



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Multiple Wh-Questions
A Typological-Syntactic Analysis

verfasst von / submitted by

Magdalena Lohninger, BA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Arts (MA)

Wien, 2019 / Vienna 2019

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 066 867

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Allgemeine Linguistik:
Grammatiktheorie und kognitive Sprachwissenschaft

Betreut von / Supervisor:

Mag. Susanne Wurmbrand, Privatdoz. PhD

English Abstract

I present a new syntactic analysis for multiple wh-constructions. Adopting Richards 1997, I assume that there are two types of languages concerning wh-movement: Such which A-move their wh-words (A-languages) and such which A'-move them (A'-languages). I expand Richard's account by assuming that in both types, wh-movement targets the CP. This is done via A'-movement, as well as via A-movement. Building on recent work on Cross-clausal A-dependencies (CCA), I take over the idea of an A-position inside CP. I use this A-position as a landing-site for wh-movement and by that enable wh-movement to be A-movement into CP. My analysis combines two former accounts in a new way and gives correct predictions about the behaviour of the two language classes (A- versus A'). Finally, assuming that CCA and certain instances of wh-movement target an (maybe even the same) A-position in CP, poses the suspicion that those two phenomena might correlate. Examining them typologically, I bring forward that there is a unidirectional implication: A-wh-movement seems to entail CCA. I explain this observation, assuming that the wh-part of CP is above the CCA-part of CP and thus conditions its quality. This assumption ties together CCA-phenomena and A-wh-movement in a new syntactic way.

German Abstract

In meiner Arbeit präsentiere ich eine neue syntaktische Analyse für multiple wh-Konstruktionen. Richards 1997 adaptierend nehme ich an, dass es, in Bezug auf wh-Bewegung, zwei Typen von Sprachen gibt: Solche, die ihre wh-Wörter durch A-Bewegung anheben (A-Sprachen) und solche, die das durch A'-Bewegung tun (A'-Sprachen). Ich erweitere Richard's Darstellung um die Idee, dass in beiden Sprachtypen wh-Bewegung die CP zum Ziel hat. Das bedeutet, sowohl A'-Bewegung als auch A-Bewegung führen in die CP. Ich verwende dazu die Idee einer A-Position innerhalb der CP, aufbauend auf rezenten Studien zu 'Cross-clausal A-dependencies' (CCA). Ich benutze diese A-Positions als Zielpunkt für wh-Bewegung und ermögliche dadurch wh-Bewegung, A-Bewegung in die CP zu sein. Meine Analyse kombiniert zwei frühere Zugänge auf neue Art und Weise und macht richtige Vorhersagen über das Verhalten der zwei Sprachgruppen (A- versus A'). Nimmt man an, dass sowohl CCA als auch bestimmte Arten von wh-Bewegung eine A-Position (vielleicht sogar dieselbe) zum Ziel haben, wirft den Verdacht auf, dass die beiden Phänomene zusammenhängen könnten. Mittels einer typologischen Untersuchung stelle ich fest, dass dies der Fall ist. Es gibt eine unidirektionale Implikation zwischen ihnen: A-wh-Bewegung scheint CCA zu bedingen. Ich erkläre diese Beobachtung durch die Annahme, dass der wh-Teil der CP über dem CCA-Teil der CP liegt und dadurch seine Qualität beeinflusst. Diese Annahme vereint CCA-Phänomene und A-wh-Bewegung auf eine neue syntaktische Art.

Multiple Wh-Questions

A Typological-Syntactic Analysis

Magdalena Lohninger

Contents

1	Introduction	1
2	Three Types of Languages On The Surface	4
2.1	All wh-words are raised	4
2.2	One wh-word is raised	5
2.3	All wh-words remain in-situ	6
3	Earlier Approaches	8
3.1	Multiply Filled SpecCP and Multiple CP Specifiers	9
3.2	Adjunction to IP/TP	13
3.3	Focus-Movement Instead of Wh-Movement	14
3.4	An Additional Functional Phrase	15
3.5	LF versus PF Movement	16
3.6	Multiple Spell-Out	19
4	Dividing Languages by Type of Movement Instead	25
4.1	Superiority	28
4.2	Wh-Islands	33
4.3	Scrambling	34
4.4	Weak Crossover	36
4.5	Summary of Data	38

5	Cross-Clausal A-dependencies	41
5.1	A-Dependencies Across CP Borders	42
5.1.1	Types of CCA	43
5.1.2	Towards an Account	45
5.1.3	Summary of Typological CCA-Phenomena	57
5.2	Analysis: An A-position in CP	58
6	A New Analysis: Wh-movement as A-movement to CP	60
6.1	The General Idea	62
6.1.1	A'-languages	62
6.1.2	A-languages	63
6.2	Superiority	65
6.3	Wh-Islands Block Movement	66
6.4	Long-Distance Questions	68
6.5	Weak Crossover	69
6.6	What Makes All Those Wh-Words Move? Five Hypotheses	71
6.7	SP and PL Reading	80
6.8	Summary	82
7	Typological predictions	85
7.1	✓ A-wh-movement, ✓ CCA	87
7.1.1	Turkish	87
7.1.2	Japanese	89
7.1.3	Greek	90
7.1.4	Hungarian	92
7.2	x A-wh-movement, x CCA	93
7.2.1	Bulgarian	93
7.2.2	English	95
7.3	x A-wh-movement, ✓ CCA	96
7.3.1	Korean	96

7.3.2	Brazilian Portuguese	98
7.3.3	Mandarin Chinese	98
7.3.4	Romanian	99
7.4	✓ A-wh-movement, ✗ CCA	100
7.5	Discussion	101
8	Open Issues	107
8.1	The Everlasting Problem: German	108
8.1.1	German as an A-language	109
8.1.2	SP and PL Readings	110
8.1.3	No Multiple Wh-Questions At All	112
8.1.4	Superiority	114
9	Conclusion	118
	Bibliography	120

1 Introduction

This thesis addresses multiple wh-questions from a typological-syntactic perspective. I bring forward a new analysis of multiple wh-constructions, based on the idea of A-movement into the CP-domain. I present, based on Huang's 1998 proposal that below the surface, all wh-words have to move to sentence-initial position, a unified account for multiple-fronting, single-fronting and wh-in-situ languages. I do so by uniting an approach by Richards 1997 who claims that at LF, wh-movement is either A- or A'-movement, and one by Wurmbrand (To Appear), who proposes an A-position inside CP for Cross-clausal A-dependencies (CCA). I use Richards' approach to divide languages into two classes: Those which A-move their wh-words and those which A'-move their wh-words. Wurmbrand's account provides me with the necessary landing-sites: Adopting her idea of an A-position in CP, I claim that some certain wh-movement is in fact A-movement into CP. Putting together these two accounts and applying them on wh-movement renders a new syntactic analysis for multiple wh-questions proposed in this thesis. This analysis leads to a short typological investigation: Assuming an A-position in CP which can be targeted by wh-words as well as CCA-phenomena, predicts a correlation between these two, which indeed seems to exist. As far as my examination covers it, there is an implicational relation between wh-movement and CCA, namely that whenever a language A-moves its wh-words, it allows CCA but not the other way around. I will explain this generalisation with the assumption that the landing-site of wh-movement is higher up in CP than the one for CCA and thus entails it. This observation seems to be of typological relevance and might provide deeper understanding of how mul-

multiple wh-questions work syntactically.

The thesis is arranged in the following way: In chapter two I present a three-way division of languages concerning wh-questions. On the surface, there are multiple wh-fronting languages, single wh-fronting languages and wh-in-situ languages. Later, I will replace this segmentation with one proposed in Richards 1997. In chapter three, I present earlier syntactic approaches on multiple wh-questions. Those illustrate the variety of explanatory attempts that have been made in order to explain the different behaviour of languages and support the relevance of a unified account. In chapter four, I introduce the account brought forward in Richards 1997. He proposes that languages do not part into classes which are defined by how many wh-words are moved on the surface but rather pattern together based on their covert movement behaviour. He claims that there are only two types of languages: Those which A-move their wh-words and those which A'-move their wh-words. I adopt this claim and from thereon, divide languages into two classes only. In section five, I introduce recent work on CCA, mainly based on Wurmbrand (To Appear). She analyses CCA-phenomena via an A-position inside CP. An exact description of her proposal is important since I will adopt this assumption for wh-movement. In chapter six, I specify my approach: I bring forward a new analysis of multiple wh-movement. Focussing on languages which A-move their wh-words, I claim that this A-movement targets an A-position inside CP which is similar to the account used for CCA-phenomena. I present five detailed hypotheses of how this movement could take place, what implications it would have and what predictions it makes for the typical behaviour of A-wh-languages. Eventually, I arrive at a hypothesis which analyses all wh-movement to be A-movement to a focus-position or to be scrambling first and from there, movement upwards into CP. This analysis seems to make correct predictions concerning the observed behaviour of A-languages: not obeying Superiority, blocked movement by wh-islands, different behaviour of long-distance questions and avoidance of weak crossover effects. In section seven, I investigate the idea that if both, A-wh-movement and CCA target an A-position in CP, then there

might be a correlation between them. I examine a small set of languages with the result of observing an apparent correlation. A-wh-movement seems to entail CCA. I formalize the observation with a typological generalization and give a possible syntactic explanation why the the entailment is unidirectional. In chapter eight, I list issues and research questions which will be left open for further investigation. Eventually, I present the difficult situation of German and how it resists to be fit into my framework. In section nine, I conclude.

2 Three Types of Languages On The Surface

Languages behave differently in how they treat multiple wh-questions. The crucial difference (on the surface) between them is how many of their wh-words they move (or not move) to sentence initial position. There are languages that move all wh-words to sentence-initial position. A lot of syntactic analyses have been struggling to find proper landing positions for all those raised wh-words. Hence, there are a lot of different approaches to the phenomenon, mostly dealing with Slavic languages. Then, there are languages that move one wh-word to sentence initial position and leave all others in-situ. These are languages which are easy to describe in 'usual' syntactic theory, since very well examined languages like English belong to this group. Finally, there are languages that leave all their wh-words in-situ like Chinese or Japanese. These languages face explanatory issues as well, since scope-properties of wh-words need to be explained differently here than by simple overt raising. In this chapter, I will quickly look at the three language-types and give typological examples.

2.1 All wh-words are raised

Bulgarian, Bosnian-Croatian-Serbian (henceforth on B/C/S)¹, Romanian and Polish belong to the type of language which raises all their wh-words. Those languages,

¹In most literature, the union of languages spoken in Bosnia, Croatia and Serbia (which are now seen as three separate languages - politically but to some extent also syntactically) is labeled

however, do not behave uniformly when it comes to Superiority of the wh-words, raising wh-words out of islands, and how they deal with long-distance raising. I will explain those differences and some more (proposed in, amongst many others, Rudin 1988, Bošković 1997, Richards 1997) in the following chapters. What is important for this moment is the raw observation that all wh-words are raised:

- (1) Na kogo kakvo dade Ivan
whom what give-3PS-SG Ivan
‘What did Ivan give to whom?’
Bulgarian (Bošković 2002: p. 351)
- (2) Kogo komu przedstawiłeś?
whom to-who introduced-2.SG
‘Who did you introduce to whom?’
Polish (Rudin 1988: p. 474)
- (3) Ko si koga tvrdio da je istukao?
who AUX whom claimed that AUX beaten
‘Who did you claim beat whom?’
B/C/S (Richards 1997: p. 31)
- (4) Cine cui ce ziceai că i-a promis?
who to-whom whta you-were saying that to-him has-promised
‘Who did you say promised what to whom?’
Romanian (Comorovski 1986: p. 171)

2.2 One wh-word is raised

Into this class fall those languages which are best documented in literature so far: English, German and several others. They raise one wh-word to sentence-initial position and leave all others in their base-position.

‘Serbocroatian’. This was the official notation before the Yugoslavia war. Nowadays, these three languages are separately referred to as Bosnian, Croatian and Serbian. I will therefore subsume them here under the term B/C/S.

- (5) What did John give to who?
English (Bošković 2002: p. 351)
- (6) Wer hat Hans was gegeben?
 who has John what given
 'Who gave what to John?'
German

2.3 All wh-words remain in-situ

There is a fairly big discussion going on about wh-in-situ languages. Those leave all their wh-words in their base-position and do not (at least not overtly) move anything anywhere. To this class belong, amongst others, Mandarin Chinese, Japanese, Tibetan and Turkish.

- (7) John gei-le shei shenme
 John give-PERF-3.PS-SG whom what
 'What did John give to whom?'
Mandarin Chinese (Bošković 2002: p. 351)
- (8) Bkrashis-lags-gi gagi garebyadnas gzigs-gnang-pa-red
 Tashi-HON-ERG which why buy-HON-PAST-AGR
 'Why did Tashi buy what?'
Tibetan (Richards 1997: p. 32)
- (9) John-ga nani-o naze katta no?
 John-NOM what-ACC why bought Q
 'What did John buy why?'
Japanese (Richards 1997: p. 31)
- (10) Tamer ne-yi nere-ye koy-du?
 Tamer.NOM what-ACC where-DAT put-PAST.3SG
 'What did Tamer put where?'
Turkish (İşsever 2009: p. 107)

For a long time, the distinction made by the surface appearance of languages served as the main distinction criterion. However, those three types seem to have more in

common than one may think. In my thesis, this trinity of languages will eventually be replaced by an approach which does not divide languages by their surface behaviour but by their covert wh-movement. This idea will be discussed in section 4. First, I will present some selected syntactic approaches dealing with multiple wh-questions.

