) * Whose kissing John₁ does he₁ object to?

Chomsky (1993) suggested that these effects arise because movement produces “copies” of the phrase being moved and that certain of these copies are interpreted semantically. On this view, the representation that (3) would get is (4), and the principles of semantics require that all of the lower copy be interpreted.


³ And this is indicated by the ability of examples such as They asked which pictures of each other she liked to meet the locality requirements that hold between reciprocals and their antecedents. See Barss (1986) and Fox and Nissenbaum (2004).
I adopt here the convention of indicating that a copy is unspoken by putting it into a shaded font. The disjoint reference effect arises in (3), then, because one copy of John is c-commanded by he in (4) and this is the configuration that triggers disjoint reference effects. This “copy theory” of movement appears to be the most successful account of such phenomena, so I will adopt it.

But it raises some questions. There are two I will try to answer in this paper, and they are:

(5) Why is only one of the copies pronounced?

(6) How are the two copies interpreted so that one binds the other?

The present answer to (6) is in Fox (1999, 2002, 2003) (see also Sauerland (1998)) and it has the consequences for the treatment of determiners that I am interested in. It is expressed in terms of a semantics like that in (1) which treats determiners as relations between predicates of individuals. I’ll begin by reviewing Fox’s proposal, modifying it slightly so that we get an answer to (5). Then I will translate it into an event-based semantics that makes use of an approach to quantification in Elbourne (2005). The
result is a view that splits up the meaning of determiners, and in this respect has elements in common with the papers in Szabolcsi (1997), and also Williams (1986, 1988), Beghelli (1993, 1995), Sportiche (2003), Hallman (2000), Butler (2004), and Kratzer (2005). It’s different in its details, however, and I think better equipped to make sense of the “apparent” resumptive pronouns discussed in Aoun, Choueiri and Hornstein (2001) (see section 5) as well as the cases in which copies are claimed to be doubled. We’ll begin by considering how to interpret copies so that reconstruction effects are captured.

1 Trace Conversion

Another context where the sort of reconstruction displayed in (3) is found are cases of Quantifier Raising, like that in (7).

(7) * A different student told her\( _{1} \) every story about Diana\( _{1} \)’s parents.

The every story DP can have the subject in its c-command domain. We know that because the subject in (7) can get an interpretation that’s only available when different is c-commanded by the universally quantified DP, as (8) shows.

(8) a. Every woman talked to a different student.
   b. * [Her visit to every woman] disturbed a different student.
   c. * A different student cried after every woman left yesterday.

But at the same time, the every story-DP must be within the c-command domain of her; that we know because there is a disjoint reference effect between Diana and her. We must therefore let movement create a representation for (7) like that in (9).
As in (3), this representation places one copy of Diana within the c-command domain of her, and therefore correctly produces the environment for disjoint reference. The only difference between (3) and (9) is which copy gets pronounced. In (9) it is the lower copy, while in (3) it is the higher one. Throughout the rest of this paper, I will focus on cases like (9), ones in which Quantifier Raising occurs.

A feature of this representation that will be important for what follows is that the higher, unspoken, copy contains the NP that we see in the lower, spoken, copy. On some accounts of these cases, the higher copy contains only the quantifier and not the NP. I will therefore report the argument from Fox (2002) for this feature of the proposal.

Fox’s argument comes from a phenomenon discovered by Fiengo and May (1994). There are situations where the disjoint reference effect that (7) illustrates are overcome. If the name is within a relative clause that is forced by ellipsis to be interpreted outside the phrase that contains the coreferent pronoun, as in (10), then the disjoint reference effect is modulated.

(10) ?I told her₁ every story that Diana₁ asked me to Δ.

Δ = tell her x
The ellipsis in (10) appears to lie within the VP that serves as its antecedent, and this produces well-known problems. Thus, both the disjoint reference effects and the presence of ellipsis indicate that the copy theory of movement cannot have its normal outcome in this case:

(11)

There is a parallel phenomenon in cases of movement that form questions. Unlike (3), where a name inside a moved *wh*-phrase is interpreted in its lower position, examples where a name is within a relative clause in a moved *wh*-phrase do not trigger parallel disjoint reference effects; (12) is such a case.

