

Syntax of Free Relatives and the Matching Phenomenon: A Preliminary Study*

Sakumi Inokuma
University of Tokyo

inokumas@l.u-tokyo.ac.jp

Abstract

The matching phenomenon has been regarded as an idiosyncratic property of free relatives (FRs) and thus as a crucial cue to investigating the syntax of FRs. In the first part of this article I establish (i) that the matching phenomenon is a phenomenon prevalent in argument-sharing constructions, and (ii) that the matching phenomenon is a result of phonological deletion rather than of syntactic agreement. In the second part of the article I present a rough sketch of a new approach to the syntax of FRs that incorporates Nunes' (2001, 2004) sideward movement and Chomsky's (2004) theory of adjunction. I also discuss the derivational ordering and the notion of "simultaneity" in a broader theoretical context at some length.

Keywords: matching phenomenon, free relatives, argument-sharing, parallel merge, sideward movement

1. Introduction

In this paper I will investigate the nature of so-called the Matching Phenomenon (Bresnan & Grimshaw (1978)). The matching phenomenon in free relatives (henceforth FRs) is illustrated by German examples in (1).¹

(1) German

- a. Wer_{NOM} nicht stark ist, muss klug sein.
who not strong is must clever be
'Who isn't strong must be clever.'
- b. *Wen_{ACC}/wer_{NOM} Gott schwach geschaffen hat, muss klug sein.
whom/who God weak created has must clever be
'Who God has created weak must be clever.' (Groos & van Riemsdijk (1981))

* I am grateful to Akira Watanabe and Noriko Imanishi for insightful comments and suggestions. All remaining errors and shortcomings are mine.

¹ I will focus on the case matching phenomenon in the present paper, putting aside another aspect of the matching phenomenon found in FRs, i.e., categorial matching.

(2) German

- a. Ich nehme, wen_{ACC} du mir empfiehlst.
I take whom you me recommend
'I take whom you recommend to me.'
- b. *Ich nehme, wem_{DAT}/wen_{ACC} du vertraust.
I take whom/whom you trust
'I take whom you trust.'

(ibid.)

In (1a), the *wh*-element which functions as the subject within the FR clause receives nominative Case. The FR clause as a whole, *wer nicht stark ist*, also functions as the subject of the main clause and presumably receives nominative Case. When the two cases are not the same, as in (1b), where *wh*-element receives accusative Case within the embedded clause and the FR clause receives nominative Case within the matrix clause, the sentence results in ungrammaticality. In other words, the case which *wh*-element receives in the FR clause and the one which the FR clause itself receives in the matrix clause must "match." (2a,b) show the cases in which the FR clause functions as the object in the matrix clause. There are no asymmetries among the cases (NOM, ACC, DAT) in matching phenomena.

Headed relative clauses do not manifest the matching phenomenon, as illustrated by (3). The head noun and the relative pronoun can carry different cases without any problems (*der Mensch*_{NOM} vs. *den*_{ACC}; *den Kandidaten*_{ACC} vs. *dem*_{DAT} in (3a) and (3b) respectively).

(3) German

- a. Der Mensch_{NOM}, den_{ACC} Gott schwach geschaffen hat, muss klug sein.
the person whom God weak created has must clever be
'The person whom God has created weak must be clever.'
- b. Ich nehme den Kandidaten_{ACC}, dem_{DAT} du vertraust.
'I take the candidate whom you trust.'

Syntactic investigations into FRs which have accumulated since the late 1970's often focus on explaining the nature of the matching phenomenon (Bresnan & Grimshaw (1978), Groos & van Riemsdijk (1981), Hirschbühler & Rivero (1983), Suñer (1984) among others). This article offers a new perspective on the nature of the matching phenomenon. It will be shown that the phenomenon is not an idiosyncratic property of FRs but the one generally found in "constituent-sharing" constructions.

This paper is organized as follows: section 2 observes the case-matching phenomenon of FRs and the theoretical complications that they pose. Section 3 turns to other constructions that show the same phenomenon, and establishes that the matching phenomenon is not a

syntactic but rather morpho-phonological phenomenon in nature. Section 4 surveys the recent approaches to the matching phenomenon and the syntax of FRs, and points out their problems in theoretical and empirical terms. Building upon these discussions, I will suggest an alternative approach to the syntax of FRs that would explain the matching phenomenon while avoiding the idiosyncrasy of their derivations in section 5. The approach pursued here is characterized as a combination of (a revised version of) Nunes' (2001, 2004) sideward movement theory and Chomsky's (2004) theory of adjunction. In the end, the matching phenomenon is argued to be a result of PF-licensing condition on chains, rather than of syntactic agreement.

2. Case Matching in Free Relatives

Citko (1998) observes that Polish FRs also show the matching phenomenon. Sentences in (4) and (5) are in complete parallel with German examples in (1) and (2). Here too, the *wh*-element and the FR that it introduces must match in case.

(4) *Polish*

- a. [Kto_{NOM} chce]_{NOM} dostanie present na gwiazdkę.
 who wants will get a gift for Christmas
 'Who wants will get a gift for Christmas.'
- b. * [Komu_{DAT} się teraz przyglądasz]_{NOM} jest naszym szefem.
 Who REFL now look at is our boss
 'Who you are now looking at is our boss.'

(Citko (1998))

(5) *Polish*

- a. Zrobię [co(kolwiek)_{ACC} Maria zrobiła]_{ACC}.
 I will do whatever Maria did
 'I will do whatever Maria did.'
- b. *Przepytam [kto(kolwiek)_{NOM} pierwszy przyjdzie]_{ACC}.
 I will ask whoever first comes
 'I will ask whoever comes first.'

(ibid.)

The matching phenomenon has been regarded as evidence for the Head Hypothesis since Bresnan & Grimshaw (1978). Under the Head Hypothesis, the *wh*-element occupies the head noun position rather than COMP (SpecCP) position, going through an ordinary case-assigning operation within the matrix clause. Under the Head Hypothesis, the matching condition is stated as some version of a licensing condition on an empty element (whether null operator or PRO) which lies in the embedded relative clause and is bound by this overt *wh*-element.

(6) [DP wh_i [CP/IP ... e_i]]

Groos & van Riemsdijk (1981), on the other hand, claim that the *wh*-element occupies the COMP (SpecCP) position, in the same fashion as interrogatives and overtly headed relatives. The head noun position is occupied by empty *pro* or PRO. Under this COMP Hypothesis, the matching condition is stated in terms of accessibility to COMP of the matrix operations.

