

Edges and Linearization

by

Tue H. Trinh

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

ARCHIVES

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2011 [September 2011]

© Massachusetts Institute of Technology 2011. All rights reserved.

Author

↓ Department of Linguistics and Philosophy
// August 10, 2011

Certified by
Noam Chomsky
Professor
Thesis Supervisor

Certified by
Danny Fox
Professor
Thesis Supervisor

Certified by
.....
Irene Heim
Professor
Thesis Supervisor

Certified by ...
.....
David Pesetsky
Professor
Thesis Supervisor

Accepted by
.....
David Pesetsky
Department Head
Linguistics Section

Edges and Linearization

by

Tue H. Trinh

Submitted to the Department of Linguistics and Philosophy
on August 10, 2011, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis is concerned with how grammar determines the phonological consequence of syntactic dislocation. It centers on a hypothesis regarding the linearization of movement chains - the Edge Condition on Copy Deletion, eventually named the Edge Condition in the last chapter, when it receives its final formulation. The empirical phenomena under investigation include (i) predicate cleft constructions in German, Dutch, Hebrew, Vietnamese, Swedish and Norwegian, (ii) NP-split constructions in Vietnamese and (iii) cross-linguistic variation in head ordering patterns.

Thesis Supervisor: Noam Chomsky
Title: Professor

Thesis Supervisor: Danny Fox
Title: Professor

Thesis Supervisor: Irene Heim
Title: Professor

Thesis Supervisor: David Pesetsky
Title: Professor

Acknowledgments

I owe special thank to the members of my committee: Noam Chomsky, Danny Fox, Irene Heim and David Pesetsky. I am convinced that they are among the best people on earth, both with respect to their intelligence, and with respect to their humanity. What they have taught me goes way beyond the pages of this dissertation, and will be with me for the rest of my life.

I am grateful to Luka Crnić, my classmate and faithful friend, for the many hours we spent on laughing at the same things instead of working on our “joint project” as we were supposed to.

During the last five years, I have received support and encouragement from many people. I would like to especially mention Gennaro Chierchia, Michelle DeGraf, Kai von Fintel, Sabine Iatridou, Shigeru Miyagawa, Norvin Richards and Hedde Zeijlstra. My fellow students have been helpful and generous, and also extraordinarily tolerant. I thank them all sincerely.

Last but not least, I thank those who give me the security of unconditional love, without which waking up in the morning is pointless. They know who they are.

Contents

1	Introduction	9
1.1	Two patterns of pronunciation	9
1.2	The content of the base position	11
1.2.1	Trace Theory	11
1.2.2	Copy Theory	15
1.3	The Edge Condition on Copy Deletion	17
1.4	Overview of the dissertation	19
2	Predicate cleft constructions	21
2.1	Introduction	21
2.2	Type 1: Hebrew and Vietnamese	31
2.2.1	Hebrew	32
2.2.2	Vietnamese	37
2.2.3	Interim summary	40
2.3	Type 2: German and Dutch	41
2.3.1	German	42
2.3.2	Dutch	47
2.3.3	Interim summary	49
2.4	Type 3: Swedish and Norwegian	50
2.4.1	Swedish	51
2.4.2	Norwegian	57
2.4.3	Interim summary	58
2.5	Head-to-head movement	59
2.5.1	The typology	59
2.5.2	Head-to-head movement as a PF operation	61
2.5.3	The [$\pm$ V doubling] parameter	70
2.6	Conclusion	72
3	NP-Split constructions	73
3.1	Introduction	73
3.2	Relational vs. non-relational nouns	78

3.2.1	Cases of non-optionality	78
3.2.2	Cases of optionality	81
3.3	Measure words	98
3.3.1	Semantics	98
3.3.2	Split and modification	101
3.4	Mandarin Chinese	105
3.5	Conclusion	107
4	Constraining headedness	123
4.1	Introduction	123
4.2	The Final-Over-Final Constraint	126
4.2.1	Biberauer et al (2010)	126
4.2.2	The Head Ordering Generalization	130
4.2.3	Deriving the HOG	131
4.3	Conclusion	135

Chapter 1

Introduction

1.1 Two patterns of pronunciation

The ultimate aim of this thesis is to deepen our understanding of what is sometimes called the “displacement property” of natural language (cf. Chomsky 1995, 2000, 2004). The basic observation behind this designation is that linguistic expressions can be pronounced at one place and function as if they are in others. Illustration is provided by the English sentence **whom do you like**,¹ where the following can be observed of the word **whom**: (i) it is assigned accusative case, (ii) it bears the thematic role of the liked person, i.e. the person receiving the addressee’s affection, and (iii) it satisfies the requirement that **like** have a direct object, in the sense that its absence would cause the sentence to be deviant in the same way that ***you like** is deviant. These are just the attributes associated with the position of **her** in **you like her**. One informal way to describe the facts, then, is to say that **whom** is present at two places: the post-verbal position, where it acquires the aforementioned attributes, and the clause-initial position, where it is pronounced.

Such “double existence” phenomena are attested in every language, and have been given various theoretical treatments in the course of generative grammar’s history. Uniting them, nevertheless, is the idea that sentences are derived by successive application of rules mapping one syntactic object to another, and that principles of grammar may apply to representations constructed at different points of the derivation. The theory proposed in Chomsky (1965), for example, distinguishes between the deep structure of a sentence, which determines its meaning, and the surface structure, which determines its sound. It is on the basis of the deep

¹Object language expressions will be underlined in the text.

structure of **whom do you like**, which is approximately **you like whom**, that **whom** is identified as the direct object of **like**, assigned accusative case and given the appropriate thematic role. Application of syntactic rules to **you like whom** will yield **whom do you like**, the surface structure, which serves as input to phonetic interpretation.² The mapping from deep to surface structure, in this case, has an effect on **whom** which warrants the term “movement,” or “displacement”: **whom** disappears from one position (its base position) and reappears in another (its derived position).

(1) you like whom → whom do you like

Thus, “displacement” is the name given to a phenomenon – i.e. one of linguistic expressions being pronounced in one place and performing functions dedicated to another – which reveals how this phenomenon is modeled in the theory, or more precisely in Chomsky (1965) (cf. also Chomsky 1955, 1957, 1964). Thirty years after the publication of this work, the phenomenon is still called “movement,” but its conceptualization has undergone a change. Instead of (1), we have (2).

(2) you like whom → whom do you like whom → whom do you like ~~whom~~

This, of course, is the Copy Theory of Movement (CTM), which analyzes movement of a constituent X as a sequence of two separate operations (cf. Chomsky 1993, 1995, Gärtner 1998, Sauerland 1998, 2004, Fox 1999, 2000, 2002, Corver and Nunes 2007). The first, call it Form Chain, copies X into the derived position, forming a chain (α, β) where α is the higher (i.e. c-commanding) and β the lower (i.e. c-commanded) copy of X. The second operation, Copy Deletion, maps (α, β) into (α, β) : it deletes the lower copy, making it invisible to the phonology.

A question that arises naturally in the context of the CTM is then whether cases exist in which Form Chain applies but Copy Deletion does not, i.e. cases where a constituent exhibits properties of moved elements and at the same time is pronounced at both the derived and the base position. Several recent works have concluded that this question is to be answered in the affirmative (cf. Nunes 2003, 2004, Fanselow and Mahajan 1995, Fanselow 2001, Grohmann 2003, Grohmann and Nevins 2004, Grohmann and Panagiotidis 2004, Hiraiwa 2005, Martins 2007, Cheng 2007, Vicente 2005, 2007, 2009, Kandybowicz 2006, 2007, 2008, 2009, among others). The conclusion is backed by examples of “doubling” such as the following Spanish sentence, taken from Vicente (2007:7).

²The rules are (i) wh-movement, (ii) T-to-C movement (Subject-Auxiliary-Inversion), and (iii) do-support.

- (3) Jugar Juan suele jugar al futbol los domingos
 play.INF Juan HAB.3SG play.INF at football the sundays
 ‘As for playing, Juan usually plays football on Sundays’

Vicente shows that topicalization of verbs in Spanish is subject to the same locality conditions as, say, wh-movement in English. Nevertheless, the moved verbs are pronounced twice, at [Spec,C] and inside TP. Similar examples of doubling have been provided for other languages in the works cited above. Now given that doubling exists, a second question poses itself: when does it exist? This thesis proposes a partial answer to this question. The rest of this chapter will be devoted to setting up the theoretical background for the formulation for this answer, as well as presenting the answer itself. The chapters that follow justify it with empirical observations.

1.2 The content of the base position

One would think that doubling, which turns out to be attested in many languages, should have been the first class of data pointing linguists to the Copy Theory of Movement. What indicates the existence of copies more clearly than the fact that we hear them? Curiously, this is not the case. It requires sophisticated arguments to reach the conclusion that movement leaves something at the base position, and several years later, that that thing is a full-fledged copy of the moved constituent. These arguments, interestingly, did not involve facts about doubling at all. Let us consider some of them.

1.2.1 Trace Theory

The hypothesis that the movement leaves something behind was first advanced in the context of the debate on whether UG should contain “global rules.” A global rule maps a structure σ to another structure σ' under the condition that σ has a certain derivational history. More formally, the domain of global rules consists not of syntactic representations, but of sequences of these. Addition of these rules would greatly increase the expressive power of grammars, thereby magnifying the problem of learnability, hence requires empirical justification. Lakoff (1970) provides several cases to this end, which in general are intended to show that

mapping of one structure to another can depend on how the first is derived. One is the paradigm in (4).³

- (4) a. who do you want to succeed
b. who do you wanna succeed

The string **want to** has been contracted to **wanna** in (4b). Lakoff observes that **succeed** in (4a) is ambiguous between ‘be successful’ and ‘replace,’ which is due to the fact that (4a) can be derived by wh-movement either from (5a) (the ‘be successful’ meaning) or from (5b) (the ‘replace’ meaning).

- (5) a. you want who to succeed
b. you want to succeed who

The same ambiguity is not detected in (4b), in which **succeed** only has the meaning of ‘replace.’ This is evidence that **want to** cannot be contracted to **wanna** if **who** intervenes between **want** and **to** at some point in the derivation. Thus, the rule of wanna-contraction must take the transformational history of the structure containing **want to** into account: it is a global rule. Another similar datum presented in Lakoff (1970) in support of global rules is the contrast in (6).

- (6) a. I know that the concert’s at two o’clock
b. *I know where the concert’s at two o’clock

The observation is that the auxiliary **is** cannot be reduced to **s** (and cliticized onto the preceding word) if at some point in the derivation it is followed by a phrase which subsequently undergoes movement. The reason (6b) is ungrammatical, then, is that at one point in its derivation, **where** follows **is**. The rule of auxiliary reduction, Lakoff concludes, must therefore be a global rule, as it must refer to the structure which underlies the structure it affects, i.e. the transformational history of the latter.

In their reply to Lakoff (1970), Baker and Brame (1972) suggest an account for wanna-contraction and auxiliary reduction which does not necessitate global rules. Specifically, they propose that movement should be represented in such a way that it “leave a special feature or boundary symbol behind in the place formerly occupied by the moved constituent. Lowering of stress on the auxiliary and the contraction of **want to** are then blocked by the presence of the feature in question.” This idea, of course, is the origin of Trace Theory, according to which formatives of a special sort, (indexed) traces, are left at the base position of moved

³Lakoff actually uses examples involving relative clauses instead of questions, but the point made is the same.

elements. Application of wh-movement to **you want who to succeed**, in which **succeed** has the meaning of ‘be successful,’ and to **you want to succeed who**, where it means ‘replace,’ would yield (7a) and (7b), respectively.

- (7) a. who₁ do you want *t*₁ to succeed
 b. who₁ do you want to succeed *t*₁

While **want** and **to** can contract to **wanna** in (7b), the trace of **who** prevents this from happening in (7a). This is why when wanna-contraction does apply, **succeed** can only be construed as ‘replace.’ The ungrammaticality of (6b) can be explained similarly: the trace of **where** blocks reduction (and cliticization) of the auxiliary **is**. Trace Theory, then, allows a way to capture “global” facts while keeping the Markovian character of syntactic transformations. It does this by providing the possibility for structures to retain aspects of their transformational history, specifically the information on which constituents have moved from where.

The attempt to enhance the explanatory power of linguistic theory also includes simplifying the rules of particular grammars by extracting generalizations from their complexities and reformulating these as principles of UG. The idea is to minimize what has to be learned (i.e. what needs to be written into the individual rules) by maximizing what does not (i.e. what is attributed to UG). Trace Theory turns out to be conducive to this enterprise as well. Take the Specified Subject Condition (SSC), proposed in Chomsky (1971, 1976), for example.

- (8) No rule can involve X and Y in the configuration ... X ... [_{α} ... Y ...]
 where α is a cyclic node ($\bar{S}$ or NP) containing a specified subject

The SSC is proposed as a universal condition on rules of particular grammars. It simplifies these in the sense that it relieves their formulation of the specification that they do not apply under the said circumstance. Two rules which Chomsky (1976) discusses are Reciprocal Interpretation (RI) which “assigns an appropriate sense to sentences of the form NP ... each other,” and Disjoint Reference (DR) which “assigns disjoint reference to a pair (NP, pronoun).” The first rule applies in **the men like each other**, establishing the anaphoric relation between **the men** and **each other**. The second rule applies in **the men like them**, preventing coreference of the two noun phrases.⁴ Evidence that RI and DR are constrained by the SSC is the fact that anaphoric relation between **the men** and **each other** is impossible in **the men want John to like each other**, showing that RI does not apply, and coreference between **the men** and **them** is

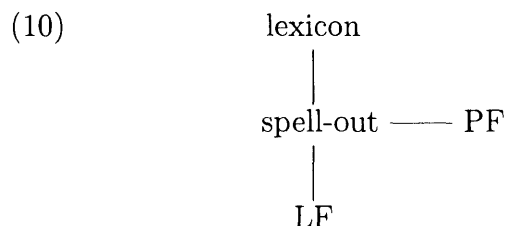
⁴These rules are, of course, the historical antecedents of Condition A and Condition B of the Binding Theory (cf. Chomsky 1981).

not impossible in **the men want John to like them**, showing that DR does not apply. In both cases, the specified subject **John** of the embedded sentence blocks application of the relevant rule. Now consider the pair of sentences in (9).

- (9) a. John₁ seems to the men [_{t₁} to like each other]
 b. The men₁ seem to John [_{t₁} to like them]

It can be observed that RI and DR are blocked in these examples as well: **each other** cannot be related to **the men** in (9a), and coreference of **the men** and **them** is possible in (9b). A plausible explanation is that there is a specified subject in the embedded S of (9a) and (9b), namely the trace of the moved noun phrase. Trace Theory thus “permits otherwise valid conditions to apply, [...] overcoming cases of misapplication of rules [...]” (Chomsky 1976). At the same time, the stipulation that traces are phonologically empty accounts for the fact that the moved element is not heard at the base position.

Trace Theory also reconciles the idea that meaning is read off from a single level of syntactic representation with the “duality of semantics,” i.e. the fact that interpretation of sentences seems to involve two sorts of information, one concerning thematic notions such as agent, theme, location etc, the other notions such as topic, focus, scope. Chomsky (1965) takes deep structure to be the input to semantic interpretation, but it was recognized that surface structure is needed for information of the second kind. Subsequent works assumed that both deep and surface structure contribute to meaning (e.g. Chomsky 1972). Traces allow thematic relations to be conserved throughout the derivation, hence opens the way to models of grammatical description in which both sound and meaning are constructed from a single syntactic structure. One of these is the “minimalist” model proposed in Chomsky (1993, 1995) and elaborated elsewhere.



According to this model, at some point in the derivation the syntactic object which has been constructed will undergo two separate sets of rules, mapping it to syntactic representations interpretable by semantics and phonology, i.e. logical form (LF) and phonetic form (PF), respectively. This will be the model we assume in the discussion that follows.

1.2.2 Copy Theory

While the silence of traces is a stipulation, it is not an unnatural one, given the assumption that traces are “impoverished” constituents, with no internal morphological or syntactic structure. How would t_2 be pronounced? However, this assumption is challenged by a combination of both empirical and theoretical considerations. The former concern facts which show how a subpart of the moved phrase can have effects that it would have if it were in the position of the trace. Consider the following paradigm.

- (11) a. John likes pictures of himself
b. *John’s sister likes pictures of himself
- (12) a. [Which picture of himself]₂ does John like t_2
b. *[Which picture of himself]₂ does John’s sister like t_2

The contrast in (11) derives from Condition A, but how Condition A can account for the contrast in (12), which clearly parallels that in (11), is not obvious, if all there is at the base position of the moved wh-phrase is a monolithic t_2 . One account might be that Condition A applies with respect to the base positions of constituents, but this hypothesis is untenable in the face of such examples as **John₁ seems to himself t_1 to be sick**. The proposal was made, then, that the wh-phrase “reconstructs” into the trace position at LF, where Condition A applies. But this does not seem much more than a reformulation of the problem.

Another set of data which show how the internal structure of moved phrases can be relevant for the evaluation of constituents containing their traces pertains VP-Ellipsis. This phenomenon is exemplified by (13).⁵

- (13) First, John talked to Mary. Then, Bill did.

The VP in the second sentence is elided. Note that this sentence cannot be understood as anything other than ‘Bill talked to Mary.’ This suggests that at LF, the VP of the second sentence must be the same as that of the first. In other words, the LF of the second sentence must be **Bill did <talk to Mary>**, where the angled brackets indicate that the VP inside them is deleted at PF. Fox (2000, 2003) propose the principle of Parallelism: VP₁ can be deleted only if the discourse contains a pronounced VP₂ (the antecedent VP) such that VP₂ is syntactically identical to VP₁. Now consider (14).

⁵The example is taken from Fox (2003).

- (14) a. I know which cities Mary will visit and now I want to know the cities Sue will
 b. *I know which cities Mary will visit, and now I want to know the lakes Sue will

If we assume that the base positions of moved elements contain primitive traces, the LF's of (14a) and (14b) are essentially (15a) and (15b), respectively.⁶

- (15) a. I know which cities₁ Mary will visit t_1 , and now I want to know the cities₂ Sue will <visit t_2 >.
 b. *I know which cities₁ Mary will visit t_1 , and now I want to know the lakes₂ Sue will <visit t_2 >

It is then hard to explain why **visit** t_1 can license the deletion of **visit** t_2 in (15a) but not in (15b). The facts just reviewed will become less of a puzzle if we assume that instead of a trace, movement leaves a copy of the moved constituent at the base position, which is deleted at PF but is present at LF. Thus, the LF's of the sentences in (12) are as follows.

- (16) a. [Which picture of himself]₂ does John like [which picture of himself]₂
 b. *[Which picture of himself]₂ does John's sister like [which picture of himself]₂

The contrast between these two sentences can then be accounted for in terms of the difference between them that emerges when traces are replaced with copies: in the first sentence one copy of **himself** is bound, in the second no copy of this anaphor is. Similarly, an explanation becomes feasible for the contrast between the two sentences in (14), whose LF's are given in (17).

- (17) a. I know which cities₁ Mary will visit which cities₁, and now I want to know the cities₂ Sue will <visit the cities₂>.
 b. *I know which cities₁ Mary will visit which cities₁, and now I want to know the lakes₂ Sue will <visit the lake₂>

Parallelism can then be defined in such a way that **visit which cities₁** can license the ellipsis of **visit the cities₂**, but not **visit the lake₂**.⁷ This task seems, and

⁶The LF's in (i) also requires assuming a specific theory of relative clauses which we will not be able to discuss. See Sauerland (1998) for arguments in favor of this theory.

⁷See Fox (1999, 2002) and Sauerland (1998, 2004) for proposals as to how lower copies are interpreted.

is, much more executable than that of explaining the relevant contrast without copies.⁸

The conceptual argument for copies and against traces consists, essentially, in the idea that syntax should perform operations on lexical items and combinations thereof, without introducing any new element in the course of the derivation. In other word, LF should contain nothing that cannot be retrieved from the lexicon. This is the Inclusiveness Condition of Chomsky (1995). Since traces are not lexical items but entities that spring into existence when movement takes place, they constitute a violation of the Inclusiveness Condition.

1.3 The Edge Condition on Copy Deletion

We have briefly sketched the path from Chomsky (1965) to the Copy Theory of Movement, which shows how the base position of a moved constituent gradually gets filled with content, first with a trace, then with a copy of the moved element. What is remarkable is that this path is totally free of facts about doubling, even though these point towards the CTM in a particularly clear way. Let us now come back to the question at the end of section 1.1: when does doubling exist? Two plausible approaches to this question seem to be the following: (i) we assume that Copy Deletion must apply if it can and ask what blocks it; (ii) we assume that Copy Deletion can apply only if it must and ask what forces it. To the best of my knowledge, the first approach has always been the one chosen. In other word, the presupposition has been that Copy Deletion applies by default, and that cases of doubling are those where this rule is blocked for some reasons. For concreteness, let us assume, for the discussion that follows, the following rule.⁹

- (18) Pronunciation Economy (PE)
Copy Deletion must apply when it can

Consequently, attempts at explaining doubling have been limited to specifying conditions for Copy Deletion and showing that doubling is attested when one or more of these conditions fail to obtain. This thesis is no exception: it proposes a condition on Copy Deletion. This condition is given in (19).

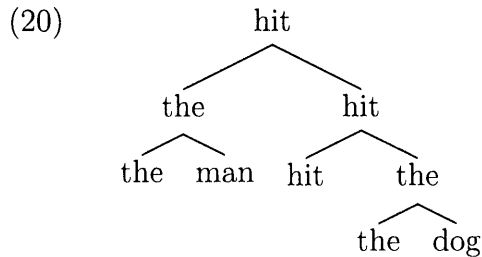
⁸See Sauerland (1998, 2004) for details. These works also provide an account of the acceptability of (i), which undoubtedly has posed a question to the reader by this point.

(i) I know which cities Mary will visit, and now I want to know which lakes she will

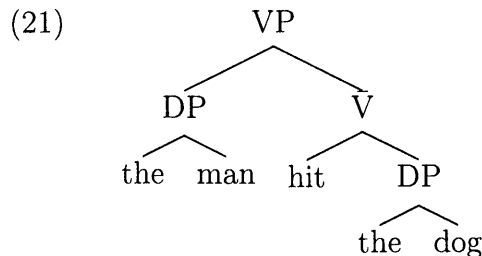
⁹As speaking involves cognitive effort, Pronunciation Economy might be derived from the general principle of minimizing computation (Chomsky p.c.).

- (19) Edge Condition on Copy Deletion (ECCD)
 For any chain (α, β) where α is the higher and β the lower copy, deletion of β requires that β ends an XP

In the following chapters, a range of facts will be discussed with the aim of showing that they support the ECCD, i.e. that they follow from the ECCD, given other independently motivated claims. First, however, it is necessary to clarify the terms used in (19), in particular the notion of an XP. I will take an XP to be a maximal projection in the relational sense, as proposed in Muysken (1982), Chomsky (1994). Following Chomsky (1993, 1994, 1995) and subsequent works, I assume that syntactic structures are built by applying Merge, which maps two syntactic objects α and β to a new syntactic object $\delta = [\gamma \ \alpha \ \beta]$, where γ is the label of δ and is either the label of α or the label of β , with every lexical item being its own label. As an example, consider the following structure.¹⁰



A maximal projection is a constituent which does not project, i.e. a constituent whose mother has a different label than the one it has. Thus, **hit the dog** is not a maximal projection, since its mother, **the man hit the dog**, has the same label as its own. In contrast, **man** is a maximal projection, since it does not project: the label of its mother is **hit**. In what follows, I will continue to use the more traditional notation, switching to the bare phrase structure notation only when it facilitates presentation. Thus, I will often represent (20) as (21), assuming the reader's ability to construct one from the other.



¹⁰The choice of projection labels, presumably, is determined by interface conditions. In other word, it is assumed that the wrong label will lead to crash, for one reason or another.

For now, I assume that a constituent K “ends an XP,” or is “XP-final,” if and only if the last morpheme of K coincides with the last morpheme of an XP. For example, **the dog** and **hit the dog** are XP-final in (20), but not **hit** or **the**.

1.4 Overview of the disseration

The content of this dissertation can actually be summarized in one phrase: to provide evidence for the ECCD. The next chapter discusses predicate clefts in six languages: Hebrew, Vietnamese, German, Dutch, Swedish and Norwegian. It will be shown that variations exist among the different languages as well as among constructions within a single language. For example, doubling is attested in Hebrew but not in German, and internal to Hebrew, doubling is obligatory when the clefted predicate is a transitive verb, optional when it is an intransitive. It will be argued that these variations follow from the ECCD.

Chapter 3 deals exclusively with Vietnamese, specifically with NP-split constructions in this language. Intricate relationships obtain between how pronunciation patterns of these constructions and the semantics of the NP in question. Again, it will be argued that the facts receive a natural account in terms of the ECCD in conjunction with a specific theory of the meaning of nouns in classifier languages. Chapter 3 contains an appendix on the semantics of nominals in Vietnamese and Mandarin Chinese which presents this theory in more details.

The last chapter, chapter 4, argues for a slight revision of the ECCD. Also, a reformulation of this principle is also proposed, motivated by a specific reconceptualization of overt movement. The reformulation is to be the final version of the ECCD in this dissertation, and is named the Edge Condition. It is argued that the Edge Condition provides an elegant account of the Final-Over-Final Constraint (FOFC), introduced in Biberauer et al. (2010) as a condition imposed on phrase structure configuration. Specifically, it is shown that the Edge Condition explains why the FOFC holds where it holds, and why exceptions to the FOFC are found where they are found.

Chapter 2

Predicate cleft constructions

2.1 Introduction

The purpose of this chapter is to provide empirical arguments for the Edge Condition on Copy Deletion (ECCD), introduced in section 1.3. of chapter 1.

- (1) Edge Condition on Copy Deletion (ECCD)
For any chain (α, β) where α is the higher and β the lower copy of the moved constituent, deletion of β requires that β end an XP

The arguments are based on observations concerning predicate clefts in a number of languages. I use the term “predicate clefts” descriptively: it refers to constructions in which a single verb occupies the left-peripheral $\bar{A}$ -position in the clause, and remains neutral as to what the associated semantic/pragmatic effects are as well as whether the verb in question is syntactically a V or a remnant VP. Nevertheless, it would perhaps be useful to start the discussion by considering the original use of this term. To the best of my knowledge, it originated in the discussion of African languages in which it denotes clauses bearing some resemblance to English clefts (cf. Koopman 1984: 37–38, 153–186, Koopman 2000a: 357–374). Consider the following Vata sentences (Koopman 1984: 157).

- (2) a. ml̄ wà ml̄
leave they left
‘They LEFT’
b. nyĒ à nyĒ à nŌ dàlà
give we give our mother money
‘We GAVE money to our mother’

The capitalization of the main verb in the translation of (2a) and (2b) represents “contrastive verb focus,” in the sense that the verb “is understood as contrasting with some verb implicit in the discourse” (Koopman 2000b: 357). Semantically, then, Vata predicate clefts are similar to a cleft sentence in English: (2a) can be rendered as ‘it is leaving that we did.’ With respect to syntactic form, (2a) and (2b) also parallel English clefts in that the focused element is left-peripheral. What distinguishes these sentences from their English counterparts is of course the fact that there are two overt instances of the clefted constituent. In what follows, I will call these V1 and V2, with V1 being the clause-initial instance and V2 the other.

Koopman (1984: 158–161) shows that the relation between V1 and V2 in Vata predicate clefts exhibits properties typical of $\bar{A}$ -dependencies. For example, it can cross a finite clause boundaries, as evidenced by (3), and it is subject to familiar locality conditions such as the Complex Noun Phrase Constraint (cf. Ross 1967) and the Non-bridge Verb Island Constraint (cf. Erteschik-Shir 1977), as can be seen from the ungrammaticality of (4) and (5), respectively.

- (3) yĒ ñ gūgū nā àbà pà wĪ nā ñ yĒ ngÚa yé é
 see you think NA Aba throw voice NA you saw them PART Q
 ‘Do you think that Aba announced that you SAW them?’
- (4) *tākā [s ñ wà [NP fòtó_i [s mŪmŪ_i [s ñ tàkà 6Ó àbà e_i]]]]
 show you like picture ITIT you showed REL Aba
 (‘You like the picture that you SHOWED Aba’)
- (5) *yē kòfí pĒ mlÍ nā wà yĒ mÓ yé
 see Kofi shout PART NA they saw him PART
 (‘Kofi shouted that they SAW him’)

This leads Koopman (1984) to propose that Vata predicate clefts involve a special sort of head-movement of the focused verb to C. The movement cannot be movement to [Spec,C] since that would be a violation of the Structure Preservation Hypothesis (SPH) which bars movement of heads into specifier positions (Emonds 1964).¹ Koopman calls movement of the focus verb to the C head “ $\bar{V}$ -movement,” which she postulates as an addition to the inventory of movement types of UG. As the name suggests, $\bar{V}$ -movement is basically head-movement of V which shares properties with phrasal $\bar{A}$ -movement. The fact that V is pro-

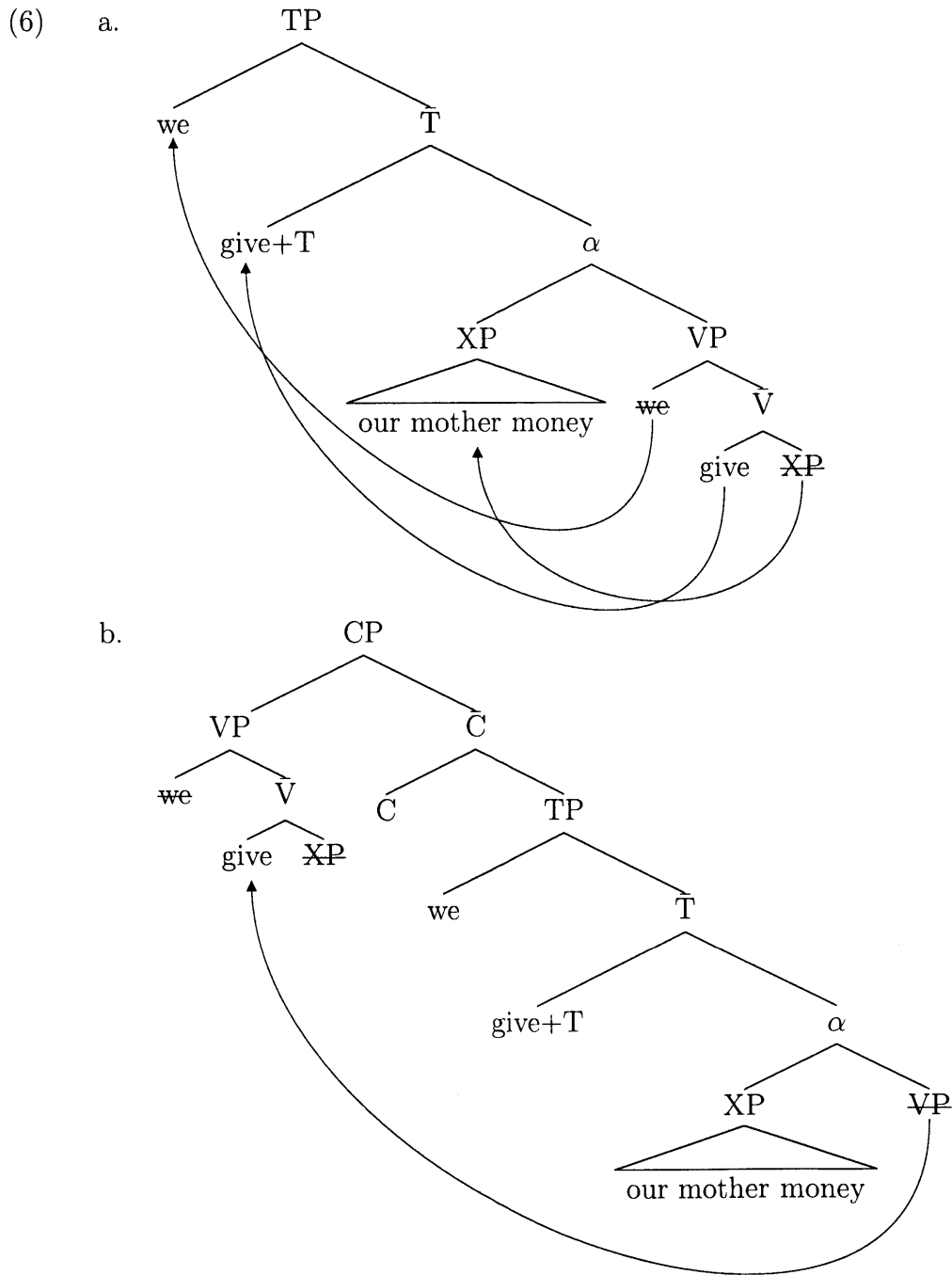
¹At the time there was no distinction between C and [Spec,C], as both were called COMP. Thus, what Koopman actually said is that the focused verb undergoes head-movement to COMP. We can, however, translate head-movement and phrasal movement to COMP as movement to C and to [Spec,C], respectively.

nounced twice is given an explanation based on the Empty Category Principle (ECP), which requires that empty categories be “properly governed.” (cf. Chomsky 1981, 1982, 1986). Koopman argues that if the “doublet” of the fronted V were a trace, which by hypothesis is an empty category, then that trace would be one which is not properly governed. For this reason, $\bar{V}$ -movement in Vata predicate clefts cannot leave behind a trace: some sort of resumptive element must fill the base position of the relevant chain, just as the resumptive pronoun **it** fills the (non-properly governed) embedded subject position in the English sentence **this is a donkey that I wonder where it lives** (this is the position from which the null relative operator has moved). However, Vata has no “resumptive pro-verb” to do the job, so it resorts to using the same phonetic material as that of the clefted predicate itself, with the result that that predicate is pronounced twice.

The analysis given in Koopman (1984) presupposes the GB framework in which the notion of government plays a central role.² Koopman (2000b) proposes another analysis of the same facts which is more in line with recent theoretical developments, one of which is the elimination of government as a unifying concept from linguistic theory (cf. Chomsky 1993, 1995). Since the mid 1990’s, observations that have been captured in terms of government are recast as emerging from the interaction of more basic processes and relations. Another development is the acceptance of “remnant movement” as a theoretical option: constituents containing traces of earlier movement operations are allowed to move, creating configurations in which traces are not bound by their antecedents (cf. Thiersch 1985, den Besten and Webelhuth 1987a, 1990, Webelhuth 1992, Kayne 1998, Müller 1998). Such configurations were considered illegitimate in earlier framework (cf. Fiengo 1974, Chomsky 1975, Fiengo 1977, among others). The final new ingredient in Koopman’s 2000 analysis of Vata predicate clefts is the copy theory of movement: “traces” are full-fledged copies of the moved constituent that have been bleached of phonetic content by an operation, Copy Deletion, distinct and in principle independent from the “movement” per se (see chapter 1 for more discussion). Against the background of these assumptions, Koopman (2000b) proposes that Vata predicate clefts are generated by a two-step process: (i) movement of things out of the VP, (ii) movement of the (remnant) VP to [Spec,C]. Take (2a), for example. This sentence would have the derivation in (6), where ~~striketrough~~ represents application of Copy Deletion.³

²Government features in accounts of such diverse phenomena as binding, case licensing, θ -role assignment and island effects.

³I leave open what the internal structure of XP the label of α in (6) is. These issues do not affect the point being made.



At the end of the derivation, we have “two overt copies of the same V, each carrying different morphology” (Koopman 2000b: 362). Support for this is the fact that V1 and V2 are actually pronounced differently: they carry different tones (nyĒ vs. nyÉ). Presumably, this is because V1 is not tensed and V2 is.

The $\bar{A}$ -properties of Vata predicate clefts now follow from the fact that these constructions are derived by regular, phrasal $\bar{A}$ -movement. It remains to explain why the sentence is pronounced the way it is. First, given the assumption that higher copies are by default overt and lower copies covert, the fact that we hear V2 follows from it being the higher copy of the chain created by V-to-T movement, and the fact that we do not hear XP or the subject at [Spec,C] follows from these being the lower copies of the chains created by A-movement of the subject from [Spec,V] to [Spec,T] and scrambling of XP, respectively. What about V1? It is actually the lower copy of the chain created by V-to-T movement, which means that it should be covert. Koopman suggests that the overtness of V1 is forced by the condition on recoverability of deletion (Chomsky 1965, Pesetsky 1998a), since “if the focused verb were silent, nothing would signal verbal focussing” (Koopman 2000b: 361).

