

RELATIVE CLAUSE DELETION

CHRIS COLLINS
New York University

1 Introduction

The goal of this paper is to investigate the ambiguity of sentences like (1):

(1) At the party, I saw three boys who I know and one girl.

On one interpretation, I propose (1) has the following structure:

(2) At the party, I saw three boys who I know and one girl <who I know>

The angled brackets <...> indicate that the enclosed string is present syntactically but not pronounced. I will call the process by which the relative clause in (2) is not pronounced *relative clause deletion*.

The interpretation of (1) with the structure in (2) can be brought out by the following situation (situation A). Suppose you go to a party, and there are twenty people: ten boys and ten girls. Furthermore, you see all twenty people. Suppose out of the twenty people, you know three boys and one girl. So in particular, you see ten girls, but you only know one of them. Then (1) with structure (2) is true in that situation.

Sentence (1) also has another structure where *one girl* is not modified by a covert relative clause. Consider situation B. Suppose you go to a party with 11 people: ten boys and one girl. Suppose that you see all eleven people. You happen to know three of the boys. Furthermore, you do not know the girl. Then (1) is also true in this situation. But in this case, the structure is different from (2). Rather, when (1) is true in situation B, there is no covert relative clause modifying *one girl*.

Crucially, (1) with structure (2) is false in situation B (since you do not know the girl in situation B). And (1) with no covert relative is false in situation A (since you see more than one girl). So this makes it easy to distinguish the two structures of (1) truth conditionally.

2 Comparatives

The example in (1) involves a coordinate structure. Another common source of deleted relatives is comparatives.

(3) At the party, there were more girls who I know than boys.

This sentence is ambiguous in much the same way as (1). On one interpretation it is equivalent to:

(4) At the party, there were more girls who I know than there were boys who I know.

On another interpretation, (3) is equivalent to:

(5) At the party, there were more girls who I know than there were boys (known or unknown).

Such comparatives with deleted relatives are easy to find on the internet:

(6) In eastern European countries, there are more boys who smoke than girls.
(<http://www.uv.es/lisis/sofia/buelgahandbook.pdf>)

(7) Cats are high on the kill list because there are many more cats who need homes than dogs.
(<http://www.sonoma-county-gazette.com/cms/pages/sonoma-col-arg1-columns-arg2-Pet%20and%20Animal%20Care.html>)

(8) There are as many boys who play with the dolls as girls.
(<http://fahnlanders.blogspot.com/>)

(9) “Even in Europe, there are many fewer boys who study ballet than girls,” Alberto observes.
(<http://balletnebraska.org/2012/a-dancers-journey-alberto-liberatoscioli/>)

In another kind of comparative structure, the head of the relative clause is not the compared element. In this case as well it is possible to find internet examples (all of which have the relevant interpretation for me):

(10) Girls who've experienced parental separation are more likely than boys to turn to the Non-Productive coping strategy of self-blame to deal with uncontrollable family stressors.
(<https://www.google.com/search?tbo=p&tbm=bks&q=isbn:0203938704>)

(11) ...and that girls who dropout are more likely than boys to return.
(<http://books.google.com/books?id=esksAAAAIAAJ>)

- (12) My figures suggested that girls who are dyslexic or dyspraxic are much more likely than boys to escape being diagnosed at school.
(<https://www.google.com/search?tbo=p&tbm=bks&q=isbn:0203857968>)
- (13) Men who are widowed are more likely than women to remarry.
(http://www.nap411.com/index.php?option=com_content&view=article&id=54&Itemid=168)
- (14) Children who are exposed to iodine-131 are more likely than adults to get cancer later in life.
(<http://newamericamedia.org/2011/03/are-there-health-risks-in-us-from-japans-nuclear-disaster.php>)
- (15) In fact, according to the Environmental Protection Agency, a child who is exposed to a carcinogen is *10 times* more likely to develop cancer than an adult.
(<http://www.babble.com/mom/home-sweet-home-a-toxic-world-for-baby/>)

I assume that the analysis of (13) involves relative clause deletion as illustrated in (16):

- (16) Men who are widowed are more likely to remarry than women <who are widowed>
<are likely to remarry>.