(12) Which story that Diana told does she now regret?

Without modification, the copy theory of movement would also wrongly give this sentence a representation in which a *Diana* falls within the c-command domain of *she* and a disjoint reference effect should consequently arise.
The solution to both cases is to allow for a derivation that involves the following steps.

(14) Late Merger
   a. Build the D+NP phrase that will move
   b. Move that D+NP
   c. Build and attach the relative clause to the higher copy only.

This is David Lebeaux's solution,⁴ and it is adopted by Fox. These derivations would give to (10) and (12) the representations in (15).

⁴ See Lebeaux (1988), and also Freidin (1986).
(15) a. TP
   └── DP
       ├── I
       │    └── T
       │       └── vP
       │           └── every
       └── vP
           └── VP
               └── V
                   └── told
                       └── DP
                           └── her
                               └── X
                                   └── DP
                                       └── every
                                           └── story

TP
   └── DP
       └── D
           └── NP
               └── CP
                   └── that Diana
                       └── asked me to
                           └── ∆
As (15) indicates, this account claims that in (10) the relative clause is not in the spoken copy of the object, but is instead part of the higher, unspoken, copy. It’s this part of the account that serves as evidence that the higher, unspoken, copy in such cases contains the NP part of the object. That is necessary because under standard assumptions, a restrictive relative clause of this kind must attach to an NP to produce the right meaning. Fox (2002) produces a variety of arguments that the relative clause in such examples is indeed not in the spoken copy of the DP.⁵ One of these is based on the contrast in (16), from Tiedeman (1995).

(16)  a. * I said that everyone you did $\triangle$ arrived.
     $\triangle = \text{say that } x \text{ arrived}$

b. I said that everyone arrived that you did $\triangle$.
     $\triangle = \text{say that } x \text{ arrived}$

(Fox 2002, (35b), (36b): 77)

Just as this proposal predicts, the relative clause containing an elided VP cannot be spoken within the antecedent VP: that’s what makes (16a) ungrammatical. Instead,

⁵ See also Baltin (1987), and Fox and Nissenbaum (1999) for arguments that “Extraposition from NP” has the syntax in (15a).
that relative clause must be spoken in a position outside the antecedent VP and, more particularly, as part of the material that determines the scope of the quantificational DP the relative clause modifies. That's what's happened in (16b), which has the representation in (17).

For this account to be complete, it requires an explanation for when late merger derivations are possible and when they are not, for otherwise all of the effects gained by the copy theory of movement will be lost. See Takahashi (2006) for many steps in this direction.⁶

We're now ready to see Fox's proposal for interpreting copies. He suggests that movement creates not only copies of the phrase that is moved, but also that it appends the same index on all of them. He then devises the following rule for interpreting structures with movement indices in them.⁷

(18) Trace Conversion

---

6 The account reviewed here for Extraposition from NP is argued in Fox and Nissenbaum (1999) to only arise when the term extraposed is an adjunct. When a PP or clause that is an argument of the NP appears in extraposed position, they argue that the clause itself has moved.

7 See also Sauerland (2004) for an examination of alternative ways of formulating this rule.
interpret $\phi$ as a function that maps an individual, $x$, to the meaning of $\phi[x/n]$. $\phi[x/n]$ is the result of replacing the head of every constituent with the index $n$ in $\phi$ with the head $\text{the}_x$, whose interpretation, $\llbracket\text{the}_x\rrbracket$, is: $\lambda P. \llbracket P \cap \lambda y. y = x\rrbracket$.

We can think of this rule as having two parts. One part expresses the standard method of interpreting a binder. It gives the phrase that the binder combines with the denotation of a lambda-abstract. That part of the rule could be expressed with (19).

(19) In $\phi'$, change the denotation of $\phi$ to $\lambda n. \phi$.

The other part rewrites the meaning of the copies inside $\phi$. The new meaning is one that is like a definite description, but with a variable bound by the lambda-operator introduced in the other part of the rule. It has the effect of turning a lower copy of “$\llbracket\text{DP}_x\text{ every story}\rrbracket$,” for example, into:

(20) $\llbracket\text{DP}_x\text{ every story}\rrbracket \rightarrow \text{the story that is } 3$.