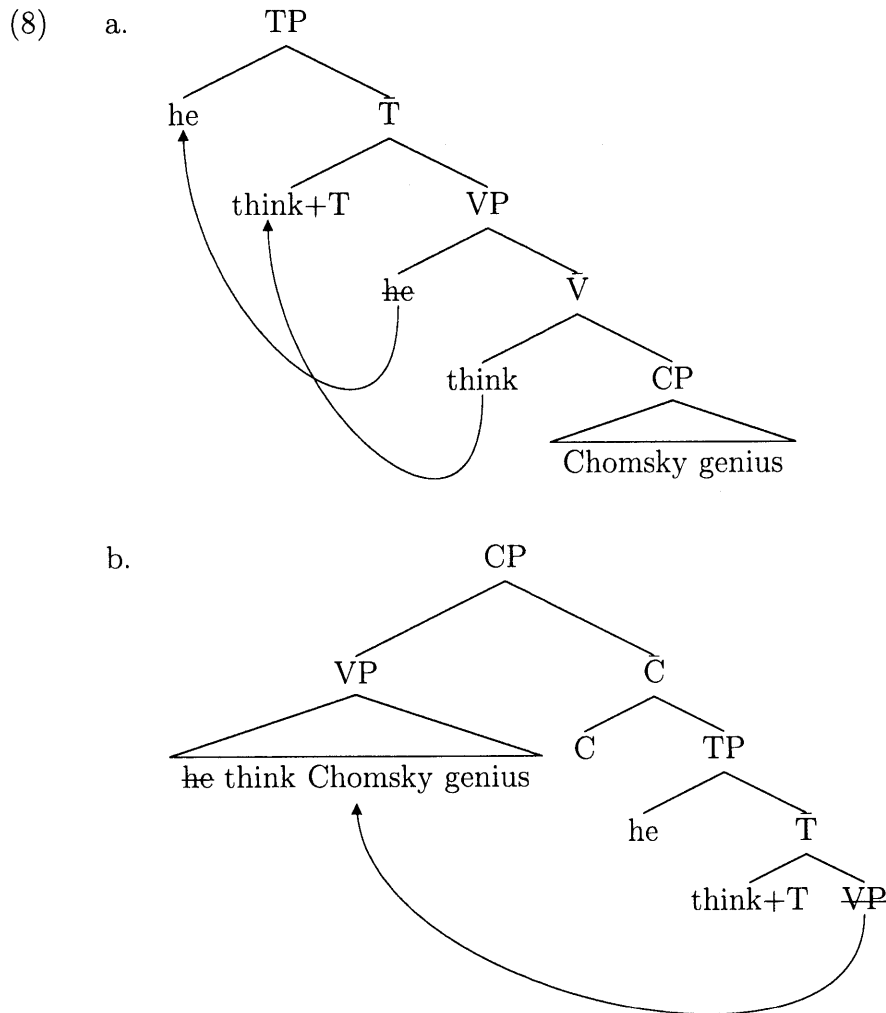
One fact about Vata predicate clefts is that “the focused verb cannot be accompanied by any of its arguments” (Koopman 1984: 155). The overt material in the clause-initial $\bar{A}$ -position can only be a single verb. In Koopman’s (1984) account, this follows from the assumption that movement of V to the C-domain is head-movement. In Koopman (2000b), however, this remains somewhat of a mystery: Koopman must essentially stipulate that internal arguments of the verb, in fact all phrasal constituents of the VP, must vacate the VP prior to VP fronting. In this connection, it is interesting to note that Abels (2001) shows Russian predicate clefts to be just like Vata predicate clefts except that the internal arguments of V do not always move out of VP. Here is an example from Abels (2001: 6).⁴

- (7) Dumat’ čto Xomskij genij on dumaet no čitat’ ego knigi ne čitaet
 think that Chomsky genius he thinks but read his book not reads
 ‘He does think Chomsky is a genius, but read his books he doesn’t’

For such sentences as those in (7), Abels (2001) proposes a derivation which is identical in form to the derivation that Koopman (2000b) proposes for Vata predicate clefts modulo the movement of internal arguments of V out of VP. The first sentence in (7), for example, has the derivation in (8).⁵

⁴Note that the meaning of Russian predicate clefts are not that of sentences where the V(P) is focused, but of those where it is a (contrastive) topic. As far as the possibility exists that the C-domain can host both focused and topicalized elements, our discussion, which aims at analyzing the syntax of these constructions, should not be affected.

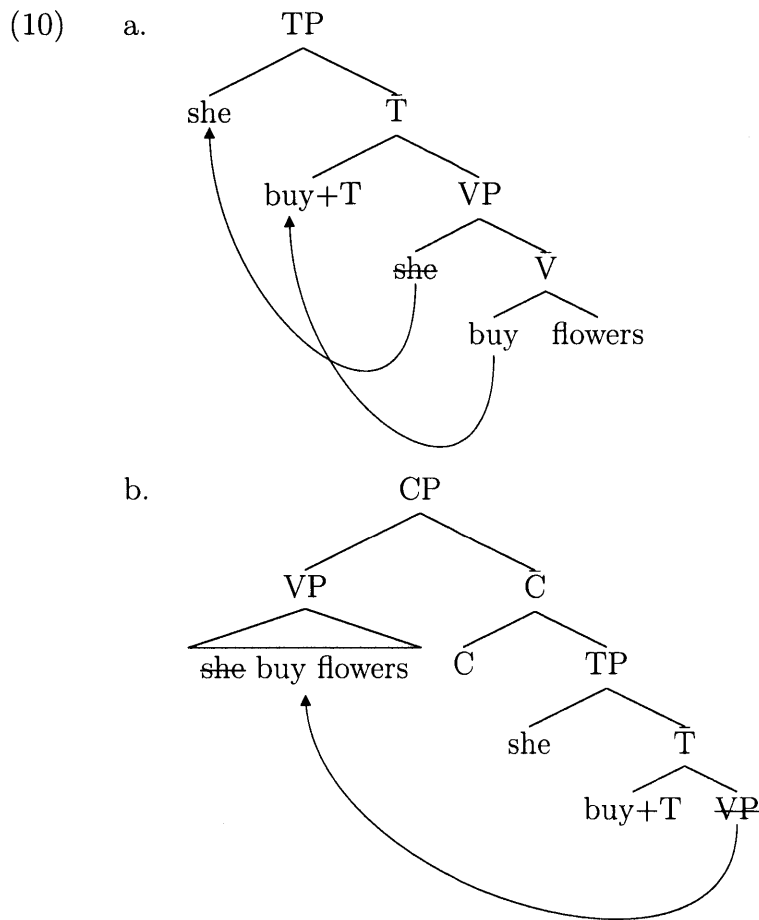
⁵For predicate clefts in Russian where the internal arguments of the clefted verb remains TP-internal, Abels proposes that these have undergone either object shifts, as in the case of DP arguments, or extraposition, as in the case of CP arguments.



The explanation in Abels (2001) for the doubling of V, i.e. the fact that V1 is overt, is similar to that in Koopman (2000b): “The pronunciation of the first occurrence of the verb in the moved VP is forced because semantically, it carries the focus, which needs an audible exponent” (Abels 2001: 15–16). In other word, deletion of V1 does not take place because it is not “recoverable.”

Landau (2006) offers a similar analysis for one variety of Hebrew predicate clefts, namely that in which the V together with its internal argument fronts. An example is given in (9) (Landau 2006: 37). From what Landau says, (10) would be the derivation of (9).

- (9) liknot et ha-praxim, hi kanta
 to-buy ACC the-flowers she bought
 ‘As for buying flowers, she bought’



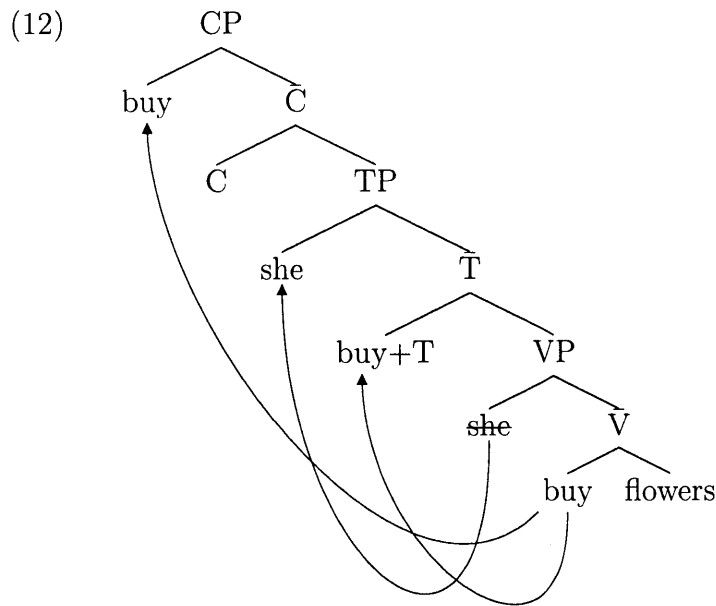
However, Landau gives a different rationale for the overttness of V1. He rejects the line of explanation adopted by Koopman and Abels on the ground that it involves a PF operation, namely the deletion of chain links, being “contingent on LF recoverability” (Landau 2006: 35), and advances the hypothesis that the application of Copy Deletion should be determined without recourse to semantic/pragmatic considerations. Specifically, he proposes that the obligatory pronunciation of V1 is due to a phonological requirement imposed by C, namely that C simply demands that its specifier have an overt head. Landau takes the “characteristic intonation of the fronted VP” to be a reflection of (the satisfaction of) this requirement.

It turns out that Hebrew predicate clefts have another variant in which [Spec,C] is filled by only a single verb, as exemplified in (11) (Landau 2006: 37).

- (11) liknot, hi kanta et ha-praxim
 to-buy she bought ACC the-flowers
 ‘As for buying, she bought flowers’

It is in the analysis of this variant of Hebrew predicate clefts that Landau differs from Koopman and Abels. Specifically, Landau assumes that it is V and not a remnant VP which undergoes $\bar{A}$ -movement to [Spec,C], the reason being that the “vacating movement” of VP-internal constituents necessary to create a remnant VP cannot be independently motivated for Hebrew, hence a remnant VP movement analysis would be justified only on the ground that it avoids violation of the SPH, or the “Chain Uniformity Condition” in more recent terminology (cf. Chomsky 1995). Landau, however, points out, correctly in my opinion, that this condition is unwarranted and should not be assumed. We come back to this issue below.

There is no explicit discussion of doubling with respect to this second variant of Hebrew predicate clefts in Landau (2006), but the assumption seems to be that what is said about the first variant will carry over to this case as well. Thus, the analysis for (11) would be (12), if I understood Landau correctly.



At this point, it is helpful to step back and make explicit an assumption underlying the three explanations of doubling given in Koopman (2000b), Abels (2001) and Landau (2006). First, note that V1 is part of two chains: one resulting from V-to-T movement, call it CH₁, and another, CH₂, resulting from movement of V/VP to [Spec,C]. The three authors mentioned above takes it for granted that Pronunciation Economy (PE), which says that lower copies are covert by default (see chapter 1), entails that V1 should be covert, and that the overtiness of this element must hence be derived from some other requirement which overrides PE. In other word, it is CH₁ which determines the way V1 is “treated” by Copy Dele-

tion. Membership in CH₂, in which V1 is actually the higher copy, or part of the higher copy, does not rescue V1 from being made covert. We can thus say that these authors assume that being the lower copy of one chain is sufficient ground for being “silenced” by Copy Deletion.

Aboh and Dyakonova (2009) offer a new perspective on predicate clefts in Russian and Gungbe which, in effect, challenges the assumption above. Aboh and Dyakonova attribute the same “two-chain” structure to the clefts they discuss as did Koopman, Abels and Landau. However, they take the pronunciation of V1, i.e. the V copy inside the fronted VP, to be entirely natural, the reason being that it is part of the higher copy of the chain created by movement of VP to [Spec,C]. Thus, Aboh and Dyakonova assume that being (part of) the higher copy of one chain is sufficient ground for default overtiness. The fact that we hear two instances of the verb in predicate clefts, one in [Spec,C] and one in T, is simply due to the fact that these are both higher copies of chains.

Another line of approach to V doubling in predicate clefts is taken by Kandybowicz (2006, 2007, 2008), with the relevant language being Nupe. In this approach, the two pronounced instances of the clefted verb are in fact related by movement, in the sense that the derivation of the sentence involves movement from the position of V2 to that of V1. Kandybowicz accounts for the overtiness of both V1 and V2 by saying that the chain linearization operation does not recognize these elements as constituting a chain. The reason is that V2 “fuses” with a low focus head in the morphological component, thus becoming subpart of a word. By hypothesis, chain linearization, or more precisely linearization in general, applies after morphology and does not affect word-internal constituents. As a consequence, Copy Deletion, which is part of linearization, does not apply to V2, resulting in doubling. Kandybowicz provides experimental evidence that V1 and V2 are phonetically different, which he takes to reflect a difference in their morphological make-up.

We have discussed predicate clefts in a number of languages and the analyses that have been given for them. Common to all these analyses is the idea that the two instances of the clefted verb are morphologically distinct. In fact, this assumption is needed to explain the overtiness of V2, thus ultimately necessary to explain verb doubling in predicate clefts. Koopman, Abels, Landau and Aboh and Dyakonova need to say that V2 is actually the higher copy of a chain, namely that created by V-to-T movement, while Kandybowicz needs to say that V2 is subpart of a word which does not include V1. Admittedly, the data considered by these authors make it plausible to assume morphological distinctness between V1 and V2, since it is often the case that a difference in overt form can be observed between these two elements. The question to be raised, then, is whether the

“morphological distinctness” theory is general enough to cover all cases of verb doubling. The answer given in Vicente (2007) is negative. In this work, Vicente investigates predicate clefts in Spanish and Hungarian, where doubling is also attested. Vicente’s main objective is to develop a unified theory of movement which does not distinguish between phrasal and head movement as two primitive types of operation, and while he considers data which involve verb doubling, Vicente makes it explicit that his aim is not to propose an account as to why there is doubling (Vicente 2007: 5–8). He does point out, however, that the “morphological distinctness” theory is unlikely to be the whole story to doubling. The main reasons for Vicente’s scepticism is the fact that doubling maybe required even when V1 and V2 are morphologically identical, as exemplified by the Spanish sentence in (13a) and the Kwa sentence in (13b).

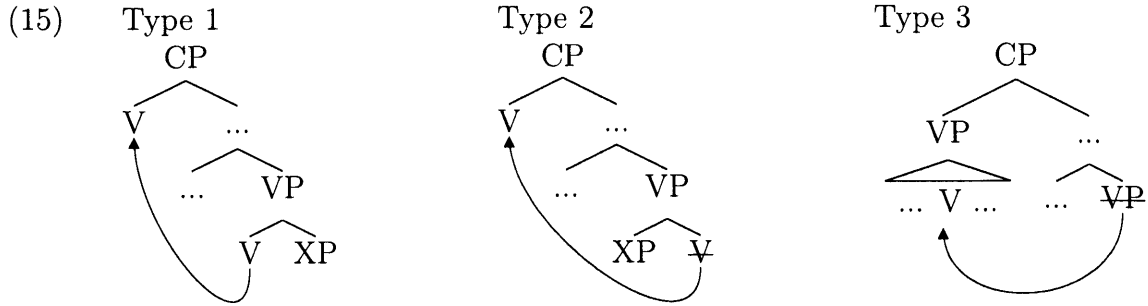
- (13) a. Jugar, Juan suele jugar al futbol los domingos
 play.INF Juan HAB play.INF at soccer the sundays
 ‘As for playing, Juan usually plays soccer on Sundays’
 b. ðù wè Séna ðù blèdì ló
 eat FOC Séna eat bread DET
 ‘Séna ATE the bread’

On the basis of facts such as these, Vicente concludes that morphological distinctness of V1 and V2 “cannot be the unique factor regulating [verb doubling],” adding, however, that he has “nothing interesting to suggest as to what those extra factors might be” (Vicente 2007: 7).

The purpose of this chapter is to show that the ECCD is one of these “extra factors.” To this end, I will discuss two sorts of predicate clefts, none of which involves morphologically distinct copies of the clefted verb. The first sort is what Vicente’s examples in (13) illustrate: clefts with verb doubling in which V1 and V2 are morphologically identical. The other sort of predicate clefts I will consider does not involve doubling at all. It simply shows a verb in the topic position, with a gap at the base position of that verb, as exemplified by the German sentence in (14).

- (14) geküsst hat Hans Maria
 kissed has Hans Maria

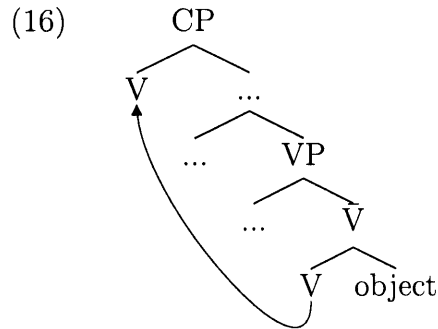
I will argue that these two variants of predicate clefts are actually of three structural types, given in (15). The first type involves moving V out of a head-initial VP into [Spec,C] and pronouncing the lower copy of the chain created. The second type involves moving V out of a head-final VP and not pronouncing the lower copy, as shown in (16). The last type involves remnant VP movement.



The languages which feature in the discussion are Hebrew, Vietnamese, German, Dutch, Swedish and Norwegian. I will try to show that predicate clefts in Hebrew and Vietnamese are of Type 1, those in German and Dutch are of Type 2, and those in Swedish and Norwegian are of Type 3, and that this typology constitutes supporting evidence for the ECCD.

2.2 Type 1: Hebrew and Vietnamese

Suppose L is a head-initial language. Then, V precedes its complement in L. This means that movement of V from its base position to [Spec,C] results in a chain in which the lower copy does not end an XP, i.e. is not at the right edge of a maximal projection: it is separated from such an edge by the post-verbal object.



The Edge Condition on Copy Deletion (ECCD), repeated in (17), predicts that both copies of this chain must be pronounced.

- (17) Edge Condition on Copy Deletion (ECCD)
 For any chain (α, β) where α is the higher and β the lower copy of the moved constituent, deletion of β requires that β end an XP

In the following, I argue that Hebrew and Vietnamese verify the prediction made for L, hence constitute evidence in favor of the ECCD.

2.2.1 Hebrew

In Hebrew, verbs can topicalize, stranding their arguments. An example is given in (18). As can be observed, the fronted verb is doubled.⁶

- (18) liknot Dan kiva liknot et ha-sefer
 buy.INF Dan hoped buy.INF ACC the-book
 ‘As for buying, Dan hoped to buy the book’

I claim that the doubling in (18) is due to a violation of the ECCD. This requires excluding the plausible hypothesis that it is **kiva**’s inability to license a gap that blocks Copy Deletion in (18).⁷ As the grammaticality of (19) shows, **kiva** can be followed by an empty category.

- (19) liknot et ha-sefer Dan kiva
 buy.INF ACC the-book Dan hoped
 ‘Dan hoped to buy the book’

The relation between V1 and V2 in Hebrew predicate clefts exhibits properties of $\bar{A}$ -dependencies: it is unbounded and subject to island constraints (cf. Chomsky 1977). Illustration for unboundedness is given in (20).⁸

- (20) liknot nidme li še-Rina amra še-Dan kiva liknot et
 buy.INF it seems to me that Rina said that Dan hoped buy.INF ACC
 ha-sefer
 the book
 ‘As for buying, it seems that Rina said that Dan hoped to buy the book’

The four examples that follow show that verb topicalization in Hebrew is subject to the Complex NP Constraint, the Subject Island Constraint, the Adjunct Island Constraint and the Factive/Non-Bridge Constraint, in that order.

- (21) *liknot Dan daxa et ha-te’ana še-hu kiva liknot et
 buy.INF Dan rejected ACC the-claim that-he hoped buy.INF ACC
 ha-sefer
 the-book

⁶I thank Asaf Bachrach, Danny Fox, Roni Katzir, and Omer Preminger for the Hebrew judgements.

⁷See King (1970), Labov (1969) for such a hypothesis with respect to contracted auxiliaries in English and the null copula in Ebonics (i.e. Black English).

⁸Out of four Hebrew speakers I have consulted, three judged (20) to be grammatical, while one gave it two question marks. I have no explanation for this variation.

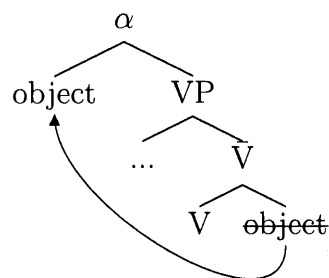
(‘As for buying, Dan rejected the claim that he hoped to buy the book’)

- (22) *liknot še-Gil yirce liknot et ha-sefer ze cafuy
 buy.INF that-Gil want.FUT buy.INF ACC the-book COP expected
 (‘As for buying, that Gil will want to buy the book is expected’)
- (23) *liknot Dan samax ki Dina kiva liknot et ha-sefer
 buy.INF Dan was.happy because Dina hoped buy.INF ACC the-book
 (‘As for buying, Dan was happy because Dina hoped to buy the book’)
- (24) *liknot Dan laxaš/hitca’er še-Dina kiva liknot et
 buy.INF Dan whispered/regretted that-Dina hoped buy.INF ACC
 ha-sefer
 the-book
 (‘As for buying, Dan *whispered/*regretted that Dina hoped to buy the book’)

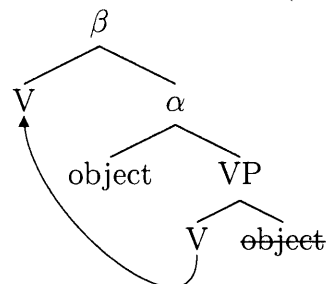
I will assume, in light of these data, that verb topicalization in Hebrew involves $\bar{A}$ -movement, and that the verbal topic is in [Spec,C] (cf. Rizzi 1997). Let us turn now to the question whether what $\bar{A}$ -moves is a V or a VP, i.e. whether Hebrew predicate clefts are to be analyzed as involving V-to-[Spec,C] movement, or remnant VP movement. This question is of crucial relevance: if the clefted predicate is a V, the observed doubling would follow from the ECCD, but if it is a (remnant) VP, doubling would have to be explained in some other way. Given the fact that we hear only a single verb in the topic position, I will take it for granted that the V-to-[Spec,C] movement analysis is the null hypothesis. In fact, the only objection against this analysis, as far as I can see, is that it violates the SPH/Chain Uniformity. This objection, however, presupposes that the SPH/Chain Uniformity is a valid principle. Since the validity of this principle is at issue in our present discussion, the objection is not really appropriate. Thus, it remains to examine how plausible the remnant VP movement analysis is. In what follows, I will argue that the remnant VP movement analysis is not plausible.

A remnant VP movement analysis requires, of course, a remnant VP, i.e. a VP out of which the object has moved. There are two ways for the object to escape the VP: by Object Shift (OS) to the left or by extraposition to the right. Suppose OS moves the object to some position X to the left of VP. As the lower V copy ends up preceding the object, OS has to be followed by movement of V to some position Y above X, which I will call “short V movement” (SVM). Thus, the derivation has to consist of the following steps.

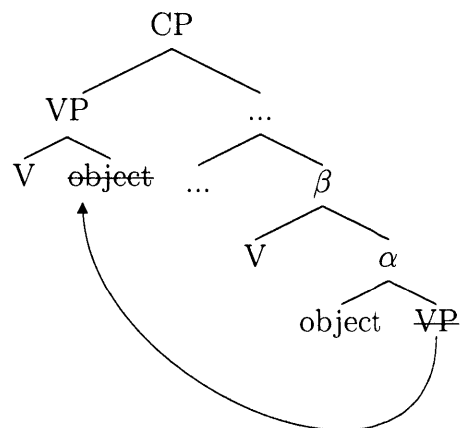
- (25) a. Object Shift (OS)



- b. Short V movement (SVM)



- c. Remnant VP movement



Evidence against the derivation in (25) includes the fact that OS is not independently attested in Hebrew (cf. Landau 2007), as the ungrammaticality of (26) shows.

- (26) *Dan kiva et ha-sefer liknot
Dan hoped ACC the book buy.INF

The remnant VP movement analysis thus faces the challenge of explaining why OS applies only if VP topicalizes. A case can be made that SVM cannot be motivated independently of VP topicalization either. First, note that the lower

copy of the SVM chain does not delete, as indicated in (25b): its deletion would deprive VP of all overt material and result in [Spec,C] being phonologically null, contrary to facts. Suppose that it is an intrinsic property of the SVM chain that it does not undergo Copy Deletion. Then, if SVM is independent of VP topicalization, either (27a), which is derived by SVM and OS, or (27b), which is derived by SMV alone, or both, would be grammatical.

- (27) a. *Dan kiva liknot liknot et ha-sefer
 Dan hoped buy.INF buy.INF ACC the book
 b. *Dan kiva liknot et ha-sefer liknot
 Dan hoped buy.INF ACC the book buy.INF

The fact, however, is that none of (27a) and (27b) is grammatical. This is evidence that SVM is not independent of VP topicalization. Now let us entertain the hypothesis that non-deletion of the lower copy of the SVM chain is conditional upon VP topicalization, say in just the way the lower copy of V-to-T movement is in Abels' account of Russian predicate clefts. Specifically, assume that the overtness of this copy is forced by recoverability of some sort, which is plausible as its deletion would take away all indication that VP topicalization has taken place. When there is no "need" for it to be overt, the lower copy in the SVM chain does delete, just as the lower copy of V-to-T movement. This assumption, indeed, would allow us to say that SVM can take place independently of VP topicalization, albeit without any phonetic effect, and that the ungrammaticality of the sentences in (27) does not undermine the derivation in (25).

There is, however, a problem with this proposal. The problem is this: fronting an intransitive verb results in optional, not obligatory, doubling of the predicate.

- (28) a. lalexet Dan kiva
 walk.INF Dan hoped
 b. lalexet Dan kiva lalexet
 walk.INF Dan hoped walk.INF

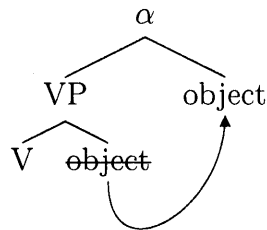
A straightforward account of this fact is that an intransitive verb can, but does not have to, undergo SVM: when it does, doubling results, and when it does not, there is no doubling. The obvious question then becomes what distinguishes transitive from intransitive verbs with respect to SVM? There seems to be no way to answer this question other than adding more stipulations to the account.⁹

⁹We will see below that the optionality of doubling of clefted intransitives receives a quite natural explanation in terms of the ECCD.

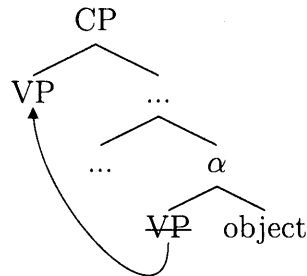
On the basis of these considerations, I conclude that the “object shift” version of the remnant VP movement analysis of Hebrew predicate clefts is untenable.

What about the extraposition version? Suppose the derivation of a predicate cleft sentence in Hebrew contains the following steps.

- (29) a. Extraposition of the object



- b. Topicalization of the remnant VP



The prediction would then be made that objects of topicalized verbs must be able to extrapose. This prediction is arguably wrong. In Hebrew, weak pronouns such as **oto** ‘it’ do not extrapose easily. Observe the contrast in (30).

- (30) a. Dan kiva liknot oto maxar
 Dan hoped buy.INF it tomorrow
 b. ??Dan kiva liknot maxar oto
 Dan hoped buy.INF tomorrow it

When the object is a full DP, there is no such contrast: full DPs have no problem extraposing.

- (31) a. Dan kiva liknot et ha-sefer maxar
 Dan hoped buy.INF ACC the-book tomorrow
 b. Dan kiva liknot maxar et ha-sefer
 Dan hoped buy.INF tomorrow ACC the-book

If Hebrew predicate cleft requires the object to have extraposed before the remnant VP undergoes movement to [Spec,C], there should be a contrast between

(32a) and (32b). In the former, the constituent which has allegedly extraposed is a full DP, while in the latter, it is a pronoun. This prediction is not born out: (32b) is perfectly acceptable.

- (32) a. liknot Dan kiva liknot et ha-sefer
 buy.INF Dan hoped buy.INF ACC the-book
 ‘As for buying, Dan hoped to buy the book’
 b. liknot Dan kiva liknot oto
 buy.INF Dan hoped buy.INF it
 ‘As for buying, Dan hoped to buy it’

I take the above considerations to be sufficient ground to prefer the V-to-[Spec,C] movement analysis of Hebrew predicate clefts over the remnant VP movement analysis.

2.2.2 Vietnamese

The aim of this subsection is to show that Vietnamese provides evidence for the ECCD in exactly the same way as Hebrew does. Topicalization in Vietnamese involves fronting the topic constituent to the left periphery, preceding the topic marker **thi** (cf. Cao 1992), which we will assume to be a C head.

- (33) a. No nen doc sach
 he should read book
 ‘He should read books’
 b. No thi nen doc sach
 he TOP should read book
 ‘As for him, he should read books’
 c. Sach thi no nen doc
 book TOP he should read
 ‘As for books, he should read them’
 d. Doc sach thi no nen
 read book TOP he should
 ‘As for reading books, he should do that’

Topicalization of verbs in Vietnamese is possible, and shows doubling in the same way it does in Hebrew.¹⁰

¹⁰Note that (33d) shows that the modal verb **nen** can license a gap. Thus, the case cannot be made that Copy Deletion fails in (34) due to requirements of **nen**.

- (34) Doc thi no nen *(doc) sach
 read TOP he should *(read) book
 ‘As for reading, he should read books’

The relation between the topic and the base position exhibits standard symptoms of $\bar{A}$ -movement: it is not clause-bound, as (35) shows, and it is constrained by islands, as shown by the four examples that follow (35), which evidence the Complex NP Constraint, the Subject Island Constraint, the Adjunct Island Constraint and the Factive/Non-Bridge Constraint, in that order.¹¹

- (35) doc thi toi nghi la no nen doc sach
 read TOP I think that he should read book
 ‘As for reading, I think that he should read books’
- (36) *doc thi toi tin chuyen no doc sach
 read TOP I believe story he read book
 (‘As for reading, I believe the story that he reads books’)
- (37) *doc thi no doc sach la tot
 read TOP he read book COP good
 (‘As for reading, that he reads books is good’)
- (38) *doc thi no vui vi toi doc sach
 read TOP he happy because I read book
 (‘As for reading, he is happy because I read books’)
- (39) *doc thi toi tiec/thi-thao la no doc sach
 read TOP I regret/whisper that he read book
 (‘As for reading, I regretted/whispered that he read books’)

The argument against a remnant VP movement analysis of predicate clefts which was made for Hebrew can be replicated for Vietnamese, as basically the same patterns are observed in the latter. Thus, the object shift version of the remnant VP movement analysis is challenged by the fact that Object Shift (OS) and Short Verb Movement (SVM) with doubling are not independently attested.

¹¹Note that subject clauses in Vietnamese are not introduced by a complementizer.

- (i) no doc sach la tot
 he read book is good
 ‘That he reads books is good’

- (40) a. *No nen sach doc
 he should book read
 b. *No nen doc doc sach
 c. *No nen doc sach doc

Assuming doubling in SVM to be conditional on VP topicalization faces the same difficulty as we have seen with Hebrew, as fronting intransitive verbs in Vietnamese also results in optional doubling.

- (41) Ngu thi no nen (ngu)
 sleep TOP he should (sleep)

The argument against the extraposition version of the remnant movement analysis is also similar to that made with respect to Hebrew: categories that do not extrapose have no problem being stranded objects in predicate cleft constructions. Bare nouns such as **sach** ‘book,’ for example, cannot extrapose. This is illustrated in (42). In (43), however, we see that **sach** can be stranded.

- (42) a. no doc sach hom-qua
 he read book yesterday
 b. *no doc hom-qua sach
 he read yesterday book
- (43) doc thi no nen doc sach
 read TOP he should read book

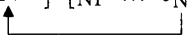
In light of these facts, I conclude that a V-to-[Spec,C] analysis of Vietnamese predicate clefts is to be preferred to a remnant movement analysis.

Intransitives

We have seen how the ECCD accounts for the obligatory doubling of $\bar{A}$ -moved transitive verbs. What about intransitives? In both Hebrew and Vietnamese, topicalization of intransitive verbs results in not obligatory, but optional, doubling.

- (44) a. lalexet Dan kiva (lalexet)
 walk.INF Dan hoped (walk.INF)
 b. Ngu thi no nen (ngu)
 sleep TOP he should (sleep)

Can we derive the difference between transitives and intransitives with respect to doubling from the ECCD? I will argue that the answer is positive, given the theory of intransitives presented in Hale and Keyser (1993, 2002) which, incidentally, was developed on the basis of considerations having nothing to do with doubling. Basically, Hale and Keyser propose that there are really no verbs without NP complement: intransitives are the result of incorporating the head of a phonologically null NP into the governing V.

$$(45) \quad [_{VP} [_{V^0} V^0 + N^0] [_{NP} \dots t_{N^0} \dots]]$$


The crucial point to note here is that the VP in (45) is homophonous with one of its subpart, namely $\alpha = [_{V^0} V^0 + N^0]$. This means that there is no difference between topicalization of VP and topicalization of α with respect to the phonological material in [Spec,C]. However, the ECCD predicts a difference between these two cases with respect to the phonology of the base position. Specifically, fronting α will result in doubling, as the lower copy of the chain created does not end an XP, whereas fronting VP will not, VP is, by definition, an XP. Assuming that both VP and α can be topic, we predict exactly what is observed: if the topic is an intransitive verb, doubling is optional.

The ECCD, then, provides a straightforward account of an otherwise quite puzzling fact: the optionality of doubling in predicate clefts when the clefted predicate is intransitive. We consider this another piece of evidence in favor of the ECCD.

2.2.3 Interim summary

I have examined predicate clefts in Hebrew and Vietnamese. What can be observed of these constructions is that the fronted verb is doubled: it is pronounced twice, at both the sentence initial position and at the base position. Given that Hebrew and Vietnamese are SVO languages, this observation follows from the ECCD under the assumption that predicate clefts in these languages involve V-to-[Spec,C] movement. I have argued that this assumption is empirically motivated. In addition, the fact that doubling of the verb is optional if that verb is intransitive is shown to follow straightforwardly of the ECCD in conjunction with the theory of argument structure proposed in Hale and Keyser (1993, 2002).

2.3 Type 2: German and Dutch

We have seen how the fact that the topicalized verb is doubled in Hebrew and Vietnamese follows from the ECCD. However, it turns out that there is another way to derive doubling in these two languages. Suppose we assume the following principle.

- (46) Uniformity Condition on Copy Deletion (UCCD)
 Copy Deletion cannot apply if the relevant chain is not uniform

A chain CH is uniform iff the two copies in CH are either both XP, or both X^0 . In the chain created by V-to-[Spec,C] movement in Hebrew and Vietnamese, the higher copy is an XP, as it does not project, while the lower one is not and XP, as it does project. The UCCD implies that Copy Deletion cannot apply, i.e. that both copies are pronounced. This is borne out by facts, as we have observed.

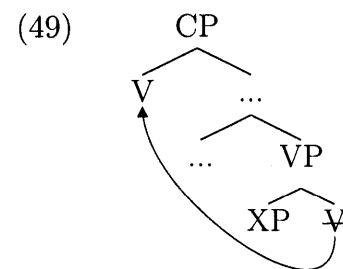
This section aims at providing evidence for the ECCD and against the UCCD. Note that the UCCD predicts doubling whenever there is V-to-[Spec,C] movement, since this creates a non-uniform chain. Thus, (47) is a consequence of this principle.

- (47) Consequence of the UCCD
 If V moves to [Spec,C], verb doubling will result

Under the ECCD, on the other hand, V-to-[Spec,C] movement alone does not prevent Copy Deletion from applying. It must also be the case that the lower V copy does not end VP. Thus, the ECCD implies something weaker than (47).

- (48) Consequence of the ECCD
 If V moves to [Spec,C] and VP is head-initial, verb doubling will result

In what follows, I will defend the ECCD against the UCCD by providing facts which falsify (47) but does not falsify (48). These would be cases where V moves to [Spec,C] from a head-final VP and no doubling results, i.e. cases which instantiate the following analytical option.



The chain created by movement of V in (49) is not uniform, as the higher copy is an XP while the lower is not. Contrary to what the UCCD predicts, the lower copy deletes. Note, however, that this copy is at the right edge of the VP. Thus, its deletion is not prevented by the ECCD, and given Copy Inequality, is expected. To show that (49) exists, then, is to show that the UCCD is wrong and the ECCD is right. In what follows, I argue that there are instantiations of (49) in German and Dutch.

2.3.1 German

It has frequently been observed that German has predicate clefts. In other word, German allows a (non-tensed) verb without any arguments to occupy the “Vorfeld,” i.e. [Spec,C] (cf. Thiersch 1985, Müller 1998, Fanselow 2002, Hinterhölzl 2002, among others). Here are two examples from Hinterhölzl (2002:127). Two examples are given in (50).

- (50) a. lieben will Hans die Maria
 love wants Hans the Maria
 ‘Hans wants to love Maria’
 b. gelesen hat Hans das Buch
 read has Hans the Buch
 ‘Hans has read the book’

It has also been assumed that predicate cleft in German is regular topicalization (cf. Thiersch 1985, den Besten and Webelhuth 1987a, Müller 1998). Indeed, the relevant characteristic features of $\bar{A}$ -movement to [Spec,C] are attested, namely clause unboundedness and island-sensitivity.

- (51) lesen denke ich wird Hans ein Buch
 read think I will Hans a book
 ‘I think Hans will read a book’

 (52) *lesen glaube ich die Geschichte, dass Hans ein Buch wird
 read believe I the story, that Hans a book will
 (‘As for reading, I believe the story that Hans will read a book’)

 (53) *lesen ist dass Hans ein Buch wird ganz überraschend
 read is that Hans a book will totally surprising
 (‘That Hans will read a book is totally surprising’)

- (54) *lesen bin ich glücklich, weil Hans ein Buch wird
 read am I happy because Hans a book will
 ('As for reading, I am happy because Hans will read a book')
- (55) *lesen bereue/flüstere ich, dass Hans ein Buch wird
 read regret/whisper I that Hans a book will
 ('As for reading, I regret/whisper that Hans will read a book')

The simplest account, so it seems, for German predicate clefts would be that it involves $\bar{A}$ -movement of V to [Spec,C], i.e. topicalization of V. Nevertheless, researchers have argued that German really has no V topicalization, and that all cases of predicate clefts in German involves movement of a remnant VP to [Spec,C] (cf. Thiersch 1985, den Besten and Webelhuth 1987a, Müller 1998). My goal in this section is to argue that German does have V-to-[Spec,C] movement. As explained above, this argument is in effect an argument against the UCCD and in favor of the ECCD. Logically, this goal can be achieved by showing that one case of predicate clefts in German does not involve remnant VP topicalization. Nevertheless, I will show three such cases. First, consider the paradigm in (56) from Hinterhölzl (2002: 132).