(7) [DP PRO_{*i*} [CP wh_i ... e_i]]

We will not enter into detailed comparisons of these two traditional approaches here.² What is important for our purpose is that both of them assume the case-matching phenomenon to be a syntactic phenomenon.

The notion of the "case" involved in case-matching in FRs, however, is not that of an abstract Case, but rather of a morphological case. Compare German examples below with (1a) and (2a).³

(8) German

- a. Was_{ACC/NOM} du mir gegeben hast, ist prächtig.
what you me given have is wonderful
'What you have given to me is wonderful.'
- b. Hast du was_{NOM/ACC} im Programm war schon kopiert?
have you what on the program was already copied
'Have you already copied what was on the program?'

(Groos & van Riemsdijk (1981))

In these examples, the abstract Cases required by the embedded and the matrix clauses are not the same one; in (8a), the former is accusative and the latter nominative, and vice versa in (8b). In particular, the Case relation in (1a) and (8a) is the same, yet (8a) is grammatical and (1a) is not.

The point becomes clearer when we look at Polish, where case morphology is richer than German. Consider Polish examples in (9) with the case paradigm of *wh*-elements, *kto* 'who' and *co* 'what' in the language (10).

² See van Riemsdijk (2005) for a survey of this field.

³ Note that Subjective/Objective/Oblique distinction of Case along the line of Chomsky (1981) does not make a right distinction.

(9) Polish

- a. Jan nienawidzi [kogo_{ACC} Maria lubi e_{ACC}]_{GEN}
Jan hates who Maria likes
'Jan hates whoever Maria likes.'
- b. *Jan nienawidzi [czego_{GEN/COACC} Maria lubi e_{ACC}]_{GEN}
Jan hates what Maria likes
'Jan hates whatever Maria likes.'

(Citko (1998))

(10) Case syncretism in Polish

Nom	kto 'who'	co 'what'
Gen	kogo	czego
Dat	komu	czemu
Acc	kogo	co
Loc	kim	czym
Instr	kim	czym

The generalization on *wh*-forms in FRs drawn from German and Polish examples, then, is that Case mismatch is allowed as long as the same morphological form is available for those distinct Cases by syncretism.⁴

The situation poses an obvious problem on the standard Case theory. First, if the structure of FRs literally contains only one *wh*-element, then we must assume that it can carry two case features, assigned within the embedded and the matrix clauses. The consideration of this line immediately leads us to (i) postulate empty categories, e.g., PRO, *pro* or some kind of null operator; or (ii) relax the Case theory so that it allows multiple Case-marking. Both directions are indeed pursued in the previous studies, to which we will return in section 4.

Related but more complicating is the fact that the relevant notion of case is morpho-phonological one. Suppose we postulate the empty category in analyzing the syntax of FRs, and interpret the matching phenomenon as agreement between this empty category and the overt *wh*-element. But what kind of agreement is this? The agreement relations usually observed are sensitive to abstract features rather than phonological forms.⁵ This situation suggests several possible paths: (i) the operation is phonological in nature, i.e., the matching phenomenon results from a deletion operation by virtue of PF identity rather than syntactic agreement; or (ii) the operation is still syntactic. In the latter case we have to revise Case

⁴ Our discussion builds mainly on German and Slavic (especially Polish) examples. This is because the morphological matching phenomenon is only apparent in languages with rich case inflection, thus excluding non-case-marking languages like English and Dutch. Needless to say, how far the present discussion generalizes must be subject to further investigation.

⁵ Note further that one of the elements involved in this relation is an empty category, i.e., an item without any phonological features.

theory so that it could allow multiple case-marking and be "loose" about abstract Case in terms of e.g. feature underspecification.

Interestingly, the pattern of case-matching in FRs is replicated in other argument-sharing constructions, which suggests that the matching phenomenon is not an idiosyncratic property of FRs as has long been assumed, and must be explained in more general terms.

3. Case Matching in Other Constructions

The previous section observes that FRs show a peculiar matching phenomenon with regard to case morphology of *wh*-elements. This section provides evidence that the case matching phenomenon is not an idiosyncratic property of FRs but that it is found rather prevalently across constructions that involve argument-sharing.

First, consider Across-the-Board (ATB) *wh*-questions. Typical ATB *wh*-questions are exemplified by (11).

- (11) Who_i does John love t_i^1 and Mary hate t_i^2 ?

Here again, the single overt *wh*-element *who* is linked to two Case positions, indicated as t_i^1 and t_i^2 without any theoretical commitments. The situation takes us back to the problem with regard to Case theory pointed out in section 2; a single overt element enters multiple Case-checking relations. The cases that the *wh*-element receives in each component clause must match, but again, the relevant notion of "case" is not a syntactic one, but morpho-phonological. This is illustrated by German examples (12).

(12) German

- a. Welcher Speise verschreibt sich der Snob *e* und enthält sich der
 which food_{DAT/GEN} devotes self the snob e_{DAT} and abstains self the
 Mönch *e*?
 monk e_{GEN}
- b. *Welches Getränks / Welchem Getränk verschreibt sich der Snob *e*
 which beverage_{GEN}/which beverage_{DAT} devotes self the snob e_{DAT}
 und enthält sich der Mönch *e*?
 and abstains self the monk e_{GEN}

(Sauerland (1996))

(13) Case paradigm of *welcher* 'which'

	MASC.SG.	FEM.SG.	NEUT.SG.	PL.
Nom	welcher	welche	welches	welche
Gen	welches	welcher	welches	welcher
Dat	welchem	welcher	welchem	welchen
Acc	welchen	welche	welches	welche

Welcher must agree with its nominal restriction, with a feminine *Speise* in (12a) and with a neuter *Getränk* in (12b). As (13) shows, *welcher* in the feminine paradigm (along with the feminine noun *Speise*) syncretizes in Genitive and Dative forms, whereas in the neuter paradigm it does not. Hence the two sentences in (12) result in different grammaticality. Notice that these two form a minimal pair. The same holds in Polish and Russian examples in (14, 15) and (16) respectively.

(14) Polish

- a. Kogo Jan lubi *e* a Maria podziwia *e*?
 who_{ACC} Jan likes *e*_{ACC} and Maria admires *e*_{ACC}
 'Who does Jan like and Maria admire?'
- b. *Kogo/Komu Jan lubi *e* a Maria ufa *e*?
 who_{ACC/DAT} Jan likes *e*_{ACC} and Maria trusts *e*_{DAT}
 'Who does Jan like and Maria trust?'