3 Earlier Approaches

I will now introduce some of the earlier approaches that have been made to analyse multiple wh-questions. Most of them deal with multiple wh-fronting languages, i.e. those which overtly raise all wh-words to sentence-initial position. Some of them deal with all kinds of multiple wh-questions. All the accounts presented here will be important for my following assumptions concerning multiple wh-questions insofar, as they highlight inconsistencies within the three groups named above. Additionally, they give a foretaste on what the problem we deal with seems to be: The syntactic landing site of wh-words (overtly or covertly). I will only introduce some of the syntactic accounts proposed in literature. There are several semantic accounts as well but since my analysis will be purely syntactical and the scope of my thesis is limited, I will not illustrate them here. Eventually, I will stick to the assumption that on LF, all wh-words are raised in every language and will no longer distinguish between multiple-, single- and no-fronting languages. This section is divided into six subsections which are named after the proposed landing sites or explanatory attempts for multiple wh-questions. It starts with two analyses locating multiple wh-words in multiple SpecCPs or multiply filled SpecCPs. Following, I will present an analysis which places certain wh-words in TP adjunctions. Then, there is the idea that certain wh-words focus-move instead of properly wh-move. Next, there is an analysis proposing an additional functional phrase between CP and TP which wh-movement can target. Last, I bring forward an analysis assuming multiple Spell-Out in order to explain multiple-wh-questions. Eventually, I will agree with none of them completely but present an approach by Richards 1997 who divides wh-

movement into A'- versus A-movement. This idea will be examined in subsequent chapters.

3.1 Multiply Filled SpecCP and Multiple CP Specifiers

Pesetsky 2000 Pesetsky 2000 bases his analysis on the surface structure of multiple wh-questions. This means, he suggests different syntactic structures, depending on how many wh-words are moved overtly. He proposes an analysis based on different SpecCP potential of languages. There are languages with only one SpecCP (German), languages with multiple SpecCPs (English, Bulgarian) and languages with no SpecCPs at all (Japanese). The idea of Pesetsky 2000 is based on overt and covert movement. To that, he adds feature-movement (first presented in Chomsky 1995) but claims, differently to Chomsky 1995, that feature-movement is not the same as covert movement but an instance of its own. In wh-questions, C bears an uninterpretable wh-feature which needs to be deleted. Languages have three options to do so by either overtly moving their wh-words to SpecCP, moving them covertly to SpecCP or moving only the interpretable features on the wh-word. Pesetsky 2000 claims that languages part into three different kinds: They either have a C_{0-Spec} , i.e. a C-head introducing no specifiers, a C_{1-Spec} , i.e. a C-head introducing one specifier or a C_{m-Spec} , i.e. a C-head introducing multiple specifiers. A language always has to fill all the provided specifiers. He claims that Bulgarian and English have C_{m-Spec} . In Bulgarian, all the wh-words move overtly and obtain Superiority. In English, one wh-word moves overtly and the rest move covertly. English shows Superiority between the wh-words, therefore it must have multiple SpecCPs. German has only one SpecCP which can be filled by overt phrasal movement. All the other wh-words only undergo feature-movement. Hence, German does not show Superiority between the wh-words. Finally, Japanese has no SpecCP at all and the uninterpretable wh-feature on its C is deleted by wh-feature movement alone. This

explains why nothing moves overtly and why no Superiority is observed.

Putting all this together, we receive a tripartition of languages by their amount of CP-Specifiers. Evidence for that comes from the so called „Intervention-Effect” also proposed in Beck 1996. The Intervention Effect describes the circumstance that feature-movement is only possible if no scope-bearing element intervenes between the wh-in-situ and the C-head. Hence, English and Bulgarian show no intervention effect, since all of their wh-movement is phrasal and not feature-movement. Due to the fact that all their wh-movement is phrasal and scope-bearing elements can block-feature movement only, no intervention effect can be seen. In the following example, *no* is the intervener and does not render the question ungrammatical:

- (11) Which person did not read which book?

English (Pesetsky 2000: p. 79)

German does show an intervention-effect since its second wh-word shows feature movement and a scope-bearing element (*no one*) blocks this feature-movement. Therefore, (12) is ungrammatical or at least strange:

- (12) ?? Welche Kinder haben niemandem welche Bilder zeigen wollen?
which children have no-one which pictures show want
'Which children wanted to show nobody which pictures?'

German (Pesetsky 2000, p.94)

Finally, due to Pesetsky 2000, in Japanese, all wh-movement is feature-movement. Hence, a present intervener (*only*) which c-commands one or both wh-words blocks feature-movement and the question gets ungrammatical:

- (13) ?* Taroo-wa [Hanako-sika dare-ni nani-o yom-ana-i to]
Taroo-TOP Hanako-only who-to what-ACC read-NEG-PRES C
Tomoko-ni itta no?
Tomoko-to said Q
'To whom did Taroo tell Tomoko that only Hanako read what?'

Japanese (Pesetsky 2000: p. 116; originally from Tanaka 2001)

To sum up, Pesetsky 2000 groups languages into three types, depending on how many specifiers they allow: Languages with zero specifiers (Japanese and Korean),

languages with one specifier (German) and languages with multiple specifiers (English and Bulgarian).

Rudin 1988 Rudin 1988 gives an analysis for multiple wh-fronting languages (such as Bulgarian or B/C/S) only. She differs between two types of multiple wh-fronting languages, namely Bulgarian-like ones and B/C/S-like ones. Bulgarian-like languages obey Superiority and – due to Rudin 1988 – the fronted wh-complex cannot be split up by adverbs or parentheticals (there has been a lot of literature contradicting this fact. For the moment, let us assume Rudin 1988’s data to be correct in order to present her analysis).

(14) ✓ SUPERIORITY

- a. Koj kogo kǎde e vidjal?
who whom where is seen
'Who has seen whom where?'
- b. Koj kǎde kogo e vidjal?
- c. * Kǎde koj kogo e vidjal?
- d. * Kogo koj kǎde e vidjal?
- e. * Kogo kǎde koj e vidjal?

Bulgarian (Krapova 2002: p.6)

(15) ✗ ADVERBS

- a. Zavisi ot tova, koj kogo prǎv e udaril.
depends on this who whom first has hit
'It depends on who hit whom first.'
- b. * Zavisi ot tova, koj prǎv kogo e udaril.
depends on this who first whom hit

Bulgarian (Rudin 1988: p. 467)

B/C/S-like languages behave exactly the opposite way with respect to the above named criteria. This means they show no Superiority-effects and the wh-complex can be split up by parentheticals and adverbs.

(16) ✕ SUPERIORITY

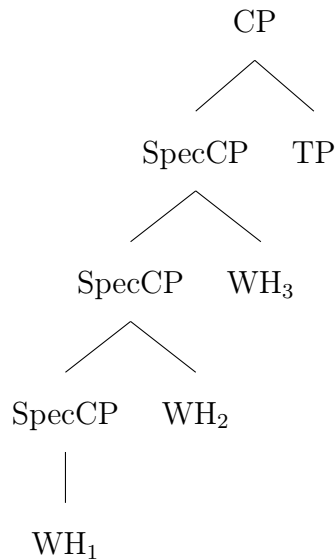
- a. Kogo komu przedstawił?
whom to-whom introduced-2s
'Whom did you introduce to whom?'
- b. Komu kogo przedstawił?
Polish (Rudin 1988: p.474)

(17) ✓ ADVERBS

- a. Ko je koga prvi udario?
who has whom first hit
'Who hit whom first?'
- b. Ko je prvi koga udario?
who has first whom hit
B/C/S (Rudin 1988: p. 467)

Besides these differences, Bulgarian-like languages do allow multiple wh-extraction from an embedded clause and wh-island violations whereas B/C/S-like languages do not. This difference will be examined in greater detail in the following chapter and is not important here but should to be mentioned for completeness.

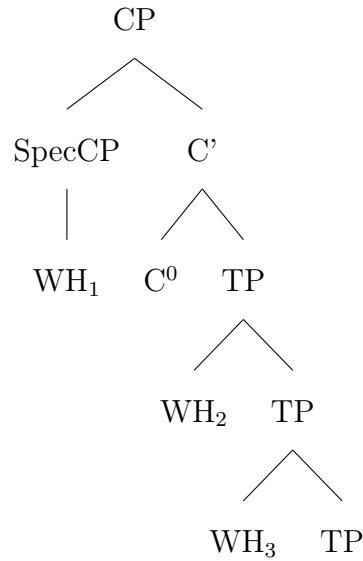
For the first type, Bulgarian-like languages, Rudin 1988 proposes an analysis with a multiply filled SpecCP. She assumes that all wh-words are raised to SpecCP and together create a constituent in there. Rudin 1988 expects that the first wh-phrase is substituted to SpecCP and that the other wh-phrases are adjoined rightward to that SpecCP. They need to be adjoined to the right, otherwise Superiority between the first wh-phrase and the others would be disturbed (see Rudin 1988, p. 480):



By that account, she also explains why wh-words can escape wh-islands: There can be several positions within each SpecCP and hence, a trace in one SpecCP-position does not block a wh-word from moving to another position within the same SpecCP and leaving a trace there.

3.2 Adjunction to IP/TP

Rudin 1988 For B/C/S-like languages, Rudin 1988 proposes an analysis based on adjunction to IP/TP. Since B/C/S-like languages allow intervening material between the first and second wh-word and do not show Superiority-effects, Rudin 1988 claims that the first wh-word is properly raised to SpecCP whereas the second and third wh-words are simply adjoined to TP:

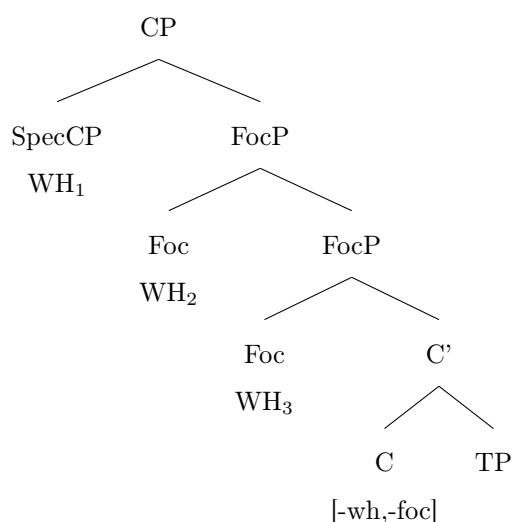


By that, the second and third wh-word move high up as well but do not enter the CP domain. Hence, there is no Superiority since it does not matter which wh-word moves to SpecCP and Adverbs can intervene since they can adjoin higher to TP.

3.3 Focus-Movement Instead of Wh-Movement

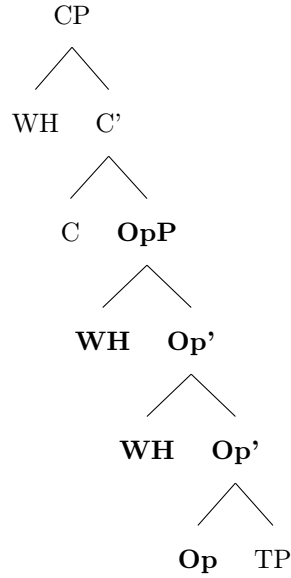
Bošković 1997/2002 Bošković 1997 (also Bošković 2002) as well only examines multiple wh-fronting languages like Bulgarian or B/C/S. His interpretation of the multiple-wh-raising phenomenon is the following: He assumes one wh-word to be raised because of wh-feature-saturation of the C head. He explains the movement of the other wh-words not as raising but focus-driven fronting. Both movements target CP. In Bulgarian, the structurally highest wh-word is raised to SpecCP (due to "Attract Closest") to saturate the feature on the C head. All wh-words contain a Focus-feature as well, which is licensed within the CP. Therefore, all wh-phrases need to go through the C-head. This means that wh_1 is properly raised whereas wh_2 and wh_3 are just focus-fronted. Bošković 1997 assumes that it does not matter whether wh_2 or wh_3 is moved first, since they have to pass through the same amount of nodes. Assuming this, he also explains the Superiority which is obtained in Bulgarian between wh_1 and the rest but not between wh_2 and wh_3 .

That means that Bošković assumes a structure where all wh-words are contained within SpecCP (just like Rudin 1988), but - contrary to Rudin 1988 - just the first wh-word is properly raised whereas the fronting of all the other wh-words is focus-driven. Assuming a split CP as proposed in Rizzi 1997, this looks probably like the following structure:



3.4 An Additional Functional Phrase

Citko 1998 Citko 1998 proposes an analysis which is based on data she takes from Rudin 1988. Remember, Rudin 1988 claims that there are two types of multiple wh-fronting languages. Bulgarian-type languages: Those where (according to her) all wh-words move to CP (e.g. Bulgarian) versus B/C/S-type languages: those where only one wh-word moves to CP and all others adjoin to TP/IP (e.g. B/C/S, Polish). This difference is of big importance in general and will accompany us for the rest of the thesis. Citko 1998 suggests that the lower wh-words of the latter class (i.e. B/C/S-like) do not adjoin to TP/IP but instead move to their very own functional phrase, called 'Operator Phrase' (OpP). OpP is located between CP and TP. (Structure from Citko 1998: (12) p.103)



In this analysis, all languages exhibit an operator phrase. [Q] on C attracts the first wh-word it can find to SpecCP. All other wh-words are attracted by [Op] on the Op-head. In B/C/S-like languages, they are then moved to SpecOP. Bulgarian-like languages, however, execute overt Op to C raising and by that, SpecOP positions are not projected. Hence all wh-words have to move to SpecCP. CP has multiple Specifier positions in this case. The second and third wh-words are nevertheless in all languages attracted by [Op] and not [Q]. The difference between Bulgarian-type languages and B/C/S-type languages is visible in this case if there are multiple SpecCPs or multiple SpecOpPs. The type of movement of the second wh-word is in both languages [Op]-driven and only differs in its landing site.

3.5 LF versus PF Movement

Several authors (amongst others L. L. S. Cheng 1997, Huang 1998, Pesetsky 1987) go away from observing what is moved on the surface to rather what is moved at LF. This is especially important for wh-in-situ languages since there is no overt movement of any wh-word. Huang 1998 claims that covertly, all wh-words in each

language are moved to some position very high up in the structure.¹ He names two reasons why all wh-words should be fronted at LF (his arguments are based on Mandarin Chinese, a wh-in-situ language): Selectional requirements of the matrix verb and the scope of wh-words. I will briefly sum up these points, based on a summary in L. L. S. Cheng 1997. Starting with selectional requirements of the matrix verb, the observation made by Huang 1998 is the following: Verbs in Mandarin Chinese select complement clauses in the same way as verbs in English. That is, 'think' and 'believe' do not select embedded questions whereas 'wonder' and 'ask' obligatorily select questions and 'know' optionally does.