3 Identity and Parallelism

In all cases so far, the deleted relative clause is syntactically identical to its antecedent (see van Craenenbroeck, Jeroen and Jason Merchant 2013 for a recent discussion of the issue of syntactic vs. semantic identity conditions on ellipsis).

(17) Syntactic Identity

A relative clause R is deleted under syntactic identity with an antecedent relative clause.

Syntactic identity accounts for the following contrast:

- (18) a. There are more women who are married than men.
b. There are more women who have a husband than men.

Given relative clause deletion, (18a) can mean that there are more married women than married men. However, (18b) does not have the interpretation “than men who have a wife”, even though that is a plausible DP to compare with “women who have a husband”. Such an interpretation for (18b) is ruled out by the identity condition in (17), since the relative clause “who have a husband” is not identical to “who have a wife”. Rather, (18b) either means (i) that there are more women who have a husband than men all together (no relative clause deletion), or (ii) that there are more women who have a husband than men who have a husband (relative clause deletion).

In addition to (17), there also seems to be a parallelism condition to the effect that the deleted relative clause and its antecedent must be found in the same syntactic environment:

(19) Parallelism

Relative clause deletion can only take place in the following structure:

[_{XP1}.....[Head1 Antecedent].....] and/than [_{XP2}.....[Head2 <Relative Clause>]]
 where Head2 is focused and XP1 is a member of F(XP2), the focus value of XP2.

The two XPs in (19) differ in the NP heads of their respective relative clauses (including modifiers): Head1 and Head2, where Head2 (and perhaps also Head1) is focused. I understand F(XP), the focus value of XP, as in Fox 1999 (see section 7 below).

Given (19), one issue to investigate is whether Head1 and/or Head2 are marked prosodically when there is relative clause deletion. In other words, in a sentence like (1) are the two interpretations distinguished prosodically? Is *one girl* stressed in a particular way when the relative clause is deleted?

When the parallelism condition is not met, relative clause deletion is not generally possible with indefinite determiners:

(20) One boy who I know was talking to three girls.
 (no relative clause deletion interpretation)

(21) Some boy who is in my class was talking to three girls on the steps.
 (no relative clause deletion interpretation)

(22) Boys who drop out usually blame it on girls.
 (no relative clause deletion interpretation).

But much more work is needed to nail down the specific conditions on when relative clause deletion is possible.

4 Deletion of other Modifiers

While the focus of this paper is on relative clause deletion, it is clear that other modifiers can be deleted in the same way. Consider the following example:

(23) Some boy from every school reads comic books, and some girl does too.

The first conjunct in (23) can have an inverse linking interpretation (see May 1977) where for each school, there is some boy who reads comic books. Furthermore, if the first conjunct has an inverse linking interpretation, then so does the second conjunct. The presence of an inverse linking interpretation for the second conjunct suggests the following underlying representation:

(24) ...and some girl <from every school> does <read comic books> too.

I have not systematically investigated deletion of prenominal adjectives, but preliminary data shows that it may sometimes be possible:

- (25) a. There were more tall girls than boys in my class.
 b. Tall girls are more likely than boys to play basketball.
 c. I saw three tall girls and one boy.
- (26) a. I taught more foreign girls than boys.
 b. Foreign men are more likely to remarry than women.
 c. At the party, I saw one foreign boy and two girls.

The modifier deletion interpretation is possible in the (a) sentences, difficult in the (b) sentences and impossible in the (c) sentences. I have no account of this variation.

5 Definite Determiners

In (1) the relevant DPs were both indefinites. Relative clause deletion is also possible with definite DPs in examples like the following:

- (27) a. I saw the boy who I know, and the girl too.
 b. I saw the one boy who I know, and the three girls too.
 c. I saw the boy who I met in Paris, and the girl too.
 d. I saw the boy who I met in Paris, but not the girl.

(27d) is derived from (28) by deleting everything following *but* other than the negation and *the girl*:

- (28) I saw the boy who I met in Paris but I didn't see the girl who I met in Paris.