(Citko (2005))

(15) Polish

- a. *Czego/Co Jan nienawidzi *e* a Maria lubi *e*?
 what_{GEN/ACC} Jan hates *e*_{GEN} and Maria likes *e*_{ACC}
 'What does Jan hate and Maria like?'
- b. Kogo Jan nienawidzi *e* a Maria lubi *e*?
 whom_{GEN/ACC} Jan hates *e*_{GEN} and Maria likes *e*_{ACC}
 'Whom does Jan hate and Maria like?'

(ibid.)

(16) Russian

- devuška, kotoroj ja byl uvlečen *e* i daval den'gi *e*, ...
 girl who_{INST-DAT} I was carried-away-with<INST> and gave<DAT> money
 'the girl who I was carried away with and gave money to'

(Franks (1992))

In fact Citko (1998), focusing on this parallelism between ATB *wh*-questions and FRs with respect to the matching phenomenon, suggests an approach that unifies ATB *wh*-movement and FRs by allowing multidominance structure in grammar. Her "ATB-Merge" approach, or more updated "Parallel Merge" approach to the matching phenomenon (Citko (2005)) suffers

from both technical and empirical problems, to which we will return in the next section. But before that, notice that the matching phenomenon in the sense discussed here can also be found in other constructions: so-called Parasitic Gap constructions in German ((17) vs. (18)) and Russian ((19) vs. (20)).

(17) German

Dieser Dame_i hätte er sich [ohne PG_i schon mal Geld angeboten zu haben]
 this lady_{GEN} had he self without PG_{DAT} already money offered to have
 niemals entsinnen können.
 never remembered could
 'He never would have been able to remember this lady if he hadn't offered her money
 before.' (Kathol (2001))

(18) German

*Hans hat seine Tochter_i [ohne PG_i Geld zu geben] unterstützen können.
 Hans has his daughter_{ACC} without PG_{DAT} money to give support could
 'Hans was able to help his daughter without sending her money.' (ibid.)

(19) Russian

deвушка, kotoroj Ivan davalden'gi e do togo, kak (on)stal izbegat' e, ...
 girl who_{DAT-GEN} Ivan_{NOM} gave money_{DAT} until (he) started to-avoid e_{GEN}
 'the girl who Ivan gave money to until he started to avoid her' (Franks (1992))

(20) Russian

*mal'čik, kotoromu/kotorogo Maša daala den'gi e do togo, kak (ona) tala
 boy who_{DAT/who}_{GEN} Maša_{NOM} gave money_{DAT} until (she) started
 izbegat' e, ...
 to-avoid e_{GEN} (ibid.)

The exact nature of apparent parasitic gap constructions like (17) in German put aside,⁶ it is clear that the condition is morpho-phonological rather than syntactic. If we 'reconstruct' the antecedent DP *seine Tochter* 'his daughter' into the parasitic gap position in (18), for instance, this DP would have to be realized in the dative form, as *seiner Tochter*. There are no syncretic forms available in this (specific) environment, thus ungrammaticality results, unlike (17), where the syncretic form *dieser Dame* is available for genitive and dative Case. Franks (1992) proposes a unified approach to ATB and PG constructions,⁷ claiming that the matching phenomenon is a reflex of (PF-) licensing condition of null operators. He suggests that null

⁶ See Kathol (2001).

⁷ To be more precise, Franks (1992) reduces ATB to PG, that is, to adjunction structure.

operators must be licensed at LF and PF independently. At LF, what he calls 'thematic prominence,' stated in terms of a feature [\pm prominent], must be identical on the antecedent DP and the null operator; at PF the phonological form must be identical on the two. His insight is interesting especially in the current framework (Chomsky (2000) and subsequent work), which attempts to derive syntactic constraints exclusively from two interface conditions. It is not clear, however, what it means to assume that a phonologically empty element must be licensed in terms of its postulated phonological shape.

Finally, consider Right Node Raising (RNR) in German, taken from Wilder (1999) and Sag (2003). Notice that RNR apparently does not involve operator movement, hence Franks' explanation in terms of null operator licensing is not applicable. Yet we find the same matching phenomenon as the one observed above.

(21) *German*

- a. Er findet und hilft Frauen.
 he finds_{<ACC>} and helps_{<DAT>} women.
 'He finds and helps women.' (Sag (2003))
- b. Ich weiß, dass Fritz Herrn Meyer begrüßte und dankte
 I know that Fritz Mr_{ACC/DAT} Mayer greeted_{<ACC>} and thanked_{<DAT>}
 'I know that Fritz greeted and thanked Mr. Mayer.' (Wilder (1999))

(22) *German*

- a. *Er findet und hilft die/den Frauen.
 he finds_{<ACC>} and helps_{<DAT>} the_{ACC/the_{DAT}} women
- b. *Ich weiß, dass Fritz dem/den Herrn begrüßte und dankte
 I know that Fritz the_{DAT/the_{ACC}} Herrn greeted_{<ACC>} and thanked_{<DAT>}

To sum up, the matching phenomenon is not an idiosyncratic property of FRs, but rather one generally found in argument-sharing constructions. This implies that if we want to analyze the syntax of FRs with critical reference to the matching phenomenon, we would also have to analyze the syntax of those constructions by the same mechanism.

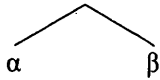
In the next section I will first examine Citko's (2005) recent account on the matching phenomenon in ATB *wh*-questions and ask whether her account can be extended to the matching phenomenon in general.

4. Parallel Merge & Sideward Movement Approach to Argument-Sharing

4.1. Citko's (2005) "Parallel Merge"

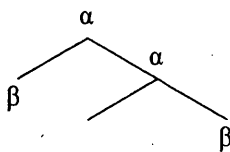
Citko (2005) introduces a new subtype of the operation Merge, called Parallel Merge, claiming that it is derivable from a combination of the properties of external and internal merge. Consider the properties of the primitive operations External Merge and Internal Merge (Chomsky (2004, 2005a, 2005b)).

(23) *External Merge*



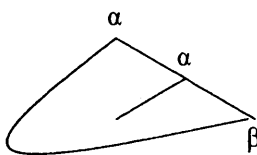
- i. Take two distinct rooted structures (α and β).
- ii. Join them into one.

(24) a. *Copy theory of movement*



- i. Take one structure (α) and its subpart (β).
- ii. COPY β .
- iii. MERGE (the copy of) β and α .

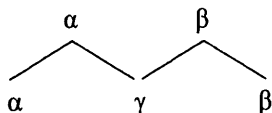
b. *Internal Merge theory of movement*



- i. Take one structure (α) and its subpart (β).
- ii. "REMERGE" α and β .