- (56) a. ??[dass Fritz Peter liebt] weiss ich nicht wer gesagt hat
 [that Fritz Peter loves] know I not who said has
 'I don't know who has said that Fritz loves Peter'
- b. *[dass Fritz t_1 liebt] weiss ich nicht wen₁ er gesagt hat
 [that Fritz loves] know I not whom he said has
 'I don't know who he said that Fritz loves'
- c. ??geküsst weiss ich nicht wen sie hat
 kissed know I not whom she has
 'I don't know whom she has kissed'

Hinterhölzl notes that (56a) shows "a mild, subjacency-like violation," while (56b) is "ungrammatical" (Hinterhölzl 2002: 131). He attributes this contrast to the Proper Binding Condition: the topic constituent in (56b) contains an unbound wh-trace, whereas no such trace is present in the topic constituent in (56a) (cf. Pesetsky 1998b). Significantly, Hinterhölzl notes that (56c) is just as good as (56a), even though the object of **geküsst** is a wh-phrase which has moved to the embedded [Spec,C]. The relative acceptability of (56c) would follow immediately if we assume that the topic constituent in (56c) is a V and not a remnant VP, i.e. if we assume that German has V topicalization. Thus, I take (56) to be evidence that German has V topicalization.

Second, note that in German, wh-phrases can be used as either interrogative pronouns or indefinites. But irrespective of their function, wh-phrases do not scramble (cf. Müller and Sternefeld 1996). The contrast in (57) shows the resistance of interrogative wh-phrases to scrambling, and that in (58) shows the same for wh-indefinites.

- (57) a. Ich weiß nicht, wem₁ der Fritz t₁ was gesagt hat
 I know not, to whom the Fritz what said has
 ‘I don’t know what Fritz said to whom’
 b. *Ich weiß nicht, wem₁ was₂ der Fritz t₁ t₂ gesagt hat
 I know not, to whom what the Fritz said has
- (58) a. Dass der Fritz wen geküsst hat
 that the Fritz whom kissed has
 ‘That Fritz has kissed someone’
 b. *Dass wen der Fritz geküsst hat
 that whom the Fritz kissed has

Let us register another fact about German. In this language, extraposed materials must follow both the main verb – if it is not in C – and any auxiliary that follows the main verb. Extraposition to a position between the main verb and a following auxiliary is not possible.

- (59) a. Dass er [von einer schönen Frau] geträumt hat
 that he of a beautiful woman dreamed has
 ‘That he has dreamt of a beautiful woman’
 b. Dass er t₁ geträumt hat [von einer schönen Frau]₁
 that he dreamt has of a beautiful woman
 c. *Dass er t₁ geträumt [von einer schönen Frau]₁ hat
 that he dreamt of a beautiful woman has

This means that the presence of an auxiliary to the right of a constituent indicates that that constituent has not extraposed. It then follows that if the object of a topicalized verb is a wh-phrase and precedes an auxiliary, it must be inside VP: it cannot have scrambled, because wh-phrases do not scramble, and it cannot have extraposed, because extraposed constituents must follow the auxiliary. But if the object has neither scrambled nor extraposed, the topicalized verb cannot be a remnant VP, since no remnant VP was created. Examples (60) and (61) instantiate just this scenario, the former with an interrogative, the latter with an indefinite wh-phrase (cf. Fanselow 2002: 101, 103).

- (60) Geküsst wüsste ich gern wer wen hat
 kissed knew I gladly who whom.ACC has
 ‘I would like to know who kissed whom’
- (61) Geküsst dürfte er schon öfter wen haben
 kissed might he already more-often whom.ACC have
 ‘He may very well have kissed somebody quite often’

The object **wen** in (61) and (60) must be VP-internal. Consequently, the verb which occupies the matrix [Spec,C] in these sentences is just that: the verb. It cannot be a remnant VP. Thus, (61) and (60) indicate that German permits V-topicalization.

Let us consider one more argument that V topicalization is possible in German. It has been observed that scrambled objects become opaque for extraction (Müller 1998). This is evidenced by the contrast in (62), taken from Müller (1998:12).

- (62) a. Worüber₁ hat keiner [ein Buch t₁] gelesen
 about-what₁ has no one a book read
 b. *Worüber₁ hat [ein Buch t₁]₂ keiner t₂ gelesen
 about-what₁ has a book no one read

Now consider (63a) and (63b), both taken from Fanselow (2002:110). In (63b), the pronoun **da** ‘there’ is extracted from the PP damit ‘therewith.’ This means that the PP has not scrambled. But it cannot have extraposed either, since it is followed by the auxiliary **haben**.

- (63) a. Er dürfte sie ja wohl kaum damit widerlegt haben
 he might her JA well barely there with refuted have
 ‘He is not really likely to have refuted her with that’
 b. Widerlegt dürfte er sie da₁ ja wohl kaum [t₁ mit] haben
 refuted might he her there JA well barely with have

It follows that [Spec,C] in (63b) cannot contain a remnant VP, since no remnant VP was created in the first place. Thus, (63b) is a case of V topicalization. I conclude – following Fanselow (2002) and Hinterhölzl (2002) – that there exist cases of predicate cleft in German which do not involve a remnant VP.¹² Given

¹²The theories which Fanselow and Hinterhölzl propose actually involve moving a VP to [Spec,C], not a V, as I have argued for. Specifically, Fanselow takes the moving V to be a VP in itself. His theory entails a revision of the traditional theta-theory. Hinterhölzl takes the fact that only a verb is pronounced at [Spec,C] to result from partial deletion. I take the empirical data that these authors present to support their analyses without adopting these analyses.

that these cases show no doubling, and that VP is head-final in German, it can be concluded that non-uniformity of chains alone does not prevent Copy Deletion, i.e. that the ECCD is to be preferred over the UCCD.

Truckenbrodt's challenge

Truckenbrodt (2009) poses a challenge to the claim that V topicalization exists in German.¹³ Assuming all VPs to be head-final in German, Truckenbrodt presents the paradigm in (64) and (65) to show that V can occupy [Spec,C] if and only if its argument can evacuate VP. Specifically, (64) is supposed to show that V cannot front if its complement cannot extrapose, and (65), that V can front if its complement can extrapose.

- (64) a. *Haben dürfte er sie geküsst
 have might he her kissed
 → **Haben** cannot front
 b. *Er dürfte sie t_{VP} haben [$_{VP}$... geküsst]
 he might her t_{VP} have kissed
 → the complement of **haben** cannot extrapose
- (65) a. ??Wollen hat er sie küssen
 want has he her kiss
 → **wollen** can front
 b. ??dass er sie t_{VP_1} hat [$_{VP_1}$ t_{VP_2} wollen] [$_{VP_2}$... küssen]
 that he her has want kiss
 → the complement of **wollen** can extrapose

I find Truckenbrodt's observations extremely interesting. However, I think there is another way to look at the facts. First, note that Truckenbrodt's analysis of (65b) implies that extraposition of VP as well as extraposition from an extraposed VP are both possible in German. This assumption, I think, is not uncontroversial. Second, it has been reported to me (Gisbert Fanselow p.c.) that the relative acceptability of (65a) is universal in German, while that of (65b) is dialectal. In other words, there are dialects of German which accept (65a) and do not accept (65b). This weakens Truckenbrodt's case, since his argument crucially appeals to the correlation in acceptability between (65a) and (65b). Third, the possibility

¹³The arguments I gave for the possibility of V topicalization in German were included in Trinh (2009). Truckenbrodt (2009) was a commentary on that paper, and the discussion of Truckenbrodt (2009) presented here is part of Trinh (2010), which is my reply to the commentaries on Trinh (2009).

of getting the complement of V out of VP is not a sufficient condition for having V in [Spec,C], as (66) shows (Gisbert Fanselow p.c.).

- (66) a. dass [sie küssen]₁ keiner [VP t₁ können] muss
 that her kiss no one can must
 → the complement of **können** can scramble
 b. *können muss sie küssen keiner
 can must her kiss no one
 → **können** cannot front

This shows that in principle, the perfect auxiliary **haben** may be prevented from topicalizing by some independent principle, just as the modal verb **können** is. Vietnamese strengthens this idea. In this language, VP do not extrapose or scramble, and fronting verbs requires V doubling. However, we see the exact same contrast between the perfect auxiliary **da** and the verb **muon** ‘want.’

- (67) a. *da thi John da gap Mary
 PERF TOP John PERF meet Mary
 b. ?muon thi John muon hon Mary
 want TOP John want kiss Mary

These facts suggest that the contrast – with respect to topicalization – between **haben** and **können**, on the one hand, and **wollen**, on the other, might be unrelated to extraposition/scrambling. The question, then, is what it is related to. I leave this question to future research.¹⁴

2.3.2 Dutch

Dutch also has predicate clefts, like German. Two examples are given in (68).¹⁵

- (68) a. kussen wil Jan een vrouw
 kiss wants Jan a woman
 ‘Jan wants to kiss a woman’

¹⁴A speculative note: it seems that V can topicalize in German (and probably in other languages too) only if its subject has a theta-role assigned by it. Thus, **scheinen** is ambiguous between ‘seem’ and ‘shine.’ Only under the latter reading can this verb undergo topicalization. Also, the **drohen** ‘threaten’ is ambiguous between an agentive and a non-agentive reading in German. However, fronting **drohen** forces the agentive reading. If this generalization is true, it would account for the ability of **wollen** to front, and the inability of **können** and **haben** to do the same, assuming that the latter two are raising verbs. I thank Gisbert Fanselow for the observations about **scheinen** and **drohen**.

¹⁵I thank Hedde Zeijlstra for providing the Dutch examples.

- b. gedronken heeft Jan een biertje
 drunk has Jan a beer
 'Jan drank a beer'

The next examples are evidence that predicate cleft in Dutch involves $\bar{A}$ -movement. Specifically, (69) shows that it is not clause-bound, and the four examples after (69) show that it is constrained by the Complex NP Constraint, the Subject Island Constraint, the Adjunct Island Constraint, and the Factive/Non-Bridge Island Constraint, in that order.

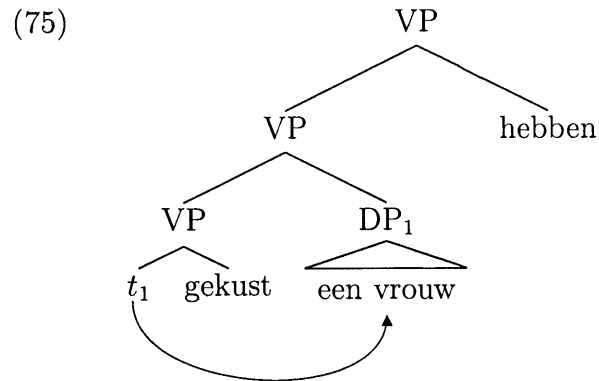
- (69) lezen denk ik dat Jan een boek wil
 read think I that Jan a book wants
 'I think Jan wants to read a book'
- (70) *lezen geloof ik dat verhaal, dat Jan een boek wil
 read believe I the story, that Jan a book wants
 ('As for reading, I believe the story that Jan wants to read a book')
- (71) *lezen is dat Jan een boek wil totaal verrassend
 read is that Jan a book wants totally surprising
 ('As for reading, that Jan wants to read a book is totally surprising')
- (72) *lezen ben ik gelukkig, omdat Jan een boek wil
 read am I happy because Jan a book wants
 ('As for reading, I am happy because Jan wants to read a book')
- (73) *lezen betreuer/fluister ik, dat Jan een boek wil
 read regret/whisper I that Jan a book wants
 ('As for reading, I regret/whisper that Jan wants to read a book')

Are there cases of predicate cleft in Dutch in which the object of the fronted verb stays inside VP? I think the answer is yes. Consider (74).

- (74) gekust wil hij vaak een vrouw hebben
 kissed will he often a woman have
 'He wants to have often kissed a woman'

The fact that the object **een vrouw** is preceded by a VP-adverb suggests that it has not scrambled out of the VP (**vaak** 'often' can modify the event of kissing). And the fact that **een vrouw** is followed by the auxiliary **hebben** is evidence that it has not extraposed. If the object in (74) had extraposed, its surface position

would have to be one between the auxiliary **hebben** and the base position of **gekust**. In other word, the derivation would have to contain the step in (75).



But step (75) is not available in Dutch: it is not possible in this language to extrapose a DP object to a position between the main verb and a sentence-final auxiliary. This is shown by the contrast in (76).

- (76)
- a. Jan wil een vrouw gekust hebben
 Jan wants a woman kissed have
 - b. *Jan wil gekust een vrouw hebben
 Jan wants kissed a woman have
 ‘Jan M have kissed a woman’

I conclude that the object **een vrouw** ‘a woman’ in (74) has neither scrambled nor extraposed out of VP. This means that what moves to [Spec,C] is not a remnant VP, but a bare V, which in turn means that Dutch has V-to-[Spec,C] movement.

2.3.3 Interim summary

It was argued that doubling in Hebrew and Vietnamese predicate clefts is predicted by the ECCD, which says that Copy Deletion requires the lower copy to be at the right edge of an XP if the higher one is. We then entertained another principle, the UCCD, which says that Copy Deletion requires the relevant chain to be uniform, and from which the facts in Hebrew and Vietnamese also follow. Our discussion of German and Dutch aims at settling the choice between the ECCD and the UCCD in favor of the former. The UCCD predicts that V-to-[Spec,C] movement in these two languages results in doubling, whereas the ECCD makes no such prediction. What we tried to argue for in the last two subsections is that

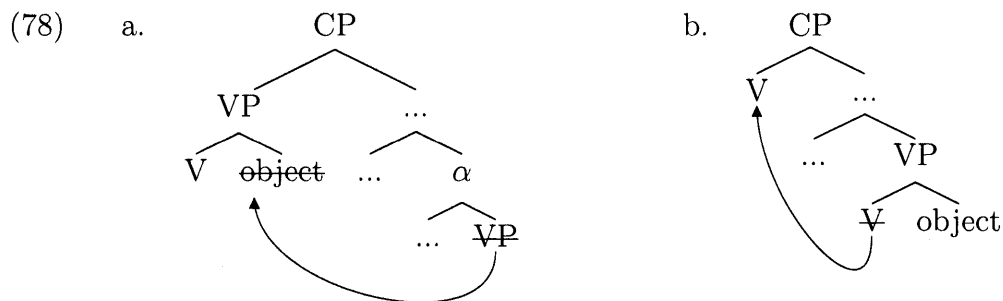
German and Dutch have V-to-[Spec,C] movement without doubling, contradicting the UCCD, but not the ECCD.

Note that while the absence of V doubling in German and Dutch is consistent with the ECCD, it is not necessitated by it. In other words, the ECCD allows, but does not require, the deletion of the lower V copy. The fact that V doubling is ungrammatical in German and Dutch, then, can be considered evidence for Pronunciation Economy (see chapter 1), repeated here.

- (77) Pronunciation Economy (PE)
Copy Deletion must apply when it can

2.4 Type 3: Swedish and Norwegian

Suppose language S is head-initial. The ECCD predicts that predicate clefting in S should show doubling of V, if it is V-to-[Spec,C] movement, since the lower copy of the resulting chain does not end an XP. Equivalently, the ECCD predicts that if predicate clefting in S does not show doubling, it cannot be V-to-[Spec,C] movement, but must be remnant VP movement. Thus, if an SVO language has predicate clefts without verb doubling, the ECCD entails that these clefts have the analysis in (78a), not (78b).



Swedish and Norwegian are SVO languages which have predicate clefts without verb doubling. In what follows, I will argue that these should be analyzed as involving remnant VP movement as in (78a), not V-to-[Spec,C] movement as in (78b). This is, in effect, arguing that Swedish and Norwegian are not counterexamples to the ECCD.

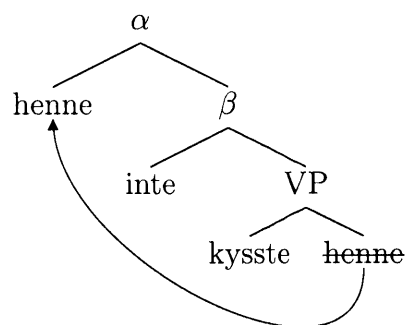
2.4.1 Swedish

Swedish is an SVO language which has predicate cleft without V doubling. Illustration is given in (79), taken from Holmberg (1999:7).

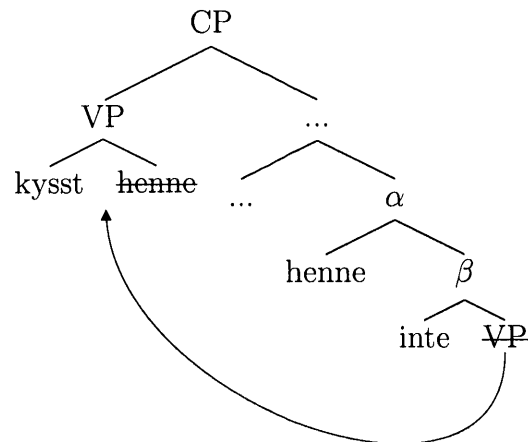
- (79) Kysst har jag henne inte (bara hållit henne i handen)
 kissed have I her not (only held her by the-hand)
 'Kissed her I haven't (only held her by the hand)'

Holmberg claims that predicate cleft in Swedish is V-to-[Spec,C] movement. But note that (79) does not force a V-to-[Spec,C] movement analysis. The object of **kysst** 'kissed' is to the left of the sentential negation **inte** 'not.' Thus, (79) could be derived by Object Shift followed by topicalization of the remnant VP, as shown in (80).

- (80) a. Object Shift



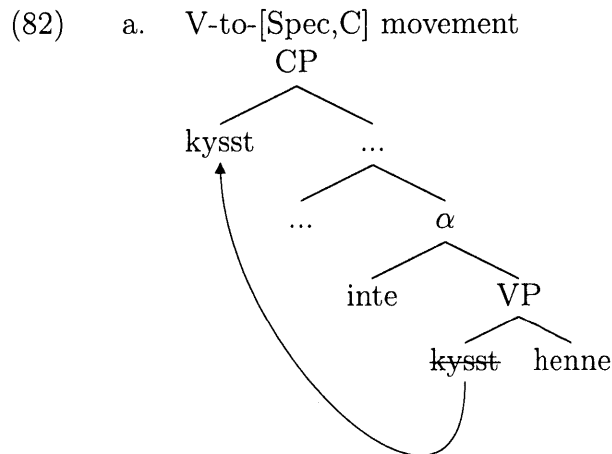
- b. Remnant VP movement



Support for (80) as the derivation of (79) is the fact that “OS is usually optional in Swedish but it is obligatory if the verb occurs in topic position” (Engels and Vikner 2009: 4), as the paradigm in (81) shows.¹⁶

- (81)
- a. Jag har inte kysst henne
I have not kissed her
 - b. Jag kysste henne inte
I kissed her not
 - c. *Kysst har jag inte henne
kissed have I not her
 - d. Kysst har jag henne inte
kissed have I her not

Holmberg, however, argues that (81) is not the derivation of (80). Instead, he proposes that (80) is derived by topicalization of V followed by counter-cyclic Object Shift, which he assumes to be a PF operation. The two relevant steps of this derivation is given in (82).

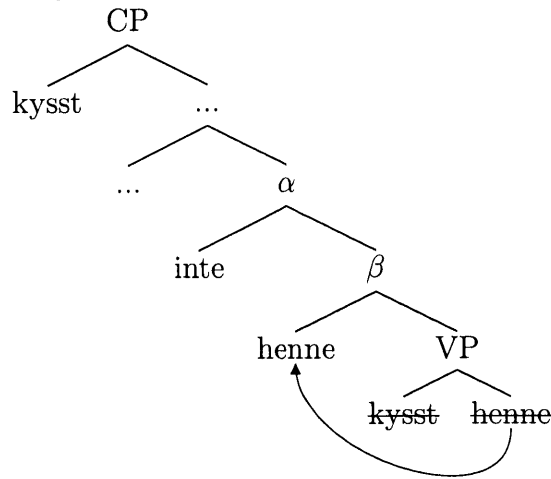


¹⁶I learned from Holmberg (2009) that this observation is first made in Erteschik-Shir (2001). Note that object shift is not possible across an unmoved verb.

- (i) Jag har henne inte kysst
I have her not kissed

This fact should not undermine the claim that predicate clefts in Swedish is remnant VP movement. In fact, it strengthens it: the same verb which, if unmoved, blocks object shift turns out to require it, when moved, which is strong evidence that object shift is the necessary condition for clefting.

b. Object Shift (at PF)



The reason for Holmberg to choose (82) as the derivation of (79) over (81) is Holmberg's Generalization (HG), given in (83).

- (83) Holmberg's Generalization (Holmberg 1999: 15)
Object Shift cannot apply across a phonologically visible category asymmetrically c-commanding the object position except adjuncts

Assuming HG, it would indeed be impossible to derive (79) by first shifting the object and then topicalizing the VP, because the first step of this derivation will violate HG: it is Object Shift applying across a phonologically visible category c-commanding the object position, namely the verb. Holmberg postulates HG on the basis of facts such as (84) (Holmberg 1999: 1–2). These show that the object cannot shift across an overt VP-internal constituent, the verb in the case of (84a), the indirect object and the verbal particle, in the case of (84b) and (84c), respectively.

- (84) a. *jag har henne₁ inte [VP kysst t₁]
I have her not kissed
b. *jag gav₁ den₂ inte [VP t₁ Elsa t₂]
I gave it not Elsa
c. *dom kastade₁ mej₂ inte [VP t₁ ut t₂]
they threw me not out

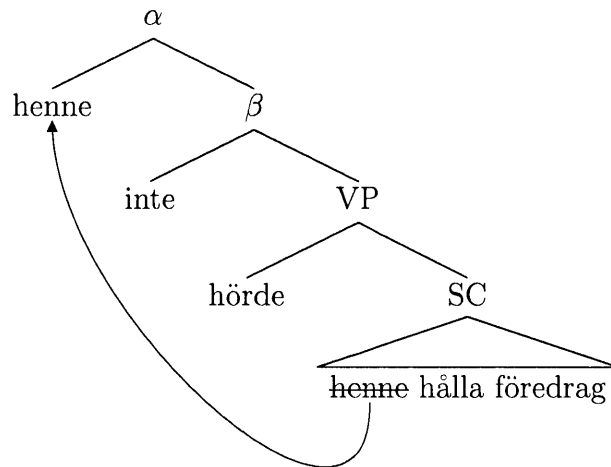
Holmberg also presents (85) as supporting evidence for HG (Holmberg 1999: 8–9).

- (85) a. jag hörde henne₁ inte [t₁ hålla föredrag]
I heard her not give talk

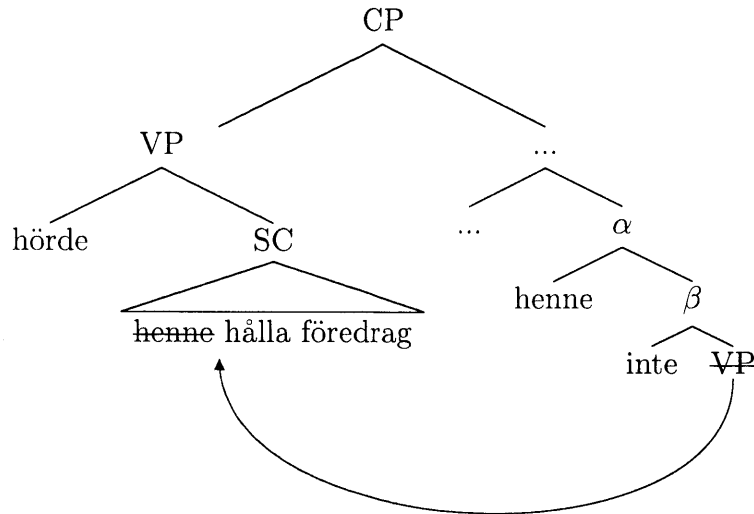
- b. [hörde henne hålla föredrag] har jag inte
 heard her give talk have I not
- c. *[hörde t_1 hålla föredrag] har jag henne₁ inte
 heard give talk have I her not

Example (85a) shows that the subject of the small clause can shift to the left of the matrix negation, while (85b) shows that fronting of a VP which contains a small clause is possible. If HG does not hold, nothing prevents us from deriving the ungrammatical (85c) by first shifting **henne** to the left of the matrix negation, then fronting the remnant matrix VP to [Spec,C], as in (86).

(86) a. Object Shift



b. VP topicalization



On the other hand, if HG is true, the derivation in (86) will not be possible, because (86a) violates HG: **henne** is shifted across **hörde**, which is phonologically visible and c-commands **henne**.

Thus, (84) and (85) led Holmberg to HG, and HG forces him to say that (79) has the derivation in (82), and consequently that Swedish has V-to-[Spec,C] movement. Now note that while HG implies the V-to-[Spec,C] movement analysis of Swedish predicate clefts, the facts in (84) and (85), being facts, do not imply HG. If there is another explanation for these facts which does not require HG, then Swedish predicate cleft can be analyzed as remnant VP movement.

Such an explanation is given in Fox and Pesetsky (2005). These authors propose a theory of syntax-phonology mapping that includes the following premises.

- (87) Fox and Pesetsky (2005)
- a. Spell-out linearizes VP and CP ('phases') cyclically
 - b. Spell-out cannot add inconsistent information
 - c. Traces are invisible to Spell-out

The reader should consult Fox and Pesetsky's original paper for details.¹⁷ For present purposes, I will just concentrate on the second assumption – (87b) – which can be stated informally as follows: if X precedes Y at one Spell-out, X precedes Y at every Spell-out. For example, if V precedes the object when VP is spelled out, then V must precede the object when CP is spelled out. This theory turns out to account elegantly for all the facts in (84) and (85). The ungrammaticality of (84a) is bad because the verb precedes the object when VP is spelled out but follows it when CP is spelled out. Similarly for (84b) and (84c), with the role of the verb played by the indirect object and the particle, respectively. The ungrammaticality of (85c) is due to the fact that when the VP containing the small clause is spelled out, the small clause subject **henne** precedes **hålla** and **föredrag**, but when the matrix CP is spelled out, it follows **hålla** and **föredrag**.

Fox and Pesetsky also present the following contrast as additional evidence for their theory and against HG.

- (88)
- a. ?[Gett henne t_1] har jag den₁ inte
 given her have I it not
 - b. *[Gett t_1 den] har jag henne₁ inte
 given it have I her not

¹⁷Fox and Pesetsky's (2005) proposal holds that 'traces' are irrelevant for determining linear order. In other word, only higher copies play a role. As far as I can see, this does not affect my proposal: Fox and Pesetsky's theory determines the linear order of higher copies of chains, and the CCD determines whether the lower copy of a single chain can be deleted.

The relative acceptability of (88a) shows that Object Shift across the verb followed by remnant movement of VP must be possible in Swedish. More significantly, the contrast between (88a) and (88b) speaks against HG, since HG predicts no contrast between these sentences. They incur the same violation of HG, namely Object Shift across the verb, and should thus be equally bad. The theory of Fox and Pesetsky, on the other hand, predicts the contrast in (88). In (88a), the relative order between **gett**, **henne** and **den** remains the same at both the VP and the CP level. In (88b), this is not the case: when VP is spelled out, the order is **gett** < **henne** < **den**, but when CP is spelled out, it is **gett** < **den** < **henne**.

Thus, the theory of Fox and Pesetsky explains the facts without invoking HG. Since HG is what forces Holmberg to reject the remnant movement analysis for Swedish predicate fronting, Fox and Pesetsky have in effect shown that this analysis is possible. These authors actually went further and show that the remnant movement analysis is necessary, i.e. that not only does Swedish have remnant VP topicalization, it also lacks V-to-[Spec,C] movement. Consider the following examples from Fox and Pesetsky (2005: 27).

- (89) a. *Hört₁ har jag henne₂ inte t₁ [t₂ hålla föredrag]
 heard have I her not give talk
 b. *Hört₁ har jag inte t₁ [Per hålla föredrag]
 heard have I not Peter give talk

If V-to-[Spec,C] were possible in Swedish, we would incorrectly expect (89a) and (89b) to be acceptable: they could be derived by moving the main verb to [Spec,C]. But if V-to[Spec,C] movement is impossible in Swedish, the ungrammaticality of these sentences might be accounted for by invoking the general impossibility of extracting the ECM infinitival from the VP containing it (cf. Fox and Pesetsky 2005: 27).¹⁸

We have seen that the theory of cyclic linearization proposed by Fox and Pesetsky can account for the paradigms in (84) and (85) without recourse to Holmberg's Generalization (HG), which is what forces the assumption that Swedish has V-to-[Spec,C] movement. It must be noted, however, that Holmberg does provide

¹⁸Note that the German counterpart of (89) is perfectly acceptable (Marie-Christine Meyer, Irene Heim p.c.).

- (i) arbeiten habe ich ihn gesehen
 work have I him seen
 'I have seen him work'

additional facts to evidence the existence of such movement in Swedish. The following examples are taken from Holmberg (1999).

- (90) a. ?Kysst har jag inte Marit
 kissed have I not Marit
 b. Sett har jag inte den idioten (men jag har talat med honom
 seen have I not the idiot (but I have talked with him
 på telefon)
 on phone)
 c. Bo ska han i Malmö, men han ska jobba i Koppenhamn
 live will he in Malmö, but he will work in Koppenhamn

These sentences show a single verb being fronted to [Spec,C], stranding its internal argument inside the sentence. If the stranded internal arguments in (90) are at their base position, then we are forced to say that Swedish has V-to-[Spec,C] movement, which is Holmberg’s claim. To counter this claim, we need to say either (i) that the internal arguments in (90) are not at their base position but, say, have extraposed to the right of the VP, or (ii) that the judgement given for (90) is not totally accurate. My argument will be a combination of both. Note, first, that among the objects in (90) – i.e. **Marit** in (90a), **den Idioten** in (90b) and **i Malmö** in (90c) – the first is likely the hardest to extrapose, the last the easiest. Among the Swedish speakers I have asked, all found (90a) and (90b) to be bad, and some found (90a) to be “very bad.” This strongly suggest that predicate clefts in Swedish is possible to the extent that the verb’s internal argument(s) can vacate the VP, either by Object Shift or by extraposition. As a matter of fact, Holmberg (2009) revised his judgement, giving (90a) a star and (90b) two question marks.¹⁹ Holmberg also accepts that the PP object in (90c) is of the type which “can, and willingly does, extrapose [...]” and concedes that the arguments against V-to-[Spec,C] movement in Swedish are “sound enough” (Holmberg 2009: 271).

2.4.2 Norwegian

I have not been able to find systematic discussion of Norwegian with respect to predicate fronting. From the data I collected, however, it appears that Norwegian lacks V-topicalization. Consider (91).

¹⁹Holmberg (2009) is a commentary on Trinh (2009), which contains the arguments against V-to-[Spec,C] movement in Swedish presented above.

- (91) a. syngi har jeg ikke
 sung have I not
 ‘I did not sing’
 b. syngi trur jeg at han ikke har
 sung believe I that he not has
 ‘I believe he did not sing’

These sentences show that Norwegian allows [Spec,C] to be occupied by intransitive verbs. However, when the verb is transitive, topicalization is possible only if the object is fronted together with the verb.

- (92) a. sett mannen har jeg ofte
 seen the-man have I often
 b. *sett har jeg oft mannen
 seen have I often the-man
 ‘I often saw the man’

This suggests that Norwegian allows VP-topicalization, but not V-topicalization. Consequently, predicate fronting in Norwegian must be remnant VP movement, which means that it is possible only if the object has evacuated the VP. This seems correct, as (93) was judged as relatively acceptable by my informant.

- (93) ?sett har jeg dem ofte
 seen have I them often
 ‘I saw them often’

In (93), the object has shifted out of the VP, as the VP-adverb to its right indicates. The surface word order results from moving the remnant VP to [Spec,C]. I tentatively conclude that Norwegian does not have V-topicalization.

2.4.3 Interim summary

Norwegian and Swedish are SVO languages. The ECCD implies that if predicate clefts in these languages involve V-to-[Spec,C] movement, they will exhibit double pronunciation of the sort observed in Hebrew and Vietnamese. As Norwegian and Swedish predicate clefts do not show double pronunciation, we expect them not to be derived by V-to-[Spec,C] movement. The goal of the previous subsection has been to argue that our expectation, contrary to claims that have been made in the literature, is not contradicted by facts.²⁰

²⁰Holmberg (1999: 11) notes that V-topicalization, i.e. V-to-[Spec,C] movement, “has not been discussed in the literature [...] in relation to the Scandinavian languages.” While the absence of

2.5 Head-to-head movement

2.5.1 The typology

The last three previous sections discussed predicate clefts in Hebrew, Vietnamese, German, Dutch, Swedish and Norwegian with the aim of defending the hypothesis that chain linearization is subject to the Edge Condition on Copy Deletion (ECCD). I have argued for a V-to-[Spec,C] analysis of predicate clefts in Hebrew, Vietnamese, German and Dutch and for a remnant VP movement analysis of predicate clefts in Swedish and Norwegian. How these arguments support the ECCD is explained in the interim summary at the end of each subsection. Given the fact that Hebrew and Vietnamese show verb doubling while the Germanic languages do not, my argument is, in effect, an argument for the following typology (the gray cells contain SVO languages).²¹

	+V doubling	–V doubling
(94) +V topicalization	Hebrew, Vietnamese	German, Dutch
–V topicalization	—	Swedish, Norwegian

Descriptively, then, the languages vary with respect to two “parameters”: [$\pm$ V topicalization] and [$\pm$ V doubling]. Without the ECCD, both parameters would be irreducible, i.e. the structure of the typology would be totally arbitrary. Assuming the ECCD allows us to do away with one parameter which, interestingly, can be either [$\pm$ V topicalization] or [$\pm$ V doubling]. Thus, suppose we take [$\pm$ V

V-topicalization in the discussion of the Scandinavian languages might be an unfortunate fact, the reason for it might just be that V-topicalization is absent in these languages. Den Besten and Webelhuth observed that “[t]here is a sharp contrast between the Germanic SVO and SOV languages with respect to sentences where a nonfinite verb is topicalized [...]” (den Besten and Webelhuth 1987b: 15). The ECCD itself does not derive this observation, but it does provide us with a way to do that, namely by assuming that the Germanic languages, for some reason, disallow verb doubling of the sort seen in Hebrew and Vietnamese. This issue will be discussed below.

²¹What has been shown in section 2, 3 and 4 is, strictly speaking, that some cases of predicate clefts in Hebrew, Vietnamese, German and Dutch involve V-to-[Spec,C] movement, and some cases of predicate clefts in Swedish and Norwegian involve remnant VP movement. What I have intended to argue for is instead the following: V-to-[Spec,C] movement is involved in all cases of predicate clefts in Hebrew and Vietnamese, some cases of predicate clefts in German and Dutch, and no cases of predicate clefts in Swedish and Norwegian. As far as I can see, there is no empirical reason not to make the inductive inference from what has been shown to what I have intended to argue for. In fact, once we abandon the SPH/Chain Uniformity, we might even assume, perhaps on consideration of economy, that all cases of predicate clefts in Hebrew, Vietnamese, German and Dutch, and no cases of predicate clefts in Swedish and Norwegian, involve V-to-[Spec,C] movement. This assumption, in fact, underlies the typology in (94).

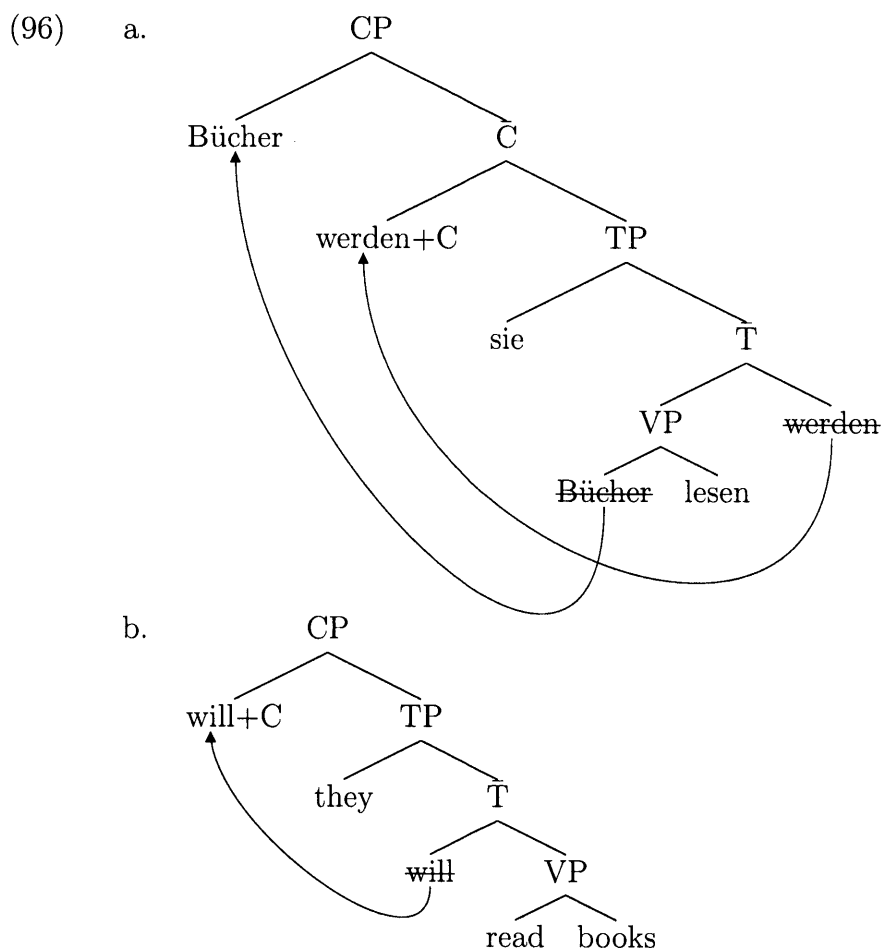
doubling] to be primitive: [+V doubling] languages are those that allow V doubling and [-V doubling] languages those that do not. Whether a language L allows V topicalization or not is then a consequence of the ECCD and the setting of [$\pm$ V doubling] for L. By hypothesis, the Germanic languages are [-V doubling]. As German and Dutch are SOV, fronting V does not necessitate V doubling, hence no problem arises. Swedish and Norwegian, on the other hand, are SVO languages, which means that topicalizing V requires doubling of V, contradicting the [-V doubling] setting. As a consequence, Swedish and Norwegian do not topicalize V. As for Hebrew and Vietnamese, the relevant parametric setting is [+V doubling]. The fact that these are SVO languages entails that fronting V requires V doubling, and since V doubling is possible, V topicalization is possible. Now what if we assume that [$\pm$ V topicalization] is primitive? Hebrew, Vietnamese, German and Dutch would then be set to [+V topicalization]. The fact that Hebrew and Vietnamese have V doubling while German and Dutch do not follows from the first two being SVO and the last two being SOV languages. As for the Scandinavian languages, they are [-V topicalization], which means they just do not allow movement of V to [Spec,C]. The absence of V doubling would then follow trivially.

