

in Charles E. Gribble, Ed.
James Daniel Armstrong: in
Memoriam, Slavica 1994.
(written 1981)

Bulgarian Relativization Strategies

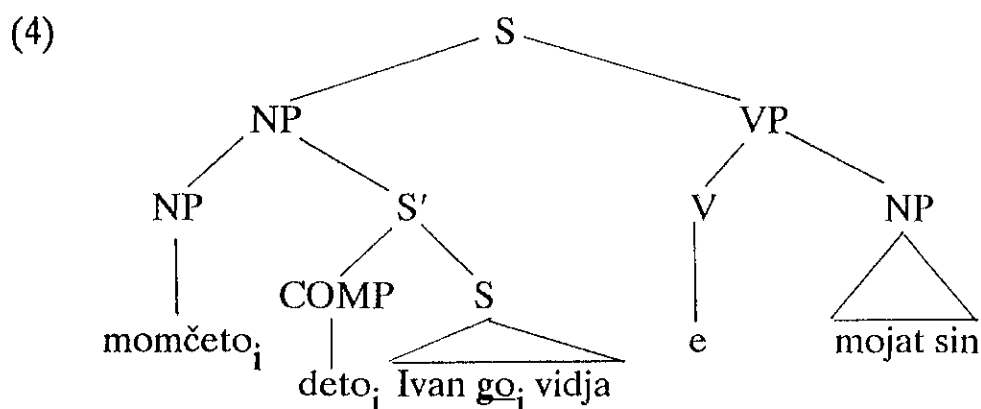
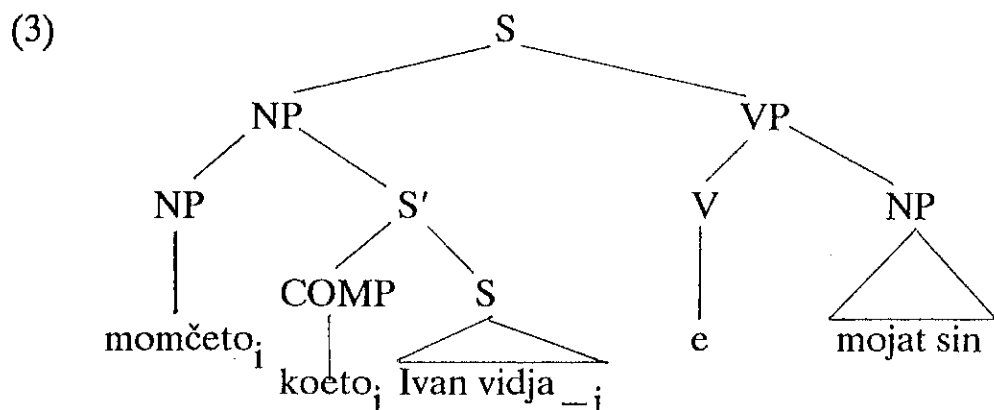
Catherine Rudin

In Modern Standard Bulgarian there are two types of relative clauses exemplified by the (a) and (b) versions of (1) and (2) below:

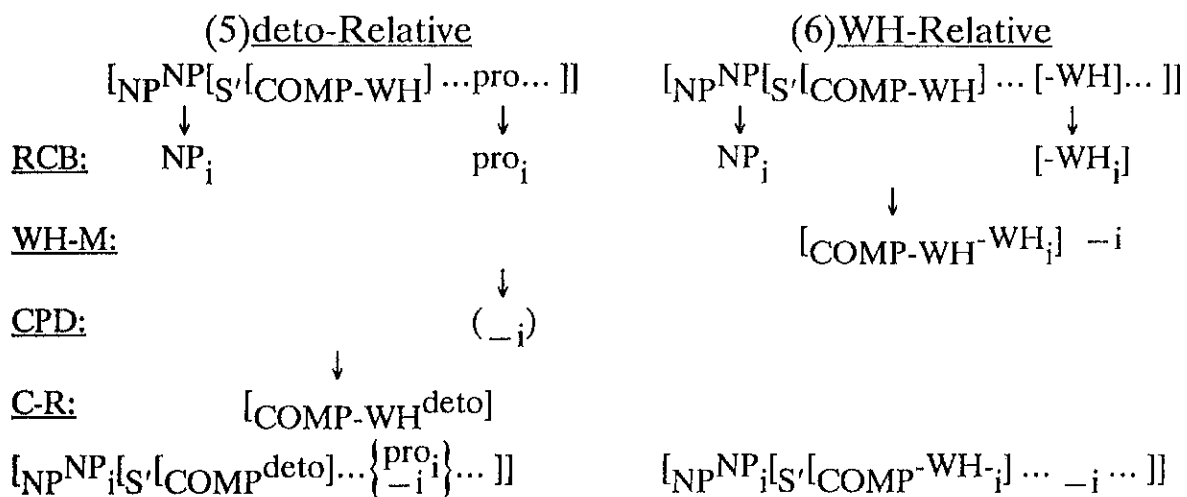
- (1) a. Tova e ženata kojato sedeše do mene.
that is woman-the who sat by me.
'That is the woman who was sitting by me.'
- b. Tova e ženata deto sedeše do mene
that is woman-the that sat by me
'That is the woman who was sitting by me.'
- (2) a. Momčeto koeto Ivan vidja e mojat sin.
boy-the whom Ivan saw is my son
'The boy whom Ivan saw is my son.'
- b. Momčeto deto Ivan go vidja e mojat sin.
boy-the that Ivan him saw is my son
'The boy whom Ivan saw is my son.'