- (18) Huangrong xiangxin guojing mai-le sheme?
Huangrong believe Guojing buy-ASP what
'What does Huangrong believe that Guojing bought?'
Lit.: *'Huangrong believes what Guojing bought.'
- (19) a. Qiaofong wen wo guojing mai-le sheme?
Qiaofong ask me Guojing buy-ASP what
'Qiaofong asked me what Guojing bought.'
b. * Qiaofong wen wo guojing mai-le shu.
Qiaofong ask me Guojing bought-ASP book
Lit.: *'Qiaofong asked me Guojing bought a book.'
- (20) Botong zhidao huangrong xihuan shei.
Botong know Huangrong like who
'Botong knows who Huangrong likes.'
'Who does Botong know Huangrong likes?'
- Mandarin Chinese (L. L. S. Cheng 1997: p. 193-195)

The conclusion by Huang 1998 is that Mandarin Chinese wh-words (i.e. in-situ wh-words) covertly move to the same position as English wh-words overtly do, namely,

¹This statement is not universally applicable. By now, it is assumed that not all wh-in-situ languages behave the same way but that some of them do not move their wh-word at LF. There are languages which create interpretability via inverse linking (May and Bale 2017, Pollard and Yoo 1998, Chang 1997 amongst many others). I am aware of this debate and therefore only include languages which are claimed to front their wh-words at LF.

SpecCP. Via this movement, the wh-word is - due to its position in the left periphery - visible for the matrix verb and makes the embedded CP selectable for question-selecting verbs.

Additional to that, wh-words always have wide scope over all other quantifiers in the sentence. It is not important, where the wh-word is generated, i.e. if it is the subject or object or anything else:

- (21) Mei-ge-ren dou mai-le sheme?
 every-CL-person all buy-ASP what
 'What did everybody buy?'

Mandarin Chinese (L. L. S. Cheng 1997: p. 196)

If one assumes that wh-words are fronted to SpecCP at LF, they would have scope over everything else in the sentence which explains the scope properties shown in (21).

Pesetsky 1987 argues similarly but claims that only certain wh-words move at LF whereas others do not. He divides wh-words into D(iscourse)-linked ('which',...) and non-D-linked ('who', 'what',...). Non-D-linked wh-words are quantifiers and move to some sentence-initial position at LF. D-linked wh-words are not quantifiers and do not move at LF. D-linked wh-words have a finite set of answers which is limited by the discourse. By that, they can take scope via a Q-morpheme which binds wh-words (see Baker 1970). Non-D-linked wh-phrases do not have this possibility and have to move at LF in order to take scope.

For the sake of the argument I will explain D-linking in a bit more detail here, even though it will not be of great importance for my analysis and will not be discussed further in my thesis. There are several differences between non-D-linked wh-words and D-linked wh-words, the most striking one amongst them being Superiority: Non-D-linked wh-words show Superiority effects whereas D-linked wh-words do not:

- (22) Non-D-linked:
 a. ? What did you persuade whom to read?
 b. * Mary asked what who read.

English (L. L. S. Cheng 1997: p. 197)

(23) D-linked:

- a. Which book did you persuade which man to read?
- b. Mary asked which book which man read.

English (L. L. S. Cheng 1997: p. 197)

The same is true for other languages such as Bulgarian:

(24) Non-D-linked:

- a. Koj kogo e vidjal?
who whom AUX seen
'Who has seen whom?'
- b. * Kogo koj kade e vidjal?
whom who AUX seen
Lit.: 'Whom has who seen?'

Bulgarian (modified from Krapova 2002: p. 6)

(25) D-linked:

- a. Koe momče koja kniga e kupilo?
which boy which book AUX bought
'Which boy bought which book?'
- b. Koja kniga koe momče e kupilo?
which book which boy AUX bought
'Which book bought which boy?'

Bulgarian (Krapova 2002: p. 4)

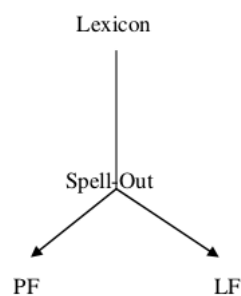
Since it would extend the scope of this thesis, I will only concentrate on non-D-linked wh-words and questions containing those. Finding a solution for D-linked wh-words is an issue of its own and will not be covered here.

3.6 Multiple Spell-Out

Turek 2012 brings forward a unifying account for both, multiple and single wh-fronting languages. She claims that the main difference between them, namely their

surface linear order, is a result of differing Spell Out. Spell Out does not apply in a uniform manner cross-linguistically but varies language-specifically. Multiple fronting languages like Polish have multiple Spell Out and therefore, Spell Out takes place at every level of a strong phase (vP, CP). Single fronting languages like English, on the other hand, undergo Spell Out only when the syntactic derivation is completed. This idea is based on the feature driven theory proposed in *The Minimalist Program* by Chomsky 1992. In Minimalism, wh-movement is feature-driven. A wh-word carries an uninterpretable Wh-feature [uWh] (making it active) as well as an interpretable Question feature [iQ]. An interrogative C-head contains an uninterpretable Question-feature [uQ] which needs to be valued. By probing down, it finds its interpretable goal in the [iQ] wh-word and triggers Agreement with it in order to become checked and interpretable. By that, the [uWh] gets valued and is removed from Narrow Syntax by the Spell Out Operation. Wh-fronting languages additionally have an EPP-feature in C which triggers wh-movement to SpecCP. (see Turek 2012, p. 133). However, according to Turek 2012, this theory does not explain why in some languages, all wh-words move to sentence-initial position whereas in others, only one or none moves.

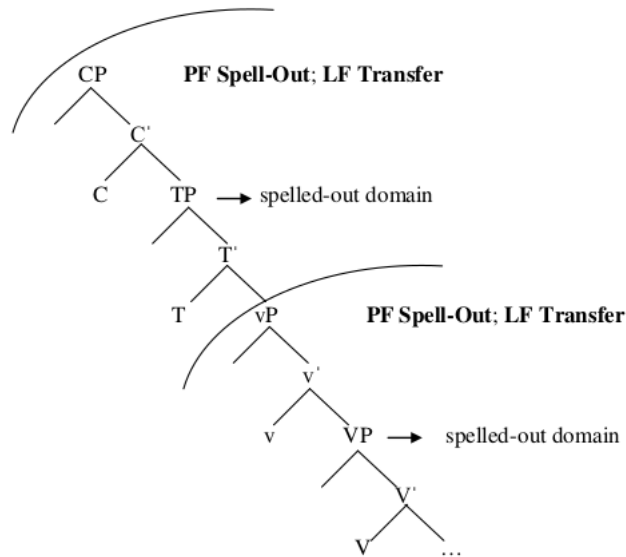
Turek 2012 tries to solve this problem by different stages of Spell Out: Spell Out is the point in derivation where LF and PF split and changes at LF are no longer influencing PF. This goes back to the T/Y model introduced by Chomsky 1995:



(Figure: Turek 2012: p. 146; originally from Chomsky 1995)

Spell Out marks the difference between overt and covert operations: Covert operations are no longer visible on the surface if they happen after Spell Out. According

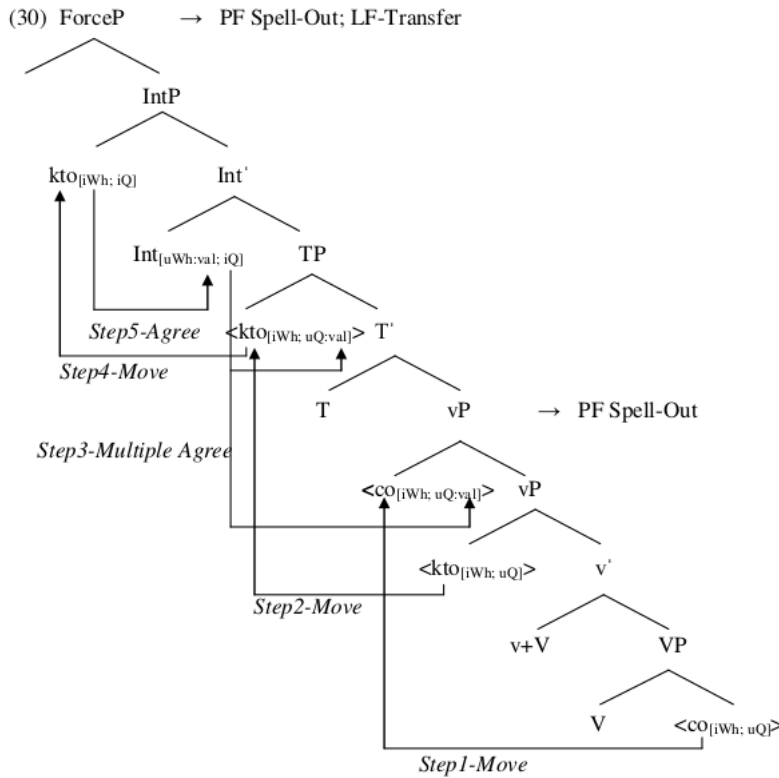
to Phase Theory (amongst others Chomsky 2000, Chomsky 2008, Uriagereka 1999, Nissenbaum 2000), multiple Spell Outs can and do take place in one single derivation. This means, it happens after each completed phase, namely vP and CP. Hence, transfer from LF to PF happens not only once in a derivation but after each phase (see Turek 2012, p. 146f.):



(Figure: Turek 2012, (18), p. 148)

Turek 2012 proposes that languages differ in whether Spell Out applies multiple times or only once. She claims that this is the main factor for differing surfaces in multiple vs. single wh-fronting languages. In a single fronting language, such as English, Spell Out happens only once after the the derivation is completed (i.e. at the CP level) whereas in multiple fronting languages, such as Polish, Spell Out happens after completion of each phase (i.e. after vP and CP). Since wh-words move cyclically through SpecvP, this makes a difference for their appearance on PF. In both languages, all wh-words carry [iWh] and require feature-checking. Turek 2012 assumes that C attracts all wh-words (bearing [iWh]) via a Multiple Agree Relation (based on the notion of Multiple Agree in Hiraiwa 2001, meaning that one probe can enter an Agree relation with multiple goals). Movement of the wh-words is then triggered by Agree, and not by an EPP-feature as originally in Chomsky 1992.

Hence, all wh-words are raised to CP. This happens cyclically through SpecvP. In multiple-fronting languages like Polish, the object-wh-phrase is attracted by C and moves through the vP-edge upwards. The subject wh-word on the other hand moves from SpecvP towards CP. The phase is completed with the object-wh on its edge. Spell Out does not care if the components of its phase are interpretable or not, at least according to Turek 2012. The first Spell Out happens after the completion of SpecvP, the structure is then transferred to PF. In this structure, the object wh-word is located on the edge of vP. The subject-wh is attracted by C and raised to SpecCP. By that, [uWh] on C is satisfied. Therefore, the object-wh does not need to move further and its copy in SpecvP is transferred to PF. This results in an LF where all wh-words are fronted but the second one is actually on the edge of vP instead inside CP. The second wh-word needs to be raised to the edge of vP in order to make it complete. Spell Out happens only to complete phases and vP would not be complete with an empty left edge. (see Turek 2012, p. 163f.):



(figure: Turek 2012, (30), p. 163)

This explains the fronting-paradigm of Polish where the second wh-word must be moved but not necessarily to an A'-position:

- (26) a. Co Ewa komu obiecała?
 what Eva whom promised
 'What did Eva promise to whom?'
 b. Co komu Ewa obiecała?
 what whom Eva promised
 'What did Eva promise to whom?'
 c. * Co Ewa obiecała komu?
 what Eva promised whom
 Int.: 'What did Eva promise to whom?'
 Polish (Turek 2012: p. 144)

In (26a) and (26b) one can see that the second wh-word can either be raised to CP or remain in verb-initial position in vP. Importantly, it needs to be raised to one of these positions and cannot remain in-situ as (26c) shows. Turek 2012 does not state how (26b) is explained in her framework: here, both wh-words clearly move to CP but I do not know if this is covered by her theory.²

In languages like English, contrary to Polish, the second wh-word stays in-situ. According to Turek 2012, this stems from the fact that in English, Spell Out only happens after the whole derivation is completed. This means, both wh-words are goals to [uWh] on C but only the higher one is attracted to move (Attract Closest) whereas the other one can stay in situ. It does not have to move to the edge of vP in order to complete the vP phase because Spell Out does not happen at this point anyways (see Turek 2012, p. 165f.).

²Turek 2012 presents very interesting data on long-distance wh-extraction in Polish: It works with the complementizer 'żeby' but not with the complementizer 'że' (see Turek 2012, p.172/173). According to her, 'że' introduces indicative clauses whereas 'żeby' introduces subjunctive complement clauses. Thus, indicatives do not allow wh-extraction, subjunctive ones do. This might be interesting to look at via the framework proposed in Wurmbrand and Lohninger 2019. However, this goes beyond the scope of this thesis and is just a further idea to investigate.



24

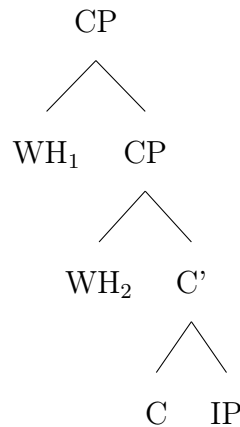
4 Dividing Languages by Type of Movement Instead

Since grouping languages by how many wh-words they raise overtly seems to pose severe explanatory issues, Richards 1997 proposes to distinguish languages not by how many wh-words they move on the surface but rather what kind of movement wh-movement is. Thereby, he includes not only languages which raise their wh-words overtly but also such which do that covertly. This is based on the assumption that at LF, all wh-words raise to sentence-initial positions for interpretability. I am aware of the debate concerning wh-in-situ languages. There are probably wh-in-situ languages which do not front their wh-words but create wh-licensing via inverse linking. (See L. L.-S. Cheng et al. 2003, May and Bale 2017, Pollard and Yoo 1998, Chang 1997 amongst others). That issue is not mentioned in Richards 1997 and since this chapter is mostly a summary of his work, I will not dive into the debate in great detail. Additionally, there might be languages which only adjoin their wh-words to TP and do not move them into CP, as proposed for Hungarian in Richards 1997.¹ In this thesis, I stick to Richard's assumption that in the languages observed, all wh-words are raised at LF. Richards 1997 claims that all languages falling into this class solely differ in which of these processes happen overtly and which covertly. What divides languages into two groups, according to Richards 1997, is whether they

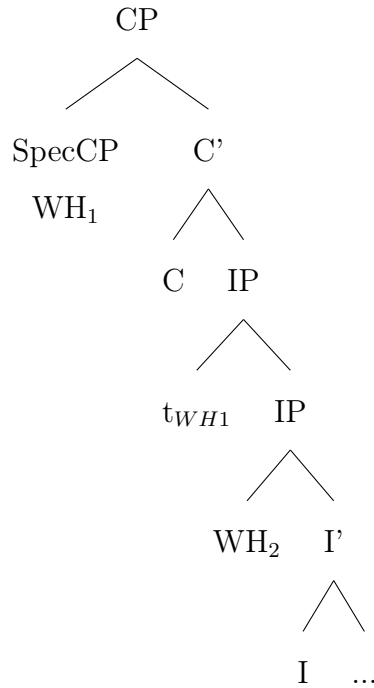
¹For my own analysis, I will assume that if there is LF movement of wh-words, it is to CP.