External merge is straightforward. Take two syntactic objects, and put them into one. As for internal merge, formerly called "Move," Citko interprets Chomsky's characterization of internal merge rather literally. She distinguishes the internal merge theory from the copy theory, and defines the former as involving a "loop" that allows an object already (externally) merged to "remerge" itself with its mother node, without creating a copy (24b).⁸ This opens up a possibility, Citko claims, that an operation "Merge" applies to an element already externally merged (as in internal merge) and has it remerge with a distinct syntactic object (as in external merge). Then we obtain "Parallel Merge," as illustrated in (25).

(25) *Parallel Merge*

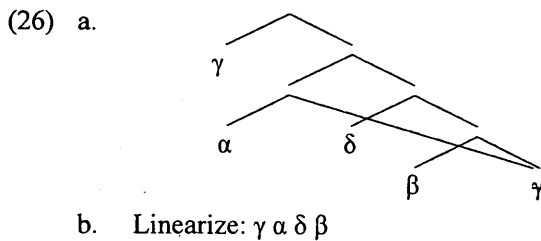


- i. Take two distinct rooted structures (α and β).
- ii. Merge one of them (β) and the subpart of the other (γ).

There need to be several clarifications about the properties of parallel merge. First, Citko

⁸ See also Chomsky (2005b: note 17) for complications and misunderstandings of the definition of Internal Merge.

assumes the Linear Correspondence Axiom (LCA) (Kanye (1994); Chomsky (1995)). The essential insight of the LCA is that linear order is determined solely by asymmetric c-command: α precedes β iff α asymmetrically c-commands β . Under this conception of the LCA, the multidominance structure like (25) is not linearizable since the c-command relation of the shared element γ cannot be unambiguously determined, irrespective of how α and β turn out to be ordered. Citko follows Chomsky (1995) in this respect, assuming that phonologically empty elements are invisible to the LCA. In other words, multidominance structure is allowed as long as the shared element is moved out of that position prior to Spell-Out.⁹ Hence the structure in (26a) is linearized as (26b) in accordance with the LCA.



Second, Citko's system crucially depends on the notion of simultaneous derivations. In comparing her own analysis with Nunes' (2004) sideward movement analysis, Citko points out that the sideward movement violates Activity Condition (Chomsky (2001)).

(27) [_{CP}[*which paper*]¹ did+Q [_{&P}[_{TP} John file [*which paper*]²][_& and [_{TP} Mary read [*which paper*]³]]]]] (Nunes (2004):3.5.1)

In the sideward movement analysis, the "shared" element *which paper* enters the derivation by merge with *read*. After undergoing ordinary checking relations within the lower TP, *which paper* moves to the complement position of *file*. Since the Case feature of this DP is checked and deleted within the lower TP, it should be inactive at the stage where sideward movement applies. Citko further claims that the problem does not arise under the parallel merge approach, since α and β *simultaneously* merge with γ in (25).¹⁰ This in turn has a theoretical implication in broad interest; the computational system can manipulate more than one tree simultaneously in the strict sense.¹¹

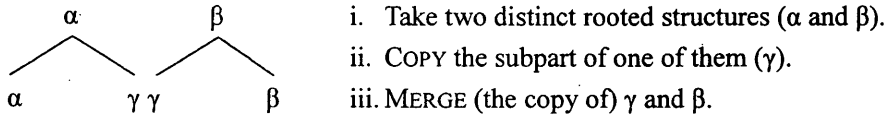
⁹ This assumption implies that an empty category such as PRO and *pro* can stay in-situ even in the multiply dominated position, a point we ignore here.

¹⁰ This assumption departs from Citko's original formulation of parallel merge, since at the moment γ merges with α and β , it is not a subpart of either α or β . Parallel merge simply reduces to simultaneous external merge.

¹¹ Conceptualization along this line leads us to Chomsky's (2004) theory of adjunction by different planes.

Another question with regard to the relation between parallel merge and sideward movement, as Wilder (1999) has already pointed out, is to what extent the two approaches are empirically different. Recall Citko's original characterization of parallel merge as a combination of internal and external merge in (25). If we interpret the operation Move as "copy theory" as in (24a) rather than as "internal merge" as in (24b) in the sense of Citko, then her "parallel merge" will result in a structure in (25').

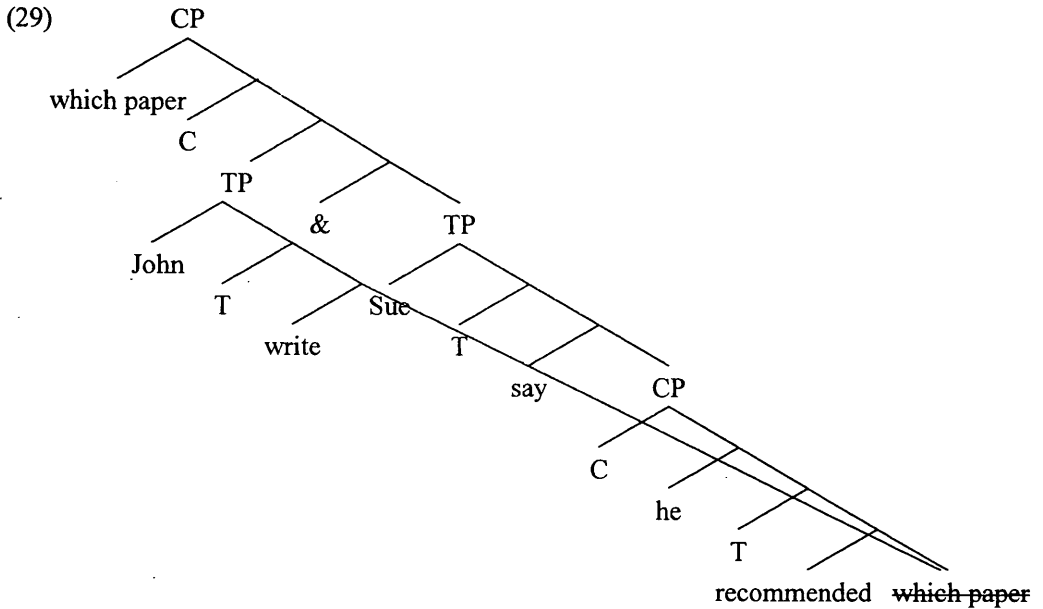
(25') *"Parallel Merge" under the copy theory: Sideward Movement*



The structure is the same as the one formed by sideward movement in its essence; take an element (γ) within a syntactic object (α) and copy it to another syntactic object (β). Activity problem discussed above put aside, there seems to be no clear empirical evidence that favors one over another. It can be argued that under the sideward movement theory, each of the two copies can go through subsequent syntactic operations independently, hence bearing different consequences for e.g., scope-taking and reconstruction phenomena. In this case we would have to allow doubly headed, or "forked" chains in grammar. But the issue largely depends on how one would implement the theory, largely a technical matter. Possible evidence that argues against the "parallel" version comes from sentences like (28).