Relative clause deletion is also possible with definite DPs in comparatives. The following example can have the interpretation "more likely to remarry than women who were widowed."

- (29) The men who were widowed were more likely to remarry than the women.

In some cases involving a definite determiner, it is unclear whether there is relative clause deletion, or a semantic process of domain restriction (on the issue of domain restriction with definite DPs see Heim 2011). Consider (30):

- (30) I met a boy and a girl in Paris. The girl was a teacher.

In this sentence, the DP *the girl* in (30) seems to have a contextually specified domain narrowing of the set of all girls (the denotation of *girl*) to the set containing just the girl who I met in Paris. One possibility is that this domain narrowing is a semantic process distinct from the syntactic process of relative clause deletion. Domain narrowing would be purely semantic, involving no deleted relative clause. If the domain narrowing approach to (30) were correct, we

would have to question whether relative clause deletion were involved at all in any of (27) and (29), and perhaps even in the cases with an indefinite determiner.

Another possibility is that (30) also involves relative clause deletion. So that the underlying structure of (30) is really:

(31) I met a boy and a girl in Paris. The girl <who I met in Paris> was a teacher.

On the relative clause deletion analysis of (31), the deleted relative clause does not have a syntactically identical antecedent, since there is no previous relative clause of the form [who I met in Paris]. For the same reason, the deletion in (31) would violate the parallelism condition (19).

A related fact is that relative clause deletion readings are much easier with definite determiners in sentences like (20-22) above:

(32) The one boy who I know was talking to the three girls.

This sentence easily admits the following interpretation:

(33) The one boy who I know was talking to the three girls who I know.

Faced with these data, one option would be to conclude that there is a semantic process of domain narrowing (distinct from the syntactic process of relative clause deletion) found with definite DPs. Alternatively, it may be the case that the conditions in (17) and (19) do not hold for relative clause deletion with definite DPs. While I am unable to resolve this issue here, the following section gives independent syntactic evidence that relative clause deletion is possible in certain cases with definite DPs.

6 Reconstruction

In this section, I provide syntactic evidence for relative clause deletion. In the following example, it appears that the reflexive in *picture of himself in the woods* is bound by *John*.

(34) The picture of himself in the woods that John took is prettier than the picture of himself on the boat.

We can understand this in terms of the copy theory of movement, as illustrated in (35a). I assume that OP is the covert equivalent of the relative pronoun *which*. Furthermore, as demonstrated in section 9, the CP complement of *the* is dominated by another projection RelP (into which the NP *picture of himself* raises), which I leave out here for simplicity. As shown in (35b), the restriction of the higher copy is deleted, and the OP of the lower copy is deleted and replaced by a variable (see Chomsky 1995: 202-212). Fox 2003 claims that there is syntactic rule (“Trace Conversion”) which converts the lower copy into a definite description. I will not pursue this issue here.

(35) a. the [_{CP} [OP picture of himself] that John₁ took [OP picture of himself₁]]

- b. the [_{CP} OP_x that John₁ took [x picture of himself₁]]

In the LF representation of (35b), the lower copy contains the reflexive, which is bound by the subject. But then by the same reasoning, the NP *picture of himself on the boat* in (34) must also be modified by a covert relative clause allowing for reconstruction.

A similar example can be constructed based on bound variable pronouns:

- (36) I saw the picture of his father that every boy took,
but not the picture of his mother.

In (36) the universal quantifier is the subject of a finite relative clause. Furthermore, the pronoun in *picture of his father* is interpreted as a variable bound by *every boy*. Lastly, there is an entailment that I saw several pictures, one for each boy. Such an entailment suggests that *every boy* undergoes QR out of the finite relative clause (see Hulseley and Sauerland 2006: 132). But QR of *every boy* over the head of the relative clause *picture of his father* would yield a WCO violation. Because of this, Hulseley and Sauerland 2005: 134 propose that such examples involve reconstruction of the head of the relative clause. The resulting structure can be represented as in (37):

- (37) [[every boy]_y [I saw the [_{CP} OP_x [that y took [x picture of y's father]]]]]

But then, given that the pronoun in *picture of his mother* in (36) is interpreted as a bound variable, by parity of reasoning *picture of his mother* should also undergo reconstruction into a relative clause (out of which QR has taken place). Such a relative clause is made available by relative clause deletion: *the picture of his mother* <*that every boy took*>.