The first type, in (1a) and (2a), is considered more literary, while the construction in (1b) and (2b) is colloquial, but both types are in common use and are essentially synonymous. I will refer to the first type as "WH-relatives" and the second as "*deto*-relatives."

WH-relatives consist of a head and a clause containing a WH word (relative pronoun)¹ coreferential both to the head and to a gap or missing constituent later in the clause. *Deto*-relatives also consist of a head and a clause, however, instead of a WH word the clause contains the relative complementizer *deto* and may contain a pronoun instead of a gap. Tree (3) is the (surface) structure of sentence (2a), and tree (4) is the structure of (2b), for example.



In this article I will show that the two relative constructions, although synonymous, have distinct deep structures and are derived by different types of rules. The deep structures and derivations I propose for *deto* and WH-relatives are outlined in (5) and (6).²



The rules involved in these derivations are as follows:

- (7) *Relative Clause Binding (RCB)*
 In a structure of the form $NP_i \overline{NP} \overline{S'}$, find an unindexed pronoun in S' and coindex it with NP_j .

Note that "pronoun" includes both WH words and personal pronoun. This rule ensures that a relative clause will always contain some element coreferential to the head; general conventions require the coindexed elements to be non-distinct in number and gender.

(8) *WH-Movement* (WH-M)

$[\text{COMP}_{\alpha\text{WH}}] \dots [\alpha\text{WH-phrase}] \rightarrow [\text{COMP}_{\alpha\text{WH}}^{\alpha\text{WH-phrase}}] \dots _i \dots$
 where α ranges over + in questions and - in relative clauses.

I assume that WH-Movement is unbounded and places the WH phrase in the COMP position; see Rudin (1981, 1982) for some justification of these assumptions.

(9) *Controlled Pronoun Deletion* (CPD)

$\text{NP}_i [\dots \text{pro} \dots] \rightarrow \text{NP}_i [\dots _i \dots]$

where " $_i$ " is an indexed empty node, i.e., a gap or "trace."

(10) *COMP-Realization* (C-R)

$[\text{COMP-WH}^e] \rightarrow [\text{COMP-WH}^{\text{deto}}]$

This is a surface "spelling out" rule. If COMP contains any lexical material (WH phrase, e.g.) at the surface level, the rule cannot apply.

Based on the small amount of data introduced so far, alternative analyses in which both types of relative clauses are derived from a single underlying structure seem just as plausible as the one proposed here; however, as we shall see, such analyses encounter serious problems as soon as more detailed data is introduced, while analysis (5-6) accounts easily for the facts of Bulgarian. Let us begin by looking at the distribution of resumptive pronouns in relative clauses.

In WH-relatives, resumptive pronouns normally do not occur at all regardless of the function of the relativized NP.³ In *deto* relatives, the surface presence or absence of a resumptive pronoun depends on the case-role of the relativized NP: the pronoun is impossible if it is the subject of its own clause, optional if it is the direct object, and obligatory if it has any oblique function. The oblique category can be divided into two types, indirect object, which requires a dative clitic as the resumptive pronoun,⁴ and object of a preposition (other than the dative *na*), which requires a non-clitic pronoun following the preposition. Examples of the four case-role types for *deto* relatives are given in (11), with the possibility or necessity of a pronoun indicated by parenthesis notation ((*X) = "X is impossible," *(X) = "X is obligatory"). The lack of a resumptive pronoun in corresponding WH-relatives is shown in (12).