With Trace Conversion, a simple example like (21) will get an interpretation like that shown.

---

8 See, for example, Heim and Kratzer (1998).
(21) She told every story.
\[ \forall x [x \text{ is a story} \rightarrow \llbracket \text{she} \] told \llbracket \text{the} [\lambda z. \ z \text{ is a story} \cap \lambda y. y = x] ] \]\n
A clumsy paraphrase for the denotation of this sentence is:

(22) For all x, if x is a story, then she told the thing that is a story and x.

Trace Conversion turns determiners of lower copies into kinds of restricted variables, then. It works for all of the cases we’ve viewed so far, and it arguably extends correctly to all cases in which a DP has moved. As noted at the outset, it’s built upon a view of determiner quantification that has them relate properties of individuals. It still requires translation into a framework in which predicates describe events. But before we do that, let’s take a closer look at it and the notion of “copies” that it relies on.
2 Trace Conversion

The second part of Fox’s Trace Conversion has the undesirable property of letting a whole class of lexical items be ambiguous. In the cases we’re examining, those lexical items are determiners. It resolves that ambiguity by syntactic rule. It claims that the meanings of determiners are not fixed, but change according to their position. It gives syntax the power to change lexical content. That’s more than syntax should be allowed. So let me offer a variant of Fox’s proposal that avoids these consequences.

This variant builds on ideas many have had about the syntax of quantification. Imagine, as in Matthewson (2001), that quantified nominals make use of two functional heads. One has the denotation of quantifiers, and the other is something with the meaning given to determiners by Fox’s Trace Conversion rule. Unlike Matthewson, but like those cited above, let’s separate these two functional heads, putting the term that expresses the quantification in the position where its scope is computed, while the definite determiner part is in construction with the np. The morphological form of the definite determiner varies depending on the quantificational term. Let’s follow Kratzer (2005) and Adger and Ramchand (2005) and let this dependency be mediated by agree. Agree will determine the morphological form of the term in the lower position, and make both heads share an index. On this view, then, there is only one determiner — the one that Fox’s Trace Conversion creates — with a morphological form that is fixed by agreeing with a silent quantifier.

(23) The only (quantificational) determiner is [the_x]. Its morphological form is determined by the silent Q it agrees with.

This proposal would give to (24) the derivation indicated. (I will use “∀” to represent the silent universal quantifier that agrees with every.)

(24) She told every story.

a. 

b. 

On this view then, there is no rule of Trace Conversion, and we can rely on the plainer binding rule in (19) to express configurations created by movement. The part that Fox credited to a special rule that re-writes the denotation of a quantificational
determiner is now taken up by a more complex syntactic theory about how quantificational determiners are represented. It requires one part of quantificational determiners to stand in the scope position, and the other to stand in the variable position, where Fox’s rule created a new denotation for a determiner. This isn’t a small rethinking. It requires a complete overhaul of the syntax, and semantics, of quantificational DPs. We need to understand, among other things, the conditions that determine how far apart the two parts can be, and how the rules of Spell Out determine which component is pronounced and how. I can’t do that overhaul here, and not just for reasons of space. Nonetheless, I hope it strikes the reader as an appealing alternative to Fox’s rule.

3 Multidominance

Let’s now consider what the copies are in the copy theory of movement. There are two criteria that any successful characterization of copies must meet.

(25)  a. Only one copy should be able to be pronounced.
     b. Every copy must be identical.

One proposal that achieves both these goals is (26).

(26) Copies of $\alpha$ are one and the same $\alpha$ in different syntactic positions.

A simple implementation of this idea is to let phrase-markers allow for multidominance, i.e., relax the requirement that a term have no more than one mother (see Nunes (2001), Starke (2001), Frampton (2004) and Fitzpatrick and Groat (2005) for recent proposals along these lines, and Citko (2005) for an application of the idea to across-the-board movement). This would give to (24) the representation in (27).
Determiners and Movement

(27)

An example involving late merger, such as (15a), will get a representation like (28).  