7 Strict and Sloppy Readings

An example of the distinction between strict and sloppy interpretations of a pronoun in an elided VP is given in (38) (see Ross 1986[1967]: 207 who first noted the distinction):

- (38) John loves his mother and Bill does too.
Sloppy: Bill loves Bill's mother.
Strict: Bill loves John's mother.

In the case of relative clause deletion, the following examples illustrate sloppy interpretations:

- (39) a. I met three freshmen who liked their professors and two sophomores.
b. I met more freshmen who like their professors than sophomores.
c. Freshmen who like their professors are more likely to attend class than sophomores.

All three sentences in (39) admit relative clause deletion with a sloppy interpretation of the deleted pronoun. For example, the relevant set of sophomores in (39b) are those who like their own professors.

Other examples show strict interpretations:

(40) I like the picture of John that was on his mother's fridge, but not the picture of Bill.

The most natural interpretation of (40) is the following:

(41) I like the picture of John that was on John's mother's fridge, but
I do not like the picture of Bill that was on John's mother's fridge.

Recall the identity condition on relative clause deletion in (17). With the exception of the data in section 5 involving definites, such a condition handles all the data given in the paper so far. For example, (39b) will have the following analysis:

(42) I met more freshmen who_1 like $their_1$ professors than sophomores $\langle who_1$ like $their_1$ professors \rangle .

Since the second relative clause is syntactically identical to the first, it can be deleted. The indices in (42) are not interpreted in terms of coreference, but rather variable binding (who_1 binds $their_1$ syntactically and $there_1$ is interpreted as a bound variable). So there is no implication that $their_1$ in the first relative clause refers to the same person as $their_1$ in the second relative clause. See Sag 1977[1976]: 74, 86-102 for a related treatment of the identity condition on VP deletion.

Consider now an example where the antecedent of the pronoun is not contained within the deleted relative clause. In this example, John and Bill are wildlife biologists who tag wild animals and do follow up investigations of the animals they tag. I report on their activities by saying:

(43) John spotted one lion that he tagged last week, and Bill two zebras.

On one interpretation, this can be true if both John and Bill spotted many lions and zebras, but only a few that they actually tagged. On that reading, there is relative clause deletion, represented as follows (the verb *spotted* is deleted by gapping):

(44) $John_1$ spotted one lion that he_1 tagged last week,
and $Bill_2$ \langle spotted \rangle two zebras \langle that he_2 tagged last week \rangle .

Crucially the index of the pronoun in the deleted relative is different from the index of pronoun in the non-deleted relative (since their antecedents $John_1$ and $Bill_2$ have different indices). The question then is whether (44) violates the syntactic identity condition in (17). There are a number of possible ways to deal with sentences like (44) that have been proposed in the literature. First, Fiengo and May (1994: 95) propose the syntactic notion of i-copy, where the dependency in the first clause is identical to the dependency in the second clause, even though the indices are different. Second, Rooth (1992) proposes to simply ignore indices for the purposes of the syntactic relation of "reconstruction" ("pronominal indices may vary"), constraining the interpretation of the deleted VP by a semantic condition on focus. Third, Fox

(1999) proposes to eliminate the syntactic identity condition all together, and to force syntactic identity through a parallelism condition on focus. On all of three of these approaches, the non-identity of the indices in (44) would not block deletion.