- (11) a. SUBJ: Tova e knigata deto (*tja) beše na masata.
 this is book-the that it was on table-the
 'This is the book that was on the table.'

- b. DO: Tova e knjigata deto (ja) kupih včera.
 this is book-the that it bought-1s yesterday
 'This is the book that I bought yesterday.'
- c. IO: Tova e ženata deto *(i) dadoh knjigata.
 this is woman-the that to-her gave-1s book-the
 'This is the woman that I gave the book to.'
- d. PP: Tova e knjigata deto govorehme *(za neja).
 this is book-the that talked-1p about it
 'This is the book that we were talking about.'

- (12) a. SUBJ: Tova e knjigata kojato (*tja) beše na masata.
 b. DO: Tova e knjigata kojato (*ja) kupih včera.
 c. IO: Tova e ženata na kojato (*i) dadoh knjigata.
 d. PP: Tova e knjigata za kojato govorehme (*za neja).

Leaving aside for the moment the distribution of resumptive pronouns in *deto*-relatives, the fact that no pronoun is permitted in WH-relatives while one usually is possible in *deto*-relatives is accounted for very naturally under the analysis proposed above. No pronoun appears in the lower clause of a WH-relative, because the relativized position, which contained a WH phrase in underlying structure, is left empty after WH-Movement moves the WH phrase to COMP. In *deto*-relatives, on the other hand, the pronoun in relativized position in the deep structure does not move, so, unless CPD applies, it remains in place in the surface form of the clause.

Alternative analyses in which *deto* and WH-relatives have a common source cannot account for this data in a principled way. I briefly examine two such analysis types here and return to them again as more data is taken into consideration.

I. *The Base Generation Hypothesis*. This is the term I will use for any analysis in which WH words are base generated in their surface position (in COMP) and RCB links a gap or pro in the relative clause with the head noun in both WH- and *deto*-relatives (see for example Engdahl's (1979) analysis of Swedish). The deep structure of both types of relatives, under this hypothesis, would be (13a) or perhaps (13b); which version is chosen makes no difference for our purposes.

- (13) a. [_{NP}NP [_S'[COMP-WH -WH-phrase] ...(pro)...]]
 b. [_{NP}NP [_S'[COMP-WH WH-phrase *deto*] ...(pro)...]]

In either version, an optional rule of WH-deletion is needed. If it applies, a *deto* relative is derived; if it does not, a WH-relative results.

The (13b) version would also need to delete *deto* in WH-relatives, while in (13a) *deto* would be spelled out by COMP-Realization after deletion of the WH phrase. So far so good – but how can the absence of resumptive pronoun in WH-relatives be accounted for on this analysis? Since nothing moves out of the relative clause at all, there is no reason for it to contain an obligatory gap in WH-relatives but a pronoun in *deto*-relatives. To capture the facts, some sort of ad hoc device must be added; perhaps a filter ruling out a bound pronoun in a clause with a surface WH word, or a condition on CPD making it obligatory when COMP contains WH. Note that the apparent redundancy of a pronoun in WH relatives is not sufficient to explain its ungrammaticality, since Bulgarian freely allows redundant pronoun copies of NPs: *Ivan go vidjal* (Ivan-him-I saw) ‘I saw Ivan.’

II. *The Movement-and-Deletion Hypothesis.* Under this type of analysis *deto*- and WH-relatives would both have the deep structure (14a/b) again, it makes little difference whether *deto* is present in the base or not.

- (14) a. [_{NP} NP [_S [_{COMP-WH}] ...-WH phrase...]]
 b. [_{NP} NP [_S [_{COMP-WH} *deto*] ...-WH phrase...]]

WH-Movement would place the WH phrase in COMP, and then optional WH-Deletion and either *deto*-Deletion or COMP-Realization would apply in the same way as in the Base Generation analysis. This is of course, basically Chomsky’s (1977) analysis of English. This analysis again provides no principled explanation for the difference in acceptability of resumptive pronouns between WH- and *deto*-relatives. If WH-Movement were formulated so as to leave a pronoun copy behind in the relativized position, that pronoun would have to be deleted or filtered out just in case a WH phrase appeared on the surface. If WH-Movement did not leave a copy, some other device would be needed to insert a pronoun if the WH phrase had been deleted.