10 Heidi Harley asks by what derivation a representation like that in (28) could be manufactured. If we assume that the basic tree building operation is MERGE, as in Chomsky (1995), then we seek a derivation that involves nothing more than bringing two terms together to form a larger term. I suggest that (28) is achieved by way of a derivation in which at one point story has two mothers: one containing every as well as story, and another that contains the relative clause as well as story. It’s this second phrase that will merge with $\forall$ to form the QP in the higher position. The derivation I’ve just described is in all relevant respects equivalent to derivations that involve “Sideward Movement” in Nunes (2001, 2004). The difference between Nunes’s representations and mine are that the terms which MERGE forms under Nunes’s scheme are sets, whereas they are mereologies in my representations.
These representations ensure that every copy has exactly the same material in it, and that the semantic contribution it makes is precisely the same in each position.

It also provides a way of deriving that only one copy may be pronounced, along the lines described in Nunes (1995, 1996, 1999). The procedure that maps syntactic structures onto strings will (sometimes) get conflicting information from phrases that have more than one mother, and this leads to conflicts that Nunes argues are resolved by failing to linearize the copy in all of its positions. For instance, consider how a representation like (29) might get linearized.

Nunes does not have representations that involve multidomiance. Instead, copies are represented as more than one phrase occupying different positions. He then stipulates that the linearization algorithm treats these phrases as one and the same item, with the consequences that will be described in a moment. I am merely removing this stipulation through a multidominant definition of "copy." The technique of using ill-formed linearizations that result from multidominant representations to produce some effect has a precedent in Citko (2005), which examines cases of across-the-board movement. In her work, multidominant relationships arise in coordinations of certain sorts, and the linearization problems they create are alleviated by movement.
Nunes adopts the linearization procedure in Kayne (1994) which involves certain complexities that are better dodged here. So, for illustrative purposes, I'll adopt an algorithm that incorporates some of the cyclic nature of procedures described in Epstein and Seely (2002) and Fox and Pesetsky (2004), and expresses linear information in terms of strings.

(30) Let \( y \) be a node in a phrase-marker projected from \( \alpha \), and \( \beta \) be an immediate daughter of \( y \). Map each \( y \) to a string that is formed from the strings associated with \( \alpha \) and \( \beta \), so that:

a. If \( y \) is an \( X^0 \), then the string associated with \( y \) is the terminal \( y \) dominates, and

b. If \( y \) has only one daughter, \( \alpha \), then the string associated with \( y \) is the string associated with \( \alpha \), and

c. The string associated with \( y \) is formed by concatenating the string associated with \( \beta \) to the left of the string associated with \( \alpha \) when \( \beta \) is in Specifier of \( y \), otherwise
d. The string associated with \( \gamma \) is formed by concatenating the string associated with \( \alpha \) to the left of the string associated with \( \beta \).

A proper understanding of (30) would require unpacking what is meant by "specifier." This is one of the foundational problems for linearization theories, and not usefully engaged here. I’ll stipulate that the DPs \( I \) and \( her \) are in Specifier positions in (29).

When (30) interprets (29), it will require that \( story \) precede \( yesterday \) (because \( VP^* = \alpha \) and \( AdvP = \beta \) for \( VP \)) and it will require that \( yesterday \) precede \( story \) (because \( TP^* = \alpha \) and \( QP = \beta \) for \( TP \)). This, Nunes assumes, is an illicit result as it disobeys the well-formedness constraint on linearizations in (31).

(31) Consistency
If \( \alpha \) precedes \( \beta \), then \( \beta \) cannot precede \( \alpha \).

Representations that have multiply dominated phrases in them will (often) violate (31) in just the way that (30) does. If these phrase markers are to have licit linearizations, (30) must be modified.

As it stands, (30) achieves the goal of making sure that every terminal in a phrase-marker is part of the resulting string by requiring that the string associated with every daughter phrase is part of the string associated with the phrase immediately dominating. If that requirement were replaced with something else that ensured every terminal gets into the linearization, it would be possible to avoid producing an output that violates Consistency in situations where something is multiply dominated. Lifting that requirement would allow the linearization scheme to ignore those phrases that lead to violations of Consistency. That will also have the effect of (usually)\(^{12} \) ensuring that a multiply dominated phrase be linearized in only one of the positions it occupies.