I will only examine those languages which move their wh-words at LF and leave languages creating interpretability via inverse-linking out. My assumption will be that whenever there is LF-movement, it is to CP.

move their wh-words via A-movement or A'-movement (if they are moved to an A-, or A'-position.) Richards 1997 claims that typologically, one can observe exactly this distinction. He separates languages into two groups: CP-absorption languages (e.g. Bulgarian, Rumanian, Chinese, Tibetan) and IP-absorption languages (e.g. B/C/S, Polish, Japanese, Korean). The difference between CP- and IP-absorption languages, according to Richards 1997, lies in the type of movement their wh-words undergo (covertly or overtly) and therefore the landing sites of the wh-words. In CP-absorption-languages, wh-movement is A'-movement and targets SpecCP. There are multiple SpecCPs which serve as landings sites for the wh-words. All wh-words are attracted by one attractor (C-head) and moved to the multiple SpecCPs above. (Structure: Richards 1997, p.26)



In IP-absorption languages, wh-words undergo A-movement. This means, they are attracted by I and adjoin to IP via A-movement. Wh-movement in these languages resembles Scrambling in that it has A-properties. In some languages, one of those IP-adjoined wh-words is then attracted by [wh] on C and moved to SpecCP. This movement then is A'-movement. All other wh-words remain in IP. According to Richards 1997, there are also languages which leave all of their wh-words adjoined to IP. Importantly, this is just the case for local wh-movement. Long-distance wh-movement is A'-movement, since IP-absorption languages show different properties in local wh-movement and long-distance wh-movement. I will come back to that later. (Modified structure: Richards 1997, p.26):



Richards 1997 proposes several criteria on the basis of which CP- and IP-absorption languages may be distinguished, i.e. why in some cases, wh-movement is A'-movement whereas in others, it is A-movement. Richards 1997 does not longer differ languages by how many wh-words they overtly move, but assumes that all languages behave the same at LF: They either A'- or A-move their wh-words. Therefore, there are no longer three classes of languages (multiple-, single-, non-raising) but only two (A'-moving vs A-moving). Hence, there will always be examples for all three surface-classes (multiple, single, none), dividing them into the new, two-folded system (A' vs. A). The criteria used to make this division are: Superiority effects, wh-movement out of islands, allowance of A-scrambling, WCO-effects and interacting wh-dependencies. Richards 1997 uses languages from each conventional raising class to demonstrate his idea: For multiple-raising languages he examines Bulgarian, Polish, Hungarian and B/C/S, for single wh-fronting languages he uses German and English and for wh-in-situ-languages he looks at Chinese, Japanese, Korean, Tibetan and Hindi.

4.1 Superiority

As described in earlier chapters, CP-absorption languages show Superiority-effects.

This means, that the sequence of their wh-words is not random but determined.

- (27) a. Koj kogo vižda?
who whom sees
'Who sees whom?'
- b. * Kogo koj vižda?
whom who sees
Bulgarian (Richards 1997: p. 30; originally from Rudin 1988)
- (28) a. Bkrashis-lags-gi gagi garebyadnas gzigs-pang-pa-red?
Tashi-HON-ERG which why buy-HON-PAST-AGR
'Why did Tahi buy what?'
- b. * Bkrashis-lags-gi garebyadnas gagi gzigs-pang-pa-red?
Tashi-HON-ERG why which buy-HON-PAST-AGR
Tibetian (Richards 1997: p. 32)
- (29) a. Who bought what?
- b. * What did who buy?
English (Richards 1997: p. 46)

IP-absorption languages, on the other hand, do not show Superiority in local wh-fronting:

- (30) a. Ko je koga vidjeo?
who AUX whom seen
'Who saw whom?'
- b. Koga je ko vidjeo?
whom AUX who seen
B/C/S (Richards 1997: p.30; originally from Bošković 1997)
- (31) a. John-ga nani-o naze katta no?
John-NOM what-ACC why bought Q
'What did John buy why?'
- b. ? John-ga naze nani-o katta no?
John-NOM why what-ACC bought Q

Japanese (Richards 1997: p. 31)

- (32) a. Wer hat was gekauft?
who has what bought
'Who bought what?'
b. Was hat wer gekauft?
what has who bought

German (Richards 1997: p.46; originally from S.-S. Kim and Sternefeld 1997)

They do, on the other hand, show Superiority effects in long-distance wh-fronting:

- (33) a. Ko si koga tvrdio da je istukao?
who AUX whom claimed that AUX beaten
'Who did you claim beat whom?'
b. * Koga si ko tvrdio da je istukao?
whom AUX who claimed that AUX beaten

B/C/S (Richards 1997: p. 51; originally from Bošković 1997)

- (34) a. Mary-ga [John-ga nani-o naze katta to] omotteiru no?
Mary-NOM John-NOM what-ACC why bought that thinks Q
'What did Mary think John bought why?'
b. * Mary-ga [John-ga naze nani-o katta to] omotteiru no?
Mary-NOM John-NOM why what-ACC bought that thinks

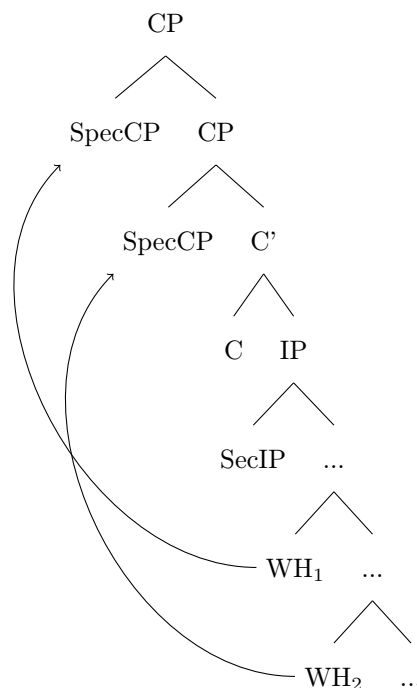
Japanese (Richards 1997: p. 31)

- (35) a. Wer t glaubt, dass Hans wen gesehen hat?
who believes that John whom seen has
'Who believes that John saw who?'
b. ?? Wen glaubt wer, dass Hans t gesehen hat?
whom believes who that John seen has
'Whom does who believe that John saw?'

German (Richards 1997: p. 46; originally from S.-S. Kim and Sternefeld 1997)²

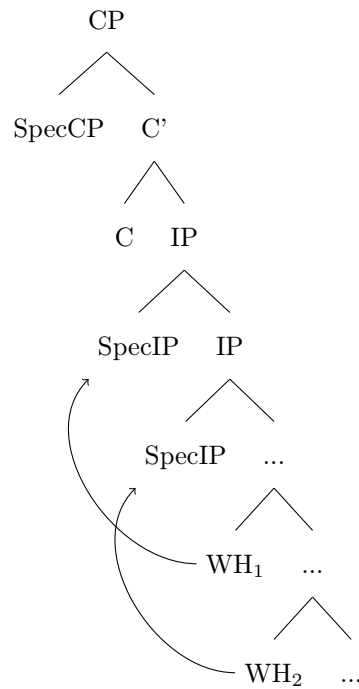
²This is not the best example since it differs from the B/C/S and Japanese example. Here, the first wh-word (*wer*) originates in the matrix clause whereas the second one (*wen*) originates in

Richards 1997 explains these Superiority-effects in the following way: In CP-absorption languages, the highest wh-word is attracted by C and must move first due to „Attract Closest” and is moved to the highest SpecCP. Then, the second wh-word is attracted by C and tucked in below the first SpecCP. (C exhibits Multiple Agree, meaning it keeps searching for goals even after it found one). By that, a second SpecCP is created. This means, wh-movement obtains both, „Attract Closest” and „Shortest Move”. Given that, their paths cross whilst moving. This means, all wh-words share one attractor (C) and move up to SpecCP with crossing paths.

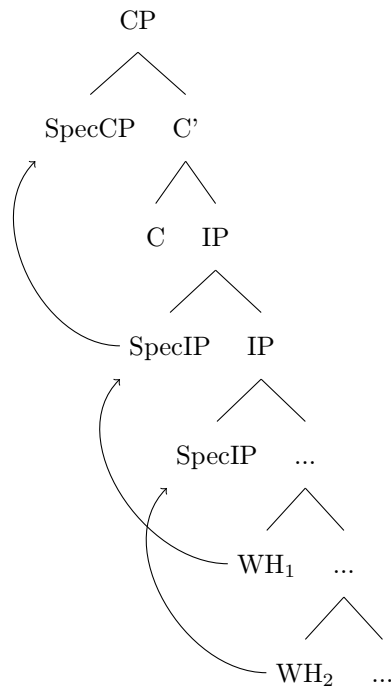


In IP-absorption languages, the wh-elements are not A'-moved to CP but A-moved to some extension of IP, i.e. they are adjoined to IP. This process rather resembles scrambling than traditional wh-movement. What exactly triggers this process, is not specified in detail, but is for sure something else than C. Richards 1997 claims that they must be in some way attracted by I. (This proposal is only true for local wh-questions. Long-distance wh-movement is - according to Richards 1997 - always A'-movement.) (Modified structure from Richards 1997, p. 26)

the embedded clause. In the other two examples, both wh-words originate in the embedded clause.



According to Richards 1997, in most languages (like B/C/S) one of the A-moved wh-words is then attracted by C and moved to SpecCP via A'-movement. All other wh-words remain adjoined to IP. This means, the main distinction between CP-absorption languages and IP-absorption languages is the number of attractors for movement: In CP-absorption languages there is only one attractor (C) whereas in IP-absorption languages, there is more than one attractor, namely C and I. I attracts all wh-words and adjoins them to IP, then C attracts one of those adjoined wh-words and raises it up to SpecCP. (Modified structure from Richards 1997, p. 26)



Which wh-word is attracted by C after adjunction to IP is not predicted since if adjoined to IP, all wh-words are equidistant from C. Therefore, no Superiority is exhibited.

Let us quickly remember the fact that Bulgarian does not show Superiority between the second and third wh-word. Following Richards 1997, this is due to an „Attract Closest Tax” which states that by movement of the first wh-word, C pays an „AC-tax” and afterwards, either wh₂ or wh₃ may be raised. This AC-tax is part of his „Principle of Minimal Compliance” which states that constraints on movement only apply to the first movement operation that targets a given head. I.e. „once an instance of movement to α has obeyed a constraint on the distance between source and target, other instances of movement to α need not obey these constraints.” (Pesetsky 2000, p.36).

4.2 Wh-Islands

CP-absorption languages can raise wh-words out of wh-islands. Richards 1997 claims that this is the case because CP-absorption languages have multiple SpecCP-positions. I.e., there are multiple A'-positions a wh-word can move to and via that escape a CP without being blocked by another wh-word occupying this position. Each wh-word has its own SpecCP-position with A'-properties and can escape the CP from there. This results in the fact that a wh-word can move over a CP boundary even if there is another wh-word in (another) SpecCP-position. Since there are multiple SpecCPs, there are multiple landing sites for wh-words. This means, Subjacency can be disobeyed and wh-islands do not stop movement. (See Richards 1997, p.24-25)³:

- (36) a. Judou xiang-zhidao shei mai-le sheme.
Joudou want-know who buy-ASP what
'Judou wonders who bought what.'
i. 'For which y, y a thing such that Judou wonders who bought y.'
ii. 'For which x, x a person such that Judou wonders what x bought.'
Chinese (L. L. S. Cheng 1997: p. 195)
- b. ? Koj ot tezi knigi se čudiš koj znae koj prodava?
which of these books REFL wonder-2SG who knows who sells
'Which of these books do you wonder who knows who sells?'
Bulgarian (Rudin 1988: p. 457)

In IP-absorption languages, there is only one SpecCP. If it is filled with a wh-word, it cannot be crossed by any other wh-word and thereby "traps" wh-words in wh-

³Interestingly and not mentioned in Richards 1997 this is not true for English: English wh-words are sensitive to islands. Raising them over another wh-word is not grammatical (see Ross 1967):

- (1) * Whom did you wonder, when Bill met t?
(2) * Who did you wonder, where t met Bill?

This difference is a problem for Richard's idea. Possibly, two groups of languages are not enough to describe the typological variety.

islands. This means, Subjacency is obeyed and wh-words cannot escape wh-islands (Richards 1997, p. 24-25)⁴:

- (37) a. ?? John-wa [Mary-ga nani-o katta-ka dooka] siritagatte-iru
 John-TOP Mary-NOM what-ACC bought whether know-want
 no?
 Q
 'What does John want to know whether Mary bought?'
Japanese (Richards 1997: p. 24; originally from Watanabe 1992)
- b. * Sta si me pitao ko moše da uradi?
 What AUX-2SG me asked who can to do
 'What have you asked me who can do?'
B/C/S (Rudin 1988: p. 459)
- c. * Wen weißt du nicht, wann Brutus t erstochen hat?
 Whom know you not when Brutus stabbed has
 'For which person don't you know when Brutus stabbed this person?'
German (D'Avis 1996: p. 92)

4.3 Scrambling

Due to IP-absorption languages have local A-Scrambling which fixes weak crossover violations (see Richards 1997, p. 28):

- (38) a. ?? Njegov_i susjedi ne vjeruju nijednom politicaru_i.
 his_i neighbours not trust no politician_i
 'His_i neighbours trust no politician_i.'
- b. [Nijednom politicaru_i]_j njegov_i susjedi ne vjeruju t_j.
 no politician_i his_i neighbours not trust
 'His_i neighbours trust no politician_i.'
- B/C/S (Richards 1997: p. 28)
- (39) a. ?* Soitu_i-no hahaoya-ga dare_i-o aisiteiru no?
 guy-GEN_i mother-NOM who-ACC_i loves Q
 'Who_i does his_i mother love?'