Another possible analysis, “Movement-or-Deletion,” would also posit deep structure (14), but with different rules for the two types of relatives: WH-Movement would produce WH-relatives, while *deto*-relatives would be derived by deletion in place of the WH phrase, with no movement. (This is like Huang’s (1977) analysis of English relativization; also Grimshaw’s (1974) of Middle English.) As far as predictions concerning resumptive pronouns are concerned, “Movement-or-Deletion” is indistinguishable from “Movement-and-Deletion.”

So far we have looked only at sentences in which the relative clause is a single S. More complex examples provide additional evidence for hypothesis (5-6) and also make it possible finally to explain the conditions on Controlled Pronoun Deletion.

Both relativization strategies can reach into indefinitely deeply embedded clauses, at least in principle. As in simpler cases, WH-relatives contain a gap, while *deto*-relatives often contain a pronoun:

- (15) a. Knigata_i kojato_i ti mi kaza [S, če Ivan misli [S, če
 book-the which you me told that Ivan thinks that
 Petūr iska [S, da kupi _i]]]...
 Peter wants to buy
- b. Knigata_i deto ti mi kaza [S, če Ivan misli [S, če
 book-the that you me told that Ivan thinks that
 Petūr iska [S, da (ja_i) kupi]]]...
 Peter wants to it buy

'The book ^{which} _{that} you told me Ivan thinks Peter wants to buy...'

However, there are certain contexts into which only *deto*-relativization is possible; these are the so-called syntactic "islands." Examples of two types of islands follow:

- (16) a. Tova e momčeto deto trjabva da namerim negovata majka.
 this is boy-the that must to find-1p his mother
- b. *Tova e momčeto (na) koeto trjabva da namerim (negovata)
 this is boy-the of whom must to find-1p his
 majka.
 mother.
- c. *Tova e momčeto čijato trjabva da namerim (negovata) majka.
 this is boy-the whose must to find-1p his mother
 '*This is the boy who we have to find his mother'

- (17) a. Tova e momčeto deto misūlta če mu dadohme
 this is boy-the that thought-the that him gave-1p
 bonboni jadosva lekara.
 candy angers doctor-the
- b. *Tova e—momčeto na koeto misūlta če mu dadohme
 this is boy-the to whom thought-the that him gave-1p
 bonboni...
 candy
 '*This is the boy who the thought that we gave him candy makes
 the doctor mad.'

The sentences in (16) represent relativization of the possessive determiner of an NP, that is, of "POSS" in the structure [NP[POSS][NP]].

Those in (17) show relativization from a complex NP, that is, from inside a clause dominated by NP; "X" in [_{NP}[_S ...X...]]. Both of these positions evidently resist extraction in Bulgarian as in many languages: WH-Movement is impossible, as the ungrammaticality of (16b-c) and (17b) illustrates. *Deto*-relativization into these island contexts is permitted, however, note that the resumptive pronoun is obligatory, even if it is a direct object⁵ (recall that accusative resumptive clitics are optional in simple *deto*-relatives). (18) shows that the accusative clitic in a complex NP is obligatory.

- (18) Vidjah edna kniga deto faktūt če *(ja) prodavat me iznenada.
 saw-1s a book that fact-the that it sell-3p me surprises
 'I saw a book that the fact that they sell it surprises me.'

This contrasts sharply with the optionality of the accusative resumptive clitic in a non-island context:

- (19) Vidjah edna kniga deto (ja) prodavat eftino.
 saw-1s a book that it sell-3p cheap
 'I saw a book that's for sale cheap.'

This fact is easily accounted for given the analysis (5-6); clearly CPD is unable to operate into syntactic islands just as WH-Movement is, so the pronoun coindexed with the head in (18a) as well as (16a, 17a) cannot be deleted. This is a totally expected result, since movement and controlled deletion are subject to the same constraints in many (perhaps all) cases (See e.g. Bresnan & Grimshaw (1978).)

Another construction which is an island in many languages is the embedded question ("WH Island"). However, in Bulgarian, questions are not islands to relativization: WH-Movement and Controlled Pronoun Deletion can both reach into indirect questions, as (20) shows.