- (28) a. a boy who_i [_{TP} John invited e_i] and [_{TP} Sue said [_{CP} e_i came to the party]]
 b. Which paper_i did [_{TP} John write e_i] and [_{TP} Sue say [_{CP} (that) he recommended e_i]]?

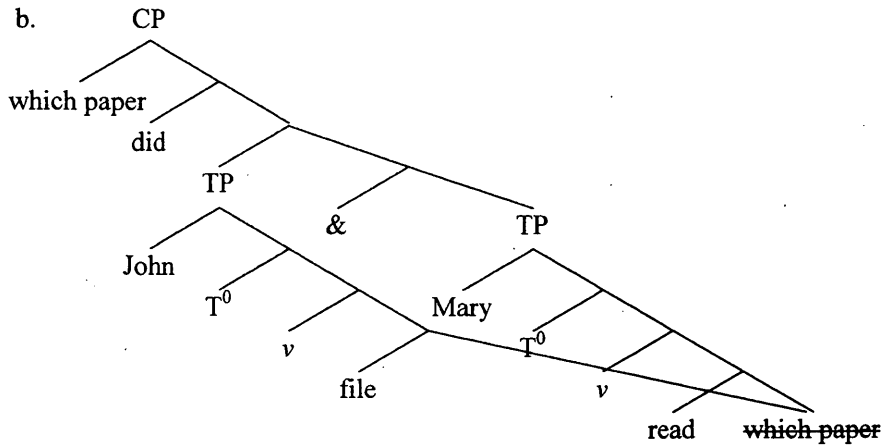
Sentences of the type (28a) show the so-called embedding effect (Franks (1992)), the nature of which is irrelevant here. What is important to the present discussion is that we must allow long distance *wh*-movement from within the second conjunct in order to maintain the parallel merge analysis. The structure of (28b) under the parallel merge approach looks like:



Since the first conjunct is a simple sentence, *which paper* moves to the matrix SpecCP in a single step (the status of the intermediate SpecvP is immaterial here; see note 12). This movement also forces the *wh*-movement from the second conjunct to occur in a single step, skipping the embedded SpecCP position. To avoid such a long-distance movement, the two *wh*-movements must occur independently, forming a split chain. Thus asymmetric ATB constructions like (28) favor the sideward movement approach.

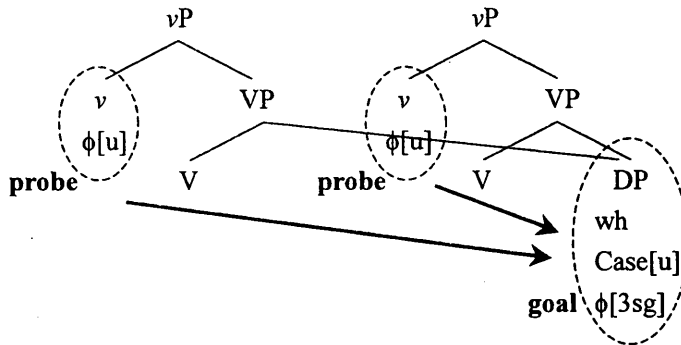
The most relevant aspect of Citko's analysis to the present discussion is the way she treats the matching effect in ATB *wh*-questions. Citko claims that her analysis can explain the case-matching phenomenon in a straightforward manner, since the two verbs *file* and *read* literally share the *wh*-element in the object position in (30).

(30) a. Which paper did John file and Mary read?



Citko adopts Chomsky's probe-goal framework and assumes that the two small *v*'s probe for *which paper* and enter into Agree relation simultaneously, schematically shown in (31).¹²

(31)



This results in an inverse multiple agree relation and allows the DP to be doubly Case-valued. In short, this amounts to stipulating that multiple valuation (of an uninterpretable feature) is allowed as long as (i) the relevant Agree relations holds simultaneously,¹³ and (ii) there are no conflicts in the resulting values. The object DP *which paper* in (30) is Case-valued by two small *v*'s, as Case[ACC/ACC], which does not cause a featural conflict, and the derivation proceeds without crashing at this stage.

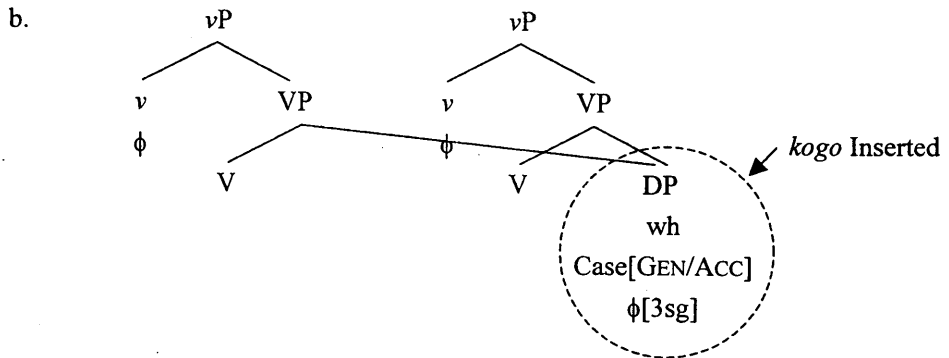
How, then, does Citko's system account for the morpho-phonological nature of the

¹² Citko (2005) does not explicitly adopt the notion of a "phase" and assumes direct movement of the *wh*-element into SpecCP position. Adoption of phase theory will necessitate "parallel internal merge" to the intermediate SpecvP position. See also note 10.

¹³ Terminological clarification. The notion of "simultaneous" used here differs from that of Chomsky's (2004). In the latter, the term refers to operations *within* a single phase; in the former, the term refers to operations *across* different phases.

matching phenomenon observed in the previous sections? Here Citko employs Distributed Morphology, and assumes that what syntax really deals with is a bundle of features with no phonological information, and that the actual morphological form is inserted after the syntactic component according to the featural values that the bundle has obtained in the course of the derivation (late lexical insertion). In principle, a single feature (e.g. Case feature [uCase]) can carry more than one value (e.g. [ACC/DAT]), regardless of whether those values are identical or not. The derivation converges if there is an appropriate form that is compatible with those values at the stage of lexical insertion; if not, it crashes. This is the morpho-phonological nature of the matching phenomenon. Thus (32b) is the partial structure of Polish ATB sentence like (32a).