In this work, I will adopt the former approach: I will assume that [$\pm$ V doubling] is primitive, and that V topicalization is in principle possible, with its absence in certain languages following from the interaction of the ECCD with the setting of [$\pm$ V doubling] in those languages. In other word, I assume that languages vary with respect to how chains are pronounced, not in what can be moved to [Spec,C]. There might be a conceptual argument in favor of this position. We have said that it is Economy that favors copy deletion. Thus, languages that allow V doubling are, in a sense, less economical than those that do not. As this sort of variation has been accepted as a theoretical option, for example with respect to wh-movement, the idea of a [$\pm$ V doubling] parameter will not be, say, as exotic as that of a [$\pm$ V topicalization] parameter, for which no parallel, in whatever sense, seems to have been entertained. Although this conceptual argument is not totally without force, it would perhaps be too bold a move to base the choice of [$\pm$ V doubling] over [$\pm$ V topicalization] as a parameter on it alone. I believe that a more substantial argument can be made to justify this choice. This argument pertains to an issue whose discussion we have deferred until this point: head-to-head movement. It is the subject of the next subsection.

2.5.2 Head-to-head movement as a PF operation

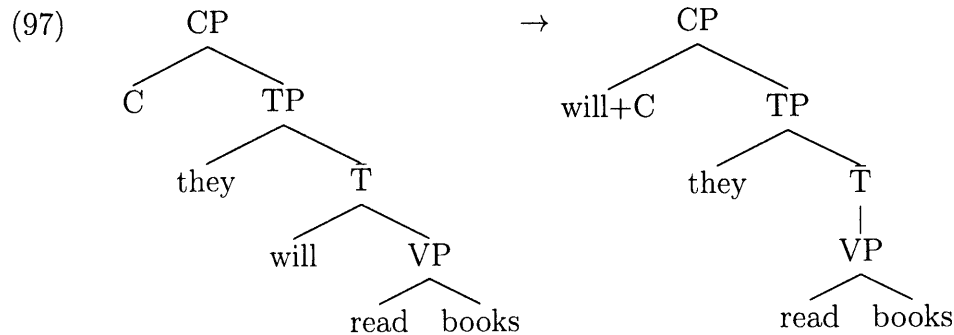
Observationally, head-to-head movement never results in double pronunciation: T-to-C movement in German behaves the same as T-to-C movement in English (cf. (95)), even though T in German ends an XP (Fanselow and Felix 1987) and T in English does not (cf. (96)).

- (95) a. Bücher werden sie lesen (*werden)
 books will they read (*will)
 b. will they (*will) read books?



The fact that the lower T copy deletes in English is a *prima facie* counter-evidence to the ECCD, since this copy does not end an XP. What I propose to account for this fact is that there is no lower T copy. In other word, I propose that head-to-head movement does not leave a copy, and consequently does not create a chain. Thus, (96b) is actually misleading as a representation of how (95b) is

derived. A more appropriate representation is perhaps the following sequence of phrase structures.



The hypothesis that head-to-head movement creates no chains sits well with the assumption that head-to-head movement takes place on the PF branch of the derivation. Lower copies are needed for the computation of meaning, but PF operations, by virtue of the grammatical model we adopt (see chapter 1), do not have any effect on semantic interpretation. Thus, there is no need, so to speak, for PF movements to leave copies at the base position. The idea that head-to-head movement is best viewed as a PF operation, driven by morpho-phonological requirements, has received a number of supporting arguments, both conceptual and empirical. Chomsky (2001), for one, presents several reasons to “suspect that a substantial core of head-raising processes [...] may fall within the phonological component.” One is that “semantic effects of head-raising in the core inflectional system are slight or non-existent [...]” For example, standard analyses place the inflected main verb in C for German, in T for French and in situ for English, but it is obvious that this has no interpretative consequences, as shown by the fact that the following three sentences mean exactly the same thing.

- (98)
- a. Tristan liebt Isolde
 - b. Tristan aime Isolde
 - c. Tristan loves Isolde

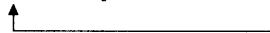
Another reason for relegating head-to-head movement to the PF component is the need, arising under the assumption that head-to-head movement is part of narrow syntax, to stipulate two sorts of movement-triggering features, namely those that drive head-to-head movement, and those that drive XP-to-Spec movement. Thus, the feature on T which requires V-to-T movement must, by stipulation, be such that it cannot be checked by movement of VP to [Spec,T], and the feature on T which requires movement of DP to [Spec,T] may not be checked by, say, D-to-T movement. However, this seemingly arbitrary difference would “follow naturally if overt V-to-T raising, T-to-C raising, and N-to-D raising are

phonological properties, conditioned by the phonetically affixal character of the inflectional categories.” The third reason has to do with the structural characteristics of head-to-head movement: it violates the Extension Condition, which requires the target of movement to be the root node of the structure to which the movement rule applies. Also, if excorporation is excluded, head-to-head movement is not successive-cyclic like phrasal A- and $\bar{A}$ -movement: after movement of X^0 to Y^0 , X^0 can only move further as subpart of $[_{Y^0} Y^0 X^0]$. This means, technically, that a head can undergo head-to-head movement only once. Speculating that “iterability is a general property of operations of narrow syntax, but these alone,” Chomsky takes this to be evidence that head-to-head movement belongs to the phonological component.

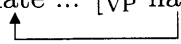
Considerations of a more empirical nature are found in the discussion of pseudogapping in Lasnik (1995, 1999a,b), Boeckx and Stjepanović (2001). Assuming the theory of Koizumi (1993, 1995) where the VO order in English results from overt Object Shift, i.e. movement of the object to $[\text{Spec, Agr}_O]$, followed by overt (head-to-head) movement of V to a higher VP shell, Lasnik proposes to analyze pseudogapping (cf. (99a)) as the combined result of Object Shift and VP ellipsis without V having moved from VP (cf. (99b)).

- (99) a. Mary hasn't dated Bill, but she has Hary
 b. $[_{TP} \text{ she has } \dots [_{\text{Agr}_O P} \text{ Hary } \dots [_{VP} \text{ dated } \text{Hary}]]]$
- 
- The diagram shows a horizontal line connecting the $[_{VP} \text{ dated } \text{Hary}]$ part of the structure in (99b) to the $[_{\text{Agr}_O P} \text{ Hary}]$ part, with an upward arrow pointing to the $[_{\text{Agr}_O P}$ position.

Lasnik's analysis raises the question as to why the verb must raise past the shifted object if VP ellipsis does not apply, i.e. why (100) is ungrammatical.

- (100) $*[_{TP} \text{ she has } \dots [_{\text{Agr}_O P} \text{ Hary } \dots [_{VP} \text{ dated } \text{Hary}]]]$
- 
- The diagram shows a horizontal line connecting the $[_{VP} \text{ dated } \text{Hary}]$ part of the structure in (100) to the $[_{\text{Agr}_O P} \text{ Hary}]$ part, with an upward arrow pointing to the $[_{\text{Agr}_O P}$ position.

The answer Lasnik gave to this problem is essentially that an unmoved V causes PF crash. Lasnik entertained different versions of this causation, each of which accommodates a specific view that Chomsky has about the motivation for overt movement. The details of these versions and the precise ways in which they differ need not concern us here. Important to note is the idea that whatever violation the unmoved V incurs, it must be such that it can be remedied by the absence of this V at PF. Thus, what drives head-to-head movement of V must ultimately be a PF requirement, and VP ellipsis, which by hypothesis is just PF deletion of VP, can eliminate the need to satisfy that requirement. Note that although this account is very close to suggesting that V moves at PF, it does not imply it: the PF crash in question may be avoided by V moving in the syntax. Boeckx and Stjepanović (2001), however, argue that there is empirical reason to assume that V moves at PF. The relevant data point is the ungrammaticality of (101a).

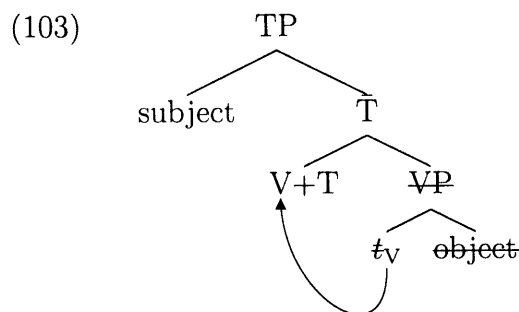
- (101) a. *Mary will like John, but Sue will hate
 b. *[TP Sue will ... [VP hate ... [VP ~~hate John~~]]]
- 

Boeckx and Stjepanović (2001) point out that nothing in Lasnik’s account forces Object Shift to precede V movement, preventing the derivation in (101b), in which V movement is followed by VP ellipsis with the object remaining in situ.²² These authors suggest that V movement be considered a PF operation, driven by purely morphological or phonological requirements. As PF operations, by definition, follow operations in the syntax, of which Object Shift is one, the derivation in (101b) will be impossible. Thus, “if head movement is a PF phenomenon, we have an answer to Lasnik’s puzzle as to why V need not raise in the pseudogapping cases” (Boeckx and Stjepanović 2001: 352). By pushing Lasnik’s idea that PF plays a crucial role in pseudogapping, then, Boeckx and Stjepanović (2001) provide an elegant solution to a problem that Lasnik’s account gives rise to.

Another very interesting argument for head-to-head movement being a PF operation is given in Schoorlemmer and Temmerman (2011). These authors draw attention to the fact that generally, an elided VP has to be lexically identical to its antecedent, except when phrases have moved from VP.

- (102) a. John should visit Mary, but Bill shouldn’t ~~visit Mary~~/*~~visit Sue~~
 b. John should visit Mary, but Sue he shouldn’t ~~visit t_{Sue}~~
 c. Mary was visited *t_{Mary}*, but Sue wasn’t ~~visited t_{Sue}~~

Thus, either [V XP] or [V *t_{XP}*] can antecede [V *t_{XP'}*], even if XP and XP’ differ lexically. Schoorlemmer and Temmerman then points out that the same does not hold when it is V that moves. They base their claim on cases of “V stranding VP Ellipsis” (VVPE), where the T head which governs the elided VP hosts not an auxiliary, but by the main verb itself, which has moved to T from VP.



²²Boeckx and Stjepanović (2001) also show that the explanation for (101) proposed in Lasnik (1999b) is inadequate.

Examples (104a) and (104b) show, for Portuguese, that $[t_V \text{ XP}]$ can antecede $[t_{V'} \text{ XP}]$ only if V and V' are one and the same lexical item.²³

- (104) a. A Ana pôs os livros na pasta, mas tu não puseste
 the Ana put the books in the case but you not put
 ‘Ana put the books in the case, but you didn’t’
 b. *Quando a Ana colocou os óculos na mesa, a Maria
 when the Ana placed the glasses on the table the Maria
 também pôs
 too put
 (“When Ana placed the glasses on the table, Maria did too”)

With respect to ellipsis licensing, then, $[t_V \text{ XP}]$ and $[t_{V'} \text{ XP}]$ behave just like $[V \text{ XP}]$ and $[V' \text{ XP}]$, i.e. as if the verb has not moved out of VP at all. While the difference between traces of XP movement and traces of V-to-T movement can be accounted for by defining the relevant notion of “identity” and such a way that, say, t_{John} and t_{Bill} are identical while t_{PLACE} and t_{PUT} are not, or by stipulating that traces of V-to-T movement must be replaced by the moved heads at LF, i.e. that head-to-head movement forces reconstruction (cf. Goldberg 2005: 175–199), Schoorlemmer and Temmerman propose that V-to-T movement be considered a PF operation, which means that at LF, where the identity condition on ellipsis holds, V and V' are inside VP, causing ungrammaticality if they differ lexically.

Another source of evidence, albeit quite indirect, for analyzing head-to-head movement as a PF operation, creating no chains, is found in the psycholinguistic work of Grodzinsky and Finkel (1998), which tested Broca’s and Wernicke’s aphasics on their ability to judge sentences involving movement. Grodzinsky and Finkel report that both groups of aphasics show “a marked deficit in structures involving XP movement [...] whereas structures and violations involving [...] X^0 movement [...] were relatively preserved” (Grodzinsky and Finkel 1998: 286). In other word, “the disturbance to receptive syntax in Broca’s and Wernicke’s aphasia is highly restricted, pertaining solely to the representation of dependencies between traces and their antecedents, provided that these antecedents are phrasal, not lexical, categories” (Grodzinsky and Finkel 1998: 287). The authors conclude that “[t]he intactness of X^0 -movement (in both judgment and comprehension) rules out a disorder overarching all types of syntactic transformations” (Grodzinsky and Finkel 1998: 288). Note that this conclusion is made under the assumption that head-to-head movement (“ X^0 movement” in Grodzinsky and

²³The reader is invited to consult Schoorlemmer and Temmerman (2011) and references therein for arguments that constructions such as (103a) should in fact be analyzed as VP ellipsis with V having moved to T, as well as facts similar to (103) in other languages.

Finkel's terminology) leaves traces just as A- and $\bar{A}$ -movement do. However, Chomsky points out, correctly in my opinion, that the experimental result "is expected if head raising is a phonological process, creating no chains" (Chomsky 2001: 38).

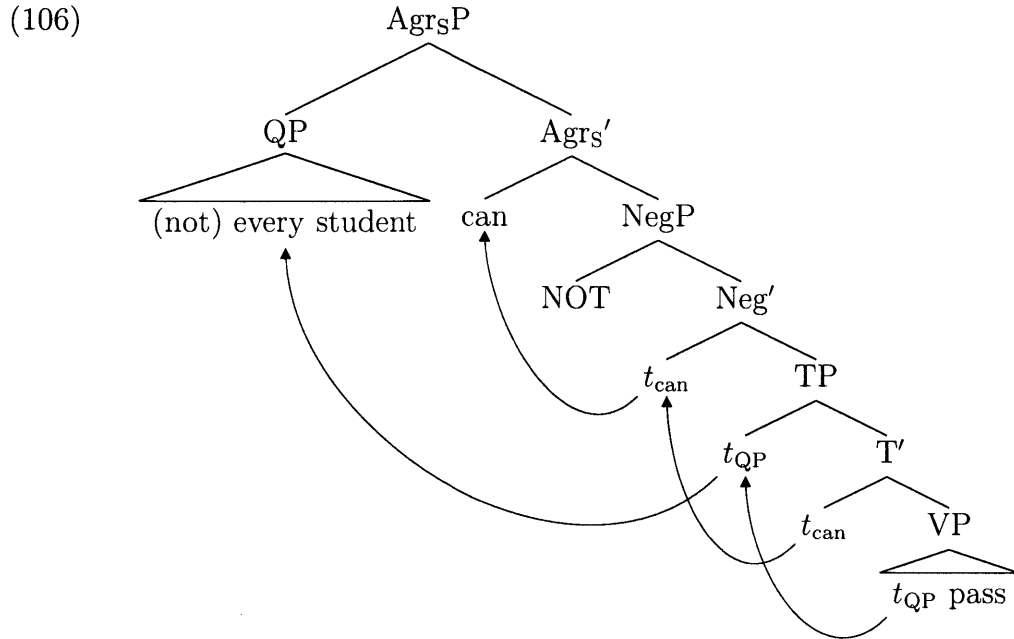
I have reviewed some arguments supporting the hypothesis that head-to-head movement is a PF operation, creating no chains and having no interpretive effects. I turn now to the opposite view, which is that head-to-head movement is not a PF operation. Matushansky (2006) and Vicente (2007), assuming a framework in which syntactic operations can intersperse with morphological ones, propose to analyze head-to-head movement as consisting of two steps: head-to-Spec movement followed by morphological merger. This approach provides a way to make sense of several special characteristics of head-to-head movement, for example the lack of "iterability" as discussed in Chomsky (2001), but it is not clear, at least not to me, how it is more empirically adequate, or less "stipulative," than the PF theory of head-to-head movement.

An elaborate argument is given in Lechner (2006) that head-to-head movement has semantic import, hence must take place in the syntax. The empirical basis of Lechner's proposal consists in sentences of the form [Negative QP + modal + VP], as exemplified by (105).

(105) Not every student can pass

Lechner notes that (105) can be understood as 'there is no possibility that every student passes,' with the scopal relation between the logical constants being 'not > can > every.' A series of auxiliary assumptions are then made. First, Lechner extends the analysis proposed in Stechow (1993), Penka and von Stechow (2001), Penka (2002) for negative quantifiers in German and claims that **not every NP** has the same meaning as **every NP**, with **not** being a morphological reflection of (syntactic) agreement between the quantifier and an abstract negation head.²⁴ Second, Lechner attributes to (105) the derivation/structure in (106), with NegP intervening between AgrsP and TP, the modal verb undergoing head-to-head movement from T to Neg to Agrs and the QP undergoing A-movement from [Spec,V] to [Spec,T] to [Spec,Agrs].

²⁴We will write "NOT" for the abstract negation head and put **not** in brackets to indicate that it is not interpreted.



These auxiliary assumptions together entail that if the subject QP cannot reconstruct into VP, the attested ‘not > can > every’ reading is possible only if **can** is interpreted at the position of the intermediate trace, below NOT but above TP, i.e. if head-to-head movement of **can** has semantic effects. What Lechner argues for in his paper, then, is that the subject in fact cannot reconstruct into VP. One argument starts from the observation that reconstruction into small clauses is impossible (Stowell 1981, Williams 1983).

- (107) A linguist_i seems [_{SC} *t_i* unhappy]
de re / *de dicto

From this observation, Lechner appeals to the “widely shared assumption” that complements of modals are small clauses (Stowell 1983, 2004) to conclude that reconstruction into the VP in (106) is impossible. The weakness of this argument, I think, is that it relies on a vague use of the term “small clause.” As far as distributional facts are concerned, the complement of **seem** in (107) just cannot be identified with the complement of **can** or any other modal verb (cf. ***John can unhappy**), and the fact that both have been called “small clause” is not, in my view, sufficient ground for making the inference that Lechner makes.

Another of Lechner’s arguments concern the fact that fronting VP results in “scope freezing” (Barss 1986) as exemplified by the lack of wide-scope reading for the universal quantifier in (108).

- (108) [_{VP} teach every student] no one will *t*_{VP}
 no one > every student, *every student > no one

Lechner then notes that “[e]ven though the concrete factors involved in the analysis of scope freezing are poorly understood, any successful theory has to [...] block reconstruction of the subject into the copy of the fronted constituent,” which gives (108), according to Lechner, the status of evidence for the inability of subjects to reconstruct into [Spec,V]. This argument, in my view, is not very convincing. It is undeniable that inverse scope is possible in English when VP is not fronted. The puzzle here is why inverse scope is not possible when VP is fronted. What Lechner says is really that because inverse scope is not possible when VP is fronted, inverse scope when VP is not fronted must not be explained in terms of subject reconstruction into VP. This seems arbitrary.

The last argument Lechner provides pertains to NPI licensing. From the paradigm in (109), Lechner concludes that universal quantifiers cannot intervene between negation and **ever**.

- (109) a. It is not possible that you will ever be on the team
 b. *It is not possible that everyone will ever be in the team

He then argues that **ever** is a TP-adverb, as follows: (i) **always** takes scope over modals, which are base generated in T, and (ii) **ever** is merged to the left of **always**, therefore (iii) **ever** must be as high as TP.²⁵ It then follows, according to Lechner, that the ungrammaticality of (110) is due to the inability of the QP to reconstruct below T.

- (110) *Not everyone can ever be on the team

This argument, as I see it, is circular, in the sense that it presupposes the conclusion. In order to say that the ungrammaticality of (110) is due to the subject’s inability to reconstruct below T, Lechner must assume that had the subject been able to reconstruct below T, the sentence would be grammatical. In other words, he assumes that intervention between negation and **ever** by a universal quantifier can be nullified by the intervening quantifier reconstructing below **ever**. But this

²⁵The claim that **always** takes scope over modals and **ever** is merged higher than **always** are backed up by (i) and (ii), respectively.

- (i) He can always count on me
 always > can, *can > always
- (ii) a. No one source is ever always authoritative
 b. *No one source is always ever authoritative

means that the premise of his argument, namely that (109b) is bad because the subject intervenes between the negation and **ever**, turns out to be saying that the subject cannot reconstruct below **ever**, which is just what he sets out to derive.

In summary, Lechner’s argument for interpretive effects of head-to-head movement consists of the following steps: (i) establishing reading R for sentence S, (ii) analyzing S in such a way that R requires interpreting an X^0 in a derived position, if subjects cannot reconstruct into VP, and (iii) providing evidence that subjects cannot reconstruct into VP. I will take it for granted that step (i) is justified by observation, and have argued above that Lechner has not been successful with respect to step (iii). But note that his step (ii) is also questionable. His analysis of (105), namely (106), is exotic, to say the least, given the fact that standard analyses situate NegP below TP (Laka 1989, 1990) and arguments, convincing in my view, have been given against Agreement Phrases (Iatridou 1990, Chomsky 1995). Thus, suppose we assume the standard [T > Neg > modal > VP] hierarchy (where the modal obligatorily raises to T), the ‘Neg > modal > subject’ reading that Lechner claims is available for (105) can still be induced if we assume that there is an intermediate position below T and the modal but above the main VP into which subjects can reconstruct (McCloskey 1997), a view which is by no means implausible and which might, on considerations of economy, be a preferable option compared to the structure in (106).²⁶

Another argument for syntactic head-to-head movement, simpler but more convincing than Lechner’s in my view, is given in Hartman (2011). In this work, Hartman takes as given that ellipsis in English can be either deletion of VP (“VP ellipsis”) or deletion of $\bar{C}$ (sluicing), and proposes a theory that predicts which kind of ellipsis is possible in which syntactic configurations. Hartman’s proposal builds on Takahashi and Fox (2006) and, greatly simplified, delivers the following result: ellipsis of a “deletable category” X – i.e. X is either VP or $\bar{C}$ – requires a constituent Y which (i) contains X, (ii) contains no deletable category larger than X, and (iii) contains no free variable. An assumption Hartman needs to make is that every instance of movement creates a lambda abstract in the sense of Heim and Kratzer (1998). Relevant to the present discussion is how Hartman explains the following paradigm.

- (111) A: Mary will leave.
 B: Do you know when (she will)? / When (*will she)?

²⁶In fact, the assumption that subjects must be able to reconstruct below T is convincingly argued for in Johnson and Tomioka (1999).

It can be observed that *wh*-adverbials allow both sluicing and VP ellipsis in embedded contexts, but allow only sluicing in matrix contexts. By hypothesis, the embedded *when*-sentence in (111b) has the structure in (112a), while the unembedded that in (112b).

- (112) a. $\text{when } [\alpha \lambda x [\bar{C} \ C \ [_{TP} \ x \ [_{TP} \ \text{she } [\beta \lambda y [\bar{T} \ \text{will } [_{VP} \ y \ \text{leave }]]]]]]]]$
 b. $\text{when } [\alpha \lambda x [\bar{C} \ \text{will } [\gamma \lambda z [_{TP} \ x \ [_{TP} \ \text{she } [\beta \lambda y [\bar{T} \ z \ [_{VP} \ y \ \text{leave }]]]]]]]]]]$

In (112a), β contains no free variable and VP is the largest deletable category in it, so VP ellipsis is possible. Sluicing is also possible in (112a) since α contains no free variable and $\bar{C}$ is the largest deletable category in α . In (112b), on the other hand, α is the only constituent which contains no free variable, and as the largest deletable category in α is $\bar{C}$, only sluicing is possible. This fits the observations exactly. Note that Hartman's explanation relies crucially on the assumption that T-to-C movement creates a lambda abstract, i.e. leaves a trace and introduces a binder, just like A- and $\bar{A}$ -movement do.

Acknowledging the elegance of Hartman's proposal, I will nevertheless take the arguments for considering head-to-head movement a PF operation to be convincing enough for us to call the whole issue an open one. In what follows, I will take the side of the people who contend that head-to-head movement takes place at PF and leaves no trace.

2.5.3 The $[\pm V \text{ doubling}]$ parameter

Recall the question to be resolved: whether we should assume $[\pm V \text{ doubling}]$ or whether we should assume $[\pm V \text{ topicalization}]$. We have seen that given the ECCD and the basic underlying word order of the languages, either parameter is sufficient to capture the typology in section 1.5.1, repeated below. The reader is referred to section 1.5.1. for discussion of this point.

	+V doubling	−V doubling
(113) +V topicalization	Hebrew, Vietnamese	German, Dutch
−V topicalization	—	Swedish, Norwegian

I have said that head-to-head movement would tip the balance towards choosing $[\pm V \text{ doubling}]$. I turn now to showing why. Under the assumption, argued for in the last subsection, that head-to-head movement takes place at PF and leaves

(114) liknot, hi kanta et ha-praxim
buy.INF she buy.PAST the flower

- The fronted infinitive is doubled by a finite form inside the sentence. The view I adopt of head-to-head movement allows us to take (114) to be a regular instance of V doubling, i.e. an instance of (V, V) being linearized at both the head and the foot position. In this case, the lower copy of (V, V) forms a proper part of a complex word, namely [V+T]. Now observe the following contrast in German.

- The derivation of (116a) is presumably identical to that of (114), modulo the subsequent T to C movement. Its result is the structure in (117).

- With [$\pm V$ doubling], it is straightforward to account for the absence of (117) in German: we just have to set German to [$-V$ doubling].²⁷ This setting entails that the lower V copy cannot be pronounced, but if it is not pronounced, T+C would be stranded, violating the Stranded Affix Filter. Thus, (117) would be “ineffable” in German. Now suppose we assume [$\pm V$ topicalization] instead of [$\pm V$ doubling]. Then we face the problem of setting [$\pm V$ topicalization] in such a way that (116a) is ruled out but (116b) is allowed. It is not obvious to me

71

how this problem can be solved. Thus, the $[\pm V \text{ doubling}]$ parameter is able to capture distinctions which cannot be accounted for by $[\pm V \text{ topicalization}]$.

2.6 Conclusion

In this chapter, I have examined predicate clefts in six languages. A range of distributional facts internal to each language, as well as the patterns of variation observed across the languages with respect to whether there is doubling of the clefted predicate, were shown to follow from the Edge Condition on Copy Deletion (ECCD), which says that phonological deletion of the lower copy of a movement chain requires that that copy be at the right edge of a maximal projection, i.e. end an XP. Head-to-head movement poses a problem to this principle, since no doubling is ever attested for this operation, and the headedness of the relevant phrases seems to play no role at all. This problem finds a solution in the assumption that head-to-head movement is a displacement operation which takes place on the PF branch, creating no chains. This assumption is in line with several recent works, even though it is challenged by some. We noted that the PF view of head-to-head movement also provides a way to choose between two possible parameters, each of which has the ability to be a primitive, given the ECCD and the basic word order of the languages in question.

Chapter 3

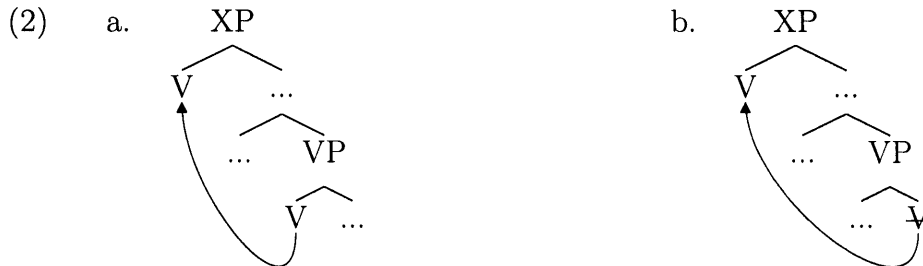
NP-Split constructions

3.1 Introduction

The previous chapter argues for the hypothesis that the linearization of syntactic chains is subject to a geometric condition on phrase structure: deletion of the lower copy of a chain necessitates that that copy end an XP, i.e. that its rightmost morpheme coincide with the rightmost morpheme of a non-projecting category.

- (1) Edge Condition on Copy Deletion (ECCD)
For any chain (α, β) where α is the higher and β is the lower copy, deletion of β requires that β ends an XP

The argument essentially consists in showing that various distributional facts about predicate cleft constructions in a diverse set of languages conform to the following interpretation: topicalizing V from a head-initial VP requires doubling of the fronted V, while topicalization of V from a head-final VP does not.



While the ECCD predicts, correctly as I have argued, that the lower V copy in (2a) must be overt, it does not predict that the lower V copy in (2b) must be covert. In other word, it only allows, but does not force, deletion of an XP-final

lower copy. What we have observed, however, is that doubling of V is impossible in (2b). I proposed to capture this with Pronunciation Economy, repeated below.

- (3) Pronunciation Economy (PE)
Copy Deletion must apply when it can

The motivation for complementing the ECCD with PE instead of strengthening it from a necessary condition to a necessary and sufficient one, say by replacing “requires” in (1) with “applies if and only if,” is that we do not want to rule out the existence of doubling cases in which the ECCD is satisfied but copy deletion, for some reason, still fails to apply. For example, we might imagine an XP-final copy which must be overt in order to support an affix. This would be an instance of doubling compelled not by the CCD, but by the Stray Affix Filter (Lasnik 1981). The CCD, thus, is intended to be one, but not the, condition on copy deletion.

Now note that the formulation of the ECCD is neutral with respect to the category of the displaced elements: α and β are to range over constituents of any type. However, the facts to which I appealed in my argument for the ECCD were limited to those concerning movement of V. The question then arises as to whether the ECCD should be reformulated in such a way that it applies only to chains created by movement of V, or even more radically, whether the distribution of copy deletion discussed so far should have an altogether different account which makes crucial reference to some property specific to verbs. The present chapter aims at convincing the reader that the answer to both of these questions is negative. It offers additional arguments for the ECCD which draw on a source of evidence entirely different from the one appealed to in chapter 2: the patterns of pronunciation and semantic interpretation of NP-split constructions in Vietnamese.

NP-Split constructions are those in which a subpart of a nominal complex is extracted from it, resulting in a “split” of the complex. (4a) and (4b) are examples from German and English, respectively (cf. Fanselow and Ćavar 2002: 65, 67). Parts of the split constituent are underlined.

- (4) a. A book appeared about Chomsky
b. Bücher hat man damals interessante in den Osten keine mitnehmen
books has one then interesting in the East no with-take
dürfen
may
'As for books, one could not take any interesting ones to the East
then'

In (4a), the DP **a book about Chomsky** is split into **a book** in the subject position and **about Chomsky** in the VP-internal position. In (4b), the DP **keine interessanten Bücher** ‘no interesting books’ is split into three discontinuous parts: **Bücher** ‘books’ at the clause-initial position, **keine** ‘no’ at the direct object position, and **interessante** ‘interesting’ at the position targeted by scrambling between **damals** ‘then’ and **in den Osten** ‘in the East.’¹

Before we go on, a few words on terminology are perhaps in order. I use the expression “nominal complex” to refer to the extended projection of N, i.e. the minimal phrase containing N and all of the functional categories associated with N such as determiners and numerals. The “NP” in “NP-Split,” then, is to be understood in the sense which dates back to the days when functional categories did not head their own projection (e.g. Chomsky (1965)). The name is thus a bit anachronistic. However, it seems to have gained enough currency as the cover term for splits of nominal complexes in general to warrant being used as it is used here.

There are several varieties of NP-split across languages other than those illustrated in (4). In this chapter, we will mostly be concerned with cases in which a single noun N is fronted from a nominal complex to the left-peripheral topic position of the sentence. Example (5) illustrates such a case, where the split nominal complex contains a numeral, a classifier (glossed as “CL”) and a noun.

- (5) sach thi no se mua hai quyen
 book TOP he will buy two CL

The fronted constituent to the left of **thi** is understood as contrastive topic: (5) means something like ‘as for books, he will buy two’ and is most naturally followed by something like ‘but as for magazines, he will only buy one.’ Thus, **thi** is often called the “topic marker” (cf. Cao 1992, for example). I will assume that it is a C head, and that the topicalized constituent is in [Spec,C] (cf. Trinh 2007). This assumption will be corroborated below.

The canonical word order of the split nominal complex in (5) is **hai quyen sach** ‘two CL book,’ which I assume has the hierarchical structure in (6a) and which is exemplified in (6b). The value of α and β in (6a), i.e. the category label of the various levels of projection above the NP, will be discussed later.

¹The adjective ‘interesting’ bears strong inflection when the DP is split (i.e. **interessante**) and weak inflection when the DP is not split (i.e. **interessanten**). For an account of this see Fanselow and Ćavar (2002: 93–94).

- (6) a. α
- hai β
- ‘two’ quyen sach
- ‘CL’ ‘book’
- b. no se mua hai quyen sach
- he will buy two CL book

To lend a bit more empirical weight to our discussion, the next example illustrates fronting of a noun other than **sach** ‘book,’ namely **meo** ‘cat.’ Note that the classifier for **meo** ‘cat’ is different from the one for **sach** ‘book.’ To avoid potential confusion, I will henceforth subscript “CL” with the (gloss of the) noun it is associated with in the sentence.²

- (7) a. no se co hai con meo
- he will have two CL_{cat} cat
- b. meo thi no se co hai con
- cat TOP he will have two CL_{cat}

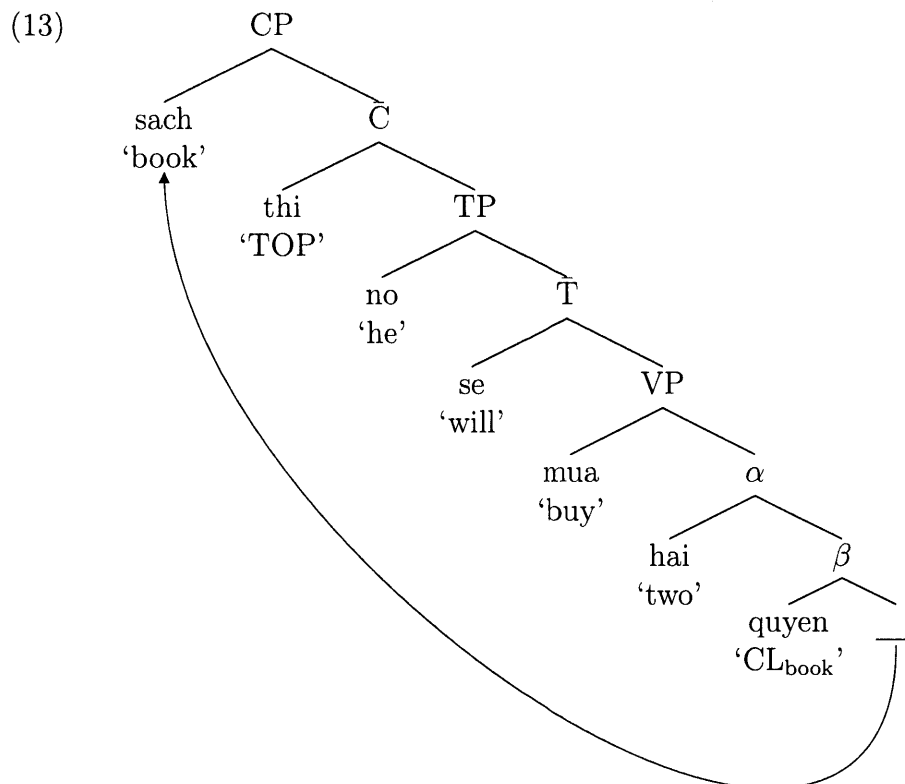
There is reason to think that such NP-Split constructions as (5) and (7b) involve $\bar{A}$ -movement of N to [Spec,C]. Specifically, the relation between the fronted noun and the stranded classifier associated with it appears to span finite clause boundaries, as evidenced by (8). At the same time, it is island-sensitive, as evidenced by (9) – (12), where the relevant islands are enclosed in square brackets.