I’ll demonstrate that by adding to the set of well-formedness constraints on a linearization something that ensures that every terminal finds itself in the resulting string (= (32)) and changing (30) to (33).

(32) Totality
Every terminal in a root node, \( \rho \), must be present in the string associated with \( \rho \).

(33) Let \( \gamma \) be a node in a phrase-marker projected from \( \alpha \), and \( \beta \) be an immediate daughter of \( \gamma \). Map each \( \gamma \) to a string that is formed from the strings associated with \( \alpha \) and \( \beta \), so that:

\(^{12} \) In situations where a phrase’s two mothers put it in positions that are linearly adjacent (as in cases of “string vacuous movement”), linearizing that phrase in both positions does not produce a violation of Consistency.
Determiners and Movement

a. The string associated with \( \gamma \) is the terminal \( \gamma \) immediately dominates, or
b. The string associated with \( \gamma \) is the string associated with one of its immediate daughters, or
c. The string associated with \( \gamma \) is formed by concatenating the string associated with \( \beta \) to the left of the string associated with \( \alpha \) when \( \beta \) is in Specifier of \( \gamma \), otherwise
d. The string associated with \( \gamma \) is formed by concatenating the string associated with \( \alpha \) to the left of the string associated with \( \beta \).

Consider how this revised linearization procedure will apply to (29). In order to meet Totality, enough phrases will have to contribute the strings they are associated with to the strings of phrases dominating them for every terminal to end up with a position. But in order to meet Consistency, not every phrase can submit the string it is associated with to the strings associated with phrases that dominate them. There are just two ways that (33) can meet those goals. In one, the string associated with \([\text{NP story}]\) is made part of the string associated with \([\text{DP every story}]\) but not the string associated with \([\text{NP story that Diana asked me to}]\). On that linearization, (29) gets mapped onto the string in (34).

\[(34) \quad \text{I told her every story yesterday that Diana asked me to.}\]

In the other linearization, the string associated with \([\text{NP story}]\) is made part of the string associated with \([\text{NP story that Diana asked me to}]\) but not the string associated with \([\text{DP every story}]\). On that linearization, (29) gets mapped onto the string in (35).

\[(35) \quad \text{I told her yesterday every story that Diana asked me to.}\]

In this example, both linearizations arise. In other cases, we will want to allow only one of the feasible linearizations. In simple constituent questions in English, for instance, a multidominant representation would allow for two linearizations parallel to the ones we’ve seen for (29): one where the moved phrase is linearized according to its higher position and another where the moved phrase is linearized according to its lower position. English, however, only permits the linearization that follows the higher position, i.e., wh-movement is “overt” in English. Whatever is responsible for this effect will not be part of what I offer. The syntactic representations I am suggesting we adopt will provide several positions in which moved phrases can be spelled out, but which of these positions are actually available will have to come from factors independent of the syntax of movement.

This, then, is the syntax I suggest we adopt for movement. It requires disassembling determiners into two parts that provide the effects of Fox’s Trace Conversion rule. And it requires modeling copies as phrases in more than one position. This second assumption can be wedded to a linearization scheme that derives the fact that
moved phrases get pronounced in only of their positions, along lines sketched by Nunes. We are now ready to turn to the task of beefing up the semantics of determiners so that they involve events.

4 Events

There is now considerable evidence that predicates like *ran* do not describe properties of individuals, as in (36), but instead relate individuals to events, as in (37).

(36) \[[\text{she}] \text{ told } [\text{it}]\]
Determiners and Movement

(37) \( \lambda e. \llbracket \text{she} \rrbracket \) told \( \llbracket \text{it} \rrbracket \) in e

\[
\begin{array}{c}
\text{TP} \\
\text{DP} \\
\lambda y. \lambda e. \text{y} \text{told} \llbracket \text{it} \rrbracket \text{ in e} \\
\Delta \\
\text{she} \\
\text{T} \\
\nuP \\
\nuP \\
\nu \\
\lambda y. \lambda e. \text{y} \text{told} \llbracket \text{it} \rrbracket \text{ in e} \\
\lambda x. \lambda y. \lambda e. \text{y} \text{told x in e} \\
\text{VP} \\
\text{DP} \\
\lambda x. \lambda y. \lambda e. \text{y} \text{told x in e} \\
\text{VP} \\
\lambda y. \lambda e. \text{y} \text{told} \llbracket \text{it} \rrbracket \text{ in e} \\
\lambda x. \lambda y. \lambda e. \text{y} \text{told x in e} \\
\text{V} \\
\text{T} \\
\text{told}
\end{array}
\]