- (32) a. Kogo Jan nienawidzi e a Maria lubi e?
 whom_{GEN/ACC} Jan hates e_{GEN} and Maria likes e_{ACC}
 'Whom does Jan hate and Maria like?'



4.2. Extension to Other Constructions

Although Citko's (2005) analysis depends exclusively on ATB questions, it has clear consequences for the analysis of other argument-sharing constructions, especially of those that show the matching effects. Let us shortly examine whether and to what extent the parallel merge approach could be extended to Parasitic Gap (PG) and Right Node Raising (RNR) constructions. I will discuss the analysis of FRs in subsequent sections.

Extension of the parallel merge approach to PG construction seems to be possible at first glance, as Nunes' (2001) primary concern is with PG construction. As discussed in the preceding section, we can reanalyze Nunes' system in terms of parallel merge. The most prominent difference between ATB and PG is that the latter involves adjunction structure.¹⁴

¹⁴ The "anti-c-command condition" between the two gaps is already replicated in ATB construction by

Under Chomsky's (2004) "different plane" theory of adjunction, sideward movement can be reinterpreted as a movement across the planes. Under the parallel merge approach, it is not clear how we can account for the simultaneity of parallel merge across the matrix and the adjunct elements. I will elaborate the sideward movement approach within the phase-based framework in the discussion on FRs below.

As for RNR, the situation is somewhat different from other constructions. While RNR typically involves a coordinate structure like ATB, it does not involve *wh*-movement, and a raising operation targets (at least superficially) the right edge of the structure. Because of these peculiarities of RNR, it is implausible to analyze RNR as a parallel merge structure that needs a subsequent movement of the shared element, and thus favor some version of a deletion approach (e.g. Wilder (1997)). Notice also that Franks' (1992) explanation of ATB and PG by null operator licensing does not account for RNR, which presumably does not involve a null operator.

5. Syntax of FRs & the PF-Nature of the Matching Phenomenon

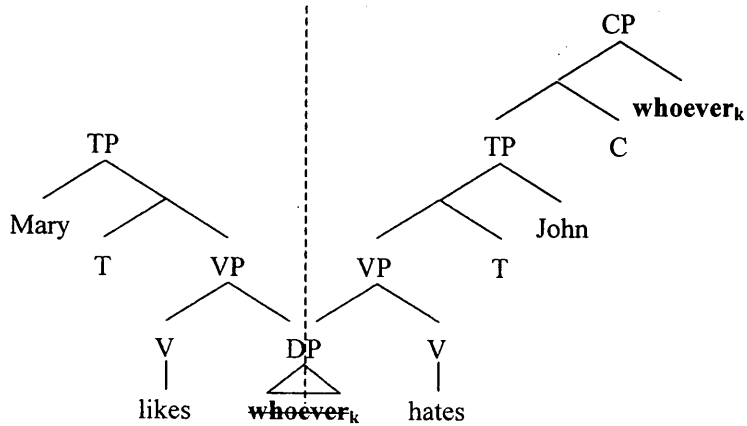
In this section I will argue that Citko's (2005) parallel merge approach does not account for the syntax of FRs and that a modified version of Nunes' (2001, 2004) sideward movement approach is viable, in combination with a "deletion" operation in terms of PF-economy.

5.1. *Parallel Merge Does Not Work for FRs*

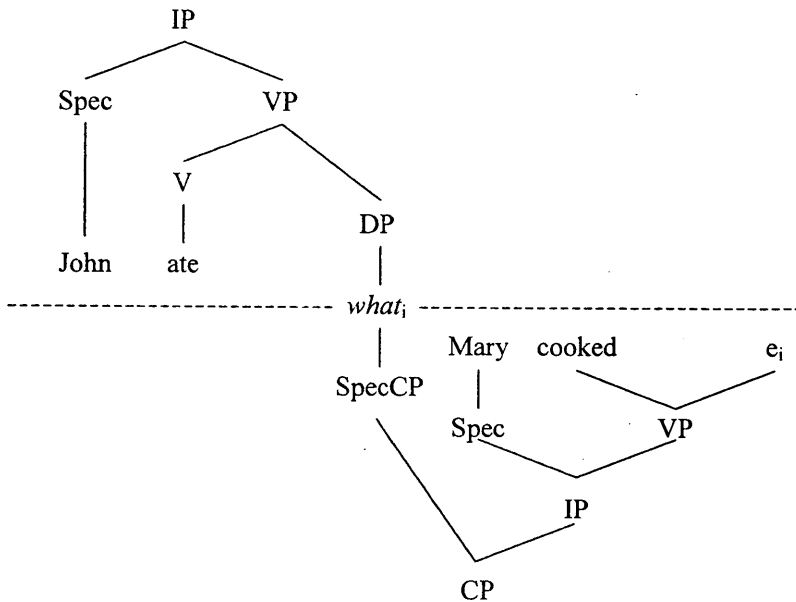
Recall that Citko's system depends on the movement of the shared element over the conjuncts. In other words, violation of the Single Mother Condition is "tolerated" as long as the structure "closes off" into the single mother node before Spell-Out. Argument-sharing constructions such as ATB, RNR and PG might be dealt with in a unified fashion in this sense since they all involve conjunctions, whether they form coordination or subordination. FRs, however, do not employ any conjunctions but rather the relative pronoun itself functions as a matrix argument. The point is illustrated by the multiple dominance structures proposed for FRs: (33) by Citko (1998) and (34) by van Riemsdijk (2005). Notice that both of the structures violate the Single Mother Condition at the end of the derivation, hence linearizing these structures needs stipulative assumptions. Citko states that which one of the clauses is linearized first is determined by clause-type. Van Riemsdijk also states that which clause counts as the root node (in his terms Stock) is somehow determined, and that the subordinate clause (called Scion or Graft) is linearized adjacent to the shared constituent (called Callus).

adopting the &P analysis.

(33)



(34)



Citko's system does not explain the linear order between the matrix PP *to the library* and the FR as in (35), because in her analysis the FR clause can never "cut into" the matrix constituents. Nor does it explain the derivation of subject FRs such as (36). The latter forms forked chains if we assume the *wh*-element is shared at the base θ -position in each clause: at the object position of *wrote* in the FR and at the subject position of *disappointed* in the matrix clause.¹⁵

¹⁵ But see e.g. Izvorski (1997) for syntactic differences between subject FRs and object FRs.