- (8) Clause unboundedness
- a. toi nghi [CP rang no se mua hai quyen/con sach/meo]
- I think that he will buy two CL_{book}/CL_{cat} book/cat
- b. sach/meo thi toi nghi [CP la no se mua hai quyen/con]
- book/cat TOP I think that he will read one CL_{book}/CL_{cat}
- (9) Complex Noun Phrase Constraint
- a. no se gap [island mot nguoi co hai quyen/con sach/meo]
- he will meet one person have two CL_{book}/CL_{cat} book/cat
- ‘He will meet a person who has two books/cats’

²Note that several different nouns may require the same classifier. For example, the noun for ‘cat’ (**meo**) and the noun for ‘dog’ (**cho**) both take **con** as their classifier. Hence, the fact that a word is glossed as “CL_X” does not mean, for example, that it cannot be glossed as “CL_Y” where $X \neq Y$.

- b. *sach/meo thi no se gap [island mot nguoi co hai
 book/cat TOP he will meet one person have two
 quyen/con]
 CL_{book}/CL_{cat}
- (10) Adjunct Island Constraint
- a. no di ve [island sau khi mua hai quyen/con sach/meo]
 he go home after buy two CL_{book}/CL_{cat} book/cat
 ‘he went home after he bought two books/cats’
- b. *sach/meo thi no di ve [island sau khi mua hai quyen/con]
 book/cat TOP he go home after buy two CL_{book}/CL_{cat}
- (11) Subject Island Constraint
- a. [island no mua hai quyen/con sach/meo] la tot
 he read two CL_{book}/CL_{cat} book/cat be good
 ‘That he read two books/cats is good’
- b. *sach/meo thi [island no mua hai quyen/con] la tot
 book TOP he buy two CL_{book}/CL_{cat} be good
- (12) Non-bridge Verb Island Constraint
- a. no thi thao [island rang no se mua hai quyen/con sach/meo]
 he whisper that he will buy two CL_{book}/CL_{cat} book/cat
 ‘He whispered that he would buy two books/cats’
- b. *sach/meo thi no thi thao [island rang no se mua hai
 book/cat TOP he whisper that he will buy two
 quyen/con]
 CL_{book}/CL_{cat}

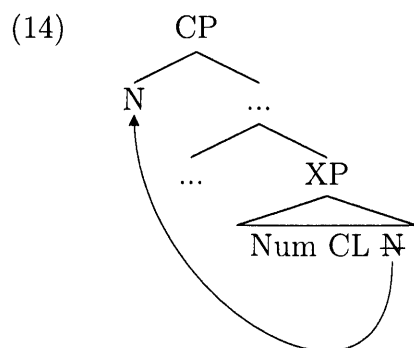
These are properties typical of $\bar{A}$ -dependencies (Ross 1967, Chomsky 1977). I will thus assume, on the basis of (8) – (12), that topicalization of N in Vietnamese is effected by $\bar{A}$ -movement of N to [Spec,C]. The sentence in (5), for example, will have the analysis in (13).



3.2 Relational vs. non-relational nouns

3.2.1 Cases of non-optionality

All cases of NP-Split we have considered involve $\bar{A}$ -movement of a noun N to $[\text{Spec}, C]$ from the right edge of a nominal complex, which I will assume is an XP , i.e. a non-projecting category. Given the ECCD (cf. (1)) and PE (cf. (3)), we predict that the lower copy of the chain (N, N) will be deleted, as depicted in (14). This means, observationally, that there is no doubling of the topicalized noun. That this is in fact the case is shown in (15).



- (15) a. sach thi no se mua hai quyen (*sach)
 book TOP he will buy two CL_{book} (*book)
 b. meo thi no se mua hai con (*meo)
 cat TOP he will buy two CL_{cat} (*cat)

Up to this point, the semantics of the topicalized noun has not featured in our discussion. It turns out that there is a complex, but storable and quite robust, correlation between a specific aspect of a noun's meaning, namely its argument structure, and the pronunciation patterns observable in split constructions where the noun is fronted. In this subsection and the next, I will show how the ECCD interacts with a set of assumptions about the syntax and semantics of various elements in the nominal complex to yield exactly the correlation in question.

My starting point will be to draw attention to the fact that the fronted nouns in (15), **sach** 'book' and **meo** 'cat,' are both non-relational nouns, i.e. ones which denote predicates of type $\langle e, t \rangle$ and do not require a complement.³

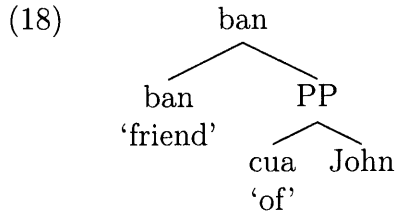
- (16) Definition
 a. $\llbracket \text{sach} \rrbracket = \lambda x(x \text{ is a book})$
 b. $\llbracket \text{meo} \rrbracket = \lambda x(x \text{ is a cat})$

Now suppose we have a relational noun of type $\langle e, \langle e, t \rangle \rangle$, for example **ban** 'friend,' which takes a complement, say **cua John** 'of John.' I will assume that the presupposition **cua** 'of' in **cua John** 'of John' has no semantic content and is there to fulfill a purely formal, presumably case-theoretic, requirement.

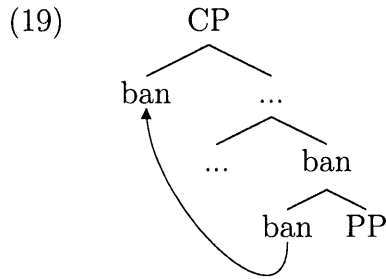
- (17) Definition
 a. $\llbracket \text{ban} \rrbracket = \lambda y(\lambda x(x \text{ is friends with } y))$
 b. $\llbracket \text{cua John} \rrbracket = \llbracket \text{John} \rrbracket = \text{John}$
 $\rightarrow \llbracket \text{ban cua John} \rrbracket = \lambda x(x \text{ is friends with John})$

I will make the standard assumption that when a head merges with its complement, it is the former that projects. As NP is head-initial in Vietnamese, merging of **ban** 'friend' and **cua John** 'of John' results in the former not being at the right edge of a non-projecting category, as the smallest category containing it which could be non-projecting, i.e. the category whose right edge it must occupy in order to occupy the right edge of any non-projecting category, has syntactic material following it, namely the PP complement.

³The language in the definition of **sach** 'book' and **meo** 'cat' in (16) suggests that these nouns denote sets which do not contain plural entities, i.e. which contain only singular books/cats. We will see later that this is not accurate, and will revise the meanings of these nouns accordingly. The same holds for the definition of **ban** 'friend' in (17).



We predict, given the CCD, that topicalization of **ban** ‘friend’ from **ban cua John** ‘friend of John’ will show doubling of **ban**, since the lower copy of (**ban**, **ban**) does not end an XP. This prediction is correct, as the contrast in (20) shows.



- (20)
- a. ban thi no se gap hai nguoi ban cua John
friend TOP he will meet two CL_{friend} friend of John
 - b. *ban thi no se gap hai nguoi cua John
friend TOP he will meet two CL_{friend} of John

The same holds for other relational nouns. Take **vo** ‘wife,’ for example, which presumably denotes the relation ‘ $\lambda y(\lambda x(x$ is female and married to $y))$.’ The contrast in (21) shows that topicalization of **vo** ‘wife,’ when this noun has a complement, also necessitates doubling, just as we observe with **ban** ‘friend’ in (20).⁴

- (21)
- a. vo thi no se gap hai nguoi vo cua John
wife TOP he will meet two CL_{wife} wife of John
 - b. *vo thi no se gap hai nguoi cua John
wife TOP he will meet two CL_{wife} of John

Of course, (21) and (20) can only be said to follow from the ECCD if the topic noun and its double are copies of a syntactic chain, i.e. that they are related

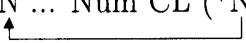
⁴Readers raised in the ideal of monogamy might feel that (21a), which means something like ‘as for wives, I met two wives of John,’ sounds “funny” for the irrelevant reason that no one, not even John, is allowed to have more than one wife. Thus, the clarity of the contrast in (21) can only obtain through abstraction from contextual information of historical accident.

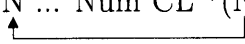
by movement. Evidence that this is the case is the fact that the same locality constraints are operative here which apply in standard, run-of-the-mill cases of $\bar{A}$ -movement. The ungrammaticality of (22) shows, for example, that fronting of **ban** ‘friend,’ which is accompanied by doubling of the same, is subject to the Adjunct Island Constraint.

- (22) **ban thi no bi om sau khi no gap mot nguoi ban cua John*
 friend TOP he got sick after he met one CL_{friend} friend of John

Other examples can be given of other relational nouns and other island constraints, but for reason of space, I will not present them here. As a matter of fact, I will proceed under the assumption that all cases of noun topicalization, whether doubling is observed or not, involve movement, and ask the reader to take me at my word that the familiar sort of evidence can be provided to show that this is in fact the case.

In this subsection, we have derived the following generalizations about NP-Split constructions in Vietnamese: (i) if a non-relational noun which ends the nominal complex is topicalized, doubling is impossible; (ii) if a relational noun which has a complement is topicalized, doubling is necessary. Schematically,

- (23) NP-Split Generalization 1
 If N is non-relational: N ... Num CL (*N)
- 

- (24) NP-Split Generalization 2
 If N is relational: N ... Num CL *(N) Compl
- 

It is evident that the ECCD is instrumental in explaining both generalizations. Thus, these can be considered evidence in favor of this principle.

3.2.2 Cases of optionality

Non-relational nouns

Modifiers follow the head noun in Vietnamese, which means that even if N is non-relational and does not have a complement, it is still possible for N not to be the last constituent of the nominal complex, as the later may contain a modifier

to the right of N. (25a) and (25b) exemplify cases in which N is followed by a PP and an AP modifier, respectively.⁵

- (25) a. no se mua hai quyen sach ve vat ly
 he will buy two CL_{book} book about physics
 ‘He will buy two books about physics’
 b. no se mua hai con meo mau den
 he will buy two CL_{cat} cat black
 ‘He will buy two black cats’

Curiously, topicalization of a non-relational noun which is followed by a modifier results in optional doubling of the fronted noun: the lower N copy can, but does not have to, delete.

- (26) a. sach thi no se doc hai quyen (sach) ve vat ly
 book TOP he will read two CL_{book} (book) about physics
 b. meo thi no se mua hai con (meo) mau den
 cat TOP he will buy two CL_{cat} black

This fact is represented schematically in (27).

- (27) NP-Split Generalization 3
 If N is non-relational: N ... Num CL (N) Mod
 ↑ |

I will show that NP-Split Generalization 3 follows straightforwardly from the ECCD in conjunction with some assumptions about the semantic interpretation of nouns, classifiers, and nominal modifiers. Let us introduce these in a brief excursus.

⁵The objection might be raised that the PP **ve vat ly** ‘about physics’ which follows the noun **sach** ‘book’ in (25a) is a complement, not a modifier. An argument that the PP is actually a modifier is the fact that it can be placed after a copula verb and serve as a predicate of the noun, as (ia) shows. This is not possible for PPs which are clearly complements of nouns, as evidenced by the degradedness of (ib).

- (i) a. quyen sach nay la ve vat ly
 CL_{book} book this is about physics
 ‘This book is about physics’
 b. *nguai ban nay la cua toi
 CL_{friend} friend this is of me
 (‘This person is my friend’)

Note that (ib) can be well-formed if we take it to mean something like ‘this person, who is a friend of someone, belongs to me.’ In this reading, the PP **cua toi** ‘of me’ is clearly not construed as the complement of **ban** ‘friend.’

Excursus. One question which emerges from data considered so far and which I have not addressed is why there is, or more precisely why there must be, a classifier intervening between a numeral and a noun in Vietnamese.

- (28) no se mua hai *(quyen) sach
 he will buy two *(CL_{book}) book

My answer to this question is essentially the one given in Chierchia (1998b). Chierchia proposes that the difference between a “classifier language” such as Vietnamese, in which the combination of a noun and a numeral requires the mediation of a classifier, and a “number marking” language such as English, where there is no such requirement, reduces to a difference with respect to the denotation of nouns.⁶ To illustrate, let us compare the noun **book** in English and its counterpart in Vietnamese, **sach** ‘book.’ Suppose there are only three books in the world, call them a, b and c. Chierchia’s theory implies the following.

- (29) Definition (in the spirit of Chierchia (1998b))
 a. $\llbracket \text{book} \rrbracket = \{a, b, c\}$
 b. $\llbracket \text{sach} \rrbracket = \{a, b, c, a+b, a+c, b+c, a+b+c\}$

The English noun **book** denotes the set of individuals which are singular books, while the Vietnamese noun **sach** denotes a superset of this set which also includes individuals that are pluralities of books ($a+b$ is a plural individual which consists of a and b, for example).⁷ In more technical terms, **book** is an “atomic” and **sach** is a “cumulative” predicate. Thus, our definition of **sach** ‘book’ in (16) is not quite correct: it should be changed to (30).

- (30) $\llbracket \text{sach} \rrbracket = \lambda x(x \text{ is a book or a plurality of books})$

The definition of other nouns of Vietnamese should be revised similarly. Now suppose that only atomic predicates can be counted (cf. Chierchia 1998b, Kratzer

⁶Languages without classifiers are called “number marking languages” because most of them have singular/plural markings on nouns, which most of the languages with classifier lack. For an account of this fact, see Chierchia (1998a,b, 2010).

⁷As an example of reference to plural individuals, consider the sentence **those are books**, uttered with the speaker pointing to a pile of books on the table. The demonstrative pronoun **those** refers in this case to the plural individual which consists of all the books on the table.

1989, among others).⁸ We then predict, correctly, that **book** can combine with a numeral while **sach** cannot.⁹

- (31) a. he will buy one book
 b. *no se mua mot sach
 he will buy one book

However, **sach** can combine with a numeral if there is a classifier between them. What Chierchia proposes, basically, is that the classifier “atomizes” the predicate it combines with: it maps a predicate into a set of singular individuals which are in the extension of that predicate.

- (32) Definition

$$\llbracket \text{CL} \rrbracket = \lambda P(\lambda x(x \text{ is a } P\text{-atom}))$$

Basically, a *P*-atom is an individual which is *P* and which has no part which is also a *P*-individual. This means, for our hypothetical case, that $\llbracket \text{CL} \rrbracket(\llbracket \text{sach} \rrbracket) = \llbracket \text{book} \rrbracket = \{a, b, c\}$, which in turn means that **quyen sach** ‘CL_{book} book,’ just like **book**, can be counted, i.e. can combine with a numeral. This is the right result.

As for nominal modifiers, I will follow Heim and Kratzer (1998) and assume that they are predicates of type $\langle e, t \rangle$ which are semantically integrated into the structure by way of Predicate Modification (PM). This rule is defined as follows.

- (33) Predicate Modification (PM)
 If α and β are both of type $\langle e, t \rangle$, $\llbracket \alpha \beta \rrbracket = \lambda x(\llbracket \alpha \rrbracket(x) = \llbracket \beta \rrbracket(x) = 1)$

To keep the discussion of NP-Split constructions in Vietnamese maximally cohesive and transparent, I have discussed the semantics of nouns and classifiers in a very brief and informal manner, leaving many questions open, and have refrained from showing how the semantics proposed for these categories can be independently motivated. In the appendix to this chapter, I will present a theory of nominal reference where the claims about nouns and classifiers made above are more properly worked out and embedded in a formal framework in which several other facts of Vietnamese are accounted for. **End of excursus.**

⁸The rationale for this is easy to see: if we are counting books, the things we are counting cannot include, besides singular books, pairs, triples, quadruples etc., of books, since we would end up not knowing how many books there are.

⁹Let us assume that the plural morpheme **s** in **he will buy two books** is there for purely formal reason and has no semantic import (cf. Krifka 1995, 2008, Sauerland 2003).

I will now show how the assumptions introduced in the above excursus can interact with the ECCD to derive NP-Split Generalization 3 (cf. (26) and (27)). First, we can conclude from the definition of PM and CL that both (34a) and (34b) are well-formed with respect to semantic type, assuming N is of type $\langle e, t \rangle$ (“Mod” stands for ‘modifier’).



In (34a), β can be interpreted by PM, since both N and Mod are expressions of type $\langle e, t \rangle$, and α can be interpreted by Functional Application (FA), since CL is of type $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$ and β of type $\langle e, t \rangle$.¹⁰ In (34b), β is interpreted by FA and α , crucially, can be interpreted by PM, since merging CL and N results in an expression of type $\langle e, t \rangle$. I will assume that both structures in (34) are available in Vietnamese. The question now is what values should we give to α and β in (34). Suppose we make the following choices.

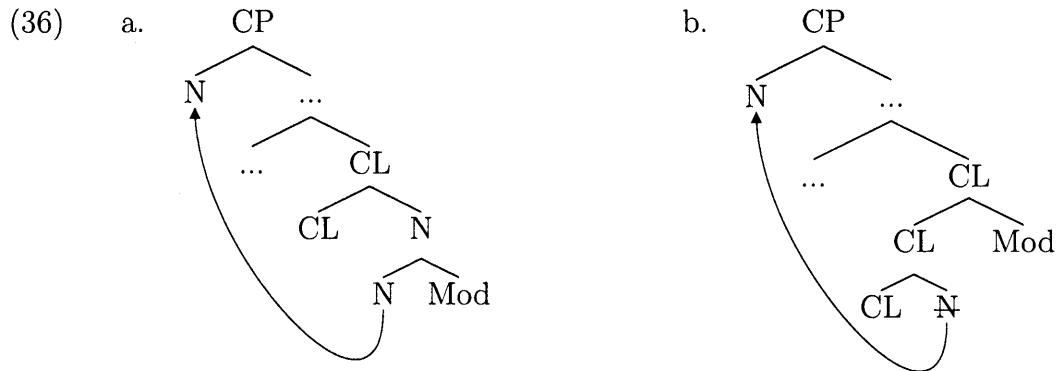


These choices are in fact based on two principles: (i) when a head H merges with a constituent K and $\llbracket H \rrbracket$ is a function which takes $\llbracket K \rrbracket$ as its argument – i.e. $\llbracket H K \rrbracket = \llbracket H \rrbracket(\llbracket K \rrbracket)$ – then H projects; (ii) what we call “modifier” does not project.¹¹ These are quite standard assumptions. They are, however, not to be considered an argument for choosing the labels in (35). Such an argument would require observations whose explanation depends on this labeling. It turns out

¹⁰The rule of Functional Application is defined as follows: If α is of type $\langle a, b \rangle$, β is of type a , and γ is the result of merging α and β , then $\llbracket \gamma \rrbracket = \llbracket \alpha \rrbracket(\llbracket \beta \rrbracket)$ (cf. Heim and Kratzer 1998).

¹¹The question can be raised as to whether the choice of defining CL as a function from predicates to predicates is arbitrary. What if we just define CL as a predicate which is true of “atomic” individuals in D_e – i.e. “absolute atoms” – and interpret the combination [CL N] by Predicate Modification? This would not be a possibility if we assume that the notion of “atoms” is property-dependent, i.e. that an atom is always an atom with respect to some property (cf. Krifka 1998, for example). The facts about doubling whose explanation depends on CL being a function from predicates to predicates, then, can be considered evidence in favor of this assumption.

that NP-Split Generalization 3 is an observation that does just this. Recall the type of constructions which this generalization concerns: a non-relational noun is topicalized from a position which linearly follows a classifier and precedes a modifier, i.e. the position indicated with “__” in [CL __ Mod]. Given what we have just said, topicalization of N from such a position can have two scenarios: (36a), in which N is extracted from (34a), and (36b), where N is extracted from (34b).



The ECCD implies that the two structures in (36) differ minimally with respect to pronunciation: in (36a), the lower copy of the chain created by N topicalization is overt, whereas in (36b) it is covert. The lower N copy in (36a) projects, and it is separated from the right edge of the NP, which is the smallest XP containing it, by the modifier. The ECCD entails that this copy cannot delete, as it does not end an XP. The lower N copy in (36b), on the other hand, does not project. It is, therefore, an XP itself. This means, trivially, that its last morpheme coincides with the last morpheme of an XP, i.e. that it ends an XP. From the CCD, we conclude that this copy can delete, and from PE, that it must delete. The availability of both (36a) and (36b) in the language, then, translates into the optionality of doubling that we observed. NP-Split Generalization 3 is derived.

In fact, we derive more than just the above mentioned optionality of doubling. Specifically, we predict that (36a) and (36b) are semantically equivalent, as they differ only in that the former contains (35a) and the latter contains (35b) and these are semantically equivalent: both (35a) and (35b) denote the set of singular individuals which are elements of $\llbracket N \rrbracket$ and of $\llbracket Mod \rrbracket$. Selecting the singularities from the intersection of $\llbracket N \rrbracket$ and $\llbracket Mod \rrbracket$, as is done in (36a), comes to the same thing as selecting the singularities from $\llbracket N \rrbracket$ and then intersecting the result with $\llbracket Mod \rrbracket$, as is done in (36b). In other words, $\llbracket CL \rrbracket(\llbracket N \rrbracket \cap \llbracket Mod \rrbracket) = \llbracket CL \rrbracket(\llbracket N \rrbracket) \cap \llbracket Mod \rrbracket$. This prediction is correct, as evidenced by the fact that (37a), which instantiates (36a), and (37b), which instantiates (36b), mean exactly the same thing, namely ‘as for books, he will buy two about physics.’

- (37) a. sach thi no se mua hai quyen sach ve vat ly
 book TOP he will buy two CL_{book} book about physics
 b. sach thi no se mua hai quyen ve vat ly
 book TOP he will buy two CL_{book} book about physics

Relational nouns

The examples of relational nouns we have considered so far all have these nouns in combination with a complement (e.g. **ban cua John** ‘friend of John’). It is a fact, however, that relational nouns such as **ban** ‘friend’ or **vo** ‘wife’ can appear in sentences without a complement.

- (38) a. no se gap hai nguoi vo
 he will meet two CL_{wife} wife
 ‘He is looking for a wife’
 b. no se gap hai nguoi ban
 he will meet two CL_{friend} friend
 ‘He will meet two friends’

Interestingly, when a relational noun without a complement is topicalized, doubling of the fronted noun is optional, just as when a non-relational noun with a modifier is topicalized.¹²

- (39) a. vo thi no se gap hai nguoi (vo)
 wife TOP he will meet two CL_{wife} (wife)
 b. ban thi no se gap hai nguoi (ban)
 friend TOP he will meet two CL_{friend} (friend)

Before we go on, let us formulate, for ease of reference, the generalization we want to derive.

¹²Some speakers report a preference for the non-doubling alternative, i.e. they give the following judgement.

- (i) a. vo thi no se gap hai nguoi (?vo)
 wife TOP he will meet two CL_{wife} (?wife)
 b. ban thi no se gap hai nguoi (?ban)
 friend TOP he will meet two CL_{friend} (?friend)

I have no explanation of this variation and will assume that whatever turns out to explain the question mark in (i), it is orthogonal to the question at hand, namely whether the ECCD is correct or not.

- (40) NP-Split Generalization 4
 If N is relational: N ... Num CL (N)

Now recall the semantics we have proposed for CL: it is a function which maps predicates into predicates. Thus, CL is an expression of type $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$, which means its sister must be of type $\langle e, t \rangle$. A relational noun, by hypothesis, denotes a relation: it is of type $\langle e, \langle e, t \rangle \rangle$. It then follows that CL cannot merge with a bare relational noun, since that would create a type mismatch. In other word, the structure in (41) is not interpretable.

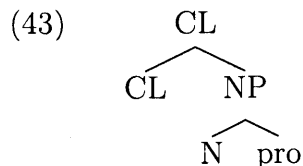
- (41)
-

Thus, we make the prediction that the sentences in (40) are semantically deviant, as they contain **nguai ban** ‘CL friend’ and **nguai vo** ‘CL wife,’ i.e. phrases of the form [CL N] where N is relational. The problem with this prediction is that it is wrong: the sentences in (41) are perceived as totally acceptable.

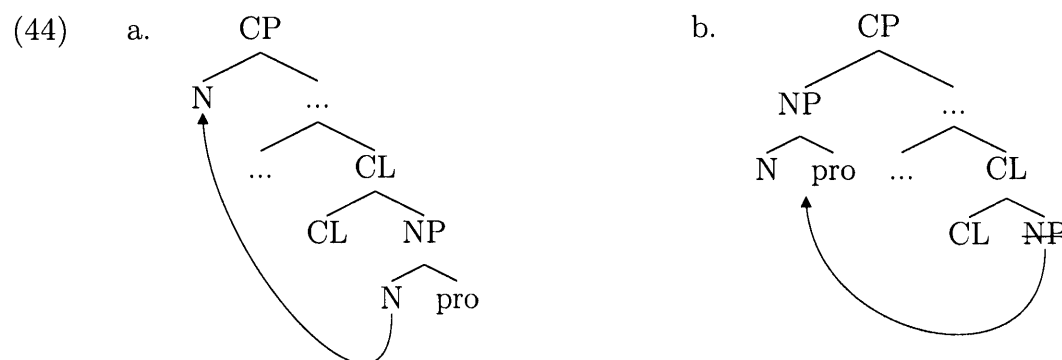
My solution to this problem is not to fix the semantics for CL or relational N, but to say that the sentences in (41) does not contain phrases of the form [CL N]. Specifically, I propose that the relational nouns **ban** ‘friend’ and **vo** ‘wife’ in (40a) and (40b), in fact all relational nouns which do not have an overt complement but which follow a classifier, have a null complement, **pro**, which is an expression of type e . This makes the sister of CL in these cases an expression of type $\langle e, t \rangle$, which is what the semantics of CL requires.

- (42)
-

I will assume that **pro** is interpreted by virtue of existential closure at the clause level. Thus, **vo pro** ‘wife pro,’ for example, will contribute a meaning of something like ‘wife of someone.’ Application of FA gives us $\llbracket \text{CL} \rrbracket (\llbracket \text{N} \rrbracket (\llbracket \text{pro} \rrbracket))$ as the meaning of α . Thus, CL and N are heads denoting functions which apply to the denotation of their sister constituent. Continuing with the assumption that such heads project, we have (43) as the labeling of the structure in (42).



As **pro** is a null pronoun, topicalization of N will fill [Spec,C] with the same phonological material as topicalization of NP = [N **pro**]. In other word, there is no difference with respect to the pronunciation of the derived position between fronting N and fronting NP. However, the ECCD predicts that there is a difference with respect to the pronunciation of the base position. Specifically, the base position of the chain (N, N) will be overt, as the lower copy is separated from the right edge of the minimal XP containing it by **pro** and hence does not end an XP. The base position of the chain (NP, NP), on the other hand, will be covert, since the lower copy is itself an XP, hence trivially ends an XP. Schematically,



In terms of pronunciation, (44a) and (44b) differ only in that the former shows doubling of the constituent in the topic position and the latter does not. The fact that in one case the topic is an N and in the other it is an NP has no phonological consequence, since NP consists of N and a null element. Assuming that both (44a) and (44b) are available possibilities in Vietnamese, we correctly predict optionality of doubling. NP-Split Generalization 4 is derived.

We see that optionality of doubling can come about in two ways. The optionality described in NP-Split Generalization 3, i.e. that of doubling modified non-relational nouns, obtains by virtue of the fact that CL allows the chain (N, N) created by topicalization to be parsed as either one in which the lower N projects or one in which it does not. The optionality of NP-Split Generalization 4, i.e. that of doubling relational nouns without an overt complement, obtains by virtue of the fact that CL forces merging of N and a null **pro**, thereby allowing the chain created by topicalization to be parsed either as (N, N) or as ([N **pro**], [N **pro**]). These two scenarios of optional doubling both have their source in the semantics

we proposed for the classifier. However, there is a crucial difference between them. In the non-relational case, the topic constituent is always N, whether there is doubling or not, but in the relational case, the topic is N when there is doubling and [N **pro**] when there is not. This should have semantic consequences. Specifically, we should expect no difference in meaning between doubling a non-relational N and not doubling it, and we have seen that this expectation is fulfilled. However, we should not expect semantic equivalence between doubling and non-doubling of a relational N: doubling indicates that the topic is N, a relation, and non-doubling indicates that the topic is [N **pro**], a predicate.

There are reasons to think that the prediction of our theory is consistent with the facts. Consider the following paradigm, where capitalization represents focus intonation.

- (45) a. Vo thi no gap HAI nguoi. Ban thi no gap BA nguoi.
 wife TOP he met two CL_{wife} friend TOP he met three CL_{friend}
 b. Vo thi no gap HAI nguoi vo. Ban thi no gap BA
 wife TOP he met two CL_{wife} wife friend TOP he met three
 nguoi ban.
 CL_{friend} friend
 c. *Vo thi no gap HAI nguoi. Ban thi no gap BA nguoi
 wife TOP he met two CL_{wife} friend TOP he met three CL_{friend}
 ban.
 friend

I believe a case can be made that the contrast between (45a-b) and (45c) follows from our account of NP-Split Generalization 4, given the theory of contrastive topics proposed in Büring (1999, 2003) (cf. also Krifka 2007). Before making my case, then, let me discuss this theory in a brief excursus.

Excursus. Note, first, that contrastive topics require focus on some other constituent of the sentence, as exemplified by (45). Thus, all the examples involving contrastive topics we have discussed so far are to be read with focus accent put on some phrase contained in the sister of the topic marker **thi**. For the sake of discussion, I will use examples where focus is on the numeral, but my argument does not depend on this particular choice and can be made with any other constituent bearing the focus accent.¹³

¹³The requirement that there is a focus in the sentence seems to have been intended by Büring to be part of the definition of contrastive topics. Büring's theory basically says that the topic value of a sentence containing a contrastive topic is a set of questions. It follows from this theory that if such a sentence contains no focus, its topic value will be a set of questions, each of which has only one answer (see below), which are arguably not "legitimate" questions.

Büring's (1995, 2003) theory of topics starts from the theory of focus proposed in Rooth (1985, 1992, 1996). According to the latter, focus on a phrase X in a sentence S induces and indicates the relevance of a set of "alternatives" to S, which are propositions obtainable by replacing X in S with an expression of the same semantic type. Rooth calls this set of propositions the "focus semantic value" of S, $\llbracket S \rrbracket^f$, in contrast to its "ordinary semantic value," $\llbracket S \rrbracket$, which is just a proposition and which is computed in abstraction from the presence of focus in S.¹⁴ For example,

- (46) a. $\llbracket \text{John}_F \text{ kiss Mary} \rrbracket = \llbracket \text{John kiss Mary}_F \rrbracket = \text{John kissed Mary}$
 b. $\llbracket \text{John}_F \text{ kiss Mary} \rrbracket^f = \{\text{John kissed Mary, Bill kissed Mary, Fred kissed Mary, ...}\}$
 c. $\llbracket \text{John kiss Mary}_F \rrbracket^f = \{\text{John kissed Mary, John kissed Sue, John kissed Anne, ...}\}$

Postulation of the focus semantic value in addition to the ordinary one makes it possible to capture several facts about language. As an example, consider the following constraint (Rooth 1992, Büring 1999).¹⁵

- (47) Question-answer constraint (focus)
 In a question-answer pair $\langle Q, A \rangle$, $\llbracket Q \rrbracket = \llbracket A \rrbracket^f$

This constraint says that the question should denote the focus semantic value of its answer. It accounts for the following paradigm.

- (48) Q: Who did John kiss?
 A: John kissed Mary_F
 A': #John_F kissed Mary

Intuitively, the question in (48) requires an answer in which focus is on the object, not the subject, of **kiss**. Assuming the proposition-set theory of questions (cf. Hamblin 1973) according to which a question denotes the set of its potential answers, the denotation of **who did John kiss** would be a set containing propositions such as 'John kissed Mary,' 'John kissed Sue,' etc. In other word, it would denote a set of propositions of the form 'John kissed x' where x is a person, which is exactly the focus semantic value of answer A in (48), hence the appropriateness

¹⁴Rooth actually write " $\llbracket \alpha \rrbracket^o$ " for the ordinary semantic value of α . I have been using " $\llbracket . \rrbracket$ " as the function mapping expressions to their ordinary semantic value, hence will stick to it for the sake of consistency. Note, also, that I have simplified Rooth's theory in order to free the discussion from unnecessary complexities.

¹⁵Again, this is not exactly what Rooth proposes. It is, however, equivalent to what Büring (1999) proposes.

of A. The focus value of answer A', however, is the set of propositions of the form 'x kissed Mary' where x ranges over persons. This set is not the denotation of the question, and this accounts for the inappropriateness of A'.

Now suppose we have a sentence where in addition to a focused constituent, there is another constituent which is interpreted as a contrastive topic (I will subscript topics with "T").¹⁶

- (49) John_T kissed Mary_F
'As for John, he kissed Mary'

To capture the contribution of topics to the interpretation of the sentence, Büring proposes another semantic value, the topic value, in addition to the ordinary and the focus value of Rooth. The topic value of (49), $\llbracket(49)\rrbracket^t$, is a set, not of propositions, but of sets of propositions.

- (50) $\llbracket(49)\rrbracket^t = \{\{\text{John kissed Mary, John kissed Sue, John kissed Anne,...}\}, \{\text{Bill kissed Mary, Bill kissed Sue, Bill kissed Anne,...}\}, \{\text{Fred kissed Mary, Fred kissed Sue, Fred kissed Anne,...}\},...\} = \{\text{who did John kiss, who did Bill kiss, who did Fred kiss,...}\}$

Topics, then, induce alternatives, just like focus. In fact, the alternatives induced by topics are just the alternatives to the focus semantic value of the sentence. In other word, the focus value of a sentence S is a question which S answers, and the topic value of S is a set of questions one of which S answers. This set of questions is constructed by substituting the topic constituent with expressions of the same semantic type. Postulation of the topic semantic value, Büring argues, enables us to capture facts such as the following.

- (51) Q: Who did Bill kiss?
A: Well, John_T kissed Mary_F
A' ≠ Well, John kissed Mary_F

The relevant constraint would be (52).

- (52) Question-answer constraint (topic)
In a question-answer pair $\langle Q, A \rangle$, $\llbracket Q \rrbracket \in \llbracket A \rrbracket^t$

¹⁶Topical interpretation typically corresponds to some sort of rising intonation (e.g. the L-H* pitch contour of Pierrehumbert (1980), the "B-accent" of Jackendoff (1972)).

The constraint says that the denotation of a question must be an element of the topic value of its answer.¹⁷ The topic value of answer A in (51) is the set described in (50), i.e. the set of questions of the form ‘who did x kiss’ with x ranging over persons. This set contains the question ‘who did Bill kiss.’ Hence, A meets the condition in (52) and is thus an appropriate answer. The topic value of A’, on the other hand, is the singleton set containing the question ‘who did John kiss,’ which means A’ does not meet (52) and is thus not an appropriate answer. The point to note here is that without the concept of topic value, the constraint in (52) cannot be formulated and consequently the contrast in (51) cannot be captured.

Büring proposes a number of other rules which make reference to the topic value of sentences. For reason of space, I will not present these. What I will present, however, is a rule which, to the best of my knowledge, has not been discussed but which will be relevant for my account of the Vietnamese facts later. It is (53).

- (53) Generalization on Consecutive Sentences (GCS)
 A sequence of consecutive sentences containing contrastive topics sounds best if these sentences have the same topic value, i.e. induce the same set of questions

The GSC makes sense of the following contrast.¹⁸

- (54) a. Books_T he will_F buy. Magazines_T he won’t_F.
 b. [Buy books]_T he will_F. [Buy magazines]_T he won’t_F.
 c. #Books_T he will_F buy. [Buy magazines]_T he won’t_F.

The two sentences in (54a) both induces the set of questions of the form ‘will he buy X,’ where X ranges over noun meanings such as ‘books,’ ‘magazines,’ ‘postcards’ etc. They both have as their topic value the set in (55).

- (55) {{he will buy books, he won’t buy books}, {he will buy magazines, he won’t buy magazines}, {he will buy postcards, he won’t buy postcards},...} = {will he buy books, will he by magazines, will he buy postcards,...}

¹⁷If the answer contains no topic-marked constituent, its topic value will be the singleton set containing the question that it answers.

¹⁸I assume that the alternative of ‘will’ is ‘won’t’ and vice versa, and the denotation of a yes/no question is the set containing its positive and its negative answer. Also, note that the absence of **buy** in the second sentence of (54a) is due to VP-elipsis.

The same holds for the two sentences in (54b): they both denote the set of questions of the form ‘will he P’ where P ranges over VP meanings such as ‘buying books,’ ‘buying magazines,’ ‘sleep’ etc, i.e. both denote the set in (56).

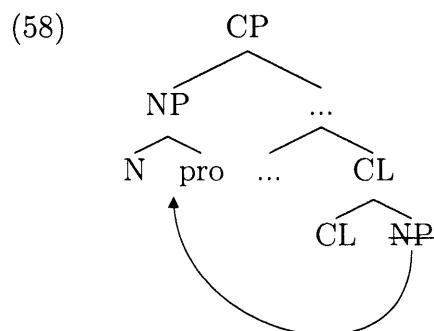
- (56) $\{\{\text{he will buy books, he won't buy books}\}, \{\text{he will buy magazines, he won't buy magazines}\}, \{\text{he will sleep, he won't sleep}\}, \dots\} = \{\text{will he buy books, will he buy magazines, will he sleep}, \dots\}$

Thus, the first two pairs of sentences in (54) obey the GSC. The last pair of sentences in (54), on the other hand, does not: the first sentence in this pair induces the set of questions in (55), while the second that in (56). The reader can verify this for himself. The fact that this pair is perceived as a strange sequence is evidence that the GSC is a valid generalization. **End of excursus.**

Let us now come back to the paradigm in (45), and ask ourselves what distinguishes the acceptable examples from the unacceptable one. Given our account of NP-Split Generalization 4, we can conclude the following: the former obey the GCS, while the latter does not. Put differently, in the acceptable cases, the sentences have identical topic values, while in the unacceptable case, they have different topic values. Consider the first pair of sentences in (45), repeated in (57).