On this view, sentences refer to the events in which the relations named by their predicates hold of the arguments those predicates combine with. The account sketched in the previous two sections is built upon a non-event view of sentence meanings and won't work with an event based semantics as it stands.

It's not hard to modify what we've got so that it fits a event-based semantics. Elbourne (2005) incorporates Fox's Trace Conversion rule to his event-based study of quantification. He provides a comprehensive semantics for definite determiners that allows them to have the very meaning that Fox suggests as a special case. On his view, as on Fox's, the movement operation will involve not only producing copies — multidominated NPs, if I am correct — but also cause the index on the lower copy to be the same as the index on the higher copy. The semantics associated with the Trace Conversion rule will then apply giving the correct interpretations.

I want to explore a different possibility here. I will use Elbourne (2005) as my model, but focus on his account of definite descriptions that, though not bound, nonetheless behave like variables related to an antecedent. One example of this sort is (38).

(38) Every man who owns a donkey beats the donkey.

In (38), the donkey is anaphoric on the donkeys described in the subject: the donkey each man beats is the donkey that he owns. This can be achieved by letting the universal quantifier associated with every quantify over the events that man who owns
a donkey and beats the donkey describe. Simplifying somewhat, Elbourne’s proposal is that $\forall$ has the denotation in (39).

$\forall = \lambda f_{<e,<_t,_,>}, \lambda g_{<e,<_t,_,>} . \lambda s . f(x) (s') = 1$, there is a minimal $s''$ such that $s' \leq s'' \leq s$ and $g(x)(s'') = 1$.

(compare: Elbourne 2005, section 2.2.4)

The variables $s, s' \& s''$ range over “situations,” which we can equate with events, and “$\leq$” is the reflexive part-of relation. (39) says that $\forall$ takes two relations between individuals and situations (qua events), $f$ and $g$, and describes those situations, $s$, in which, for every $x$, all of the smallest sub-situations of $s$ that contain $x$ and make $f$ true are part of the smallest sub-situation of $s$ that contains $x$ and makes $g$ true. What this will do in the case of (38) is give it a meaning that can be paraphrased with (40).

(40) There is a situation, $s$, such that for every $x$ and every minimal situation $s'$ in $s$ such that $x$ is a man who owns a donkey in $s'$, there is a larger minimal situation, $s''$ in $s$ that contains $s'$ such that $x$ beats the donkey in $s''$.

A looser, but perhaps more revealing, paraphrase is (41).

(41) For every $x$, all of the minimal situations of a man, $x$, owning a donkey are part of a larger minimal situation in which $x$ beats the donkey.

The notion of “minimal situation” does the work of getting the donkeys that are beaten to be the donkeys that are owned in (38). Think of a situation as being made up of individuals, relations and properties. A minimal situation can be informally described as one which is made up of only those individuals, relations and properties necessary to make some proposition true. (See Berman (1987), Schein (1993), Elbourne (2005) and especially Kratzer (1989, 1990, 2002, pear).)

(42) A situation, $s$, is a minimal situation in which $P(s) = 1$ iff there is no $s' < s$ such that $P(s') = 1$.

Therefore, a situation that makes “a man, $x$, owns a donkey” true will be a minimal situation just in case it contains only that man and one donkey and the ‘own’ relation between them. A situation that makes “$x$ beats the donkey” true will be a minimal situation just in case it contains only $x$ and the donkey and the ‘beats’ relation between them. If the ‘owns’ situation is a part of the ‘beats’ situation, then, because they are each allowed only one donkey, the donkey in both situations will be the same.

This treatment of (38) extends to cases like (43).

(43) Every man who owns a donkey beats it.