- (35) You should return [what you finished reading] to the library.
 (36) [What_i I wrote t_i] seems to have t_i disappointed you.

In addition to the theoretical problems, there is an empirical problem. As is clear from (37), the matrix argument c-commands an argument inside the FR.

- (37) a. Every student_i can invite whoever he_i likes to the party.
 b. *He_i can invite whoever John_i likes to the party.

The c-command relation across the matrix clause and the FR cannot be defined under multiple dominance structures such as (33) or (34). These considerations suggest that we should employ a more "conservative" structure with ordinary embedding relations, satisfying the Single Mother Condition.

5.2. Sideward Movement as Lexical Copying

To sum up the discussion so far, we want to build a mechanism that features the following intuitive ideas:

- (38) a. The two clauses "share" the element, but not via null operators nor by multiple dominance.
 b. The "shared" element must be able to undergo derivational steps independently in the two clauses.
 c. The derivations of the clauses must be parallel or simultaneous in a sense to be made precise.

To implement these ideas, I will present a rough sketch of a new approach that will (i) adopt Nunes' (2001) sideward movement analysis in a version reinterpreted as Lexical Copying, and (ii) extend Chomsky's (2004) theory of adjunction with different planes. The former allows forked chains in grammar, as well as the "simultaneity" of their formation. The latter guarantees that the resulting structure does not violate the Single Mother Condition. Let us look at how this system would work.

First, I adopt the probe-goal mechanism and derivation by phase, especially the notion of lexical subarray (Chomsky 2000, 2001, 2004). Lexical array (formerly called Numeration) is divided into smaller sets of lexical items, with one phase head (C and v*) per one subarray. Then the sentence in (39a) will have the lexical subarrays in (39b) at the "entrance" of the

derivation.¹⁶

- (39) a. John read what Bill wrote.
b. sPh1 {v*, Bill, wrote, what} fPh1 {C, T} :FR clause
 sPh2 {v*, John, read, ___} fPh2 {C, T} :matrix clause

I further assume that the derivational system builds up the trees in parallel from the lexical subarrays selected from the lexicon. The derivation of sPh1 goes without problems. In sPh2, on the other hand, the internal argument of *read* is absent in the subarray and will crash if we begin structure-building with this subarray. Only at this moment, I propose, is it possible to "copy" the necessary element from other subarrays; in this case *what* in sPh1. I will call this operation Lexical Copying. Lexical Copying is characterized as follows:

- (40) a. Lexical Copying occurs across lexical subarrays (i.e. phases).
b. Lexical Copying is allowed only when the original subarray is short of necessary arguments.¹⁷
c. The resulting copy has a full set of features, identical to the original copy.

Notice that Lexical Copying creates multiple occurrences of a single item, hence the identity at semantic interface is guaranteed. One of minor consequences for the theory might be that Pure Merge Condition on arguments is now stated in terms of lexical subarrays. That is, a lexical subarray must contain all the arguments needed by the predicate of that subarray prior to syntactic derivation. Note also that this operation is in effect a lexical variant of Nunes' sideward movement operation.

After Lexical Copying, derivation starts, in an ordinary fashion. At some point of derivation, the structures of the matrix and the FR clauses look like (41b). By (40c), each of the two occurrences of *what* undergoes ordinary Agree operations in its clause. Recall that we are assuming that the computational system can deal with more than one tree simultaneously.¹⁸

¹⁶ sPh stands for "substantive Phase" and fPh for "functional Phase." The distinction between the two is of little relevance here, though.

¹⁷ Obviously, this way of formulating Lexical Copying has a Last-Resort flavor. Lexical Copying might cause (though local) look-ahead problem. Clarification of these issues is a topic for future research.

¹⁸ This assumption is needed in any case, if we strictly follow the phase-based approach. Consider sentences like (i), where the subject DP contains additional phases. The derivational system must (cyclically) spell-out the phase(s) contained the subject DP before that subject DP is merged into the clausal skeleton. Whether there exists ordering (whether temporal or logical) in derivational steps (i.e. between phases) will be a topic of broader theoretical interest.

- (i) [_{TP} [_{DP} The news [_{CP} that John has passed the exam]] surprised Mary.]

(44) German

- a. Der, wen Gott schwach geschaffen hat muss klug sein.
D_{NOM} WH_{ACC} God weak created has must clever be
'Who God has created weak must be clever.'
- b. Der, wem Gott keine Kraft geschenkt hat muss klug sein.
D_{NOM} WH_{DAT} God no strength given has must clever be
'Who God has given no strength must be clever.'

Data like these can be explained straightforwardly if we assume a last resort mechanism, call it "*d*-support," which realizes mismatching case morphology and saves the derivation from crashing.

° The innovation in the present approach is the operation Lexical Copying. The approach sketched here is motivated largely by theoretical concerns. It is interesting to see, however, that FRs in Dutch and German can be introduced only by simplex *wh*-elements. Complex *wh*-elements cannot introduce FRs.

(45) German

- a. Ich esse was Maria mir serviert.
I eat what Maria me serves
'I eat whatever Maria Maria serves me.'
- b. *Ich esse welche Speisen Maria mir serviert.
I eat whichever dishes Maria me serves
'I eat whatever dishes Maria serves me.'
- (van Riemsdijk (2005))

This restriction makes sense if we assume that the *wh*-element is lexically copied. Since Lexical Copying occurs at the level of lexical (sub)arrays, it cannot copy units larger than lexical items, i.e., phrases. The simplex condition on *wh*-elements, then, can be taken as supporting evidence for the Lexical Copying approach. Note also that most of the previous approaches to FRs do not deal with this aspect of FRs.^{21,22}

²¹ But see Donati (2006) and Chomsky (2005b) for the explanation by the "move-and-project" approach to FRs. In a nutshell, the *wh*-element that introduces an FR merges in the base position inside the FR and moves to "SpecCP" of that clause, then projects to DP only when it is simplex, that is, when it is a head. This theory seems to allow label-superimposing or multiple labeling in grammar, but I will skip the detailed examination of them for reasons of time and space.

²² As for FRs in English introduced by *wh*-phrases like (i), I follow Donati (2005) in assuming that the phrasal elements (except for the simplex *wh*) base-generates in the head DP headed by *-ever* and concatenates with lexically copied occurrence of *wh*, much in the Kaynean fashion.