Consider Fox (1999) in more detail. According to Fox's parallelism condition, the antecedent (S1 = [John₁ spotted one lion that he₁ tagged last week]) has to be in the focus value of the second clause S2. The focus value of S2 is defined as the set of sentences which are alternatives to S2: $F(S2) = \{S: \exists x \exists y [S = x \text{ spotted } y \text{ that } x \text{ tagged last week}]\}$, where x ranges over DPs that are alternatives to *Bill* and y ranges over NPs that are alternatives to *two zebras*. S1 is an element of $F(S2)$, since S1 is of the syntactic form [x spotted y that x tagged last week] (assuming *John₁* and *he₁* both count as x). Therefore, parallelism is satisfied. A consequence of this approach is that indices of pronouns which are not bound by focused DPs have to match exactly (since if there were mismatched indices, S1 would not be in $F(S2)$). In other words, both syntactic identity, and the exceptions to it (as in the mismatched indices of (44)) follow from Fox's parallelism condition.

As noted, Fox argues against a syntactic identity condition, subsuming it under the parallelism condition on focus. However, syntactic identity is implicit in a number of places in his system (e.g., in the definition focus values, see footnotes 3 and 4 and the parallelism condition). But the question of whether or not deleted structures are syntactically identical to their antecedents is non-trivial (see van Craenenbroek and Merchant 2013 for an extensive survey of the issues). And different answers to the syntactic/semantic identity question could lead to different formulations of the parallelism condition, and vice versa. Therefore, I will continue to assume that there is a distinct syntactic identity condition (17), and that it should be interpreted as allowing mismatched pronominal indices only when the relevant pronouns are bound by focused DPs (allowing (44)).

8 Chomsky 1965 on Deletion

Now consider:

- (45) a. At the reception, I met three boys who liked their professors and one girl.
b. At the reception, I met one boy who liked his professor and one girl.

Both of these sentences have sloppy interpretations, even though the deleted pronoun is not identical to the pronoun in the antecedent. For example, (45b) is represented as follows:

- (46) ...one boy who₁ liked his₁ professor and one girl <who₁ liked her₁ professor>

The problem that this sentence poses for condition (17) is that the two pronouns do not have the same ϕ -features values, and hence it looks like the relative clauses are not syntactically identical. The general issue was first noted by Ross (1986[1967]: 207), who defined the identity condition needed for deletion: "Constituents are identical if they have the same constituent structure and are identical morpheme-for-morpheme, or if they differ only as to pronouns, where the pronouns in each of the identical constituents are commanded by antecedents in the nonidentical portions of the phrase-marker."

I propose to analyze facts such as (45) (and other cases of deletion where pronouns differ in phi-feature values) by applying a principle about deletion proposed by Chomsky 1965: 179 for the analysis of French examples like the following:

- (47) ces hommes sont plus intelligent-s que Marie
 these men are more intelligent-PL than Mary
 “These men are more intelligent than Mary.”

Chomsky notes: “...the deletion of the Adjective of the embedded sentence should be blocked, since it differs from the Adjective of the matrix sentence in gender and number.” From such sentences, Chomsky draws the conclusion that “...the features added to a formative by agreement transformations are not part of the formative in the same sense as those which are inherent to it or as those which it assumes as it enters a Phrase-marker.... in the case of Adjectives and the copula (also Verbs, which take part in similar rules) the inflectional features that are added by agreement transformations are apparently not considered in determining whether the item in question is strictly identical with some other item.”

Whereas Chomsky’s principle only concerns inflectional phi-features on the copula and on adjectives, in (46) the relevant phi-features are those of pronouns. In order to apply Chomsky’s principle to (46), I assume following Collins and Postal 2012: 155, that the phi-features of pronouns are derived by agreement with an antecedent. Under that approach, at the point of relative clause deletion, the two pronouns in (46) do not differ in phi-features, since they do not have any phi-features at all.

In light of the analysis of pronouns given in Collins and Postal 2012, consider the following sentence:

- (48) John met the boy who knows you, not the girl.

Collins and Postal assume that the antecedent of 2nd person pronouns is a null DP in the left periphery that they call ADDRESSEE. Similarly, the antecedent of 1st person pronouns is a null DP in the left periphery that they call AUTHOR. The phi-feature values of 1st and 2nd person pronouns are derived by agreement with these indexical DPs. If phi-features are derived by agreement, then what prevents the following representation?