The *it* in (43) is anaphoric to a *donkey* in the very same way that the *donkey* is in (38). Elbourne (2005) argues that this is because *it* is, in fact, the *donkey* with the NP containing *donkey* elided. His proposal, following Postal (1969), is that pronouns are how definite determiners are pronounced when the NP they are in construction with is elided. The transformations in (44) are all on a par.

(44)  
\begin{align*}
  & \text{a. She saw some books and he read some books. } \rightarrow \text{ She saw some books and he read some } \Delta. \\
  & \text{b. She wrote no books and he read no books. } \rightarrow \text{ She wrote no books and he read none } \Delta. \\
  & \text{c. She wrote the book and he read the book } \rightarrow \text{ She wrote the book and he read it } \Delta. \\
\end{align*}

Thus, (43) is actually (45), where *it* has the meaning of *the*, and the semantics has the same consequences for (45) that it does for (38).

(45)  
Every man who owns a donkey beats \[ [\text{DP it } [\text{NP donkey}]] \]

The work done by minimal situations in guaranteeing that the *donkey* in (38), or *it* in (43), are anaphoric to the descriptive content of the preceding quantified expression mimics the work done by indices in Fox’s Trace Conversion rule. In translating Fox’s system into an event based semantics, I propose that we dispense with the indices in his rule and exploit the minimal situation technique. Because Elbourne’s system is built on quantificational expressions binding indexed variables, I will have to change his denotation for $\forall$ as well. I propose changing it to (46).

(46)  
\[ [\forall] = \lambda f_{<e,<s,t>}. \lambda g_{<s,t>}. \lambda s. \text{ for every minimal } s' \text{ such that } s' \leq s \text{ and } \exists x f(x)(s') = 1, \text{ there is a minimal } s'' \text{ such that } s' \leq s'' \leq s \text{ and } g(s'') = 1. \]

This assumes that NPs are functions from individuals to predicates of situations (or events).

(47)  
\[ [\text{story}] = \lambda x. \lambda s. \text{ } x \text{ is a story in } s. \]

Not only will this proposal dispense with the indices in Fox’s Trace Conversion rule, it will dispense with Fox’s Trace Conversion rule entirely. Here’s an illustration.\(^{14}\)

My proposal has certain features in common with the “choice function” version of the Trace Conversion rule that Sauerland (2004) argues for. On Sauerland’s view, the definite determiner in the lower copy is a variable over choice functions that is quantified over by the higher copy. As he shows in Sauerland (1998), this approach gives an interesting explanation of certain crossover phenomena. My proposal does not obviously extend to these cases.
(48) \( \lambda s. \) for every minimal \( s' \) such that \( s' \leq s \) and \( s' \) is a story, there is a minimal \( s'' \) such that \( s' \leq s'' \leq s \) and \( \llbracket \text{she} \rrbracket \) told \( \llbracket \text{the story} \rrbracket \) in \( s'' \)

\[
\begin{align*}
\lambda s. \llbracket \text{she} \rrbracket & \text{ told } \llbracket \text{the story} \rrbracket \text{ in } s \\
\lambda g. \lambda s. \ & \text{ for every minimal } s' \text{ such that } s' \leq s \text{ and there is a story in } s', \text{ there is a minimal } s'' \text{ such that } s' \leq s'' \leq s \text{ and } g(s'') = 1
\end{align*}
\]

My proposal, then, is that lower copies are plain definite descriptions — or the pronoun version of them — which share an NP with a higher quantificational expression. The connection between the higher quantifier and the lower definite description comes by way of quantifying over situations.

Under this proposal, then, “traces” of QR are kinds of donkey pronouns.\(^{15}\) They are entirely parallel to the pronoun in (45), but whereas the pronoun in (45) is missing

\(^{15}\) Compare Boeckx (2003), who attempts to reduce donkey anaphora to movement.
its NP by way of deletion, traces share their NPs with a higher quantificational term. This achieves Chomsky (1993)’s goal of reducing the movement relation entirely to structure building; it does so in a way that is close in spirit to Starke (2001). It raises a raft of questions, of course, including how to model the familiar island conditions characteristic of movement, as well as how to distinguish pronominal binding, donkey anaphora, and traces.