⁴German is not mentioned in Richards 1997 but added by me.

- b. ? Dare_i-o soitu_i-no hahaoya-ga aisiteiru no?
 who-ACC_i guy-GEN_i mother-NOM loves Q
 'Who_i does his_i mother love?'

Japanese (Saito 1992: p. 73)

- (40) a. *... dass seine_i Kinder jeden_i lieb haben
 that his children everyone love
 '... that his children love everyone.'
 b. ... dass jeden_i seine_i Kinder lieb haben
 that everyone his children love

German (Richards 1997: p. 45)

Observed in Richards 1997, CP-absorption languages lack such Scrambling. Scrambling in CP-absorption languages is either A'-movement or completely absent.

- (41) a. * Majka mu_i običa vsemi_i čovek.
 mother his love every person
 'His_i mother loves everyone_i.'
 b. * Vsemi_i čovek običa majka mu_i.
 every person love mother his.
 'His_i mother loves everyone_i.'

Bulgarian (Richards 1997: p. 29)

- (42) a. * Tade_i mama ai meigeren_i.
 his mother love everyone
 'His_i mother loves everyone_i.'
 b. * Meigeren_i tade mama ei_i.
 everyone his mother love
 'His_i mother loves everyone_i.'

Chinese (Richards 1997: p. 29)

- (43) a. *His_i children love everyone_i.
 b. *Everyone_i his_i children love.

English (Richards 1997: p. 45)

Based on these data, Richards 1997 claims that all languages that allow A-scrambling are IP-absorption languages.

4.4 Weak Crossover

The section above leads us directly to the next one: Weak Crossover effects (WCO). WCO-effects are a common test for distinguishing A'-movement from A-movement. If an anaphoric constituent is A'-moved, it is usually moved out of the binding domain of its antecedens which renders the sentence ungrammatical. If it is A-moved, on the other hand, it remains in the binding domain and the sentence stays grammatical. Richards 1997 claims that CP-absorption languages show WCO effects with wh-words⁵.

- (44) a. Koj_i običa majka si_i ?
 who_i loves mother his_i
 'Who_i loves his mother_i?'
 b. * $Kogo_i$ običa majka si_i ?
 who_i loves mother his_i
 'Whom_i does his_i mother loves?'
 Bulgarian (Richards 1997: p. 32)

- (45) * Whom_i does his_i mother love?
 English (Richards 1997: p. 46)

- (46) * Xihuan ta_i de ren kandao shei_i?
 like he DE man saw who
 'Who_i did the person that likes him_i see?'
 Chinese (Aoun and Li 1993: p. 201)

IP-absorption languages do not show WCO effects in local wh-fronting (see Richards 1997, p. 32f.)⁶:

- (47) a. Tko_i voli svoju_i majku?
 who_i loves his_i-ACC mother-ACC

⁵The Chinese data are not mentioned in Richards 1997 but were added by me

⁶Richards 1997 also gives this example for Hindi:

- (1) Sab-ko_i unkii_i bahin pyaar kratii thii.
 everyone-ACC their sister loves do-IMP-FEM be-PAST-FEM
 'Everyone_i their_i sister loves.'
 Hindi (Mahajan 1990: p. 26)

'Who_i loves his_i mother?'

- b. Koga_i voli njegova_i majka?
who_i loves his_i-NOM mother-NOM

'Whom_i does his_i mother love?'

B/C/S (Richards 1997, p. 33)

- (48) ? Wen_i liebt seine_i Mutter?
whom loves his mother

'Who does his mother love?'

German (Richards 1997: p. 46)

Interestingly, IP-absorption-languages do show WCO-effects in long-distance wh-movement. This matches with Richards' claim that long-distance wh-raising is A'-movement whereas local wh-movement is A-movement.⁷

- (49) * Koga_i njegova_i majka misli da Marija voli?
who_i his_i-NOM mother-NOM thinks that Maria loves

'Who_i does his_i mother think that Mary loves?'

B/C/S (Richards 1997: p. 33)

- (50) * Wen_i glaubt seine_i Mutter, dass jeder t_i liebt?
whom believes his mother that everyone loves

'Who does his_i mother think that everyone_i loves?'

German (Richards 1997: p. 46)

Contrary to what is predicted, Japanese does show WCO-effects. Otherwise showing IP-absorption behaviour, this falls out of Richard's pattern.

However, the relevant argument is no wh-word but a regular quantifier and hence, I do not consider it to be a good example.

⁷Again, for Hindi, Richards 1997 brings an example without wh-word:

- (1) * Sab-ko_i uskii_i bahin-ne socaa [(ki) raam-ne dekhaa]
everyone-ACC his sister-ERG thought (that) Ram-ERG saw

'Everyone_i, his_i sister thought that Ram saw.'

Hindi (Mahajan 1990: p. 41)

- (51) ?* Soitu_i-no hahaoya-ga dare_i-o aisiteru no?
 guy GEN mother NOM who-ACC love Q
 'Whom does his mother love?'
Japanese (Richards 1997: p. 34, originally from Saito 1992)

4.5 Summary of Data

Summing up the typological data presented in Richards 1997, the following pattern emerges (Modified table taken from Richards 1997, p.45)⁸:

	Superiority	WCO	Wh-islands block movement	A-SCR
<i>Bulgarian-like</i>	✓	✓	x	x
<i>Chinese-like</i>	✓	✓	x	x
<i>English-like</i>	✓	✓	✓	x
<i>B/C/S-like</i>	x	x	✓	✓
<i>Japanese-like</i>	x	✓	✓	✓
<i>German-like</i>	x	x	✓	✓

It can be clearly seen that languages do not pattern according to how many wh-words they overtly raise. If this were the case, then Bulgarian and B/C/S (all), Chinese and Japanese (none) and English and German (one) would behave uniformly. It is

⁸Note that English behaves differently than the rest of its class concerning movement out of wh-islands. The data were added by me and is not mentioned in Richards 1997. The relevant item is marked red in the table.

striking that they do not. All of them, no matter what is overtly moved, part into two groups: On one side, those which obey Superiority, show WCO-effects, whose wh-movement is not blocked by wh-islands, which can raise all wh-words out of embedded clauses and do not show A-scrambling. On the other side, those which behave exactly the opposite concerning these criteria. This means, languages behave differently in whether their wh-words undergo A- or A'-movement. The basic discovery of Richards 1997 will be of great importance for my further work and should be kept in mind throughout my thesis.⁹ However, I have to point out some problematic parts of Richards 1997. He proposes this division into two language groups as binary typological distribution. Unfortunately, it does not seem to be that easy. For Superiority-effects and A-scrambling, the division is very clear. Unfortunately, the border between WCO-effects and wh-islands is not as clear-cut. In the table above can be seen that English wh-islands block wh-movement whereas in all other CP-absorption languages, they do not. Furthermore, Japanese does exhibit WCO effects which contradicts the predicted pattern of regular IP-absorption languages. I think a two-way split of languages looks very promising in this case but might be too easy to encompass all languages and thus be universal. There is clearly a tendency of languages showing either CP- or IP-absorption properties but the cut between them might be not as clear as proposed in Richards 1997. Possibly, it is not a discrete distinction between the two classes but a gradual scale on which languages

⁹Richards 1997 claims the following about differing CPs in English and Japanese:

'We have seen that English and Japanese treat Spec CP in crucially different ways; in Japanese, but not in English, Spec CP can be used as an intervening landing site on the way to an A-position. In the account developed here, this contrast is linked to the fact that English, but not Japanese, has overt wh-movement. The availability of overt wh-movement in a language indicates that all complementizers have a strong feature, while its absence indicates that only weak features are present on the complementizer. Thus, even complementizers which are not interrogative differ in the two languages, as desired.' (Richards 1997, p.215)

This could be interesting for further research on why the wh-surface differs in languages.

find their place. After all, the behaviour of English and Japanese could have something to do with their surface; the overt movement (or non-movement) of *wh*-words might actually influence the syntactic character. This means, the fact whether a language is *wh*-in-situ, single-fronting or multiple-fronting might be crucial for the different behaviour of English and Japanese. The assumption that LF and PF are not completely independent from each other might be the reason for the observed different behaviour of the two languages. I do not have an established theory yet how the idea of the surface influencing the LF fits into the my account or if it is relevant at all, hence I will leave it open to further investigation.

I will now proceed by presenting recent work of Susi Wurmbrand, examining Cross-clausal A-dependencies (CCA) from a typological perspective. This will be the foundation for my analysis presented later: I will use both ideas from Richards 1997 and Wurmbrand (To Appear) to form a new proposal for *wh*-landing-sites in CP. Wurmbrand (To Appear) claims that CCA are made possible by an (optional) A-position in CP. I will use this idea of an A-SpecCP to pursue the concept of A-*wh*-movement proposed in Richards 1997. Eventually, in the subsequent chapters, I will bring forward a syntactic analysis which allows *wh*-words to A-move to an A-position in CP. That the idea of an A-SpecCP additional to an A'-SpecCP is not completely new but has been discussed widely in the last years will be explained in the chapter below. I present Wurmbrand's account to analyse CCA phenomena like Hyper-ECM, Hyperraising or Long-distance Agreement via an A-position in CP in order to later use this concept for *wh*-movement as well.

5 Cross-Clausal A-dependencies

In the last section, dealing with Richards' 1997 proposal to divide languages by how they move their *wh*-words, we saw that *wh*-words can not only undergo A'-movement but also A-movement. Based on several criteria such as Superiority effects, WCO or raising out of *wh*-islands, languages seem to have two possibilities by where to move their *wh*-words: To an A'-position, as typically assumed or to an A-position somewhere lower than CP. Richards 1997 proposes that those *wh*-words which undergo A-movement adjoin to IP, since this is the highest possible A-position in the common A- vs. A'-distinction. Everything moving to CP is traditionally already A'-movement. I will eventually, contrary to Richards 1997, suggest, that there indeed is a A- vs. A' distinction in *wh*-movement but that nevertheless all *wh*-words move to the CP-domain. I reached this idea by following the latest discussion about Cross-clausal A-dependencies (CCA). That CP is a pure A'-domain has recently been challenged in several works (see among others Van Urk 2015, Wurmbrand (To Appear)). This challenge is based on the widely spread phenomenon of CCA. CCA describe cases of "typical" A-dependencies, such as Subject Raising, Case Assignment or Agreement which, in some languages, seem to be able to cross A'-boundaries. The traditional assumption of A-dependencies claims that A-relations should not be able to cross CP-borders. However, this seems to be the case in the phenomenon of CCA, where an A-relation targets an element which apparently is located in an A'-position. In recent literature, the claim arose that there might be an A-position inside CP which can be targeted by A-operations. These are A-relations coming from the matrix clause as well as A-movement coming from the embedded clause.

In this chapter I will present very recent work of Susi Wurmbrand (Wurmbrand (To Appear)) where she argues for an (optional) A-position within CP. This A-position will later serve as main pillar for my analysis of multiple wh-questions: I will claim that in some languages, namely those which show A-wh-movement, wh-words can target an A-position in CP. This idea is not completely new and is supported by the work of Wurmbrand (To Appear) and others who make use of the very same A-position to explain CCA. In the last chapter of this thesis I will compare languages allowing CCA with languages allowing A-wh-movement to see if it might be the same A-position in CP they are using. In order to present my arguments correctly in the subsequent chapters, I will now introduce the work of Wurmbrand (To Appear) in great detail. I will first illustrate two instances of CCA, Hyperraising and Hyper-ECM. Then, I will follow Wurmbrand's line of argumentation in order to explain the following striking observations about Hyper-ECM: Namely, that the accusative case must come from the matrix verb, that the embedded ACC-subject must remain in the embedded clause and that this clause is a full CP. Based on these observations, I will conclude with the analysis brought forward in Wurmbrand (To Appear). This analysis is built on the idea that there exists an (optional) A-position within full CPs which can and is targeted by A-operations. This assumption will later be the basis for my analysis of A-wh-movement.¹

5.1 A-Dependencies Across CP Borders

There are three kinds of CCA typologically. First of all, Hyper-ECM, being Exceptional Case Marking into finite clauses. Second, Hyperraising, which is Subject Raising across a clause-boundary. And third, Hyperagreement, meaning Agreement across a clause-boundary. In well-studied languages like English, those phenomena usually exist in a non-hyper-version, meaning being bound by clause boundaries

¹This section is based on Wurmbrand (To Appear) and our joint work on the topic. The work has been supported by the FWF project "Universals and variation in clausal complementation" (M 2332-G30).

(CP). Or, if they reach over a clause-boundary, such as regular local ECM or Raising, the main assumption is that there is no CP inbetween. In English ECM or Raising, for example, the embedded clauses are claimed to be only TPs or vPs. This claim stems from the observation that finite embedded clauses or embedded clauses with a complementizer in English do not allow ECM or Subject Raising and are generally very limited to certain matrix verbs.

- (52) a. I believe **her** to have won the triathlon. (*ECM*)
 b. **She** seems to have won the triathlon. (*Subject Raising*)
English (Wurmbrand (To Appear): p. 1)
- (53) a. * I believe (that) **her** won the triathlon. (*Hyper-ECM*)
 b. * **She** seems that won the triathlon. (*Hyperraising*)
English (Wurmbrand (To Appear): p. 1)

Unfortunately, these observations are only true for a very limited set of languages. In a lot of less extensively studied languages, a parallel construction to those in (53) is completely grammatical.

5.1.1 Types of CCA

Hyperraising

Hyperraising describes Subject Raising across a CP clause boundary, as it is for example possible in Brazilian Portuguese. This means, the DP in matrix subject position originates in the embedded clause and is raised to the matrix position. Subject Raising is traditionally classified as A-movement but in Brazilian Portuguese it seems to have A'-properties since it is able to escape the embedded CP.