- (57) Vo thi no gap HAI nguoi. Ban thi no gap BA nguoi.
 wife TOP he met two CL_{wife} friend TOP he met three CL_{friend}

Our account of NP-Split Generalization 4 implies that the sentences in (57) have topics of the form [N pro], since there is no doubling of the fronted relational noun. The structure of both sentences, then, must be (58).



Let us use English words to represent the topic/focus structure of these sentences.

- (59) $[s_1 \text{ He met two}_F [\text{wife of pro}]_T] . [s_2 \text{ He met three}_F [\text{friend of pro}]_T] .$

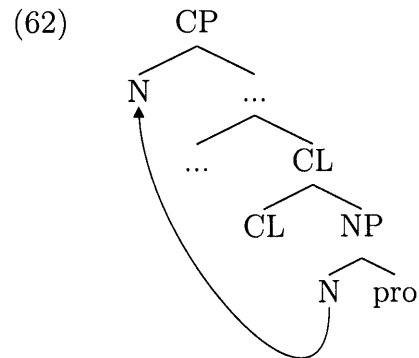
Assuming the alternatives to ‘two’ and ‘three’ are the same, as are the alternatives to ‘wife of pro’ and ‘friend of pro,’ the theory of topics presented in the excursus implies that the topic values of S_1 and S_2 in (59) are identical: both sentences have a topic value which is a set of questions of the form ‘how many P did he meet,’ where P ranges over predicates.¹⁹

- (60) $\llbracket S_1 \rrbracket^t = \llbracket S_2 \rrbracket^t = \{ \{ \text{he met one wife of pro, he met two wives of pro, he met three wives of pro, ...} \}, \{ \text{he met one friend of pro, he met two friends of pro, he met three friends of pro, ...} \}, \{ \text{he met one professor, he met two professors, he met three professors, ...} \} = \{ \text{how many wives of pro did he meet, how many friends of pro did he meet, how many professors did he meet, ...} \}$

Now consider the second pair in (45), repeated in (61). This pair differ from the first in that the fronted relational noun is doubled in both sentences.

- (61) Vo thi no gap HAI nguoi vo. Ban thi no gap BA nguoi
 wife TOP he met two CL_{wife} wife friend TOP he met three CL_{friend}
 ban.
 friend

Given the ECCD and PE, we know that the topics of these two sentences must be complementless N, i.e. N without pro. Their structure must then be that in (62).



The topic/focus structure of these sentences are as follows.

¹⁹Note that P does not have to be of the form ‘N of pro.’ Thus, we predict, correctly as it turns out, that the following sequence is also acceptable.

- (i) Vo thi no gap HAI nguoi. Giao su thi no gap BA nguoi.
 wife TOP he met two CL_{wife} professor TOP he met three CL_{friend}

(63) $[_{S_1}$ He met two_F wife_T of pro]. $[_{S_2}$ He met three_F friend_T of pro].

Again, we predict that the topic values of S_1 and S_2 in (63) are identical, since the alternatives to ‘two’ and ‘three’ are the same, as are the alternatives to ‘wife’ and ‘friend.’ Specifically, both sentences have the topic value which is the set of questions of the form ‘how many R of pro did he meet,’ where R ranges over relations.

(64) $[[S_1]]^t = [[S_2]]^t = \{\{\text{he met one wife of pro, he met two wives of pro, he met three wives of pro, ...}\}, \{\text{he met one friend of pro, he met two friends of pro, he met three friends of pro, ...}\}, \{\text{he met one daughter of pro, he met two daughters of pro, he met three daughters of pro, ...}\} = \{\text{how many wives of pro did he meet, how many friends of pro did he meet, how many daughters of pro did he meet, ...}\}$

Now consider the last pair in (45), repeated in (65).

(65) *Vo thi no gap HAI nguoi. Ban thi no gap BA nguoi ban.
 wife TOP he met two CL_{wife} friend TOP he met three CL_{friend} friend

The first sentence shows no doubling of the fronted noun. This indicates that the topic of this sentence is [N pro], as in (58). The second sentence does show doubling, so its topic must be bare N, as in (62). Here is the topic/focus structure of these sentences.

(66) $[_{S_1}$ He met two_F [wife of pro]_T]. $[_{S_2}$ He met three_F friend_T of pro].

The topic value of S_1 , then, is the set of questions of the form ‘how many P did he meet’ where P ranges over predicates. The topic value of S_2 , on the other hand, is the set of questions of the form ‘how many R of pro did he meet’ where R ranges over relations.

(67) a. $[[S_1]]^t = \{\text{how many wives of pro did he meet, how many friends of pro did he meet, how many professors did he meet, ...}\}$
 b. $[[S_2]]^t = \{\text{how many wives of pro did he meet, how many friends of pro did he meet, how many daughters of pro did he meet, ...}\}$

Thus, we see that in the acceptable cases, the two consecutive sentences have the same topic value: they induce the same set of questions and each of them answers one question from this set. In the unacceptable case, the sentences have different topic values: they induce different sets of questions. In other word, the acceptable sequences conform to the GSC, and the unacceptable one does not.

Thus, we assimilate the Vietnamese facts in (45) to the English facts in (54). This assimilation depends crucially on our account of NP-Split Generalization 4, which in turns depends on the ECCD and the semantics of classifiers.

Let us end this subsection with the discussion of a potential counterexample to what we have just said.

- (68) Vo thi no gap HAI nguoi vo. Ban thi no gap BA nguoi.
 wife TOP he met two CL_{wife} wife friend TOP he met three CL_{friend}
 ‘As for wives, he met two. As for friends, he met three.’

We predict that the sequence in (68) should be deviant: the first shows doubling, which indicates the topic is N; the second does not show doubling, which indicates the topic is [N pro]. It follows that these sentences do not have the same topic value, and should feel strange. The problem is that they are perfectly acceptable. My solution to this problem is to say that the lower copy of the fronted noun **ban** ‘friend’ is eliminated from the phonetic representation of the sentence by a process distinct from Copy Deletion and probably akin to ellipsis. I propose that the same process takes place in the second sentence of (69), where not only the lower copy of the fronted noun but also the classifier preceding it is elided.

- (69) Vo thi no gap HAI nguoi vo. Ban thi no gap BA.
 wife TOP he met two CL_{wife} wife friend TOP he met three
 ‘As for wives, he met two. As for friends, he met three.’

Evidence that (68) and (69) involves the same ellipsis rule is the fact that switching the order of the sentences in (68) and (69) leads to deviance in both cases.

- (70) a. *Ban thi no gap BA nguoi. Vo thi no gap HAI nguoi
 friend TOP he met three CL_{friend} wife TOP he met two CL_{wife}
 vo.
 wife
 b. *Ban thi no gap BA. Vo thi no gap HAI nguoi vo.
 friend TOP he met three wife TOP he met two CL_{wife} wife

Also, it should be noted that when we conjoin the two clauses in (70a) and (70b), making the sequence one big sentence, acceptability increases greatly.

- (71) a. Ban thi no gap BA nguoi, nhung vo thi no chi gap
 friend TOP he met three CL_{friend} but wife TOP he only met
 HAI nguoi vo.
 two CL_{wife} wife

- b. Ban thi no gap BA, nhung vo thi no chi gap HAI
 friend TOP he met three but wife TOP he only met two
 nguoi vo.
 CL_{wife} wife

As cataphoric ellipsis is possible only internal to one sentence, the fact that both sentences in (71) are grammatical is additional evidence that ellipsis is involved in apparent counter-examples to our prediction.

Whatever the ellipsis process in question turns out to be in details, then, it must be constrained in such a way that in a sequence of two sentences containing contrastive topics and having the same topic value, it can apply only to the second sentence. This constraint is actually important for what we said earlier. Recall how we account for the unacceptability of sequences like (71a): we said that the two sentences have different topic values. This depends on the assumption that Copy Deletion is what causes the lower copy of the fronted noun in the first sentence to be phonologically covert. This assumption, in turns, can only be maintained if ellipsis cannot apply to the lower copy of the the fronted noun in the first sentence.

3.3 Measure words

3.3.1 Semantics

Classifiers are not the only elements which mediate between Num and NP in Vietnamese. Nouns can also be made countable by expressions such as **thung** ‘box’ or **tui** ‘bag.’²⁰

- (72) a. John se mua hai thung sach
 John will buy two BOX book
 ‘John will buy two boxes of books’
 b. John se mua hai tui sach
 John will buy two bag book
 ‘John will buy two bags of books’

²⁰The reason for capitalizing the gloss for **thung** and **tui**, i.e. to write them as “BOX” and “BAG,” is to indicate that these words are being used as measure words, not as regular nouns, in which case they will be glossed as “book” and “bag.” The distinction between measure words and regular nouns will be discussed shortly.

In the literature on classifier languages, such words are variously called “mensural classifiers” (Lyons 1977), “measure words” (Li and Thompson 1989), “mass-classifiers” or “massifiers” (Cheng and Sybesma 1998, 1999b). I will call them “measure words,” because I think this term is the most transparent. There has been general agreement that measure words should be distinguished from what we call classifiers (cf. Aikhenvald 2000). The most succinct formulation of the reason for this distinction can perhaps be found in Cheng and Sybesma (1999b: 115), where it is said that measure words “create a unit of measure,” while classifiers “simply name the unit in which the entity denoted by the noun naturally occurs.” Thus, a classifier CL maps a predicate *P* into a set of “natural units” of *P* (cf. Krifka 1995, 2003), or “*P*-atoms” in our terminology, while a measure word MW maps a predicate *P* into a set of units of *P* which are measured in boxes, bags etc. In this section, I will show that the semantics of MW, in conjunction with the CCD, turns out to account for several sound-meaning correlations, as well as acceptability judgements, observable with NP-Split constructions in which the topic noun originates from a nominal complex consisting of a numeral, a measure word and an NP, as exemplified in (73).

- (73) Sach thi no se mua hai thung
 book TOP he will buy two BOX
 ‘As for books, he will buy two boxes of them’

First, there is the question of whether we need the concept of “measure word” at all. Why not just say that **thung** in (72a) and (73) is a regular noun, namely that which means ‘box.’²¹ The reason is because we can observe two clearly different uses of the word **thung**, accompanied by distinct sets of syntactic and semantic characteristics. Note, for example, that there is no classifier between the numeral **hai** ‘two’ and **thung** ‘BOX’ in (72a) and (73). There is, however, a use of **thung** in which a classifier appears between this word and a numeral.

- (74) John se mua hai cai thung
 John will buy two CL_{box} box

These two uses have different interpretive properties. Consider (75), which differs minimally from (74) in that no classifier intervenes between **thung** and the numeral.

²¹Cao (1988) takes this position. In fact, he claims that even the concept of classifiers is vacuous, and that phrases which I have analyzed as [CL N] and [MW N] are just strings of two nouns, one count and one mass. Cao’s argument is based on a number of distributional criteria. It is not clear how he would account for the facts presented here.

- (75) John se mua hai thung
John will buy two BOX

While both (74) and (75) are grammatical, they differ with respect to what they can mean. The most salient reading of the former is that John bought two empty boxes. In fact, this might be the reading of (74), since it really seems that if there are things in the two boxes, they must have gotten there by accident. On the other hand, (75) can only be understood as an elliptical sentence: it has to mean that John bought two boxes of something. This contrast would not be explained if we assume a single lexical entry for **thung**. Also, (74) entails that John ends up in possession of two (empty) physical boxes, whereas (75) can be true in a situation where John buys things which would fill two boxes but no box changes hands. For examples, John might have bought two quantities of books, each the size of a box, and the transaction might have involved the vendor putting books into two boxes, loading the books into John's car and keeping the boxes afterwards. Another interesting contrast is that between (72a), repeated in (76a), and (76b).

- (76) a. John se mua hai thung sach
John will buy two BOX book
b. John se mua hai cai thung sach
John will buy two CL_{box} box book

The minimal difference between (76a) and (76b) is that in the former there is no classifier intervening between **hai** and **thung** while in the latter there is. This turns out to correlate with a semantic difference: while (76a) means John will buy two boxes of books, i.e. two box-sized quantities of books, (76b) means John will buy two (empty) boxes of the type which is used for books, i.e. which could be called “book boxes” in English.

On the basis of facts such as these, I propose (77a) as the definition of the noun **thung** and (77b) as definition of the measure word **thung**.

- (77) Definition
a. $[[[_N \text{ thung}]]] = \lambda x(x \text{ is a box or a pluralities of boxes})$
b. $[[[_{MW} \text{ thung}]]] = \lambda P(\lambda x(x \text{ is a box-load of } P\text{'s}))$

The definition in (77b) is informal. It relies on the reader's understanding of what it means for an entity to be a box-load of something. The meta-linguistic specification in the definition, hence, is to be considered a placeholder for the formal expression that a correct theory of measure words would provide, whatever that theory may turn out to be. For present purposes, it suffices to make explicit

some consequences which I intend the definitions in (77), or more precisely their properly worked out versions, to have. First, by including pluralities in (77a), I intend to capture the fact that the noun [_N **thung**] requires a classifier to combine with a numeral (see the discussion in the previous subsection). Second, by using the phrase “a box-load of *P*’s” instead of, say, “a box filled with things that are *P*’s” in (77b), I intend to capture the fact that the measure word [_{MW} **thung**] specifies quantities, not things, as buying **hai thung sach** ‘two BOX book’ may not involve physical boxes at all. Third, by not mentioning pluralities in (77b), I intend to say that the result of merging [_{MW} **thung**] and, say, **sach** ‘book,’ is an atomic predicate. In other word, $\llbracket [\text{MW thung}] \rrbracket (\llbracket \text{sach} \rrbracket)$ is to contain no element which is the sum of more than one elements in the same. This means that [MW NP] can combine with a numeral.²²

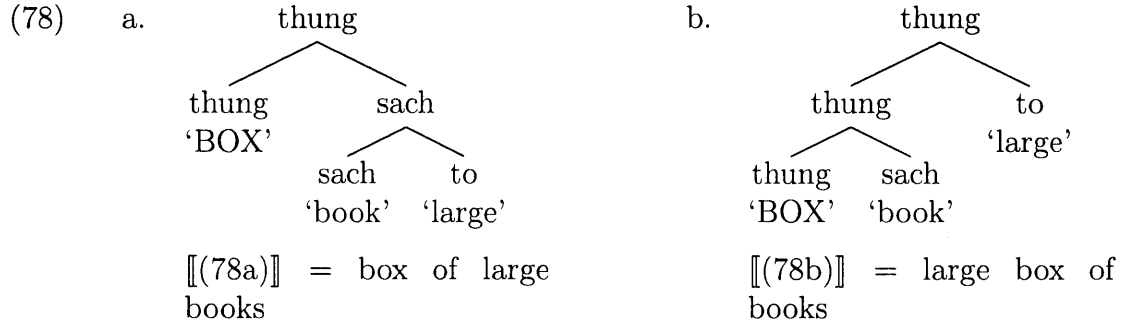
Defining [_{MW} **thung**], the measure word, as a function of predicates to predicates, i.e. an expression of type $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$, entails that (75) has to be elliptical: it must be construed as having an NP elided from the phonological representation but present for interpretation at LF. The contrast in meaning between (74) and (75) is thus accounted for. The contrast in (76) is also explained: the presence of **cai** ‘CL_{box}’ implies that the following word **thung** cannot be a measure word, i.e. cannot be [_{MW} **thung**], which in turn implies that **sach** ‘book,’ which follows **thung**, cannot be a complement of **thung**. The only choice left is to parse **thung sach** as a compound, and as compounds are left-headed in Vietnamese, we get the meaning of ‘book box,’ i.e. ‘box used for books.’

3.3.2 Split and modification

Measure words resemble classifiers in being functions of type $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$ whose range consists of atomic predicates. However, they are, in a sense, more substantive than classifiers. For example, the classifier **quyen** in effect maps books to books, while the measure word **thung** would map books to boxes of books. This difference has interesting repercussions for modification. Recall from the discussion in the previous section that a classifier-noun-modifier sequence, for

²²However, there is no guarantee, from the definition of [_{MW} **thung**], that elements of $\llbracket [\text{MW thung}] \text{ NP} \rrbracket$ do not overlap. Thus, $\llbracket [\text{MW thung}] \text{ sach} \rrbracket$ might, in fact does, include box-sized quantities of books which overlap, i.e. which contain the same books. Given the intuition that buying **hai thung sach** ‘two BOX book’ cannot mean buying two overlapping box-sized quantities of books, we will have to either refine the definition in (77b) or write into the definition of the numeral **hai** ‘two,’ assuming that it is a generalized quantifier, a specification to the effect that elements in $\llbracket \text{XP} \rrbracket \cap \llbracket \text{YP} \rrbracket$, XP being the restriction of the numeral and YP its nuclear scope, do not overlap. I will not attempt these tasks here since they will make the discussion, whose main point is to show that the ECCD is correct, unnecessarily complicated.

example **quyen sach to** ‘CL_{book} book large,’ can either be parsed as either [CL [N Mod]] or [[CL N] Mod], with no semantic consequences. This is because it follows from the definition of CL that $\llbracket \text{CL} \rrbracket (\llbracket \text{N} \rrbracket \cap \llbracket \text{Mod} \rrbracket) = \llbracket \text{CL} \rrbracket (\llbracket \text{N} \rrbracket) \cap \llbracket \text{Mod} \rrbracket$. However, the definition of measure words such as **thung** ‘BOX’ does not guarantee this equivalence. As an example, consider (78a) and (78b).²³



In (78a), the adjective **to** ‘large’ modifies **sach**, which denotes the set of books, and it is clear how the phrase labeled ‘ α ’ ends up denoting boxes of large books. In (78b), on the other hand, the same adjective modifies **thung sach**, which denotes not the set of books, but the set of box-sized quantities of books. Assuming that boxes come in different sizes, the set of box-sized quantities of books will include those the size of a small box and those the size of a large box. In other word, $\llbracket \text{thung sach} \rrbracket$ contains small quantities and large quantities of books. Intersecting $\llbracket \text{thung sach} \rrbracket$ with $\llbracket \text{to} \rrbracket$, which is what the interpretation of (78b) comes down to, will yield a subset of $\llbracket \text{thung sach} \rrbracket$ which contains only the large quantities in $\llbracket \text{thung sach} \rrbracket$. This is the set of large boxes of books.

We predict, then, that (79) is ambiguous between ‘John will buy two boxes of large books’ and ‘John will buy two large boxes of books.’ As the translation shows, this prediction is correct.

- (79) John se mua hai thung sach to
 John will buy two BOX book large
 ‘John will buy two large boxes of books / two boxes of large books’

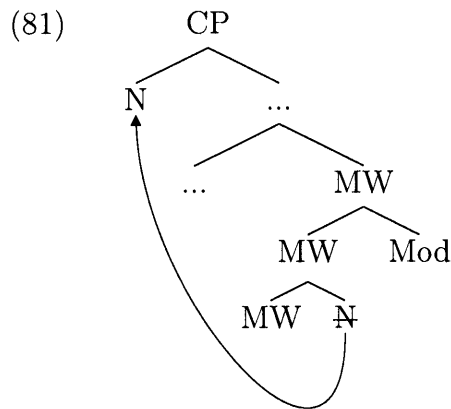
Now let us ask what prediction can be made about NP-split constructions in which the split nominal complex contains a measure word and a modifier. Consider the following sentence.

²³Note that we continue with the assumption that modifiers do not project, and heads which denote functions that take their sisters as arguments project.

- (80) Sach thi John se mua hai thung to
 book TOP John will buy two BOX large
 ‘John will buy two large boxes of books / *two boxes of large books’

The observation, then, is that (80) is non-ambiguous: it has the reading where the adjective modifies ‘boxes of books,’ but not the one where it modifies ‘books.’ It turns out that we predict this fact. Here is how.

First, the fact that there is no doubling of the fronted noun in (80) means that Copy Deletion must have applied, which means, assuming the ECCD is valid, that the lower copy of the chain created by topicalization of **sach** ‘book’ must be at the right edge of an XP. Given that (78a) and (78b) exhaust the possibilities of parsing **thung sach to** ‘BOX book large,’ and that **sach** is XP-final only in (78b), we conclude that **sach** ‘book’ in (80) must be fronted from (78b). In other word, the analysis of (80) must be (81).

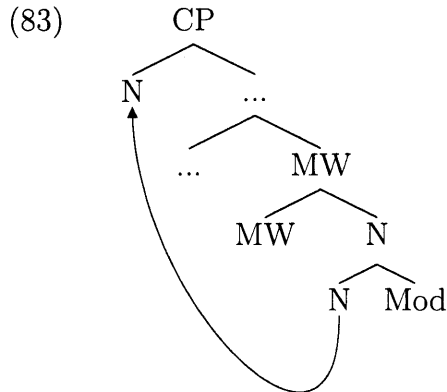


We make the correct prediction, then, that (80) must mean John will buy two large boxes of books and cannot mean John will buy two boxes of large book, since only the former meaning can be computed from (81). Now let us consider the doubling variant of (80), i.e. the sentence which differs minimally from (80) in that the topicalized noun is doubled.

- (82) Sach thi John se mua hai thung sach to
 book TOP John will buy two BOX book large
 ‘John will buy two large boxes of books / two boxes of large books’

As the translation shows, (82) is ambiguous in exactly the same way as (79) is. This fact is not what we predict. As there is doubling of the topic noun **sach** ‘book’ in (82), Copy Deletion must have been blocked, which means that the base position of the chain created by topicalization must not be XP-final. The

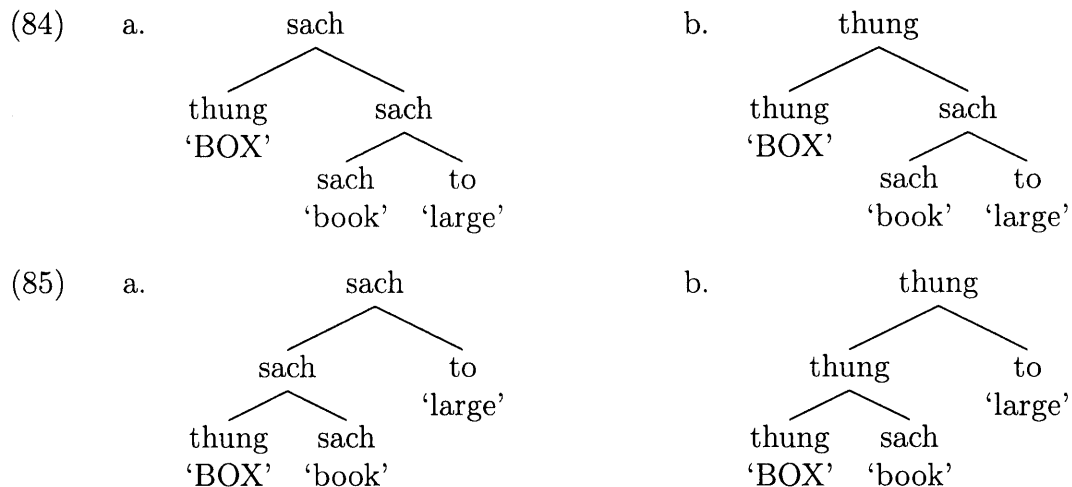
conclusion, then, is that **sach** ‘book’ must have been fronted from (78a), i.e. that (82) must have the structure in (83).



However, that means (82) can only be read as ‘John will buy two boxes of large book.’ In other word, we predict, incorrectly, that (82) is not ambiguous.

One solution to this problem is to say that Pronunciation Economy (PE) is inoperative when the lower copy is a noun following a measure word. Thus, Copy Deletion can apply in (81), but does not have to. This enables us to say that the semantic ambiguity of (82) is due to the fact that it is structurally ambiguous: it can be parsed as (81) without Copy Deletion, or as (83) where Copy Deletion cannot apply. This solution, however, requires a rather ad hoc specification about when PE can be switched off.

Another solution, which was suggested to me by David Pesetsky and which I will adopt, is to assume that when a measure word and noun merge, projection is free choice: either the measure word or the noun can project. Thus, the string **thung sach to** ‘BOX book large’ would have four possible analyses.



Among the four structures in (84) and (85), only one, namely (85b), is such that fronting **sach** from it will not result in doubling. As (85b) denotes ‘large box of books,’ we predict, correctly, that (80) is unambiguous, i.e. that it has to mean ‘John will buy two large boxes of books’: the noun **sach** ‘book’ in that sentence must have been fronted from (85b), since the lower copy of (**sach**, **sach**) has been deleted and (85b) is the only structure in which this copy is XP-final. Fronting **sach** ‘book’ from any of the remaining structures will force doubling of the fronted constituent, since in none of these structures is **sach**, in its base position, XP-final. We predict, again correctly, that (82) is ambiguous: the noun **sach** in it could be fronted from one of the structures in (84), in which case the sentence would mean ‘John will buy two large boxes of books,’ or it could be fronted from (85a), in which case the sentence would mean ‘John will buy two boxes of large books.’

3.4 Mandarin Chinese

The most impressive argument for the ECCD is perhaps showing a correlation between doubling and headedness, keeping every other factor constant. Thus, if we could find a language whose nominal complexes resemble Vietnamese nominal complexes in every relevant aspect save the fact that NP is head-final, and show that topicalization of N in that language never results in doubling of the fronted constituent, the facts about NP-Split constructions in Vietnamese which we presented in the previous sections and which we derived, ultimately, from the CCD, become more of a convincing argument for this principle, since it becomes more unlikely that these facts can be derived from, say, the semantics of classifiers. It turns out that Mandarin Chinese is such a language.²⁴ Nouns in Mandarin Chinese also need to combine with a classifier in order to be countable.

- (86) wo zhi du san *(ben) shu
 I only read three *(CL_{book}) book

In the appendix of this chapter, it will become evident that Mandarin Chinese and Vietnamese are practically identical with respect to the semantics of nouns and classifiers. An glaring difference between these two languages, however, is that NP in Vietnamese is head-initial, as we have seen, while NP in Mandarin Chinese is head-final. The next two examples show that both modifiers and complements precede the head noun in Chinese.

²⁴I thank Li Jiang (Harvard University) for providing the Chinese data.

- (87) a. wo zhi du san ben guanyu wuli de shu
 I only read three CL_{book} about physics DE book
 'I only read three books about physics'
 b. wo zhi kanjian san ge John de qizi
 I only met three CL_{wife} John DE wife
 'I only met three wives of John'

What happens when N is topicalized in Mandarin Chinese? First, let us consider cases where N has no modifier or complement.

- (88) shu wo zhi du san ben (*shu)
 book I only read three CL

Doubling is ruled out, just as we expect: the lower copy of the chain created by topicalization is at the end of NP, as well as at the end of every XP that dominates NP. Given the CCD, it can be deleted, and given PE, it must. This is exactly what happens in Vietnamese also, when unmodified, complement-less nouns are fronted. Now consider (89).

- (89) shu wo zhi du san ben guanyu wuli de (*shu)
 book I only read three CL about physics DE (*book)

In (89), a modified non-relational noun is topicalized, stranding its modifier. Recall that in Vietnamese, doubling is optional in this scenario. It is, however, impossible in Mandarin Chinese, as indicated in (89). This, of course, is just what we should expect, since the noun in question, at its base position, is XP-final: it ends not only NP, which it heads, but also VP, TP, CP etc.

Finally, consider the case where a relational noun is fronted, stranding its (overt) complement. This is the scenario in which Vietnamese forces doubling. Mandarin Chinese, on the other hand, forces deletion, as (90) shows.

- (90) qizi wo zhi kanjian san ge John de (*qizi)
 wife I only met three CL John DE

The judgement in (90) is, again, predicted by the CCD, since the lower copy of (**qizi**, **qizi**) is XP-final.

3.5 Conclusion

The previous sections were aimed at providing empirical arguments for the ECCD from facts about NP-Split constructions in Vietnamese. Specifically, I have shown how several complicated patterns of pronunciation and interpretation can be derived from the CCD, given a number of auxiliary assumptions concerning the syntax and semantics of the elements of the nominal complex, e.g. nouns, classifiers, and measure words. Mandarin Chinese was also discussed, the aim being to show that the difference in headedness of NP between this language and Vietnamese has exactly the consequences we expect it to have, given the CCD.

The auxiliary assumptions just mentioned were introduced with the objective of establishing premises for the main line of argumentation with a minimal amount of diversion. The conclusion reached at the end of the argumentation, namely that the ECCD is a valid generalization, is of course convincing only to the extent that these auxiliary assumptions can be justified on grounds independent of their role in the argument for the CCD. Some observations were presented to this end, but they in no way exhaust the amount of data supportive of the assumptions in question. For this reason, I will end this chapter with an appendix in which I present a theory of nominal reference. This theory incorporates the assumptions about nouns and classifiers which I used in my argument for the ECCD and makes it evident that these are independently needed to account for several other facts of both Vietnamese and Mandarin Chinese. That is one reason for attaching the appendix to the chapter. The other reason, which is perhaps no less important, is that a contribution will hopefully be made to the theoretical discussion on classifier languages. A general semantic theory of nominals should account for the macro distinction between classifier languages on the one hand and number marking languages on the other. It should also capture the micro differences that exist within each group. What is required, then, is detailed investigation of particular languages of either type in order to identify both the general features and the points of parametric variation between them. Several works have been devoted to the semantics of nominals in number marking languages (cf. Barker 1992, Krifka 1999, Link 1983b, Pelletier and Schubert 1989, Schwarzschild 1992, among others). A number of concrete proposals have been made on the macro variation between classifier and number marking languages, and also, on the micro variation amongst the latter (cf. Chierchia 1998b,a, 2010, Dayal 2004, Krifka 1995). To the best of my knowledge, however, the semantics of nominals in particular classifier languages have not been discussed with the same degree of explicitness, and the micro variation amongst these languages have practically

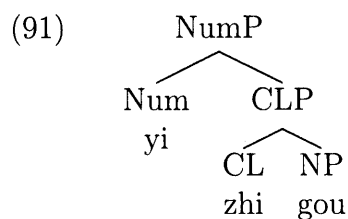
been glossed over in the semantic literature.²⁵ This appendix is intended to be a small step towards the rectification of this situation. Specifically, it develops an explicit and detailed account of the syntax and semantics of nominals in Chinese. It then extends the results of the analysis to Vietnamese and proposes an explanation for the differences between these two languages.

Thus, the appendix is somewhat of a diversion. Readers who are interested exclusively on chain linearization in general, and the ECCD in particular, can skip it without being in any way ill-prepared for chapter 4.

Appendix on nominal reference in Mandarin Chinese and Vietnamese

Mandarin Chinese

Mandarin Chinese is a “classifier” language of the East Asian variety. Thus, count nouns in this language cannot combine directly with numerals without the mediation of grammatical morphemes: the classifiers (cf. Ren 1968, Cheng and Sybesma 1999b, 2005b, Lee 1986, Li and Thompson 1989, Shyu 1995, Tang 1990, Tsai 1994, 2001, Xu 1996). For example, ‘one dog’ has to be expressed as **yi zhi gou**, where **yi** is the numeral ‘one’ and **gou** is the word for ‘dog’. The word **zhi** is the classifier which enables **gou** to combine with **yi**, so to speak. I am going to assume that the numeral c-commands both the classifier and the noun, and that the classifier c-commands the noun, as in (91) (cf. Cheng and Sybesma 1999b, 2005b, among others).



It turns out that NP, CLP and NumP can all appear independently in sentences. For example, the bare noun **gou**, the classifier phrase **zhi gou** and the numeral phrase **yi zhi gou** can all be arguments of verbs. However, these categories differ both in distribution and interpretation. This fact is captured succinctly in

²⁵Although several works have discussed the syntax of nominals in classifier languages (cf. Saito et al. 2008, Cheng and Sybesma 1999a, 2005a, Watanabe 2010, Wu and Bodomo 2009, among others).

a quote from Cheng and Sybesma (2005b:263): “In Mandarin, bare NPs can be interpreted as definite, indefinite or generic. Num-CL-NPs [numeral phrases] and CL-NPs [classifier phrases] are invariably indefinite [...] All indefinites occur in postverbal position only.”²⁶

The sentences in (92) exemplify the interpretation of NP in subject positions: (92a) shows that bare nouns (NPs) can be definite, singular or plural, and (92b-c) evidence the generic reading of NP, both as arguments of individual-level predicates such as ‘to be intelligent,’ and kind-predicates such as ‘to be extinct’ (cf. Cheng and Sybesma 1999b:510, Yang 2001:20, Rullmann and You 2006:176).

- (92) a. Gou yao guo malu
 dog want cross road
 ‘The dog(s) / *a dog / *dogs want(s) to cross the street’
 b. Gou hen jiling
 dog very smart
 ‘The dog(s) / dogs is/are intelligent’
 c. Gou juezhong le
 dog extinct ASP
 ‘Dogs are extinct’

The sentences in (93) are examples of NP in post-verbal position. We can see that NP in object positions has all the readings that NP in subject positions has, plus the indefinite reading, which is number-neutral. Thus, **kanjian gou** ‘see dog’ can mean ‘see the dog,’ ‘see the dogs,’ ‘see a dog,’ or ‘see (some) dogs’ (cf. Rullmann and You 2006:176, Cheng and Sybesma 1999b:510).

- (93) a. Wo kanjian gou le
 I see dog ASP
 ‘I saw the dog(s) / a dog / dogs’
 b. Wo xihuan gou
 I like dog
 ‘I like dogs’

²⁶To express meanings such as ‘the two dogs,’ Chinese speakers resort to the use of demonstratives. Cheng and Sybesma do not consider demonstratives and neither will I in this work. As can be seen from their examples, Cheng and Sybesma intended the term “generic” to mean both ‘kind-referring’ (or ‘D-generic’), as in **dogs will be extinct**, and ‘characterizing’ (or ‘I-generic’), as in **dogs like meat** (Krifka 1987, Krifka et al. 1995). We will use “generic” in the same way here. Note, also, that we exclude from the discussion the so-called “quantity interpretation” of numeral phrases (‘five children cannot finish ten bowls of rice’), under which these phrases can occupy subject positions (Li 1998). If the analysis proposed below is correct, numeral phrases in this reading must have more structure than that represented in (91). I leave this topic to future research.

- c. Turing faming dyannao
Turing invent computer
'Turing invented the computer'

The next two sentences, (94a) and (94b), illustrate the use of NP as predicatives (Leo Chen, T.-C. James Huang, Zhang Min personal communication). Again, we witness number-neutrality: the bare noun **gou** can be predicated of a singular subject like Fido, or a plural one like Fido and Pluto.

- (94) a. Fido shi gou
Fido be dog
'Fido is a dog'
- b. Fido he Pluto shi gou
Fido and Pluto be dog
'Fido and Pluto are dogs'

The next category, CLP [classifier phrase], can only appear in post-verbal positions, as evidenced by the contrast in (95). In addition, it can only be interpreted as a singular indefinite: (95a) has the implicature that the speaker bought one but not two books (cf. Rullmann and You 2006:175, Cheng and Sybesma 2005b:262).

- (95) a. Zuotian wo mai le ben shu
yesterday I buy ASP CL book
'I bought a book yesterday'
- b. *Zhi gou yao guo malu
CL dog want cross road
'(A/the dog wants to cross the street)'

Just like NP, CLP can function as predicatives. However, it differs from NP in that it requires a singular subject: **zhi gou** can be predicated of a singular individual such as John, but not of a sum individual such as Fido and Pluto (Leo Chen, T.-C. James Huang, Zhang Min, personal communication).

- (96) a. John shi zhi gou
John be CL dog
'John is a dog'
- b. *Fido he Pluto shi zhi gou
Fido and Pluto be CL dog
'(Fido and Pluto are dogs)'

The numeral phrase has basically the same syntax and semantics as the classifier phrase: it can only appear in post-verbal positions, and interpreted as an

indefinite, as (97) shows (cf. Yang 2001:133, Li 1998:694, Cheng and Sybesma 2005b:262).

- (97) a. Wo kan le san ben shu
I read ASP three CL book
'I read three books'
- b. *San ge xuesheng zai xuexiao shoushang le
three CL student at school hurt ASP
'Three students were hurt at school'
- c. *yi zhi gou xihuan chi rou
one CL dog like eat meat
'A dog (generally) likes to eat meat'

NumP can also appear post-copula, in which case the subject must match the predicative in number. (98) shows that the phrase **liang zhi gou** 'two CL dog' can be predicated of Fido and Pluto, but not of Fido alone.

- (98) a. Fido he Pluto shi liang zhi gou
Fido and Pluto be two CL dog
'Fido and Pluto are two dogs'
- b. *Fido shi liang zhi gou
Fido be two CL dog

The facts just described are summarized in (99). The generalization is that all nominal categories can be indefinites or predicatives, but only bare nouns can be generic or definite. Furthermore, Mandarin Chinese does not allow indefinite subjects.