5 Open questions

The account given in section 3 for why only one copy is pronounced is not complete. To see this, consider a case of wh-movement in which, unlike the cases of quantifier raising I have concentrated on, the moved phrase is allowed to be linearized in only one of its two positions. A case like (49), for example, gets a representation like (50) under the proposals here.

(49) Which story did she tell?

(50)

\[ \begin{array}{c}
\text{CP} \\
\text{DP} \quad \text{CP} \\
D \quad C \quad TP \\
\text{which} \\
did \\
\text{TP} \\
\text{she} \\
\text{T} \\
\text{VP} \\
\text{V} \\
\text{tell} \\
\text{D} \\
\text{the} \\
\text{NP} \\
\text{story} \\
\end{array} \]

The lower copy of the wh-phrase is not pronounced in (50) for the following two reasons. First: the shared NP can be pronounced in only one of its positions, and something independent of our story determines in this case that it will be the higher position. Second: the determiner in this example is silent. For both these reasons, then, the lower, “trace,” position of wh-movement has nothing in it that is pronounced.
But these two reasons are only accidentally combined. It does not seem accidental that the position in which the NP cannot be pronounced is always the position in which the determiner happens to be silent. My proposal does not connect these two facts.

This is a problem. But I am encouraged by the existence of cases that are just what would be expected if the determiner were pronounced in cases like (50). If Elbourne’s analysis of pronouns as determiners is correct, we should expect a determiner that is in construction with a silent NP to have the form of a pronoun. Therefore if (50) were to arise with an overt determiner, it should look like a resumptive pronoun but with the reconstruction effects characteristic of movement. As Elbourne (2005, section 3.5.3) notes, this is just what we find. Aoun et al. (2001) report such cases for Lebanese Arabic. An illustrative example from their paper is (51).

(51) tâlmiiz-a₁ lkósleen ma baddna nyâţabî [wala m’âllme]₁ ḏonno huwwë
      student-her₁ the-bad NEG want.1P tell.1P [no teacher]₁ that he
      zaţbar b-l-fâhîs
      cheated.3SM in-the-exam

‘her bad student, we don’t want to tell any teacher that he cheated on the exam.’

(Aoun et al. 2001, (25b): 381)

The phrase ‘tâlmiiz-a₁ lkósleen’ (her bad student) contains a pronoun that is bound to a lower quantifier (‘wala m’âllme’ (no teacher)). This is made possible by the phrase ‘tâlmiiz-a₁ lkósleen’ (her bad student) being related to the resumptive pronoun ‘huwwë’ (he) through movement. These cases, then, suggest that the two ingredients my proposal uses to make “traces” silent do, indeed, sometimes come separately. If these two ingredients are truly independent, or are interdependent in some way, is an open question.

Another open question, and one that is very likely to harbor problems for the proposal made here, concerns the class of things that move. All the cases examined in this paper involve movement of DPs, and the account is especially designed for the reconstruction effects that arise when DPs move. But it appears that many other sorts of things can move: heads of certain sorts, PPs, VPs, APs, and, perhaps, clauses. The reconstruction effects that arise when these other sorts of things move are frequently different from what is found with DPs. Reconstruction effects are largely manufactured by the syntactic representations movement generates, if the conclusions reached here are correct, and yet it’s syntactic representations that are also responsible for the fact that a moved term can be pronounced in only one of its positions. For the most part, this latter constraint on movement holds irrespective of the kind of thing that moves, and yet the reconstruction effects vary considerably. That’s a puzzle that gets no easier as a result of this paper. Indeed, because my re-
doing of Fox’s Trace Conversion rule would only apply to cases where a determiner is involved in what moves, there is no straightforward way of extending this system to phrases of other types.

6 Conclusion

We have looked at a narrow range of cases, but if the conclusions reached here can be extended to all examples of movement, then we have the following results.

(52) Results

- There is only one determiner, pronounced many ways, and it is *the*.
- Indices are not part of what relates a moved phrase with its “trace.”
- (Phrasal) Movement arises when one term has more than one position in a phrase marker, and nothing more.
References


Determiners and Movement