- (54) Os meninos parecem que fizeram a tarefa.
 the boys seem.3PL that did.3PL the homework
 'The boys seem to have done their homework.'
Brazilian Portuguese (Nunes 2009: p. 5)

That this movement is indeed A-movement and not A'-movement such as topicalization can be derived from the fact that weak pronouns and quantifiers may be hyperraised as well (55a) whereas they cannot undergo topicalization (55b), (55c):

- (55) a. Cê / Alguém parece que está doente.
 you.WEAK / someone seem that is sick
 'You seem/ Someone seems to be sick.'
- b. * Cê, João me disse que vai ser aprovada.
 you.WEAK, João me told that will be approved.FEM
 'You, John told me will be approved.'
- c. * Alguém, João me disse que seria aprovado.
 someone, João me told that would.be approved.MASC
 'Someone, John told me will be approved.'
- Brazilian Portuguese (Ferreira 2009: p. 24)

Hyper-ECM

Hyper-ECM are instances of ECM across CP borders. ECM (Exceptional Case Marking) describes a phenomenon where the embedded subject receives ACC marking from the matrix verb. ECM is well known from English (see (52a)) but there, limited to certain contexts, namely non-finite complement clauses. The standard analysis for English ECM assumes a reduced complement clause, either a TP or a vP. Since those do not create a sentence barrier, the ACC from the matrix verb can easily be assigned to the embedded subject. But taking a look at other languages, there is also ECM into CPs, meaning an ACC coming from the matrix verb and being assigned to the embedded subject below the CP boundary. This phenomenon exists, amongst many others, in Mongolian, Korean, Japanese, Turkish and Buryat.

- (56) Bat [margaash Dulmaa-g nom unsh-n gej] khel-sen.
 Bat [tomorrow Dulmaa-ACC book read-N.PST COMP] say-PST
 'Bat said that Dulmaa will read a book tomorrow.'

Mongolian (Fong 2019: p. 2)

- (57) Cheli-nun wonswungi-lul/ka banana-lul cal meknunta-ko
 Cheli-TOP monkey-ACC/NOM banana-ACC well eat-COMP

sayngkakhanta.

thinks

'Cheli thinks monkeys love to eat banana.'

Korean (Yoon 2007: p. 630)

- (58) Ooku-no nihonzin-wa dareka-o rosiago-ga dekiru to
many-COP Japanese-TOP someone-ACC Russian-NOM be-able COMP
omou.
think

'Lots of Japanese think that someone (specific) can speak Russian.'

Japanese (Horn 2008: p. 232)

Wurmbrand (To Appear) indicates that in all of the above examples, there is an overt complementizer which gives a clue about the embedded clause being a CP. Additionally, in all of them, the embedded subject is marked ACC. They look like instances of ECM with the only difference that their embedded clause seems to be a full CP (see Wurmbrand (To Appear), p.5). Wurmbrand (To Appear) shows first, why the ACC on the embedded subject must come from the matrix verb and cannot be assigned by the embedded verb. The next assumption one could make is that the embedded subject moves out of the embedded clause and receives its ACC in some higher position. She challenges this idea as well. Finally, there have been assumptions of the CP being deficient (by whatever means) in order to let an A-relation such as Case Assignment in. Wurmbrand (To Appear) also argues against a deficient CP domain. All the arguments put forward in the following are not mine but mostly adapted from Wurmbrand (To Appear).

5.1.2 Towards an Account

Taking Hyper-ECM as an example, Wurmbrand (To Appear) brings forward several observations concerning the involved case- and movement-operations: Why the accusative case must come from the matrix verb, why the ACC-DP must be located in the embedded clause, and finally why the embedded clause must be a full CP and cannot be deficient. These observations lead to her analysis of an A-position in

CP which serves as a connection between the matrix and embedded clause. Since the arguments leading to this claim are of high relevance in order to completely understand the analysis that will follow, I will explain them briefly.

ACC Comes From The Matrix Verb

That ACC-case must be assigned by the matrix predicate and not by the embedded predicate can be seen in passives: In Turkish, ACC is not possible with a passive verb, as seen in (59a). However, ACC is possible if the passive clause is embedded in a non-passive matrix clause as in (59b). This means, the ACC on the embedded subject in (59b), *pasta*, must come from the matrix verb, *hear*. On the other side, if the embedded clause is not passive but the matrix clause is, then ACC on the embedded subject is not possible (see (59c) and (59d)).

- (59) a. Makarna-Ø/*y₁ ye-n-di.
pasta-NOM/*ACC eat-PASS-PST
'Pasta was eaten.'
- b. John [makarna-y₁ ye-n-di diye] duy-du.
John.NOM [pasta-ACC eat-PASS-PST COMP] hear-PST
'John heard that pasta was eaten.'
- c. [Pelin-Ø Timbuktu-ya gi-ti diye] bil-in-iyor.
[Pelin-NOM Timbuktu-DAT go-PST COMP] know-PASS-PRS
'Pelin is known to have gone to Timbuktu.'
- d. *[Pelin-i Timbuktu-ya gi-ti diye] bil-in-iyor.
*[Pelin-ACC Timbuktu-DAT go-PST COMP] know-PASS-PRS
'Pelin is known to have gone to Timbuktu.'

Turkish (S. Şener 2011: p. 2-3)

The same restriction holds for Mongolian:

- (60) [Dulmaa(*-g) sain seheetin gej] khel-gd-sen.
[Dulmaa(*-ACC) good noble COMP] say-PASS-PST
'It was said that Dulmaa is good and noble.'

Mongolian (Fong 2019: p. 7)

Additionally, in Mongolian, a similar observation can be made: ACC is not possible in root clauses, especially in possessive-constructions like (61a). But as soon as (61a) is embedded into another clause, ACC can be assigned:

- (61) a. Sharlovan(*-g) Bat-id baigaa.
 carrot(*-ACC) Bat-DAT COP.PRES
 ‘Bat has a carrot.’
- b. Dulmaa chang-aar [sharlovan-g Bat-id baigaa gej]
 Dulmaa loud-INSTR [carrot-ACC Bat-DAT COP.PRES COMP]
 khel-sen.
 say-PST
 ‘Dulmaa said loudly that Bat has a carrot.’
- Mongolian (Fong 2019: p. 7)

Those examples show that the ACC on the embedded subject is necessarily assigned by the matrix verb and not by the embedded verb. This raises the next question: If the case comes from the matrix verb then maybe the embedded subject rises out of the embedded clause and into the matrix clause in order to receive its ACC there. Why this cannot be the case either will be discussed in the next paragraph.

The ACC-Subject Must Be Generated in the Embedded Clause

If ACC-case is indeed assigned by the matrix verb, then the next assumption might be that the embedded subject moves to or is base-generated in the matrix clause. This means that Hyper-ECM actually would be an instance of prolepsis. Prolepsis means that a matrix argument is coreferential with an embedded pronoun as in (62).

- (62) I know of Leo that he left.
 English (Wurmbrand (To Appear): p. 5)

If the language in consideration assigns ACC to its argument *Leo* and additionally has null-pronouns, this would result in a structure like (63) which on the surface looks exactly the same as Hyper-ECM.

- (63) I know Leo.ACC [_{CP} *pro* left.]
 English (Wurmbrand (To Appear): p. 5)

However, there is evidence that Hyper-ECM cannot be reduced to prolepsis (prolepsis is one but not the only option in many languages). Amongst these other options are word order, idiomatic readings, the impossibility of embedded pronominal subjects, licensing by embedded negation, scope under the matrix predicate, clefts or island sensitivity (all taken from Wurmbrand (To Appear))². I will not explain all of them in detail here since this is not relevant but break it down to two arguments: Word order and idiomatic readings. Let us start with word order: In many Hyper-ECM languages, it is possible that the embedded ACC subject is preceded by another lexical item such as an adverb. Adverbials cannot undergo movement, hence they must be in the embedded clause and therefore, everything below should not be moved to the matrix clause neither.³

- (64) Tünügün Ahmet manga [ete Aygül-ni ket-idu]
 yesterday Ahmet 1.SG.DAT [tomorrow Aygül-ACC leave-IPFV.3]
 di-di.
 say-PST.3
 'Yesterday Ahmet said that Aygül would leave tomorrow.'
Uyghur (Shklovsky and Sudo 2014: p. 388)

Besides that, idiomatic readings show a similar effect. Zeller 2006 found that in several languages with Hyper-ECM, idiomatic readings of the ACC subject with the embedded predicate are preserved. In order to receive idiomatic reading, the components of the idiom have to be clausemates. For example, in Nguni, a Bantu language (which, by the way, shows instances of Hyperraising as well), the idiomatic reading of an embedded clause disappears if the subject is not in the same clause as the idiom-predicate:

- (65) a. Izandla zi-ya-gez-an-a
 hand8 SM8-FOC-wash-REC-Fv
 Lit.: 'The hands wash each other.'

²These properties are not all explained in Wurmbrand (To Appear) in great detail; I use a lot of data and work we have been gathering together and which is still unpublished.

³Besides from Uyghur, the same is true for Buryat (Bondarenko 2017), Turkish (S. Şener 2011), Japanese (Hiraiwa 2001), Nez Perce (Deal 2017) and Mongolian (Fong 2019).

Idiomatic reading: Two people do something for each other, implying benefits for cooperation

- b. Izandla zi-thembis-a uku-gez-an-a
hand8 SM8-promise-Fv INF-wash-REC-Fv
'The hands promise to wash each other.'

Not idiomatic: The hands would have to be capable of giving a promise.

Nguni (Zeller 2006: p. 262-263)

Therefore, in order to preserve idiomatic meaning, all parts of an idiom need to be clausemates. If Hyper-ECM would be an instance of prolepsis, then they should never have an idiomatic reading. If the ACC subject is in the matrix clause and the embedded predicate in the complement clause, then they are not clausemates and hence, cannot form an idiom together. However, as observed in Wurmbrand (To Appear), cases of Hyper-ECM do preserve the idiomatic reading of the complement clause and therefore, the ACC subject cannot be moved to or base-generated in the matrix clause.

- (66) a. Toqquz qiz-ning tolghaq teng kel-di.
nine girl-GEN labor together arrive-PST.3
Lit.: 'Nine girls' labor pains came all at once.'

Idiomatic reading: 'Times are hard.'

- b. Tursun [toqquz qiz-ning tolghaq-ni teng kel-di]
Tursun [nine girl-GEN labor-ACC together arrive-PST.3]
di-di.
say-PST.3

Idiomatic reading is preserved: 'Tursun said that times are hard.'

Uyghur (Shklovsky and Sudo 2014: p. 388)

- (67) a. Dorj chang-aar Bat-iin nüd(-iig) oree deer-ee
Dorj loud-INSTR Bat-GEN eye(-ACC) top on-REFL.POSS
gar-san gej khel-sen.
climb-PST COMP say-PST

Lit.: 'Dorj said loudly that Bat's eyes climbed on top of themselves.'

Idiomatic reading: 'Dorj said loudly that Bat was very surprised.'

Mongolian (Fong 2019: p. 4)

The last observation why the ACC-subject must be generated in the embedded clause is licensing by embedded negation. We are assuming that the ACC subject is base-generated in the embedded clause. If this is the case, then it should be affected by embedded negation. This is especially important if one looks at NPIs. NPIs (Negative Polarity Items) need a negation as clausemate in order to be licensed. In (68a), for example, the NPI 'hichkim-ni' (*nobody*) must be licensed by a negation within the same clause in order for the sentence to be grammatical. Looking at Hyper-ECM, the following can be observed: The NPI is licensed by a negation inside the embedded clause, therefore clausemate with this negation and by that must be base-generated in the embedded clause and not in the matrix clause.

- (68) a. Men hichkim-ni kör-*(mi)-dim.
 1.SG.NOM nobody-ACC see-*(NEG)-PST.1.SG
 'I didn't see anybody.'
- b. Ahmet [hichkim-ni ket-mi-di] di-di.
 Ahmet [nobody-ACC leave-NEG-PST] say-PST.3
 'Ahmet said that nobody left.'
- Uyghur (Shklovsky and Sudo 2014: p. 388)

All the data given above give strong evidence that the ACC-DP is located in the embedded clause and not in the matrix clause via an instance of prolepsis.

The Embedded Clause Must Be a Full CP

Restructuring The next thing one could argue - considering the arguments above that the ACC-case comes from the matrix verb and the embedded ACC subject must be base-generated in the embedded clause - is absence of CP. In several languages, A-movement from embedded clauses to matrix clauses can be observed (e.g. pronoun fronting, A-scrambling, long passive, etc.). This is called Restructuring and is usually associated with reduced clause-size of the embedded clause (see Bondaruk 2004, Marušič 2005, Dotlačil 2007, Wurmbrand 2001, 2014, 2015 amongst many others). It means, a CP is considered to be a clause boundary and by that a barrier for any kind of A-relation. If one assumes that the embedded clause is

not a full CP but a TP or bare VP, then those barrier-violations for A-relations disappear. If that is the case, restructuring phenomena can occur. Hence, assuming there to be no CP in certain embedded clauses explains why certain CP-bound relations (like Restructuring) can target or leave the embedded clause. Additionally, verbs select for different clause sizes and by that enable or prohibit A-relations with the embedded clause (see for example Wurmbrand and Lohninger 2019 concerning universal hierarchies of clause sizes determined by semantic qualities of the matrix verb). However, the works on Hyper-ECM have concluded that Hyper-ECM is no instance of Restructuring but operates on full CP clauses. This idea comes from several sources: First of all, ECM (let alone Hyper-ECM) is possible in other contexts than Restructuring effects. All these assumptions made about Cross-clausal A-dependencies are a smaller part of a bigger project by Susi Wurmbrand dealing with a universal complement clause hierarchy (in semantic and syntactic terms) mentioned before. I will now briefly sum up the main idea, going in detail only at those segments that are relevant for my account here. Wurmbrand and Lohninger 2019 propose an implicational hierarchy of complement clauses. This hierarchy is determined by the semantics of the matrix predicate as well as the syntactic size of the embedded clause. Semantics and syntax are inseparable components and together form a matrix-complement structure. Wurmbrand and Lohninger 2019 divide matrix verbs into three semantic groups which join together with complement clauses of certain syntactic size. Those semantic groups are *Attitude*, *Irrealis* and *Tenseless* and they stand in an implicational hierarchy with each other. This means, complement clauses of *Attitude* predicates are universally more independent and complex as well as less transparent and less integrated whereas complement clauses to *Tenseless* predicates are the exact opposite. Stating an implicational complementation hierarchy in between those (Wurmbrand and Lohninger 2019 call this the *ICH*) means that in no language, a complement clause to a *Tenseless* predicate can be more complex, more independent, less transparent or less integrated than a complement clause to a predicate higher on the hierarchy (i.e. *Attitude* or *Irrealis*). They claim

this following hierarchy to be universal (scale taken from Wurmbrand and Lohninger 2019, p.15)

MOST INDEPENDENT		LEAST INDEPENDENT
LEAST TRANSPARENT	Attitude » Irrealis » Tenseless	MOST TRANSPARENT
LEAST INTEGRATED		MOST INTEGRATED
MOST COMPLEX		LEAST COMPLEX

Typologically, there can be differences in the clause size of the complement clause. A Tenseless clause for example does not always have to be a reduced clause. Languages vary in what clause size they allow for the three types. What is universal, though, is that the size of an *Attitude* complement clause must not be smaller (due to the criteria in the above table) than the size of a *Tenseless* complement clause. These findings are insofar interesting for our concerns here, as languages differ with which of these verb classes they allow certain transparency phenomena (like Restructuring).