- (99) Generalizations about Chinese NP, CLP and NumP
- a. Indefinite and predicative: all categories
- b. Definite and generic: NP
- c. Subjects cannot be indefinite

Analysis

We turn now to the analysis. The starting point will be a domain of quantification U which contains both singularities and pluralities ($U = \{a, b, c, \dots, a+b, b+c, a+b+c, \dots\}$), as assumed in many previous works (Chierchia 1998b, Landman 1989, Link 1983a, Schwarzschild 1996). The sum operator $+$ which maps singular to plural individuals, and the "part of" relation $\leq$ which partially orders U , are understood in the usual way (i.e. $x + x = x, x + y = y + x, (x + y) + z =$

$x + (y + z), x \leq y \leftrightarrow x + y = y$). It would help to define two notions which will feature in the analysis to be developed below. First, given any subset P of U , an atom of P , or a P -atom, will be a P -individual which has no proper part which is itself a P -individual. In (100) and all subsequent definitions, free variables are to be understood as universally quantified over.

$$(100) \quad x \in AT(P) \text{ iff } x \in P \wedge \forall y((y \in P \wedge y \leq x) \rightarrow (y = x))$$

‘ x is a P atom iff x is P and has no proper part which is P ’

Second, the maximal element of P , $MAX(P)$, will be that individual in P which has every element of P as part.²⁷

$$(101) \quad MAX(P) = SUP(P) \text{ if } SUP(P) \in P, \text{ undefined otherwise}$$

‘The maximal element of P is that individual in P which has every individual in P as part’

To illustrate, suppose $P = \{a, b, a + b\}$. Then, $MAX(P) = SUP(P) = a + b$. If $P = \{a\}$, then $MAX(P) = SUP(P) = a$. If $P = \{a, b\}$, $SUP(P) = a + b$ but $MAX(P)$ is undefined.

A consequence of (101) is that if $P = AT(P)$, then $MAX(P)$ is defined only if P is a singleton. To see this, let $P = AT(P)$ and $MAX(P)$ be defined. It follows from the definition of MAX (101) that there is some x such that (i) $MAX(P) = x$, (ii) $x = SUP(P)$, and (iii) $x \in P$. Given that $P = AT(P)$, it follows that (iv) $x \in AT(P)$. Now suppose P is not a singleton. Then there is some y such that (v) $y \in P$ and (vi) $y \neq x$. Given the definition of SUP (i), it follows from (ii) and (v) that (vii) $y \leq x$. Given the definition of AT (100), it follows from (iv), (v) and (vii) that $y = x$. But (vii) contradicts (vi). Thus, P is a singleton.

Following Chierchia (1998b) and Chierchia and Turner (1988), I assume that U contains a set G of kinds, and that kinds are intensional entities, or more precisely, they are individual concepts, i.e. $G \subseteq U^W$. The idea is that each kind k is an individual correlate of a property P : it maps each possible world w to the sum individual which encompasses all individuals which are P in w . This will be made more precise below.

²⁷The definition of $MAX(P)$ employs the notion ‘supremum of P ’, $SUP(P)$, which is defined in (i).

(i) $x \in P \rightarrow x \leq SUP(P)$ and $\forall y(y \in P \rightarrow y \leq z) \rightarrow SUP(P) \leq z$
‘The supremum of P is the smallest individual that has every element of P as part’

I turn now to the semantics of nominals in Mandarin Chinese. Following Chierchia (1998b), I assume that nouns in classifier languages are “cumulative” predicates, and that the function of CL is to make predicates “atomic.” The definitions of “atomic” and “cumulative” are given in (102). Basically, X is an atomic predicate if the extension of X necessarily consists of atoms of X , and X is a cumulative predicate iff the extension of X is necessarily a set closed under the sum operation (cf. Krifka 1989, Quine 1960).

- (102) X is an atomic predicate iff $\llbracket X \rrbracket^w = AT(\llbracket X \rrbracket^w)$
 X is a cumulative predicate iff $\llbracket X \rrbracket^w = +AT(\llbracket X \rrbracket^w)$

We write ‘ $+P$ ’ to denote the closure of P under $+$. For example, if $P = \{a, b, c\}$ then $+P = \{a, b, c, a + b, b + c, a + c, a + b + c\}$.²⁸ The classifier CL is defined as in (103). It denotes the atomizing function AT , which applies to a set and yields the atoms of this set.

- (103) $\llbracket CL \rrbracket^w \in D_{\langle \langle e, t \rangle, \langle e, t \rangle \rangle}$
 $\llbracket CL X \rrbracket^w = AT(\llbracket X \rrbracket^w)$

From the definition of CL in (103), we can derive the theorem that classifier phrases are atomic predicates.

- (104) Theorem 1
 CL X is an atomic predicate

We prove Theorem 1 by proving that $\llbracket CL X \rrbracket^w = AT(\llbracket CL X \rrbracket^w)$. Given the definition of CL (103), this means proving that $AT(\llbracket X \rrbracket^w) = AT(AT(\llbracket X \rrbracket^w))$, or more generally that $AT(P) = AT(AT(P))$, i.e. that $x \in AT(P)$ iff $x \in AT(AT(P))$. Now it follows from the definition of AT (100) that if $x \in AT(AT(P))$ then $x \in AT(P)$. The same definition implies that we can prove the other direction by showing that if $x \in AT(P)$ then $((y \in AT(P) \wedge y \leq x) \rightarrow (y = x))$, i.e. that if (i) $x \in AT(P)$, (ii) $y \in AT(P)$ and (iii) $y \leq x$, then $y = x$. Given, again, the definition of AT (100), it follows from (ii) that (iv) $y \in P$, and from (i), (iv) and (iii) that $y = x$. QED.

We can now explain the predicative use of NP and CLP. Since NP is cumulative, its extension includes both singularities and pluralities, which means it can be true of both singular and plural individuals, as shown in (105a-b). Since CLP is atomic, it can be true of only singular individual, as shown in (105c-d). Assuming

²⁸More formally, $+P$ is the smallest set such that if $x \in P$ then $x \in +P$, and if $x, y \in P$ then $x + y \in P$.

that analytically false sentences are ungrammatical (cf. Von Stechow 1993, Gajewski 2003, Abrusán 2007) we explain the contrast seen in (96).

- (105) a. $\llbracket \text{Fido shi gou} \rrbracket^w = 1$
iff $\llbracket \text{Fido} \rrbracket^w \in \{a, b, c, a + b, b + c, a + c, a + b + c\}$
b. $\llbracket \text{Fido he Pluto shi gou} \rrbracket^w = 1$
iff $\llbracket \text{Fido} \rrbracket^w + \llbracket \text{Pluto} \rrbracket^w \in \{a, b, c, a + b, b + c, a + c, a + b + c\}$
c. $\llbracket \text{Fido shi zhi gou} \rrbracket^w = 1$ iff $\llbracket \text{Fido} \rrbracket^w \in \{a, b, c\}$
d. $\llbracket \text{Fido he Pluto shi zhi gou} \rrbracket^w = 1$
iff $\llbracket \text{Fido} \rrbracket^w + \llbracket \text{Pluto} \rrbracket^w \in \{a, b, c\}$, i.e. iff $\perp$

As for the numeral phrase NumP, we follow Ionin and Matushansky (2006) and assume that only individuals of the same cardinality can be counted. One way to flesh out this idea is to require that the predicate P which is the complement of a numeral necessarily contain only individuals of the same number of P -parts. This is written into the definition of numerals, as exemplified by the definition of **liang** in (106). We write $|x|^P$ to denote the number of P -parts of x , i.e. $|x|^P = \#\{y | y \in P \wedge y \leq x\}$.

- (106) $\llbracket \text{liang} \rrbracket^w \in D_{\langle \langle e, t \rangle, \langle e, t \rangle \rangle}$
 $\llbracket \text{liang X} \rrbracket^w$ is defined iff $\exists n (\forall w' (\forall u (u \in \llbracket X \rrbracket^{w'} \rightarrow |u|^{\llbracket X \rrbracket^{w'}} = n)))$
If defined, $\llbracket \text{liang X} \rrbracket^w = \lambda x (x \in +\llbracket X \rrbracket^w \wedge |x|^{\llbracket X \rrbracket^w} = 2)$

From (106) we can derive the theorem that the complement of **liang** must be an atomic predicate, and also that the numeral phrase itself is an atomic predicate.

- (107) Theorem 2
 $\llbracket \text{liang X} \rrbracket^w$ is defined iff X is an atomic predicate

Theorem 2 is proved as follows. Suppose X is not atomic. From the definition of “atomic” (102), it follows that for some w' , $AT(\llbracket X \rrbracket^{w'}) \neq \llbracket X \rrbracket^{w'}$, hence $AT(\llbracket X \rrbracket^{w'}) \subset \llbracket X \rrbracket^{w'}$. This means that there exists some $u \in \llbracket X \rrbracket^{w'}$ such that $u \notin AT(\llbracket X \rrbracket^{w'})$, which means there is some $v \in \llbracket X \rrbracket^{w'}$ such that $v < u$, i.e. such that $|v|^{\llbracket X \rrbracket^{w'}} \neq |u|^{\llbracket X \rrbracket^{w'}}$. It follows from (106) that $\llbracket \text{liang X} \rrbracket^w$ is not defined. Now suppose X is atomic. Then for all w' , $\llbracket X \rrbracket^{w'} = AT(\llbracket X \rrbracket^{w'})$, hence for all w' , $|u|^{\llbracket X \rrbracket^{w'}} = 1$ for all $u \in \llbracket X \rrbracket^{w'}$, which means that $\llbracket \text{liang X} \rrbracket^w$ is defined. QED.

- (108) Theorem 3
 $\llbracket \text{liang X} \rrbracket^w$ is an atomic predicate

Here is the proof. Suppose $\llbracket \text{liang } X \rrbracket^w$ is not an atomic predicate. Then for some w' , $\llbracket \text{liang } X \rrbracket^{w'}$ contains v and u such that $v \neq u$ and $v + u = u$. By assumption, $|v|^{\llbracket X \rrbracket^{w'}} = |u|^{\llbracket X \rrbracket^{w'}} = 2$. As $v + u = u$, $|v + u|^{\llbracket X \rrbracket^{w'}} = |u|^{\llbracket X \rrbracket^{w'}} = 2$, which means $v = u$. This contradicts our assumption. QED.

The predicative use of NumP follows: as **liang zhi gou** ‘two CL dog’ denotes a set of pluralities of dogs, or more precisely a set of duos of dogs, only a plural individual like Fido and Pluto can be in that set. A singular individual like Fido cannot be in the extension of **liang zhi gou** ‘two CL dog.’ Again, assuming that analytically false sentences are ungrammatical, we explain the contrast seen in (98).

- (109) a. $\llbracket \text{Fido he Pluto shi liang zhi gou} \rrbracket^w = 1$ iff $\llbracket \text{Fido} \rrbracket^w + \llbracket \text{Pluto} \rrbracket^w \in \{a + b, b + c, a + c\}$
 b. $\llbracket \text{Fido shi liang zhi gou} \rrbracket^w = 1$ iff $\llbracket \text{Fido} \rrbracket^w \in \{a + b, b + c, a + c\}$, i.e. iff $\perp$

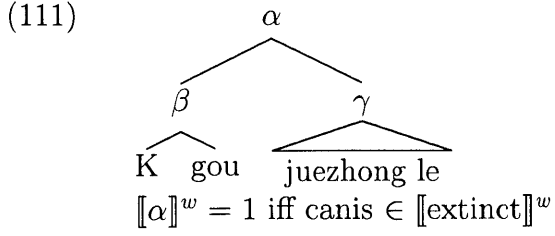
We now come to the generic reading of nominals. First, let us consider D-genericity, i.e. kind-predication exemplified by sentences such as ‘dogs are extinct’ or ‘dogs are related to wolves.’

Chierchia (1998b) advances a theory of kind reference which include the following assumptions. (i) There is a linguistic operator – which we will symbolize as “K” – that maps nominal predicates (i.e. expressions of type $\langle e, t \rangle$) into names of kinds (i.e. expressions of type e). (ii) Kinds are “individual concepts of some sort [...] functions from worlds [...] into pluralities, the sum of all instances of the kind [...]” (iii) The operator is a partial function, which means that some nominal predicates are not in its domain, i.e. “not all individual concepts are going to be kinds” (Chierchia 1998b:349-350). We adopt these assumptions, and flesh out Chierchia’s idea in the following definition of K.

- (110) $\llbracket K \rrbracket^w \in D_{\langle \langle e, t \rangle, \langle s, e \rangle \rangle}$
 $\llbracket K X \rrbracket^w = \lambda w (MAX(\llbracket X \rrbracket^w))$ if $\llbracket K X \rrbracket^w \in G$, undefined otherwise

Thus, K combines with a predicate X and yields an individual concept, a function from each possible world w to the maximal X-individual in w . Furthermore, the individual concept denoted by $\llbracket K X \rrbracket$ must be a kind: $\llbracket K X \rrbracket$ is undefined if it does not denote a kind.

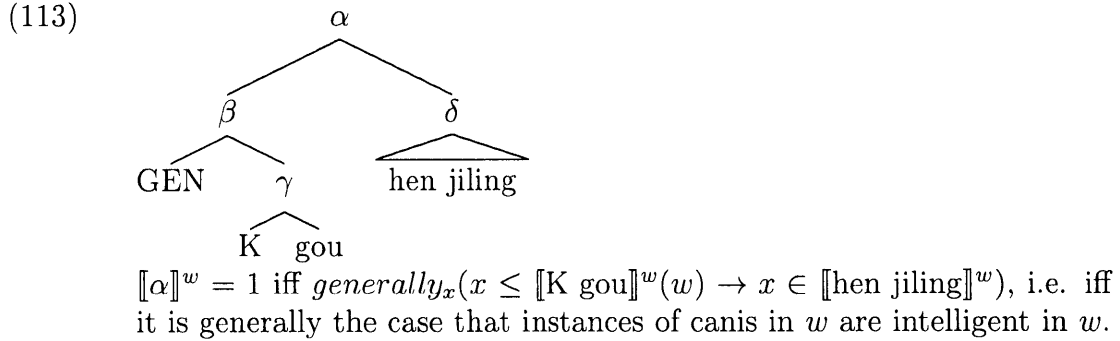
Given the definition of K, it is clear how to generate kind-predication sentences, i.e. the D-generic reading. The LF of ‘dog extinct’ will be something like (111), which will have the meaning that the kind dog, or canis, is extinct.



Let us now turn to the I-generic reading, as exemplified by sentences such as “dogs are intelligent.” We will assume, following several works, that the I-generic reading comes about via a generic operator, GEN, which takes a kind and returns a generalized quantifier (Krifka 1987, Krifka et al. 1995, Chierchia 1998b). Basically, GEN takes a kind and a predicate and returns true iff instances of the kind generally fall under the predicate. The definition of GEN is given in (112).²⁹

(112) $\llbracket \text{GEN } X \rrbracket^w$ is defined iff $\llbracket X \rrbracket^w \in G$
 If defined, $\llbracket \text{GEN } X \rrbracket^w = \lambda P_{\langle e, t \rangle} (\text{generally}_x (x \leq \llbracket X \rrbracket^w(w) \rightarrow x \in P))$

So a sentence such as ‘dogs are intelligent’ will have the LF in (113).



We have explain how bare nouns can have the generic reading, i.e. how they can denote kinds and restrict GEN. It remains to explain why classifier and numeral phrases cannot be generic. Again, we will base the explanation on an idea in Chierchia (1998b: 350), namely that “something which is necessarily instantiated by just one individual [...] would not quality as a kind” (Chierchia 1998b: 350). We explicate this idea by postulating the constraint in (114), which basically says that $\llbracket K X \rrbracket$ would denote a kind only if for some world w' , the sum individual which represents $\llbracket K X \rrbracket$ in w' is plural. We have defined $\llbracket K X \rrbracket$ as a function from

²⁹The generic operator assumed in Chierchia (1998b) also selects a kind as its restrictor, even though this is not stated explicitly in Chierchia’s paper (I thank Gennaro Chierchia for pointing this out to me). For how the word “generally” in the definition of GEN is to be understood, see Krifka (1987), Krifka et al. (1995).

worlds to $MAX(X)$. This means that $[K X]$ is a kind only if $MAX(X)$ consists of more than one X -atoms in some possible world.

$$(114) \quad \llbracket K X \rrbracket^w \in G \text{ only if for some } w', |MAX(\llbracket X \rrbracket^{w'})|^{AT(\llbracket X \rrbracket^{w'})} > 1.$$

Thus, a predicate like **gou** ‘dog’ can combine with K because there is a possible world where the maximal dog consists of more than one dog-atom. But a predicate like ‘being identical to Gennaro Chierchia’ will not be able to combine with K , because in every world, the maximal element in this predicate consists of exactly one Gennaro atom. Given (114), we can prove that K cannot combine with an atomic predicate, because if X is an atomic predicate, $MAX(X)$ is either undefined or contain just one X -atom.

$$(115) \quad \text{Theorem 4} \\ \llbracket K X \rrbracket^w \text{ is undefined if } X \text{ is an atomic predicate}$$

Theorem 4 is proved as follows. Let X be an atomic predicate and w' be a world. Given the definition of “atomic predicate” (102), $\llbracket X \rrbracket^{w'} = AT(\llbracket X \rrbracket^{w'})$. We have proved above that $MAX(\llbracket X \rrbracket^{w'})$ is defined only if $\llbracket X \rrbracket^{w'}$ is a singleton. Thus, $|MAX(\llbracket X \rrbracket^{w'})|^{AT(\llbracket X \rrbracket^{w'})} = n$ only if $n = 1$. Since w' is an arbitrary choice, no w'' is such that $|MAX(\llbracket X \rrbracket^{w''})|^{AT(\llbracket X \rrbracket^{w''})} = n$ and $n \neq 1$. This means that no w'' is such that $|MAX(\llbracket X \rrbracket^{w''})|^{AT(\llbracket X \rrbracket^{w''})} > 1$. Given the constraint on G (114), $\llbracket K X \rrbracket^w \notin G$, and given the definition of K (110), $\llbracket K X \rrbracket^w$ is undefined. QED.

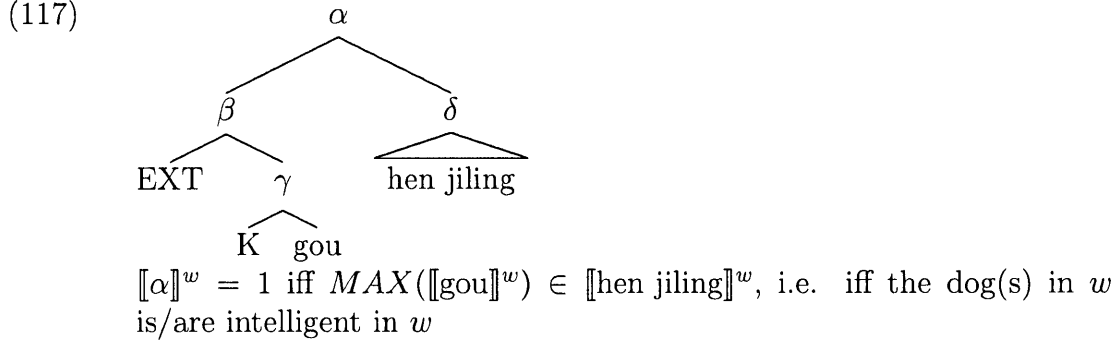
Because genericity is expressed via kind-reference, and kind reference requires cumulative predicates, it follows that classifier phrases and numeral phrases cannot have a generic interpretation, because as we have proved, both of these categories are atomic predicates.

We come now to the definite reading of nominals. Recall that in Chinese, only bare nouns can be definite. Given that only bare nouns can denote kind, this fact suggests that definiteness is also expressed via kind-reference in Chinese. It turns out that there is a very natural way that this can be done. We have defined kinds as function from worlds to maximal individuals, and it is run of the mills to analyze definiteness in terms of maximality (cf. Kadmon 1990, Roberts 2003, Sharvy 1980). So all we have to do is to define an operator EXT which takes an individual concept and applies it to the evaluation world.³⁰

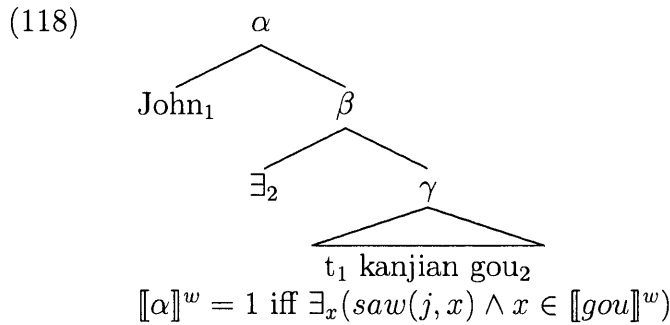
$$(116) \quad \llbracket EXT X \rrbracket^w = \llbracket X \rrbracket^w(w)$$

³⁰In this sense, EXT has the same function as the operator $\vee$ of Montague (1973).

This means that EXT combined with [K X] will give us the meaning of ‘the X.’ It also means that only [K X] can combine with EXT, because only [K X] denotes an intension. And since only bare nouns can combine with K, only bare nouns can be definite. So the LF in (117a) will give us the meaning ‘the dog is intelligent’ in a world where there is exactly one dog, and the meaning ‘the dogs are intelligent’ in a world where there are more than one dogs.



Finally, let us discuss the indefinite reading of nominals in Mandarin Chinese. This reading is only available to objects. To account for the possibility of indefinite objects, we assume that verbs and objects in Chinese can compose via the rule of Restrict (Chung and Ladusaw 2004), and Existential Closure applies at the VP level, binding free variables in it (Heim 1982, Diesing 1992a). To account for the impossibility of indefinite subjects in Chinese, we assume that subjects in Chinese cannot reconstruct into VP (Tsai 2001), hence cannot be existentially closed. The LF of **John kanjian gou** ‘John saw dog’ is and its truth conditions is given in (118).



By hypothesis, $\llbracket \text{gou} \rrbracket^w$ contains both singular and plural dogs, which means **John kanjian gou** is true iff John either saw a single dog, or he saw a plurality of dogs. This is the result we want (cf. (3)). If instead of **gou** ‘dog,’ we have the classifier phrase **zhi gou** ‘CL dog’ or the numeral phrase **yi zhi gou** ‘one CL dog’, we

predict the sentence to have the implicature that John saw a single dog (cf. Zweig 2009).

Vietnamese

Vietnamese is a classifier language which resembles Chinese in many respects. The most obvious similarity between the two languages is the fact that nouns need classifiers to combine with numerals. However, Vietnamese does differ from Chinese in a rather curious manner. Let us first take note of the facts.

Bare nouns, i.e. NPs without CL or NUM, in Vietnamese are interpretable as names of kinds and as generics. In this respect, they resemble bare plurals in English (Carlson 1977, Krifka 2003).

- (119) a. Cho se tuyet-chung
Dog will extinct
'Dogs will be extinct'
b. Cho thich an thit
Dog like eat meat
'Dogs like to eat meat'

And like English bare plurals, bare nouns in Vietnamese can also be interpreted as indefinites (Carlson 1977, Diesing 1992b). There are, however, two differences. First, the indefinite interpretation is limited to object position.

- (120) Cho an thit
Dog eat meat
'Dogs eat meat' / *'Dogs are eating meat'

Second, there is no implicature of plurality (Zweig 2009). In other word, bare nouns in Vietnamese are truly 'number neutral'.

- (121) a. John co cho
John has dog
'John has dogs/a dog'
b. John khong co cho
John not has dog
'John has no dog(s)'

Number-neutrality is also evidenced by the fact that when a bare noun is used as a predicative, it is compatible with both a singular and a plural subject.

- (122) a. Pluto la cho
 Pluto is dog
 b. Pluto va Goofy la cho
 Pluto and Goofy are dog

The combination of a classifier and an NP, which we call a ‘classifier phrase’ (CLP), is interpreted as a singular definite in Vietnamese when it occurs in an argument position.

- (123) a. John mua mot con cho. Con cho rat dep.
 John bought one CL dog. CL dog very beautiful.
 ‘John bought a dog. The dog is very beautiful.’
 b. *John mua hai con cho. Con cho rat dep.
 John bought two CL dog. CL dog very beautiful.
 (‘John bought two dogs. The dogs are very beautiful.’)

When a CLP is used as a predicative, the subject must be singular.

- (124) a. Pluto la con cho
 Pluto is CL dog
 b. *Pluto va Goofy la con cho
 Pluto and Goofy are CL dog

A CLP can also be interpreted as a singular indefinite. However, this interpretation is limited to object position.

- (125) a. John muon mua con cho
 John want buy CL dog
 ‘John want to buy the/a dog’
 b. Con cho muon can John
 CL dog want bite John
 (‘The/*a dog wants to bite John’)

The combination of a numeral with a classifier phrase - a ‘Numeral Phrase’ (NumP) - is interpreted as a definite, an indefinite, or a predicative.

- (126) a. John da doc hai quyen sach noi-tieng nhat cua Chomsky
 John PERF read two CL book famous most by Chomsky
 ‘John has read the two most famous books by Chomsky’
 b. John co hai quyen sach, tham-chi ba
 John has two CL book, in fact three
 ‘John has two books, in fact three’

- c. Pluto va Goofy la hai con cho
 Pluto and Goofy are two CL dog
 ‘Pluto and Goofy are two dogs’

In pre-verbal position, NumP can only have the definite interpretation.

- (127) Hai quyen sach nam tren ban
 Two CL book lie on table
 ‘The two books lie on the table’ / *‘two books lie on the table’

From the above we can conclude that Chinese and Vietnamese resemble each other with regard to every aspect of nominal syntax and semantics save the expression of definiteness. Recall that in Chinese, NPs (bare nouns) can be definite, while classifier phrases (CLP) and numeral phrases (NumP) cannot. Vietnamese differs from Chinese in a rather bizarre way: it shows the exact opposite. Classifier and numeral phrases can be definite in Vietnamese, while bare nouns cannot.

Our account of this difference has two components. The first is the assumption that Chinese and Vietnamese differ with respect to lexical resource: instead of EXT, Vietnamese has THE, which is defined in (128).

- (128) $\llbracket \text{THE} \rrbracket^w \in D_{\langle \langle e,t \rangle, \langle e,t \rangle \rangle}$
 $\llbracket \text{THE X} \rrbracket^w = \text{MAX}(\llbracket \text{X} \rrbracket^w)$

This allows CLP’s and NumP’s to have the definite reading. The LF for (127b), for example, would be that in (129).

- (129)
-
- ```

 α
 / \
 β δ
 / \ / \
 THE γ thich an thit
 / \
 con cho

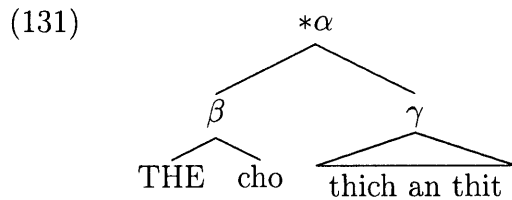
```

The second component of the account is a preference principle which says that when both K and THE can be used, i.e. when neither of them causes type mismatch, K must be used.<sup>31</sup>

<sup>31</sup>Chierchia (1998b) proposes the same preference of the kind operator over the definite article. Chierchia’s framework makes it possible to motivate this preference. The account developed here is incapable of this task. Thus, we will leave (130) as a primitive for the present.

- (130) Preference Principle  
Prefer K to THE!

The LF in (131) would then be ill-formed. This explains why bare nouns cannot be definite in Vietnamese.



## Conclusion

Research on how the mass count distinction plays out in different languages promises to inform our understanding of the relation between grammar, cognition and the physical world. Investigation of the contrast between number marking and classifier languages, and of the micro variation among languages of both types, should be of special relevance. A vast amount of work in the semantic literature has been devoted to the meaning of noun phrases in number marking languages. Analyses of classifier languages, however, have been fewer and less rigorous, and the micro variation between them has not received much attention. In this appendix, we attempt to take a small step toward eliminating this discrepancy: we present a set of facts concerning the distribution and interpretation of nominals in two classifier languages – Mandarin and Vietnamese – and derive these facts from precisely formulated assumptions. Our proposal builds entirely on suggestions that have been made in previous works. Thus, we contribute no “new idea.” Our aim is rather to show which old ideas can be selected – and explicated in certain ways – to capture the observations, and what implications this has for the parametric theory of language.

# Chapter 4

## Constraining headedness

### 4.1 Introduction

The last two chapters have used various sources of evidence to support the Edge Condition on Copy Deletion (ECCD), repeated below.

- (1) Edge Condition on Copy Deletion  
For any chain  $(\alpha, \beta)$  where  $\alpha$  is the higher and  $\beta$  the lower copy, deletion of  $\beta$  requires that  $\beta$  ends an XP

A perhaps curious feature of the ECCD is that the higher copy is totally irrelevant to the application of Copy Deletion. In other word, whether Copy Deletion applies to a chain  $(\alpha, \beta)$  is determined completely by the position of  $\beta$ . In this chapter, I will propose a slight revision of the ECCD which makes Copy Deletion sensitive to the position of the higher copy also. Recall that all the cases which I claimed follow from the ECCD have involved movement of a constituent, which is either an  $X^0$  or a (remnant) XP, to [Spec,C]. Thus, all cases have involved chains in which the higher copy is an XP. This means that the arguments which have been given to support the ECCD might as well have been arguments for the following principle, which differs minimally from the ECCD in one if-clause.

- (2) Edge Condition on Copy Deletion (Conditional)  
For any chain  $(\alpha, \beta)$  where  $\alpha$  is the higher and  $\beta$  the lower copy, deletion of  $\beta$  requires that  $\beta$  ends an XP, if  $\alpha$  ends an XP

The difference between the ECCD and the ECCD(C) is small but can be tricky. To see it clearly, let A stand for the proposition ‘ $\alpha$  ends and XP’ and B stand for the proposition ‘ $\beta$  ends an XP.’ What the ECCD says is that Copy Deletion

requires B. What the ECCD(C) says, in contrast, is that Copy Deletion requires  $A \rightarrow B$  (if A then B). Now suppose that A is true, then  $A \rightarrow B$  is true if and only if B is true. In other word, the truth of A makes B and  $A \rightarrow B$  equivalent. With respect to cases where the higher copy ends an XP, then, the ECCD and the ECCD(C) makes identical predictions: the lower copy must be XP-final in order to delete. I have argued that this prediction is born out, using sentences in which the higher copy is an XP, hence trivially ends an XP.

The last chapter saw a discussion of head-to-head movement. I have advocated the view that this operation takes place at PF and leaves no copy, thus forming no chain. Given the fact that head-to-head movement never results in doubling of the moved element, this position saves the ECCD from being contradicted by cases in which the base position of the moving head is not XP-final. What is PF-movement? I have assumed, simply, that it is displacement in the true sense of the word: it takes the relevant constituent away from its base position and puts it in its derived position.

$$(3) \quad [\dots \quad \dots X \dots] \quad \rightarrow \quad [\dots X \dots \quad \dots \underline{\quad} \dots]$$


In (3), the derived position is “created” by the movement itself, in the sense that X does not move into the position of any other constituent. Suppose, now, that it does: X moves into the position of some constituent Y, replacing Y.

$$(4) \quad [\dots Y \dots \quad \dots X \dots] \quad \rightarrow \quad [\dots X \dots \quad \dots \underline{\quad} \dots]$$


Obviously, if Y and X are made up of distinct elements in the Numeration, such an operation as that in (4) would not be tolerated, since it would amount to unrecoverable deletion of Y. But what if X and Y are not made up of distinct elements in the Numeration? What if, in other word, X and Y are two copies of the same constituent, generated by an earlier movement step in the syntax. Then there is no reason for X not being able to replace Y. Thus, I propose that the operation in (4) is allowed by grammar, if X and Y are copies of each other. PF-movement of X, then, is just displacement of X, either to a new position or to a position occupied by Y, thereby replacing Y.

This opens the possibility of analyzing overt syntactic movement as copying and PF-movement, instead of copying and deletion. For example, we can take movement of **what** in (5) to consist of (i) copying **what** into [Spec,C'] and (ii) replacing the higher copy of **what** with the lower copy, i.e. PF-movement of the lower copy of **what** to [Spec,C].

- (5) a. Step 1: copying (syntactic)  
 you saw what → what you saw what  
 b. Step 2: PF-movement  
 what you saw what → what you saw \_

The ECCD(C), then, can be reformulated as conditions on PF-movement of the “replacement” kind. Let us reformulate the ECCD(C) as follows.

- (6) Edge Condition on PF-movement (first version)  
 PF-movement of X from position A to position B occupied by a copy of X requires that A is XP-final, if B is XP-final

It is evident that all the arguments for the ECCD in the previous two chapters are, in effect, arguments for the principle in (6), once we conceptualize copy deletion as PF movement of the lower copy into the position of the higher one. In fact, those arguments can be considered arguments for (7), which is (6) generalized to all kinds of PF movement, not just the replacement kind.

- (7) Edge Condition on PF-movement (final version)  
 PF-movement of X from position A to position B requires that A is XP-final, if B is XP-final

The difference between the first and the second version of the Edge Condition on PF-movement is, descriptively, that while the former imposes  $A \rightarrow B$  on non-doubling syntactic movement only, the latter imposes the same on both non-doubling syntactic movement and head-to-head movement, which by hypothesis is PF-movement. As an example, consider the scenario in (8), where a head undergoes head-to-head movement from a non-XP-final position to an XP-final position.

- (8)
- 

The first version would allow this movement, since it is not movement to a position occupied by a copy of  $X^0$ . The second version, on the other hand, would rule out the scenario in (8), because it requires of all instances of PF-movement that the lower position is XP-final if the higher one is. This chapter argues that we

should adopt the second version of the Edge Condition on PF-movement, which I will henceforth call the Edge Condition. The argument is based on a specific interpretation of a set of typological data presented in Biberauer et al. (2010). I turn to this work in the next section.

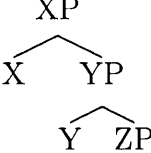
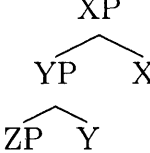
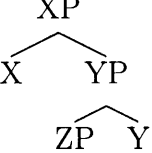
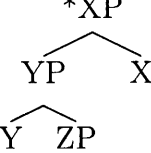
## 4.2 The Final-Over-Final Constraint

### 4.2.1 Biberauer et al (2010)

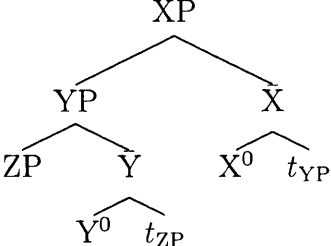
The following constraint on phrase structure configurations is introduced in Biberauer et al. (2010: 3).

- (9) The Final-Over-Final Constraint (first version)  
 If  $\alpha$  is a head-initial phrase and  $\beta$  is a phrase immediately dominating  $\alpha$ , then  $\beta$  must be head-initial. If  $\alpha$  is a head-final phrase, and  $\beta$  is a phrase immediately dominating  $\alpha$ , then  $\beta$  can be head-initial or head-final.

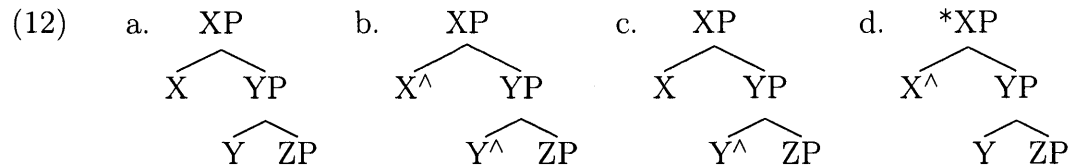
What the Final-Over-Final Constraint (FOFC) says, basically, is that among the four possibilities in (10), one of them, (10d), is not attested in natural language.

- (10) a.  b.  c.  d. 

Biberauer et al. (2010) make the Kaynean assumption that syntactic structure is by default Spec-Head-Complement, which is to say that phrases are head-initial in the base form. Head-final configurations are the result of “roll up movement,” i.e. movement of the complement to the specifier position of the governing head. Thus, (10b) is to be considered shorthand description of (11), for example.

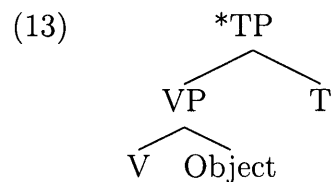
- (11) 

Biberauer et al. postulate a feature which, on a head, requires the complement of that head to become its specifier, i.e. to undergo roll up movement. They designate this feature with  $\wedge$ . Given this feature, the structures in (10) can be taken to have the following underlying form.



We see that FOFC can be reformulated as follows: if a head X bears  $\wedge$ , then the head Y governed by X must bear  $\wedge$  also.

Biberauer et al. (2010) draw the strongest support for the FOFC from data concerning the relative order of T, whose position is indicated by the finite auxiliary (Aux), the main verb (V), and the complement of the main verb (O), in the following languages: German, Dutch, Afrikaans, West Flemish, Zürich German, Old English, Modern English, Middle Dutch, Old High German, Old Norse, Finnish, Latin, and Basque. What they claim to have established is that “[i]f we consider the three elements Aux, V and O, we find all possible permutations of these, with one very striking exception: the order V-O-Aux is not found.” The evidence these authors provide to back this claim is, in my opinion, quite strong, and I will not dispute what they claim is observed among these languages. What I will propose, however, is an interpretation of their result which involves going a bit beyond pure observation of surface word order. It is this: there is no language among the ones just mentioned in which TP is head-final and VP is head-initial.



It is perhaps not surprising that the FOFC, as formulated in (9), will find exceptions. A standard sentence of German such as (14) already provides us with two counter-examples.

- (14)    a.    dass [TP [DP der Mann] [VP [DP das Buch] gelesen] wird]  
                   that            the man                    the book read        will  
                   ‘that the man will read the book’

In (14), TP and VP are head-final, but they immediately dominate head-initial DPs, i.e. the subject and the object. German PPs are also head-initial, thus also posing a problem for the FOFC when they are complements of V.