- (20) a. Tova e knigata deto se čudja dali (ja) prodavat.
 this is book-the that wonder-1s whether it sell-3p
 '*This is the book that I wonder whether they sell (it).'
- b. Tova e knigata kojato se čudja koj prodava.
 this is book-the which wonder-1s who sells
 '*This is the book that I wonder who sells.'

It is significant that **both** WH-Movement and CPD are possible here; the controlled deletion rule can apply exactly (and only) where the movement rule can. This is a striking confirmation that movement and controlled deletion are subject to the same constraints.

The fact that CPD obeys island constraints suggests the explanation for the pattern of optional and obligatory resumptive pronouns with *deto*. (Recall from example (11) that in simple *deto* relatives a resumptive pronoun is optional if accusative and obligatory if dative or object of a preposition.) If we take the statement of island contexts in Bulgarian to be (21), all of the restrictions on the occurrence of resumptive pronouns are accounted for except for nominatives, which are discussed separately below.

- (21) No rule may involve X and Y in the structure:
 ... X ... [α ... Y ...] where α = NP or PP.

The effect of (21) is to disallow movement or controlled deletion from inside an NP or PP. Assuming CPD to be optional, the island status of NP accounts for the fact that resumptive pronouns in *deto*-relatives cannot be deleted from complex NP and Det of NP contexts (as well as from inside another relative clause), while the fact that PP is an island explains the non-deletion of pronouns which are object of a preposition. Non-deletion of datives also falls under the PP-island restriction, since datives are formally PPs (*na nego* 'on him' = 'to him'). Datives do behave unlike other PPs in some ways, for example, in having a clitic form (*na nego* = *mu* (clitic)); it is possible that those speakers who accept deletion of dative clitics in *deto*-relatives have analyzed datives as "not real PPs" in some sense and therefore not islands (see footnote 4).

As far as WH-relativization is concerned, (21) explains the impossibility of WH-Movement from Complex NP and Det of NP, since it would involve a WH phrase inside an NP and COMP outside the NP. It also accounts for the obligatoriness of "pied piping," since movement of an NP out of PP, stranding the preposition, would violate the PP island. Prepositions can never be stranded in Bulgarian, by WH-Movement (e.g. (22)) or by any other rule:

- (22) a. *Deteto s koeto igraehme...*
 child-the with whom played-1p
 'The child with whom we were playing...'
- b. **Deteto koeto igraehme s...*
 child -the whom played-1p with
 'The child whom we were playing with ...'

In order to relativize an NP dominated by PP, WH-relativization must move the entire PP, as in (22a). Similarly, in order to relativize an NP dominated by another NP, it must move the whole higher NP. Compare (23) to (16b,c) for instance.

- (23) Tova e momčeto čijato majka trjabva da namerim.
 this is boy-the whose mother must to find-1p
 'That is the boy whose mother we have to find.'

WH-Movement is thus able to avoid the island constraints in some cases by "pied piping" the context along with the relativized NP, but it cannot move anything out of an island.

Unfortunately, the obligatory deletion of most subject resumptive pronouns (e.g. (11a)) is **not** accounted for by (21), nor am I aware of any other principled way to exclude them. The correct explanation will almost surely involve stress or focus, and not island conditions, since a strongly focused nominative resumptive pronoun is sometimes acceptable:

- (24) Tova e čoveka, deto *samo toj* znae otgovora.
 this is man-the that only he knows answer-the

The island and pied piping data presented above provides strong support for the analysis of relativization I am arguing for here, since this analysis, with WH-Movement only in WH-relatives and Controlled Pronoun Deletion only in *deto*-relatives, in conjunction with the island statement (21), accounts for all of the facts in a very simple and well-motivated manner. The distribution of resumptive pronouns in both simple and complex *deto*-relatives, the ungrammaticality of WH-Relativization from a complex NP, and the obligatory pied piping in WH-relatives are all explained by the single constraint (21), given (5-6) as the derivations of the two types of relatives. The only relevant fact not explained by this analysis is the behavior of subject resumptive pronouns, as noted above.