Having this proposal by Wurmbrand and Lohninger 2019 in mind, the works on Restructuring also observe the following: Restructuring phenomena are very often present with verbs lower in the hierarchy which is logical. If Restructuring is a phenomenon linked to smaller clause size and if clause-size statistically degrades from left to right in the verbal hierarchy of Wurmbrand and Lohninger 2019 *Attitude* > *Irrealis* > *Tenseless*, then Restructuring should rather be possible with *Tenseless* verbs than with *Attitude* verbs. This indeed is the case. If (Hyper) ECM was an operation targeting reduced clauses, argued by Wurmbrand (To Appear), then it should be possible with verbs lower in the hierarchy as well. Accordingly, if a language allows this operation with an *Attitude* verb then it should allow it with *Irrealis* and *Tenseless* complements as well (since we deal with an implicational hierarchy). This is, surprisingly and contrary to the ECM phenomena above, not the case, as stated in Wurmbrand (To Appear):

	Attitude, speech	Future	Tenseless
Restructuring: Romance-type	x	x	✓
Restructuring: Germanic/Slavic-type	x	✓	✓
ECM: English/Icelandic	✓	x	x

Table 1 (taken from Wurmbrand (To Appear), p.7)

As can be seen in the table above, ECM behaves contrary to all other restructuring phenomena and, even more, against the implicational hierarchy. Wurmbrand (To Appear) claims that if ECM were only possible in reduced complement clauses, and it is, in a given language, possible with "Attitude" predicates, then it should be possible with all other predicates as well since they are lower on the implicational hierarchy. This is not the case, as can be seen above. Wurmbrand (To Appear) states that there are now two options: Either the implicational hierarchy idea is abandoned. Or, one could let go of the assumption that Hyper-ECM is a phenomenon in reduced clauses but take up the idea that it might reach across CP boundaries. The latter is exactly what Wurmbrand (To Appear) proposes and I will stick to this idea as well.

Shifted Indexicals Another argument supports the idea of Hyper-ECM into full CPs, according to Wurmbrand (To Appear): Shifted Indexicals. Shifted Indexicals are observable in some languages where an embedded 1st or 2nd person pronoun is "shifted" and interpreted as referring to the matrix subject instead of the discourse participant. In English, Indexicals in embedded clauses cannot shift, therefore a 1st or 2nd person pronoun always refers to the discourse speaker or hearer:

(69) Leo said that I (= speaker) left.

English (Wurmbrand (To Appear): p. 10)

In languages allowing Indexical shift, on the other hand, the embedded 1st person pronoun in (69) could refer to the subject of the matrix clause, i.e. *Leo*:

(70) Leo said that I (= Leo) left.

English (adapted from Wurmbrand (To Appear): p. 10)

Take for example Buryat:

- (71) Sajənə **bi** tɜrgə ʒmdəl-3-b gʒʒə mʒd-3.
 Sajana **1.SG.NOM** cart break-PRT1-1.SG C know-PRT1
 'Sajana found out that I (=speaker) broke a cart.'
 'Sajana found out that I (=Sajana) broke a cart.'
Buryat (Bondarenko 2017: p. 19)

In (71), the embedded 1st person pronoun is (optionally) shifted: it can either refer to the discourse speaker or the matrix subject (*Sajana*). It is crucial, that in (71), the embedded pronoun receives NOM. If it receives ACC instead (in case of Hyper-ECM, which, as we saw before, exists in Buryat), shifting the indexical is not possible anymore and the only available interpretation is coindexation with the discourse speaker:

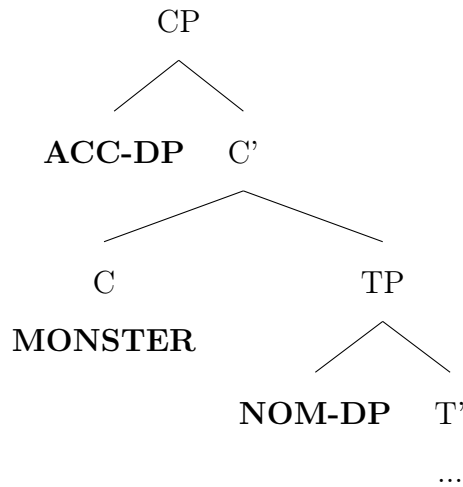
- (72) Sajənə **naməjə** tɜrgə ʒmdəl-ə(*-b) gʒʒə mʒd-3.
 Sajana **1.SG.ACC** cart break-PRT1(*-1.SG) C know-PRT1
 'Sajana found out that I (=speaker) broke a cart.'
 * 'Sajana found out that I (=Sajana) broke a cart.'
Buryat (Bondarenko 2017: p. 19)

A very similar thing is true for Uyghur:

- (73) Ahmet | sen ket-ting | di-di.
 Ahmet | 2.SG.NOM leave-PST.2.SG | say-PST.3
 'Ahmet said that you (=original hearer/*discourse hearer) left.'
- (74) Ahmet | seni ket-ti | di-di.
 Ahmet | 2.SG.ACC leave-PST.3 | say-PST.3
 'Ahmet said that you (=original hearer/discourse hearer) left.'
- Uyghur (Shklovsky and Sudo 2014: p. 386)

The analysis for Shifted Indexicals is commonly based on a shifting operator, called "Monster" operator (see Anand and Nevins 2004, Anand 2006, Sudo 2012, Sundare-

san 2012, 2018, Shklovsky and Sudo 2014, Podobryaev 2014, Messick 2016, Wurmbrand (To Appear)). The Monster operator changes the context so that embedded indexicals refer to the context of the matrix clause and not the discourse-context. It is assumed to be located on the C-head, shifting everything below. This leads to two conclusions: First, if the shifting operator is located on C-head then there must be a CP. And second, if everything below the C-head is shifted, and we take the observations from above, then nominative DPs must be below the CP whereas accusative DPs must be above it. Wurmbrand (To Appear) proposes that this is indeed the case: ACC subjects and NOM subjects do not have the same location in the structure: Whilst NOM subjects remain inside the TP, ACC subjects are moved to some instance of SpecCP. How this works and not violates any A'/A-constraints will be explained later. What is important for the moment, is that Wurmbrand (To Appear) proposes the following structure for Hyper-ECM clauses/ Shifted Indexicals:



If the ACC-DP were lower in the embedded clause, meaning below C, then it should shift as well. Since it never shifts, it must be above C. This account proposes that shifting is not a property or function of the embedded pronouns but triggered by the syntactic context. There are two more arguments to support this idea. First, in Buryat, a possessive pronoun embedded in a nominative DP can shift whereas a possessive pronoun embedded in an accusative DP cannot.

- (75) a. badmə [pro.1-SG ba:b3-mni] jab-a g3žə
 Badma.NOM [pro.1.SG father.NOM-1.SG] go-PST COMPMP
 m3də-nə.
 know-PRS
 'Badma_i knows that his_i father has left.'
 'Badma knows that my father has left.'
- b. badmə [pro.1.SG ba:b3-jə-mni] jab-a g3žə m3də-nə.
 Badma.NOM [pro.1.SG father-ACC-1.SG] go-PST COMP know-PRS
 *'Badma_i knows that his_i father has left.'
 'Badma knows that my father has left.'
- Buryat (Wurmbrand (To Appear): p. 11)

Taking all these data into consideration, I will stick with the analysis proposed by Wurmbrand (To Appear), namely that NOM-DPs are situated inside TP, whereas ACC-DPs are moved higher, presumably to some instance of SpecCP. Wurmbrand (To Appear) will give an detailed description of how this SpecCP is constituted and I will as well adopt this assumption. But before going on to explain her theory in detail, I will take a look at Hyper-ECM analysis which assumes the embedded CP to be deficient, by whatever means.

The Embedded CP Cannot Be Deficient

Finally, there have been a lot of arguments proposing that if A-relations actually cross CP borders then those CPs which allow this are not real CPs but in some way deficient. Wurmbrand (To Appear) argues against this by bringing forward the claim that CPs which allow A-relations into themselves do maintain their A'-status and are not transparent for any A-relation but only for specific ones such as *Hyper-ECM*. Wurmbrand (To Appear) presents a Japanese example to argue for that: Japanese, as seen before, allows *Hyper-ECM*. Additionally, Japanese allows A-scrambling which can change binding relations. In (76b), the scrambled *Nissan* and *Honda* can bind the quantifier *each other's spies*.

- (76) a. ??Otagai_i-no supai-ga [Nissan-to Honda-ni]_i kuwasii.
 each.other_i-s spy-NOM [Nissan-and Honda-with]_i familiar

'Each other' s spies are familiar with Nissan and Honda.'

- b. [Nissan-to Honda-ni]_i otagai_i-no supai-ga kuwasii.
 [Nissan-and Honda-with]_i each.other_i-s spy-NOM familiar
 Lit.: 'With Nissan and Honda, each other' s spies are familiar.'
 'Nissan and Honda are familiar with each other's spies.'

Japanese (Tanaka 2004: p. 5)

This means that scrambling within one clause shows A-binding. On the other hand, there is also A'-scrambling in Japanese, i.e. scrambling across a clause border. In these cases, A-binding is not possible anymore:

- (77) a. [Nissan-to Honda-ni]_i Toyota-no supai-ga [_{CP} John-o
 [Nissan-and Honda-with]_i Toyota-s spy-NOM [_{CP} John-ACC
 hoka-no dono-meekaa-yori t_i kuwasii-to] omot-teiru.
 other-s whichever-maker-more.than t_i familiar-COMP] think-PROG
 'Toyota's spy thinks of John as more familiar with Nissan and Honda
 than any other manufacturers.'
- b. ?? [Nissan-to Honda-ni]_i otagai_i-no supai-ga [_{CP} John-o
 [Nissan-and Honda-with]_i each.other_i-s spy-NOM [_{CP} John-ACC
 hoka-no dono-meekaa-yori kuwasii-to] omot-teiru.
 other-s whichever-maker-more.than familiar-COMP] think-PROG
 'With Nissan and Honda, each other's spies think of John more familiar
 than any other manufacturers.'

Japanese (Tanaka 2004: p. 5)

Crucially, the embedded subject in (77), (*John*), is ACC-marked. ACC in this situation cannot come from inside the embedded clause, hence this is an instance of *Hyper-ECM*. It indicates that CP is amenable for certain A-phenomena but remains a full CP for others.

5.1.3 Summary of Typological CCA-Phenomena

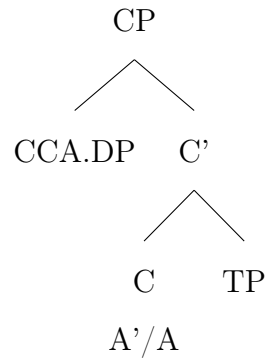
What can we conclude so far from the observations proposed and collected in Wurmbrand (To Appear)? The main proposal is that A-dependencies may cross CP-boundaries in certain cases. The best example for this is *Hyper-ECM*: It shows

an instance of case-assignment (a typical A-phenomenon) into a CP (a typical A'-domain). Wurmbrand (To Appear) elaborates the following in great detail: ACC in *Hyper-ECM* must come from the matrix predicate, the ACC-receiving subject must be generated in the embedded clause and is not prolepsis, the embedded clause must be a full CP and is not only a TP or *v*P and finally, the embedded CP cannot be deficient. It remains a full CP for other operations such as scrambling. The ACC-receiving DP must be moved to some above point C (since it would undergo indexical shift if it were below) but still need to remain inside the CP. The CP is an A'-domain and should not allow A-movement of a DP into itself. That the ACC-DP is in an A-position and not in an A'-position within CP can be seen in the *Improper A after A'-constraint* (Wurmbrand (To Appear), p.15). The *Improper A after A'-constraint*, also known as *Improper Movement Ban* rules out A-movement of an element which has been A'-moved before. Since this Constraint is assumed to be a universal and in some languages the ACC-DP moves on to the matrix clause via A-movement, the operation bringing it into CP cannot be A'-movement but must be A-movement. This leaves us with the puzzle: We have an A-moved element inside CP, either A-moving on to the matrix clause or receiving case (an A-relation) from the matrix verb. The CP, on the other hand, is not deficient and remains a clause-boundary for other phenomena. How can A-phenomena suddenly target A'-positions? Which leads us directly to the analysis proposed in Wurmbrand (To Appear) and will be explained in the following section.