- (15) Hans ist [VP [PP nach Berlin] gefahren]  
Hans is to Berlin gone  
‘Hans has gone to Berlin’

Dutch presents the same problem as German, as this language shows the same headedness for the categories as German. Several other challenging facts are acknowledged by Biberauer et al. themselves. Thus, nominal complexes in many languages show DemP, i.e. the phrase which is headed by demonstratives and which is presumably highest in the nominal complex, being head-final and at the same time dominating head-initial phrases such as DP or NP. Examples from Scottish Gaelic and Vietnamese are given in (16a) and (16b).

- (16) a. na trì leabhraichean mòra seo  
the three books big these  
b. ba guyen sach to nay  
three CL book big these

Exceptions to the FOFC are, as pointed out by Biberauer et al., also found in the domain of word formation. Consider a word such as **beheading**. Presumably, it has the following structure.

- (17)
- ```

      ing
     /  \
    be   ing
   /  \
  be  head

```

These and other cases which contradict the FOFC, as formulated in (9), force Biberauer et al. to qualify the FOFC and fine-tune the definition of the terms used in this qualification. The end result is a rather elaborate theory of phrase structure, which dictates that every node in a syntactic tree must be a node in the “Extended Projection” of a lexical head, defined as follows.

- (18) The Extended Projection of a lexical head L (EP(L)) is the sequence of categories $EP = \{\alpha_1 \dots \alpha_i \dots \alpha_n\}$ such that:
- (i) α_i is in the spine defined by L
 - (ii) for each pair of heads $\langle H_i, H_{i+1} \rangle$ in EP, H_i c-selects H_{i+1}
 - (iii) each pair of heads $\langle H_i, H_{i+1} \rangle$ is RM-compliant in formal features

We need not go into details of this definition. It amounts, in my view, to a circuitous way of saying this: among C, T and V, a head with $\wedge$ forces the one immediately below it to have $\wedge$ also. Thus, head-final TP and VP are allowed to be on top of head-initial DP or PP, and word-internal constituency, as well as structure inside nominal complexes, are exempted, so to speak. The counter-examples discussed above are hence “accounted for.” Biberauer et al. include C because they claim, on the basis of data from several languages, to have shown that there is no head-final CP dominating head-initial TP. However, throwing C into the mix causes problem, again with German. Suppose, for example, that we have the following base structure, where T c-selects VP and V c-selects CP.

$$(19) \quad (\text{dass}) \left[\text{TP} \dots \text{T} \left[\text{VP} \dots \text{V} \left[\text{CP} \dots \text{C} \left[\text{XP} \dots \right] \right] \right] \right]$$

Since TP ends up being head-final in German, T must bear $\wedge$, which means V must bear $\wedge$, which means C must also. The prediction is then that the CP complement will be head-final, which is not correct: CP is always head-initial in German. Biberauer et al. solve this problem by postulating that clausal complements are actually DP with a null D head.

$$(20) \quad \left[\text{VP} \text{ V} \left[\text{DP} \text{ D} \left[\text{CP} \text{ C XP} \right] \right] \right]$$

Because DP is not headed by C, T or V, the presence of $\wedge$ does not force D to bear this feature, and consequently the C which heads the complement of D does not have to either. Suppose V bears $\wedge$, then DP will undergo roll up movement to [Spec,V], forming the chain (DP, DP). Biberauer et al. then propose that it is the CP in the trace position which is spelled out.

$$(21) \quad \left[\text{DP} \text{ D CP} \right] \left[\bar{\text{V}} \text{ V} \left[\text{DP} \text{ D CP} \right] \right]$$

The order V-CP, with CP being head-initial, is then derived. This solution, upon closer inspection, turns out to be inadequate. It entails that CP complements are spelled out in the complement position of V, which precedes T. But this is ungrammatical in German.

- (22) a. dass Hans glauben wird dass Maria kommt
 that Hans believe will that Maria comes
 b. *dass Hans glauben dass Maria kommt wird
 that Hans believes that Maria comes will

Thus, Biberauer et al. will have to assume that the DP complement undergoes two movement operations from the same position: roll up movement to [Spec,V] and extraposition to the right of T. And it will be the CP inside the higher

copy of the chain created by the extraposition that is spelled out. Alternatively, Biberauer et al. must assume that the CP inside the lower copy of the chain created by the roll up movement extraposes out of the containing DP to the right of T. While all of these options are in principle available, it seems to me that they are more or less stipulations that have to be made to fix the problems caused by Biberauer et al.'s definition of Extended Projection.

The inadequacy of this definition is also illustrated by discourse particles, which are standardly analyzed as C heads (cf. Cheng (1991)). In many languages which are head-initial with respect to TP and VP, question particles appear sentence-finally. Examples from Chinese and Vietnamese are given in (23a) and (23b), respectively.

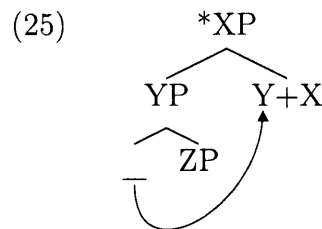
- (23) a. No se doc sach a?
 he will read book Q
 b. Ni yao kan she ben shu ma?
 you want read this CL book Q

4.2.2 The Head Ordering Generalization

I claim that if we examine the data in Biberauer et al. (2010) closely, we will be lead to the following generalization.

- (24) The Head Ordering Generalization (HOG)
 If there is head-to-head movement of Y to X in language L, then $[_{XP} \dots [_{YP} \dots] \dots]$ obeys the FOFC in L. Equivalently, if $[_{XP} \dots [_{YP} \dots] \dots]$ does not obey the FOFC in L, then there is no head-to-head movement of Y to X in L

What the HOG says is that the following scenario is ruled out by grammar.



In other word, FOFC turns out to be a description of structures which involve head-to-head movement. This is why Biberauer et al.'s strongest evidence has to do with the order of C, T, V in languages with V-to-T and T-to-C movement. As soon as we move away from cases where there is head-to-head movement,

we begin to find counter-examples in abundance. Thus, there is no head-to-head movement from DP, CP or PP complements to V, so head-final VP can dominate head-initial DP, CP or PP. The same holds for TP and the DP subject. Similarly, there is no head-to-head movement from T to C in isolating languages like Chinese or Vietnamese, so head-final CP can dominate head-initial TP. In fact, if head-to-head movement is driven by inflectional morphology, we might entertain the hypothesis that there is no head-to-head movement in Chinese and Vietnamese. This means that the order V-O-Aux, which Biberauer et al. claim is non-existent, is in principle possible. And it turns out that this order is attested in these languages, as exemplified by (26a) (Chinese) and (26b) (Vietnamese).

- (26) a. John kan shu le
 John read book has
 ‘John has read books’
 b. John doc sach duoc
 John read book can
 ‘John can read books’

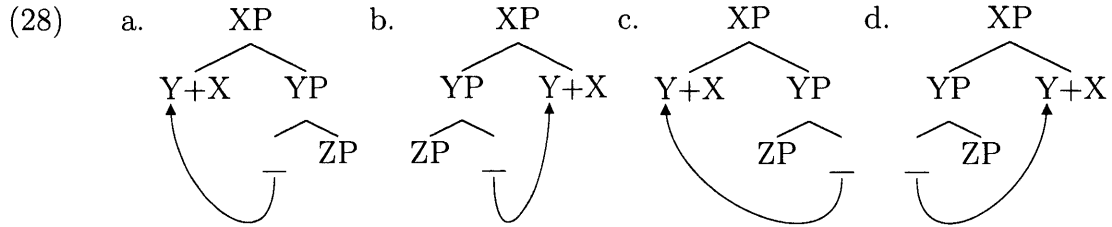
In (26a), the main verb precedes the object, while the perfect auxiliary follows the object resulting in the V-O-Aux order. In (26b), the same situation obtains but with the possibility modal verb instead. As far as I can see, Biberauer et al. are forced to say that the auxiliary in these two sentences “extrapose” to the right, which is extremely unlikely.

4.2.3 Deriving the HOG

Suppose we say that XP-finality is word-based. More precisely, if α is part of a word which is XP-final, then α counts as XP-final.

- (27) Definition
 α ends an XP, i.e. is XP-final, iff the last morpheme of α is part of the last word of an XP

Thus, Y in (26) would be XP-final, or to use another expression, would end an XP. From this assumption, the HOG follows straightforwardly. We have said that head-to-head movement is PF-movement. Suppose a head H moves from position A to position B. The Edge Condition dictates that A must be XP-final if B is XP-final. There are four possibilities: (i) Neither A nor B are XP-final, (ii) Both A and B are XP-final, (iii) A but not B is XP-final, and (iv) B but not A is XP-final. These are illustrated in (28a-d).



As we can see, the last possibility constitutes a contradiction of the Edge Condition: the higher position is XP-final but the lower position is not. This possibility is hence ruled out. Thus, we derive the HOG, thereby explaining why the exceptions to Biberauer et al.'s FOFC are found where they are found. Moreover, the principle which gives this result also explains several facts about predicate cleft and NP-split constructions, as shown in the previous two chapters.

There is another way to account for the HOG, suggested to me by Hedde Zeijlstra, who came to the same conclusion as I did about Biberauer et al.'s data (Zeijlstra personal communication). Specifically, Zeijlstra also takes these data to show that FOFC is obeyed whenever head-to-head movement is involved. However, he proposes that we account for this by assuming that there is no movement to the right, as hypothesized in Abels and Neeleman (2006). This approach would rule out not just (28d), but also (28b). Thus, it derives something stronger than the HOG, call it the HOG⁺.

- (29) The HOG⁺
 If there is head-to-head movement of Y to X in language L, then [XP ... [YP ...] ...] obeys the FOFC⁺ in L. Equivalently, if [XP ... [YP ...] ...] does not obey the FOFC⁺ in L, then there is no head-to-head movement of Y to X in L

The notion FOFC⁺ used in the formulation of HOG⁺ would be defined as follows. The wording is chosen in such a way as to highlight the difference between the FOFC and the FOFC⁺, which is also underlined.

- (30) The FOFC⁺
 If α is a head-initial phrase and β is a phrase immediately dominating α , then β must be head-initial. If α is a head-final phrase, and β is a phrase immediately dominating α , then β must also be head-initial.

An argument against this account and in favor of mine would involve showing a case of head-to-head movement of an XP-final position to another XP-final position. Such a case, I believe, is found in German. In fact, standard analyses

of German takes both VP and TP to be head-final, and posits V-to-T movement (cf. Fanselow and Felix 1987). Specifically, the clause in (31a) is given the analysis in (31b).

- (31) a. dass Hans Maria liebt
that Hans Maria loves
- b.
-
- ```

graph TD
 CP --> dass
 CP --> TP
 TP --> Hans
 TP --> T
 T --> VP
 T --> liebtT["liebt+T"]
 VP --> t1["t"]
 VP --> V
 V --> Maria
 V --> t2["t"]
 t2 --> liebtT
 t1 --> Hans

```

However, a case can be made, as pointed out by Zeijlstra (p.c.), that the finite verb is actually inside VP and that T has lowered to V, as in (32).

- (32)
- 
- ```

graph TD
    CP --> dass
    CP --> TP
    TP --> Hans
    TP --> T
    T --> t1["t"]
    T --> VP
    VP --> t2["t"]
    VP --> V
    V --> Maria
    V --> liebtT["liebt+T"]
    t2 --> liebtT
    t1 --> Hans
    liebtT --> T
  
```

This analysis faces difficulties when it comes to sentences in which the finite verb is clause-initial, such as the yes/no question in (33).

- (33) liebt Hans Maria?
loves Hans Maria

‘Does Hans love Maria?’

Assuming that **liebt** is in C will force us to assume that this verb moves directly from V to C, by passing T and thereby violating the Head Movement Constraint (Travis 1984). Assuming that **liebt** is in T entails that **Hans** is in [Spec,V], hence nothing is in [Spec,T]. However, there is reason to assume that [Spec,T] in German must be filled, just as in English. Thus, when there is no thematic subject, an expletive must be inserted.

- (34) a. *Regnet
 rains
 b. Es regnet
 it rains

Suppose we get rid of the HMC and allow V-to-C movement, then there is still a data point which speaks against head-initial TP in German, namely the distribution of the future auxiliary **werden**. It is a characteristic of this auxiliary that it is always tensed. If **werden** appears in an untensed position, say after **zu**, it must be construed as the verb ‘to become,’ or as the passive auxiliary ‘be.’

- (35) Ich verspreche, gut/gesehen/*kommen zu werden
 I promise good/seen/*come ZU WERDEN
 ‘I promise that I will become good/be seen/*come’

Thus, there is good reason to assume that future **werden** is a T head. This element also occupies the clause-initial position in yes/no questions, suggesting that it moves to C from T.

- (36) Wird Hans Maria küssen?
 will Hans Maria kiss

The crucial evidence for V-to-T movement, with TP being head-final, in German comes from two facts: (i) when C is occupied, **werden** is clause-final, and (ii) **werden** is in complementary distribution with finite verbs.

- (37) a. dass Hans Maria küssen wird
 that Hans Maria kiss.INF werden.3SG
 b. dass Hans Maria küsst wird
 that Hans Maria kiss.3SG werden.3SG

These facts follow straightforwardly from the assumption that TP is head-final in German. On the other hand, they are quite puzzling under the assumption that TP is head-initial in German.

4.3 Conclusion

This chapter proposes a slight revision of the Edge Condition on Copy Deletion: it adds an if-clause to it. In addition, reformulates this condition as one which is imposed on PF-movement, given the hypothesis, also advanced in the chapter, that overt syntactic movement consists in copying and PF-movement of the lower copy into the position of the higher one, replacing the latter. The change has no effect on how the ECCD, which is called the Edge Condition in its new formulation, accounts for the facts concerning predicate cleft NP-split constructions in chapter 2 and 3.

It is then argued that the Final-Over-Final Constraint introduced in Biberauer et al. (2010) finds several counter-examples, and that the attempt made in Biberauer et al. (2010) to fix these problems falls short of achieving its objective. A new generalization, the HOG (Head Ordering Generalization), is proposed which qualifies the FOFC in such a way that the counter-examples are excluded. It is then shown that the HOG is derivable from the Edge Condition.

Bibliography

- Abels, K., and A. Neeleman. 2006. Universal 20 without the lca. Ms., University of Tromsø and University College London .
- Abels, Klaus. 2001. The predicate cleft construction in Russian. In Annual Workshop on Formal Approaches to Slavic Linguistics: The Bloomington meeting, 2000, 1. Univ of Michigan/Michigan Slavic.
- Aboh, Enoch Olad, and Marina Dyakonova. 2009. Predicate doubling and parallel chains. Lingua 119:1035–1065.
- Abrusán, Martha. 2007. Contradiction and Grammar: the Case of Weak Islands. Doctoral Dissertation, MIT.
- Aikhenvald, Alexandra. 2000. Classifiers: a typology of noun categorization devices. Oxford University Press.
- Baker, C. L., and Michael Brame. 1972. “Global Rules”: A rejoinder. Language 48:51–75.
- Barker, C. 1992. Group terms in english: representing groups as atoms. Journal of Semantics 9:69.
- Barss, Andrew. 1986. Chains and anaphoric dependence: On reconstruction and its implications. Doctoral Dissertation, Massachusetts Institute of Technology.
- den Besten, Hans, and Gert Webelhuth. 1987a. Adjunction and remnant topicalization in the Germanic SOV-languages. In Proceedings of GLOW 42.
- den Besten, Hans, and Gert Webelhuth. 1987b. Adjunction and remnant topicalization in the Germanic SOV-languages. In GLOW 42.
- den Besten, Hans, and Gert Webelhuth. 1990. Stranding. In Scrambling and barriers, ed. Günther Grewendorf and Wolfgang Sternefeld, 77–92. Amsterdam: John Benjamins Publishing Company.

- Biberauer, Theresa, Anders Holmberg, and Ian Roberts. 2010. A syntactic universal and its consequences. A revised version will appear in *Linguistic Inquiry*.
- Boeckx, Cedric, and Sandra Stjepanović. 2001. Head-ing toward PF. *Linguistic Inquiry* 32:345–355.
- Büring, Daniel. 1999. Topic. In *Focus — linguistic, cognitive, and computational perspectives*, ed. Peter Bosch and Rob van der Sandt, 142–165. Cambridge University Press.
- Büring, Daniel. 2003. On d-trees, beans, and b-accents. *Linguistics and Philosophy* 26:511–545.
- Cao, Xuan Hao. 1988. The count/mass distinction in vietnamese nouns and the concept of "classifier". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 41:38–47.
- Cao, Xuan Hao. 1992. Some preliminaries to the syntactic analysis of the vietnamese sentence. *Mon-Khmer Studies* 20:137–151.
- Carlson, G.N. 1977. A unified analysis of the english bare plural. *Linguistics and philosophy* 1:413–457.
- Cheng, L., and R. Sybesma. 1998. Yi-wan tang, yi-ge tang: Classifiers and massifiers. *Tsing Hua Journal of Chinese Studies, New Series* 28:385–412.
- Cheng, Lisa Lai-Shen. 1991. On the typology of wh-questions. Doctoral Dissertation, Massachusetts Institute of Technology.
- Cheng, Lisa Lai-Shen. 2007. Verb copying in Mandarin Chinese. In *The copy theory of movement*, ed. Norbert Corver and Jairo Nunes, 151–174. John Benjamins Publishing Company.
- Cheng, Lisa Lai-Shen, and Rint Sybesma. 1999a. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30:509–542.
- Cheng, Lisa Lai Shen, and Rint Sybesma. 2005a. Classifiers in four varieties of Chinese. In *The Oxford handbook of comparative syntax*, ed. Guglielmo Cinque and Richard Kayne, 259–292. Oxford: Oxford University Press.
- Cheng, L.L.S., and R. Sybesma. 1999b. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30:509–542.
- Cheng, L.L.S., and R. Sybesma. 2005b. *Classifiers in four varieties of Chinese*, 259–292. Oxford University Press.

- Chierchia, G., and R. Turner. 1988. Semantics and property theory. Linguistics and Philosophy 11:261–302.
- Chierchia, Gennaro. 1998a. Plurality of mass nouns and the notion of “semantic parameter”. In Events and grammar, ed. Susan Rothstein, 53–104. Kluwer Academic Publishers.
- Chierchia, Gennaro. 1998b. Reference to kinds across languages. Natural Language Semantics 6:339–405.
- Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. Synthese 174:99–149.
- Chomsky, Noam. 1955. Logical Structure of Linguistic Theory. Manuscript. Harvard University and MIT.
- Chomsky, Noam. 1957. Syntactic structures. Mouton.
- Chomsky, Noam. 1964. A transformational approach to syntax. In The structure of language, ed. Jerry Fodor and Jerry Katz, 211–245. Prentice Hall.
- Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, Massachusetts: M.I.T. Press.
- Chomsky, Noam. 1971. Conditions on transformations. The Indiana University Linguistics Club, Bloomington, Indiana .
- Chomsky, Noam. 1972. Studies on Semantics in Generative Grammar. Mouton.
- Chomsky, Noam. 1975. Reflections on language. New York: Pantheon Books.
- Chomsky, Noam. 1976. Conditions on rules of grammar. Linguistic analysis 2:303–351.
- Chomsky, Noam. 1977. On wh-movement. In Formal syntax, ed. P. Culicover, T. Wasow, and A. Akmajian. New York, New York: Academic Press.
- Chomsky, Noam. 1981. Lectures on government and binding. Dordrecht, The Netherlands: Foris Publications.
- Chomsky, Noam. 1982. Some concepts and consequences of the theory of government and binding. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 1986. Barriers. Cambridge, Massachusetts: MIT Press.

- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In The view from Building 20, ed. Ken Hale and Jay Keyser, 1–52. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 1994. Bare phrase structure. MIT Working Papers in Linguistics .
- Chomsky, Noam. 1995. Categories and Transformations. In The minimalist program, chapter 4. MIT Press.
- Chomsky, Noam. 2000. New horizons in the study of mind and language. Cambridge University Press, Cambridge.
- Chomsky, Noam. 2001. Derivation by phase. In Ken Hale: A Life in Language, ed. Michael Kenstowicz, 1–52. MIT Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. In Structures and beyond: The cartography of syntactic structures, ed. Adriana Belletti, volume 3, 104–131. Oxford: Oxford University Press.
- Chung, S., and W.A. Ladusaw. 2004. Restriction and saturation. The MIT Press.
- Corver, Norbert, and Jairo Nunes, ed. 2007. The Copy Theory of Movement. John Benjamins Publishing Company.
- Dayal, Veneeta. 2004. Number marking and (in)definiteness in kind terms. Linguistics and Philosophy 27:393–450.
- Diesing, M. 1992a. Indefinites. MIT Press.
- Diesing, Molly. 1992b. Indefinites. Cambridge, Massachusetts: MIT Press.
- Emonds, Joseph. 1964. Root and Structure Preserving Transformation. Doctoral Dissertation, MIT.
- Engels, Eva, and Sten Vikner. 2009. Scandinavian Object Shift, remnant VP-topicalisation and Optimality Theory. Manuscript, SyntaxLab, University of Cambridge.
- Erteschik-Shir, Nomi. 1977. On the nature of island constraints. Doctoral Dissertation, Massachusetts Institute of Technology.
- Erteschik-Shir, Nomi. 2001. P-syntactic motivation for movement: imperfect alignment in Object Shift. Working Papers in Scandinavian Syntax 68:49–73.
- Fanselow, Gisbert. 2001. Partial movement. University of Potsdam .

- Fanselow, Gisbert. 2002. Against remnant VP-movement. In Dimensions of movement: From features to remnants, ed. A. Alexiadou, E. Anagnostopoulou, S. Barbiers, and H.-M. Gaertner, 91–125. Benjamins.
- Fanselow, Gisbert, and Damir Ćavar. 2002. Distributed deletion. In Theoretical approaches to universals, ed. Artemis Alexiadou, 65–107. John Benjamins Publishing Company.
- Fanselow, Gisbert, and Sascha Felix. 1987. Sprachtheorie 2: Die Rektions-und Bindungstheorie. Francke.
- Fanselow, Gisbert, and Anoop Mahajan. 1995. Partial movement and successive cyclicity. In Papers on Wh-Scope Marking, ed. Uli Lutz and Gereon Müller, 131–161. University of Tübingen: Sprachtheoretische Grundlagen für die Computer Linguistik.
- Fiengo, Robert. 1974. Semantic Conditions of Surface Structure. Doctoral Dissertation, Massachusetts Institute of Technology.
- Fiengo, Robert. 1977. On trace theory. Linguistic Inquiry 8:35–61.
- Fox, D., and D. Pesetsky. 2005. Cyclic linearization of syntactic structure. Theoretical linguistics 31:1–45.
- Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. Linguistic Inquiry 30:157–196.
- Fox, Danny. 2000. Economy and Semantic Interpretation. Cambridge, Massachusetts: MIT Press.
- Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. Linguistic Inquiry 33:63–96.
- Fox, Danny. 2003. On logical form. In Minimalist syntax, ed. Randall Hendrick, 82–123. Oxford: Blackwell Publishers.
- Gajewski, J. 2003. L-analyticity in natural language. Manuscript, MIT .
- Gärtner, Hans-Martin. 1998. Review of “the copy theory of movement and linearization of chains in the Minimalist Program”. GLoT International 8:16–20.
- Goldberg, Lotus Madelyn. 2005. Verb-stranding vp ellipsis: a cross linguistic study. Doctoral Dissertation, McGill University.

- Grodzinsky, Y., and L. Finkel. 1998. The neurology of empty categories: Aphasics' failure to detect ungrammaticality. Journal of Cognitive Neuroscience 10:281–292.
- Grohmann, Kleanthes. 2003. Prolific domains: On the anti-locality of movement dependencies. John Benjamins.
- Grohmann, Kleanthes, and Andrew Nevins. 2004. Echo reduplication: When too-local movement requires pf-distinctness. University of Maryland Working Papers in Linguistics 13:84–108.
- Grohmann, Kleanthes, and Phoevos Panagiotidis. 2004. Demonstrative doubling in Greek. University of Maryland Working Papers in Linguistics 13:109–131.
- Hale, Ken, and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. MIT Press.
- Hale, Ken, and Samuel Jay Keyser. 2002. Prolegomenon to a theory of argument structure. MIT Press.
- Hamblin, C. 1973. Questions in Montague English. Foundations of Language 41–53.
- Hartman, J. 2011. The semantic uniformity of traces: Evidence from ellipsis parallelism. Linguistic Inquiry 42:367–388.
- Heim, I.R. 1982. The semantics of definite and indefinite noun phrases. Doctoral Dissertations, UMass Amherst .
- Heim, Irene, and Angelica Kratzer. 1998. Semantics in Generative Grammar. Blackwell.
- Hinterhölzl, Roland. 2002. Remnant movement and partial deletion. In Dimensions of movement: From features to remnants, ed. A. Alexiadou, E. Anagnostopoulou, S. Barbiers, and H.-M. Gaertner, 127–151. Benjamins.
- Hiraiwa, Ken. 2005. Predicate clefts in Bùlì: Phase and category. Linguistic Analysis 32:544–583.
- Holmberg, Anders. 1999. Remarks on holmberg's generalization. Studia Linguistica 53:1–39.
- Holmberg, Anders. 2009. V-movement in affirmative replies and copy deletion. Theoretical Linguistics 269–276.

- Iatridou, S. 1990. About agr(p). Linguistic Inquiry 551–577.
- Ionin, T., and O. Matushansky. 2006. The composition of complex cardinals. Journal of Semantics 23:315.
- Jackendoff, Ray. 1972. Semantic Interpretation in Generative Grammar .
- Johnson, Kyle, and Satoshi Tomioka. 1999. Lowering and Mid-Size Clauses. Proceedings of the 1997 Tübingen Workshop on Reconstruction .
- Kadmon, N. 1990. Uniqueness. Linguistics and Philosophy 13:273–324.
- Kandybowicz, J. 2007. Fusion and pf architecture. In Proceedings of the 30th Annual Penn Linguistics Colloquium, volume 13, 85–98.
- Kandybowicz, Jason. 2006. On fusion and multiple copy spell-out: The case of verb repetition. In The copy theory of movement on the PF side, ed. Jairo Nunes and Norbert Corver. Oxford University Press.
- Kandybowicz, Jason. 2008. The grammar of repetition: Nupe grammar at the syntax-phonology interface. John Benjamins Publishing Co.
- Kandybowicz, Jason. 2009. Embracing edges: syntactic and phono-syntactic edge sensitivity in Nupe. Natural Language and Linguistic Theory 27:305–344.
- Kayne, R. 1998. Overt versus covert movement. Syntax 1:128–191.
- King, Harold V. 1970. On blocking rules for contraction in English. Linguistic Inquiry 1:134–136.
- Koizumi, M. 1993. Object agreement phrases and the split vp hypothesis. MIT working papers in linguistics 18:99–148.
- Koizumi, Masatoshi. 1995. Phrase structure in minimalist syntax. Doctoral Dissertation, Massachusetts Institute of Technology.
- Koopman, Hilda. 1984. The syntax of verbs: From verb movement rules in the kru languages to universal grammar. Foris.
- Koopman, Hilda. 2000a. The Syntax of Specifiers and Heads: Collected Essays of Hilda j. Koopman. Routledge.
- Koopman, Hilda. 2000b. Unifying predicate cleft constructions. In The Syntax of Specifiers and Heads: Collected Essays of Hilda j. Koopman, 357–374. Routledge.

- Kratzer, Angelika. 1989. An investigation of the lumps of thought. Linguistics and philosophy 12:607–653.
- Krifka, M. 1987. An outline of genericity. Partly in collaboration with C. Gerstner. SNS-Bericht, Universität Tübingen .
- Krifka, M. 1989. Nominalreferenz und zeitkonstitution: Zur semantik von massentermen, pluraltermen und aspektklassen. W. Fink.
- Krifka, M. 1995. Common nouns: a contrastive analysis of english and chinese. The generic book 398–411.
- Krifka, M. 1998. The origins of telicity. Events and grammar 197:235.
- Krifka, M. 1999. Mass expressions. Concise Encyclopedia of Grammatical Categories. Amsterdam: Elsevier 221–223.
- Krifka, M. 2007. Basic notions of information structure. Interdisciplinary Studies of Information Structure 6:13–56.
- Krifka, M., F.J. Pelletier, G.N. Carlson, A. ter Meulen, G. Chierchia, and G. Link. 1995. Genericity: an introduction. The generic book 1–124.
- Krifka, Manfred. 2003. Bare NPs: Kind-referring, indefinites, both, or neither? In Proceedings of SALT XIII, ed. Robert B. Young and Yuping Zhou, 180–203. Cornell University, Ithaca, NY: CLC Publications.
- Krifka, Manfred. 2008. Different kinds of count nouns and plurals. Handout presented at “Syntax in the World’s Languages III,” FU Berlin .
- Labov, W. 1969. Contraction, deletion, and inherent variability of the english copula. Language 45:715–762.
- Laka, I. 1989. Constraints on sentence negation. MIT working papers in linguistics 10:199–216.
- Laka, Itziar. 1990. Negation in syntax. On the nature of functional categories and projections. Doctoral Dissertation, MIT.
- Lakoff, George. 1970. Global rules. Language 46:627–639.
- Landau, Idan. 2006. Chain resolution in Hebrew V(P)-fronting. Syntax 9:32–66.
- Landau, Idan. 2007. Constraints on partial VP-fronting. Syntax 10:127–164.
- Landman, F. 1989. Groups, I. Linguistics and philosophy 12:559–605.

- Lasnik, H. 1981. Restricting the theory of transformations: A case study. Explanation in linguistics 152–173.
- Lasnik, H. 1999a. On feature strength: Three minimalist approaches to overt movement. Linguistic Inquiry 30:197–217.
- Lasnik, Howard. 1995. A note on pseudogapping. MIT Working Papers in Linguistics 27:143–163.
- Lasnik, Howard. 1999b. Pseudogapping puzzles. In Fragments: Studies in ellipsis and gapping, ed. Shlom Lappin and Elabbas Benmamoun, 141–174. Oxford: Oxford University Press.
- Lechner, Winfried. 2006. An interpretive effect of Head Movement. In Phases of Interpretation, ed. Mara Frascarelli. Mouton de Gruyter.
- Lee, T.H. 1986. Studies on quantification in Chinese .
- Li, C.N., and S.A. Thompson. 1989. Mandarin chinese: A functional reference grammar. Univ of California Pr.
- Li, Y.A. 1998. Argument determiner phrases and number phrases. Linguistic Inquiry 29:693–702.
- Link, G. 1983a. The logical analysis of plural and mass nouns: a lattice theoretic approach. Meaning, use and interpretation of language, de Gruyter, Berlin 302–323.
- Link, Godehard. 1983b. The logical analysis of plural and mass nouns: a lattice theoretic approach. Meaning, Use and Interpretation of Language 21:302–323.
- Lyons, John. 1977. Semantics. Cambridge University Press.
- Martins, Ana Maria. 2007. Double realization of verbal copies in european portuguese emphatic affirmation, 77–118. John Benjamins.
- Matushansky, Ora. 2006. Head movement in linguistic theory. Linguistic Inquiry 37:69–109.
- McCloskey, James. 1997. Subjecthood and subject positions. In Elements of grammar, ed. Liliane Haegeman. Kluwer Academic Publishers.
- Montague, R. 1973. The proper treatment of quantification in ordinary English. Approaches to natural language 49:221–242.

- Müller, Gereon. 1998. Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German. Kluwer.
- Müller, Gereon, and Wolfgang Sternefeld. 1996. A-bar chain formation and economy of derivation. Linguistic Inquiry 27:480–512.
- Muysken, Pieter. 1982. Parametrizing the notion ‘head’. Journal of Linguistic Research 2:57–75.
- Nunes, Jairo. 2003. Head movement, remnant movement, and phonetic realization of chains, 161–177. UCLA.
- Nunes, Jairo. 2004. Linearization of chains and sideward movement. Linguistic Inquiry Monographs. Cambridge, Massachusetts: MIT Press.
- Pelletier, F.J., and L.K. Schubert. 1989. Mass expressions. Handbook of philosophical logic 4:327–407.
- Penka, D. 2002. Kein muss kein problem sein. MA, Universität Tübingen.
- Penka, D., and A. von Stechow. 2001. Negative indefinita unter modalverben. Modalität und Modalverben im Deutschen 263–286.
- Pesetsky, D. 1998a. Some optimality principles of sentence pronunciation. Is the best good enough? Optimality and competition in syntax 337–383.
- Pesetsky, David. 1998b. Phrasal Movement and Its Kins. MIT Press.
- Pierrehumbert, Janet. 1980. The phonetics and phonology of english intonation. Doctoral Dissertation, MIT.
- Quine, W.V.O. 1960. Word and object. The MIT Press.
- Ren, C.Y. 1968. A grammar of spoken Chinese. University of California Press.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Elements of grammar, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer.
- Roberts, C. 2003. Uniqueness in definite noun phrases. Linguistics and Philosophy 26:287–350.
- Rooth, Mats. 1985. Association with focus. Doctoral Dissertation, UMASS-Amherst.
- Rooth, Mats. 1992. A Theory of Focus Interpretation. Natural Language Semantics 1:75–116.

- Rooth, Mats. 1996. Focus. In The handbook of contemporary semantic theory, ed. Shalom Lappin, 271–297. Oxford: Blackwell Publishers.
- Ross, John. 1967. Constraints on Variables in Syntax. Doctoral Dissertation, MIT.
- Rullmann, H., and A. You. 2006. General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. Klaus v. Heusinger et al.(eds.), Where semantics meets pragmatics. Elsevier, Amsterdam .
- Saito, Mamoru, Jonah T.-H. Lin, and Keiko Murasugi. 2008. N'-ellipsis and the structure of noun phrases in Chinese and Japanese. Journal of East Asian Linguistics 17:247–271.
- Sauerland, Uli. 1998. The Meaning of Chains. Doctoral Dissertation, MIT.
- Sauerland, Uli. 2003. A New Semantics for Number. In Proceedings of SALT XIII, ed. Robert Young and Yuping Zhou. Cornell University, Ithaca, NY: CLC Publications.
- Sauerland, Uli. 2004. The interpretation of traces. Natural Language Semantics 12:63–127.
- Schoorlemmer, Erik, and Tanja Temmerman. 2011. Head movement as a pf-phenomenon: Evidence from identity under ellipsis. Handout presented at WCCFL 29.
- Schwarzschild, R. 1992. Types of plural individuals. Linguistics and Philosophy 15.
- Schwarzschild, R. 1996. Pluralities. Springer.
- Sharvy, R. 1980. A more general theory of definite descriptions. The philosophical review 89:607–624.
- Shyu, S. 1995. The syntax of focus and topic in Mandarin Chinese. Doctoral Dissertation, University of Southern California.
- Stechow, A. 1993. Die aufgaben der syntax. Syntax. Ein internationales Handbuch zeitgen"ossischer Forschung 1:1–88.
- Stowell, Timothy. 1981. Origins of Phrase Structure. Doctoral Dissertation, MIT.
- Stowell, Timothy. 1983. Subjects across categories. The Linguistic Review 285–312.

- Stowell, Timothy. 2004. Tense and modals. In The Syntax of Time, ed. J Guéron and J Lecarm. 621–637: MIT Press.
- Takahashi, Shoichi, and Danny Fox. 2006. MaxElide and the re-binding problem. In Proceedings of Semantics and Linguistic Theory, ed. Effi Georgala and Jonathan Howell, volume 15, 223–240. Ithaca, NY: CLC Publications.
- Tang, C.C.J. 1990. Chinese phrase structure and the extended X'-theory .
- Thiersch, Craig. 1985. VP and Scrambling in the German Mittelfeld. University of Tilburg.
- Travis, Lisa. 1984. Parameters and effects of word order variation. Doctoral Dissertation, MIT.
- Trinh, Tue. 2007. The case for no Case. MIT.
- Trinh, Tue. 2009. A constraint on copy deletion. Theoretical Linguistics 35:183–227.
- Trinh, Tue. 2010. Edges and linearization: A reply. Theoretical Linguistics 36:93–110.
- Truckenbrodt, H. 2009. Comments on the syntax and prosody of german verbs. Theoretical Linguistics 35:307–317.
- Tsai, W.T.D. 1994. On economizing the theory of A'-dependencies. MIT Dissertation .
- Tsai, W.T.D. 2001. On subject specificity and theory of syntax-semantics interface. Journal of East Asian Linguistics 10:129–168.
- Vicente, Luis. 2005. Towards a unified theory of movement: an argument from spanish predicate clefts. Leiden Papers in Linguistics 2:43–67.
- Vicente, Luis. 2007. The syntax of heads and phrases: A study of verb (phrase) fronting. Doctoral Dissertation, Universiteit Leiden, Leiden, The Netherlands.
- Vicente, Luis. 2009. An alternative to remnant movement for partial predicate fronting. Syntax 12:158–191.
- Von Stechow, K. 1993. Exceptive constructions. Natural Language Semantics 1:123–148.
- Watanabe, A. 2010. Notes on nominal ellipsis and the nature of no and classifiers in japanese. Journal of East Asian Linguistics 19:61–74.

- Webelhuth, G. 1992. Principles and parameters of syntactic saturation. Oxford University Press, USA.
- Williams, Edwin. 1983. Against Small Clauses. Linguistic Inquiry 287–308.
- Wu, Y., and A. Bodomo. 2009. Classifiers ≠ determiners. Linguistic Inquiry 40:487–503.
- Xu, Lie-jiong. 1996. Limitation on subjecthood of numerically quantified noun phrases: A pragmatic approach .
- Yang, R. 2001. Common nouns, classifiers, and quantification in chinese. Doctoral Dissertation, Rutgers, The State University of New Jersey.
- Zweig, E. 2009. Number-neutral bare plurals and the multiplicity implicature. Linguistics and philosophy 32:353–407.