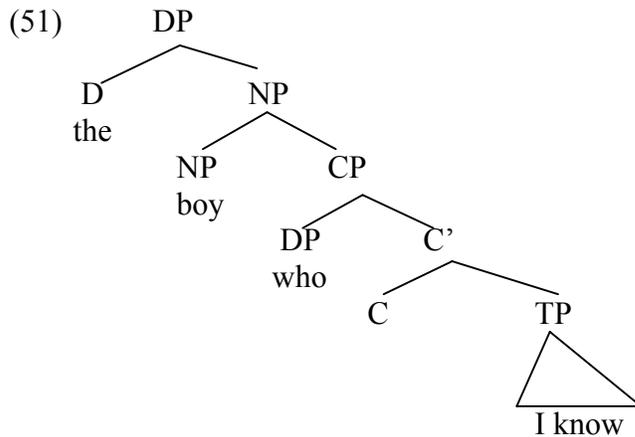
- (49) John met the boy who knows you₁, not the girl <who knows me₂>

The pronouns *you* and *me* in (49) are not co-indexed, and hence the deleted relative clause has no syntactically identical antecedent. Therefore, (49) violates (17) and (19).

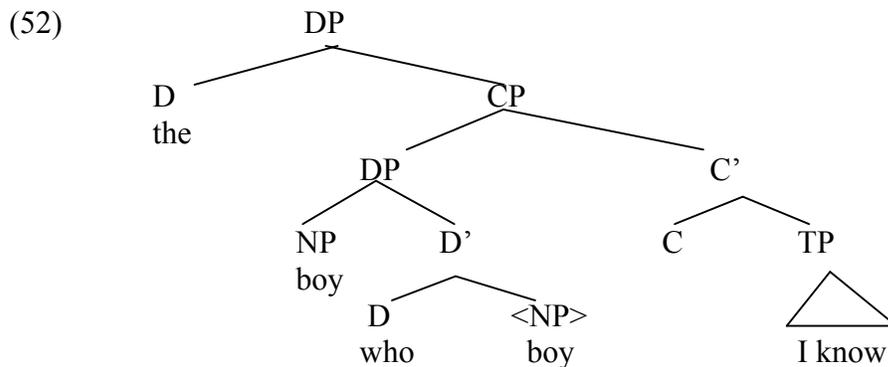
9 Structure of Relative Clauses

Lastly, consider the consequences of relative clause deletion for the structure of relative clauses. On the traditional approach, a relative clause modifies a NP. The structure of the relative clause in (50) is given in (51):

- (50) I saw the boy who I know.

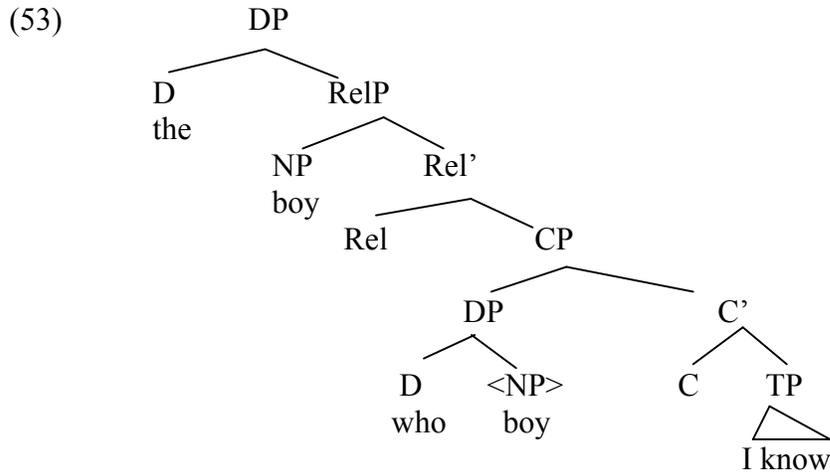


On this analysis, relative clause deletion is simply deletion of the CP. However, formulation of relative clause deletion within the head raising analysis of Kayne 1994 runs into a problem. The structure of the relative clause in (50) according to Kayne 1994:90 is the following:



In this structure, the Spec of CP is a DP, headed by the D *who*. The NP moves from the complement of that DP to the specifier. The problem for (52), from the point of view of relative clause deletion, is that the string *who I know* is not even a constituent. Combining this with the assumption that only constituents can delete, structure (52) wrongly predicts that there should be no relative clause deletion.