The alternative analyses in which the two relative clause types are underlyingly identical, far from deriving any support from the island facts, are even further weekend by them. Under the Base Generation hypothesis, an island condition like (21) would correctly account for the obligatory presence of a resumptive pronoun in *deto*-relativization from a Complex NP, Det of NP, or PP, given that a pronoun was present in the deep structure. However, it would predict exactly the same for WH-relatives, which of course is false. That is, if both (17a) and (b) were derived from (25), there would be no reason for (26) (= (17b) with the resumptive pronoun present) to be ungrammatical.

- (25) [_{NP}momčeto [_S [_{COMP}^{na} koeto] [_{NP}misŭlta [_S če *mu* dadohme ...]]]]

- (26) *Tova e momčeto na koeto misülta če *mu* dadohme
 this is boy-the to whom thought-the that him gave-1p
 bonboni jadosva majka mu.
 candy angers mother his
 ‘*This is the boy who the thought that we gave him candy angers
 his mother.’

I am aware of no principled reason why WH-Deletion should make binding into an island context possible, nor why non-deletion of WH should make such binding impossible.

The Movement-and-Deletion hypothesis would derive (17a,b) from (27):

- (27) [_{NP}momčeto [_S [_{COMP}deto] [_{NP}misülta [_S če mu dadohme ... na koeto]]]]

In order to achieve a correct description of the facts given this underlying structure and the statement of island contexts, WH-Movement would somehow have to be made to obligatorily leave a pronoun copy when moving out of an island, and the WH phrase would be required to delete if it had been moved out of an island. (WH-Deletion of course would be optional in non-island contexts, or no WH-relatives could ever be derived at all.) Both of these restrictions could be formulated as filters or conditions on rules, but not, so far as I can see, in any explanatory way; they would simply be ad hoc complications of the grammar. Similarly, in the Movement-or-Deletion variant of this analysis, WH-Deletion would have to be marked obligatory and obligatorily leave a copy behind in islands: a very peculiar situation, since free deletion rules are by definition not subject to island constraints.

As we have noted several times, analysis (5-6) accounts for all of the conditions on resumptive pronouns in both simple and complex relative clauses. The slight complication of the grammar involved in having separate underlying forms for the two types of relatives is clearly outweighed by the elegant and principled explanation of the data it makes possible, in contrast to the numerous ad hoc conditions or filters required to derive both types from a common source.

Wayne, Nebraska

Notes

¹The relative WH words are: *kojto*, *kojato*, *koeto*, *koito* ‘who-m/f/n/pl’, *kogoto* ‘whom-m’, *komuto* ‘to whom (arch.)’, *čijto*, *čijato*, *čieto*, *čiito*, ‘whose’, *kakŭvto*, *kakvato*, *kakvoto*, *kakvito* ‘what kind’, *kakvoto* ‘what’, *kŭdeto* ‘where’, *kogato* ‘when’. Interrogative WH words are the same but lack the definite suffix *-to*.

²In keeping with most recent generative work, I assume that pronouns are present in the base. [-WH] identifies relative COMPs and WH- opposed to interrogative ([+WH]) ones.

³Resumptive pronouns do occasionally show up in WH-relatives, especially unusually long and/or complex ones; such cases are probably best treated as ungrammatical though marginally acceptable.

⁴Some speakers find *deto*-relativization of a dative acceptable with no pronoun; these speakers (11c) would have (i) rather than *(i).

⁵Nominative resumptive pronouns are not obligatory even here, since F Drop can delete subject pronouns everywhere.

References

- Bresnan, Joan, and Jane Grimshaw. 1978. "The Syntax of Free Relatives in English." *Linguistic Inquiry* 9.3:331-91.
- Chomsky, Noam. 1977. "On WH-Movement" in Culicover, Akmajian, and Rizzi (eds) *Formal Syntax*. New York: Academic Press. 71-132.
- Engdahl, Elisabet. 1979. "Binding Rules in Swedish: Relativization, Complement Formation and Topicalization" MS.
- Grimshaw, Jane. 1974. "Evidence for Relativization by Deletion in Chaucer's Middle English" *Papers from the Fifth Annual Meeting, North Eastern Linguistic Society*:216-224.
- Huang, Phoebe Kan. 1977. *WH-Fronting and Related Processes*. PhD Dissertation, University of Connecticut.
- Rudin, Catherine. 1981. "Who what to whom said?: An Argument from Binding Theory against Successive Cyclicity" *CLS* 17.
- _____. 1982. *Complementizers and WH Constructions in Bulgarian*. PhD Dissertation, Indiana University.