- (i) a. I will read whatever book you write.
b. [_{DP} what-ever book [_{CP} ~~what~~ C⁰ you wrote ~~what~~]]

6. Conclusion

In this paper I have argued first that the case matching phenomenon is not an idiosyncratic property of FRs but rather one generally found in argument-sharing constructions. The morpho-phonological nature of the matching phenomenon, as well as syntactic differences among these constructions, makes it implausible to analyze them in terms of a syntactic mechanism such as feature agreement. In the light of these considerations I have examined recent approaches to the matching phenomenon and the syntax of FRs both in theoretical and in empirical terms. An alternative approach is sketched that would overcome the problems, incorporating insights of Nunes' (2001, 2004) sideward movement theory and of Chomsky's (2004) theory of adjunction.

Before closing, let us evaluate our approach with previous approaches. Theoretically speaking, the approach presented here introduces only one operation, Lexical Copying. This operation is restricted by the conditions in (40). Citko's (2005) parallel merge is less restricted in this respect, since it allows multiple dominance relation anywhere in principle, as long as the appropriate phonological form is available at lexical insertion. Under the present approach, we can maintain the probe-goal mechanism driven by phase heads without recourse to the multiple valuation of uninterpretable features in Agree relation, which leaves one-to-one correspondence between ϕ -agreement and Case-valuation as in the standard form.

The explanation of the simplex condition on *wh*-elements of FRs in terms of Lexical Copying, however, misses the cross-categorical generalization on the matching phenomenon in turn. Since argument-sharing constructions such as ATB, PG, and RNR allow complex elements to be shared, they cannot be generated by Lexical Copying, at least in a straightforward manner. We might overcome the problem by elaborating the theory of copying across the trees to include both phrasal and lexical elements, which involves a reevaluation of Nunes' original theory of sideward movement, with a modification of the Activity problem discussed in section 4. Or we might simply state that they are in fact computed differently in the syntactic component, and that the matching phenomenon is solely derived from PF economy considerations on chains. The latter path seems to me to be the better way to pursue, at this stage of the investigation.

References

- Bresnan, Joan and Jane Grimshaw (1978) "The Syntax of Free Relatives in English," *Linguistic Inquiry* 9, 331-391.
- Chomsky, Noam (1981) *Lectures on Government and Binding*, Mouton de Gruyter, Berlin.
- Chomsky, Noam (1995) *The Minimalist Program*, MIT Press, Cambridge, MA.

- Chomsky, Noam (2000) "Minimalist Inquiries: The Framework," *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michaels and Juan Uriagereka, 89-155, MIT Press, Cambridge, MA.
- Chomsky, Noam (2001) "Derivation by Phase," *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 1-52, MIT Press, Cambridge, MA.
- Chomsky, Noam (2004) "Beyond Explanatory Adequacy," *The Cartography of Syntactic Structures vol.3: Structures and Beyond*, ed. by Adriana Belletti, 104-131, Oxford University Press, Oxford.
- Chomsky, Noam (2005a) "Three Factors in Language Design," *Linguistic Inquiry* 36, 1-22.
- Chomsky, Noam (2005b) "On Phases," ms., MIT.
- Citko, Barbara (1998) "ATB Analysis of Free Relative Clauses," *CLS 34: The Main Session*, 69-81.
- Citko, Barbara (2004) "On Headed, Headless, and Light-Headed Relatives," *Natural Language & Linguistic Theory* 22, 95-126.
- Citko, Barbara (2005) "On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge," *Linguistic Inquiry* 36, 475-496.
- Donati, Caterina (2006) "On *Wh*-Head-Movement," *Wh-movement on the Move*, ed. by Lisa Cheng and Norbert Corver, 21-46, MIT Press, Cambridge, MA.
- Franks, Steven (1992) "A Prominence Constraint on Null Operator Constructions," *Lingua* 88, 1-20.
- Groos, Anneke and Hend van Riemsdijk (1981) "Matching Effects in Free Relatives: A Parameter of Core Grammar," *Theory of Markedness in Generative Grammar*, ed. by Adriana Belletti, Luciana Brandi and Luigi Rizzi, 171-216, Scuola Normale Superiore, Pisa.
- Hirschbühler, Paul and Maria-Luisa Rivero (1983) "Remarks on Free Relatives and Matching Phenomena," *Linguistic Inquiry* 14, 505-520.
- Inada, Shun'ichiro (2006: this volume) "Two Types of Relative Clauses in Germanic and Romance Languages," *Linguistic Research* 22, University of Tokyo English Linguistic Association.
- Izvorski, Roumyana (1997) "Subject Free Relatives in Null-Subject Languages: Evidence from Slavic," *Formal Approaches to Slavic Linguistics*, ed. by Wayles Brownie, Ewa Dornisch, Natash Kondrashove and Draga Zec, 232-245, Mich, Ann Arbor.
- Kanye, Richard (1994) *The Antisymmetry of Syntax*, MIT Press, Cambridge, MA.
- Kathol, Andreas (2001) "On the Nonexistence of True Parasitic Gaps in Standard German," *Parasitic Gaps*, ed. by Peter W. Culicover and Paul M. Postal, 315-338, MIT Press, Cambridge, MA.
- Nunes, Jairo (2001) "Sideward Movement," *Linguistic Inquiry* 32, 303-344.
- Nunes, Jairo (2004) *Linearization of Chains and Sideward Movement*, MIT Press, Cambridge,

MA.

- van Riemsdijk, Henk (2005) "Free Relatives," *The Blackwell Companion to Syntax: vol. II*, Martin Everaert and Henk van Riemsdijk, 339-382, Blackwell, Oxford.
- Sag, Ivan (2003) "Coordination and Underspecification," *The Proceedings of the 9th International Conference on HPSG*, 267-291.
- Sauerland, Uli (1996) "The Late Insertion of German Inflection," ms., MIT.
- Suñer, Margarita (1984) "Free Relatives and the Matching Parameter," *The Linguistic Review* 3, 363-387.
- de Vries, Mark (2002) *The Syntax of Relativization*, Ph.D. Dissertation, University of Amsterdam.
- Wilder, Chris (1997) "Some Properties of Ellipsis in Coordination," *Studies on Universal Grammar and Typological Variation*, ed. by Artemis Alexiadou and T. Allan Hall, 59-107, John Benjamins, Amsterdam.
- Wilder, Chris (1999) "Right Node Raising and the LCA," *WCCFL* 18, 586-598.