I propose a slightly different version of the head raising analysis, which allows for relative clause deletion. In this analysis, the NP *boy* moves into Spec RelP.



In this structure, CP can be deleted to the exclusion of the raised noun. I will not be able to justify this structure any further here.

10 Conclusion

In this paper, I have investigated relative clause deletion and illustrated the consequences it has for the structure of relative clauses (section 9).

It remains to revisit other areas in the syntax of relative clauses and DPs from the perspective of relative clause deletion: stacked relative clauses, amount relative clauses, infinitival relative clauses, non-restrictive relative clauses, relative clauses with pied-piping, reduced relative clauses, other post-nominal modifiers, pre-nominal modifiers of various sorts, etc.

Comparative work on relative clause deletion will also be of great interest. For example, do prenominal relative clauses undergo deletion? The “relative deletion” discussed by van Craenenbroek and Lipták (2006) is different from the relative clause deletion discussed in this paper. Rather, van Craenenbroek and Liptak claim that the “relative deletion” is a kind of sluicing where the IP of a relative clause is deleted. In the data discussed in this paper, the whole relative clause is deleted. Whether or not relative clause deletion, in the sense of this paper, is available in any other language has yet to be investigated.

Acknowledgements:

I would like to thank Richard Kayne, Stephanie Harves, Jason Merchant, Paul Postal and Andrew Radford for useful discussion of the ideas and data in this paper. I also thank the students of Syntax II (Spring 2014, NYU) where I gave a lecture on the material.

References

- Collins, Chris and Paul Postal. 2012. *Imposters*. Cambridge, MA: MIT Press.
 Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
 Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

- Craenenbroek, Jeroen van and Anikó Lipták. 2006. The Crosslinguistic Syntax of Sluicing: Evidence from Hungarian Relatives. *Syntax* 9: 248-274.
- Craenenbroeck, Jeroen van and Jason Merchant. Ellipsis Phenomena. 2013. In Marcel den Dikken (ed.), *The Cambridge Handbook of Generative Syntax*, 701-745. Cambridge: Cambridge University Press.
- Fiengo, Robert and Robert May. 1994. *Indices and Identity*. Cambridge, MA: MIT Press.
- Fox, Danny. 1999. Focus, Parallelism and Accommodation. In Tanya Matthews and Devon Strolovitch (eds.), *Proceedings of SALT 9*, 70-90. Ithaca, NY: Cornell University.
- Fox, Danny. 2003. On Logical Form. In Randall Hendrick (ed.), *Minimalist Syntax*, 82-123. Oxford: Blackwell.
- Heim, Irene. 2011. Definiteness and Indefiniteness. In Klaus von Stechow, Claudia Maienborn and Paul Portner (eds.), *Handbook of Semantics*, 996-1025. Berlin: de Gruyter.
- Hulsey, Sarah and Uli Sauerland. 2006. Sorting out Relative Clauses. *Natural Language Semantics* 14: 111-137.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- May, Robert. 1977. *The Grammar of Quantification*. Doctoral dissertation, MIT.
- Merchant, Jason. 2001. *The Syntax of Silence*. Oxford: Oxford University Press.
- Rooth, Mats. 1992. Ellipsis Redundancy and Reduction Redundancy. In Steve Berman and Arild Hestvik (eds.), *Proceedings of the Stuttgart Ellipsis Workshop*, report series Sprachtheoretische Grundlagen für die Computerlinguistik, no. 29, SFP 240, IBM Heidelberg.
- Ross, John Robert. 1969. Guess Who? In *Papers from the Fifth Regional Meeting of the Chicago Linguistics Society*, 252-286. Chicago, Illinois: Department of Linguistics, University of Chicago.
- Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT. Published as *Infinite syntax*, Norwood, NJ: Ablex (1986).
- Sag, Ivan. 1976. *Deletion and Logical Form*. Doctoral dissertation, MIT. Reproduced by the Indiana University Linguistics Club (1977).