5.2 Analysis: An A-position in CP

In order to explain these observations, Wurmbrand (To Appear) brings forward the following analysis: Based on a composite probe approach by Van Urk 2015 which states that heads may have mixed A and A'-features, Wurmbrand (To Appear) proposes that C-heads can have an A-feature additionally to their A'-feature. If a C-head has A-features, a DP agreeing with it inherits these A-features. A-movement

from a mixed A/A'-position is possible (contrary to A-movement from a pure A'-position) and by that, the *Improper Movement Ban* is obeyed. Languages differ in whether they have A-features on their C-head or not. Languages allowing CCA do have A-features on C, languages disallowing CCA do not. The following structure shows the general idea (Wurmbrand (To Appear), p.15):



By that, a SpecCP can have A-properties as well, thus allowing A-movement and A-relations to target it. This is the most important claim made by Wurmbrand (To Appear) and the one, that I will use for my further work. I will claim that this optional A-position can not only serve as a landing site for ACC-DPs but also for A-wh-movement or, if they target separate positions, that they at least influence each other.

6 A New Analysis: Wh-movement as A-movement to CP

Comparing the discussed syntactic approaches to multiple wh-questions, there seems to be one common issue: The landing site of the second wh-word. Very often, the authors realize that the second wh-word (at least in some languages) targets other positions than the regular SpecCP. Or, put differently, in some languages the landing-site and way of movement of the first wh-word seems to differ from the other wh-words whereas in other languages, the operations appear to be of the same kind. Several analyses struggle with capturing this issue properly and globally. Especially for those languages, where the first wh-movement is of different quality and apparently targets a different position than the non-initial wh-words pose a puzzle: Bošković 2002 claims that in these cases, movement of the second wh-word leads to a Focus position, Rudin 1988 localizes the second wh-word within TP in some languages, Citko 1998 inserts a new phrase between CP and TP which serves as landing site and finally Richards 1997 claims that some instance of wh-movement are actually A-movement. What unifies these accounts, is the problem of a missing position somewhere between CP and TP. A position, that is on the one hand high enough to serve as a landing site for wh-words in order to get them into or close enough to the CP domain but that, on the other hand, is different from typical CP-positions. It seems to have different properties, is neither properly inside nor outside CP and makes the whole multiple-wh-problem so hard to solve. As explained before, Richards 1997 interprets some instances of wh-movement as A-movement

but is forced to locate these wh-words inside the TP-domain since otherwise, the whole traditional A'/A-framework would crash. The authors examining CCA are confronted with a very similar issue. They face the problem that there seem to be A-relations targeting objects in A'-positions. As explicated in chapter five, the targeted elements are clearly within the embedded CP and not lower but at the same time, the operation targeting them must be an A-operation. An A-operation targeting an A'-element is and should not be possible. Hence everyone is struggling with the problem of how to get the embedded element high enough to be reached by the A'-relation from the matrix clause but still keep it inside the embedded clause. Amongst others, Wurmbrand (To Appear) proposes an A-position inside the embedded CP to solve this issue. Wurmbrand (To Appear) claims that certain languages have this A-position in their CP whereas others do not. Languages having an A-position in their CP allow CCA whereas those which only have A'-positions in their CP do not. Or, put differently, an A-position in CP is a requirement for CCA to be possible. I will propose that wh-movement which can be A-movement, according to Richards 1997, can target an A-position in CP, similar or maybe even identical to the one, CCA targets. On the following pages, I will first explain my exact idea about an A-position in CP as a landing site for wh. I will then try to explain Superiority, movement-blocking islands and long-distance wh-questions in this framework and finally present five hypotheses, why or how this wh-movement could be executed. The latter is not elaborated in perfect detail and still faces some serious explanatory problems. This thesis is rather about the idea of A-landing sites for wh-words in CP than about explaining how they move there exactly. Still, I put a lot of thought in how the process could take place and explicate my thoughts in those five hypotheses. Eventually, I make a remark on Single-pair and Pair-list readings, which seem to pattern with the A'- and A-wh-languages. From here on, I will refer to those languages which A'-move their wh-words as 'A'-languages' and to those which A-move their wh-words as 'A-languages'. Those classes are identical to what Richards 1997 refers to as 'CP-absorption-languages' and 'IP-absorption-languages'.

6.1 The General Idea

I adapt the idea of Richards 1997 that all wh-words are moved to the CP-domain at LF, independently from what is moved or not moved on the surface. I also follow Richards 1997 in his idea that what divides languages is not the amount of overtly moved wh-words but the type of movement they undergo. Richards 1997 claims that in some languages, wh-movement is A'-movement whereas in others, it is A-movement. However, unlike Richards 1997, I propose that in both cases, all wh-words move to the CP-domain. This means, both, A'- and A-movement can and do target CP. This is made possible by assuming an A-position inside CP, proposed, amongst others, by Wurmbrand (To Appear).

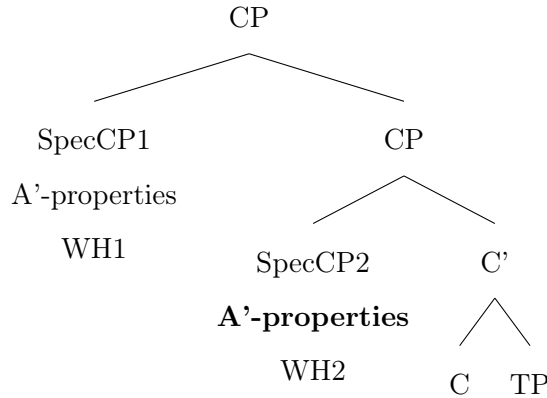
My idea is the following: There are two types of languages. Those, which move all of their wh-words via A'-movement and those, which move only the first wh-word by A'-movement but all others by A-movement. All of these movement-operations target the CP-domain which is, in the latter case, equipped with an A-position.

6.1.1 A'-languages

As described in great detail in Richards 1997, there are languages which move all their wh-words via A'-movement (such as Bulgarian-like, Chinese-like and English-like). He demonstrates these cases by finding similarities across languages with respect to the following properties: They show Superiority effects and WCO-effects, wh-islands block movement and raising all wh-words out of an embedded clause is possible. If they allow Scrambling, then it is A'-scrambling. (Modified table taken from Richards 1997, p.45):

	Superiority	WCO	Wh-islands block mvt	A-Scrambling	Interacting wh- dependencies
<i>A'-languages</i>	✓	✓	x	x	✓
<i>A-languages</i>	x	x	✓	✓	x

These properties are typical for A'-movement and therefore, Richards 1997 assumes that in A'-languages languages, all wh-movement is A'-movement. I adopt this assumption. Similar to Richards 1997, I propose a structure where all wh-words move to SpecCP. This means, I assume multiple CP-specifiers which serve as landing sites for several wh-words.



All SpecCPs in CP have pure A'-properties and receive one wh-word each. Importantly, I do not assume, like Pesetsky 2000, that there are only as much specifiers as overtly raised wh-words. I claim that this structure is purely at LF and independent from what is moved on PF. Additionally, unlike Rudin 1988, I do not claim that the wh-words cluster in some way and land all together on one SpecCP. My claim is that there are multiple SpecCP positions, all A'-positions, and each wh-word A'-moves to one of them. This is the case for all A'-languages, i.e. Bulgarian-like, Chinese-like and English-like.

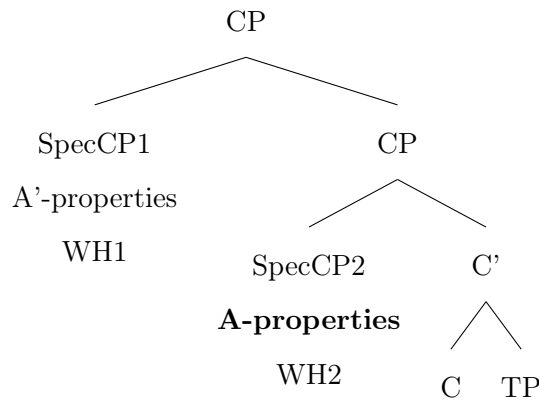
6.1.2 A-languages

A-languages (B/C/S-like, Japanese-like and German-like) are the more interesting ones. Those are the languages where the second and third wh-words do not properly A'-move to CP but are also not properly inside TP. Additionally, it is not really clear what happens to the first wh-word. In some analyses, it A-moves first before A'-moving to CP (Richards 1997), in others, it A'-moves to CP whereas all others move somewhere else (Bošković 1997). In A-languages, following the data of Richards

1997, there is no Superiority between the wh-words, there are no WCO-effects, wh-islands do block movement and raising all wh-words out of an embedded clause is not possible. Additionally, they exhibit A-scrambling. (Modified table taken from Richards 1997, p.45):

	Superiority	WCO	Wh-islands block mvt	A-Scrambling	Interacting wh- dependencies
<i>A'-languages</i>	✓	✓	x	x	✓
<i>A-languages</i>	x	x	✓	✓	x

There have been several accounts proposing solutions for this issue, most of them listed in chapter 2. I will, again, follow the general idea of Richards 1997, namely, that in these languages, wh-words undergo A-movement. Richards 1997 proposes that all wh-words first adjoin to TP and then one of them A'-moves up to CP. The other wh-words remain in TP. Differently to that, I do not think that the second (and third) wh-words do simply adjoin to TP or move to some TP-specifiers. I propose that they move into CP as well - namely into an A-position within CP, which, according to Wurmbrand (To Appear), does exist in certain languages. This position is structurally a SpecCP but has A-properties. These properties make it a target for A-operations, and by that A-movement.



I claim that the first wh-word moves via A'-movement to the highest SpecCP (from here on named the 'first' SpecCP). Whether it moves there directly via A'-movement

or stops somewhere inbetween via A-movement (e.g. to an adjoined TP-position) is left open for the moment and will be discussed later. The second wh-word, contrary to the first one, A-moves to a lower SpecCP which, unlike its higher brother, has A-properties. Importantly, the highest SpecCP always needs to have A'-properties in wh-constructions. From the second SpecCP on, either A'- or A-properties are possible. Everything below the second SpecCP has the same properties as the second one. Hence, if the second SpecCP has A-properties, all SpecCPs below have A-properties as well. On the other hand, if any SpecCP other than the highest has A'-properties, all ones the above have A'-properties as well. I am not sure if it is also necessary to claim that if the second SpecCP has A'-properties then all of the below must have A'-properties too: There could be a change later to A-properties as well. Importantly, what is excluded for sure is the case where a lower SpecCP has A'-properties whereas a higher one has A-properties.

Thereby, the only difference between A-languages and A'-languages lies in the second (and third and following) SpecCP (counting from top to bottom). In A'-languages, all of the SpecCPs are A'-positions whereas in A-languages, only the highest SpecCP is an A'-position whereas all others are A-positions.

6.2 Superiority

Superiority seems to be one of the major issues in multiple wh-questions. As proposed amongst others by Rudin 1988, Bošković 1997 and Richards 1997, Superiority exists in A'-languages here whereas it does not in A-languages. If we assume that all wh-words are located in different SpecCPs, which can have A'- or A-qualities, then the difference of Superiority can be explained. A'-languages have, as described above in detail, two (or more) CP Specifiers with A'-properties. Hence, all wh-words are in an A'-position and undergo A'-movement. Since the type of movement is the same, they compete against each other and whatever it is that moves the wh-words to the CP-domain will raise the highest wh-word first. In A'-languages, Superiority

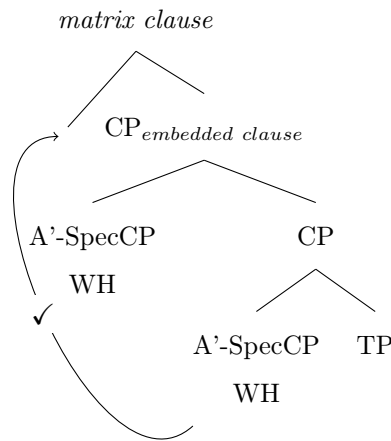
of wh-words reflects the base-generated order. This means, the subject is the highest and so on.¹ In A-languages, there seems to be no Superiority in local multiple wh-questions. I claim this stems from the fact that the wh-words undergo a different kind of movement and do therefore not compete against each other. Only one wh-word is A'-moved to the highest SpecCP - which one this is, does not matter. The others undergo A-movement to lower SpecCPs and do therefore not interfere with the A'-movement of one wh-word. This idea will be analysed in a more fine-grained way later. For the moment, a raw assumption like this could be made: Similar to the proposal in Richards 1997, all wh-words could first adjoin to TP (via A-movement) and then move up to CP (the first via A'-movement, the others via A-movement). If this is the case, C A'-attracts one wh-word, it does not need to be the highest but can be each one. If the wh-words are first A-moved into TP then they are equidistant from C and thus each can be attracted first. Hence, no Superiority arises. All other wh-words are A-moved to CP afterwards. On the other hand, it could also be the case that the first wh-word does not undergo A-movement first but A'-moves straight into CP. In section 6.6, I will bring forward five hypotheses of how this movement could take place exactly and how Superiority can be explained in each hypothesis.

6.3 Wh-Islands Block Movement

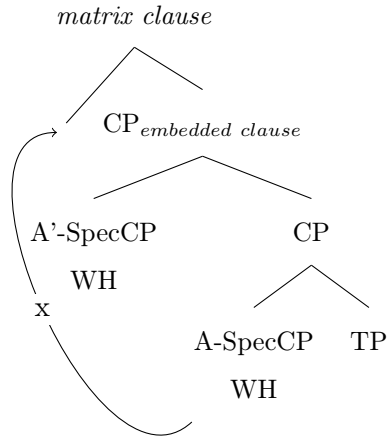
In A'-languages, wh-islands do not block other wh-movement whereas in A-languages, they do. This can be explained with the idea that in A-languages, there is only one A'-SpecCP whereas in A'-languages, there is multiple A'-SpecCPs. If a wh-word

¹There is an interesting phenomenon, mentioned in Rudin 1988 and Richards 1997, namely that there is Superiority between the first and the second wh-word but not between the second and the third. Krapova and Cinque 2005 claim that there actually is a hierarchy between the second and third wh-word as well which is simply not as strong as the one between the first and the second wh-word. In one of my hypotheses proposed in the next chapter, I will mention a possible explanation for this phenomenon.

wants to move out of an embedded clause, it needs to escape the embedded CP via a left-edge position. This position is usually SpecCP, having A'-properties and therefore allowing things to move through it to another clause. In A'-languages, there are multiple of these A'-SpecCPs, one for each wh-word. They all serve as left-edge position since they are equal in their A'-qualities. I am assuming here that each SpecCP counts as an edge if it contains A'-properties. If the second wh-word wants to move to the higher clause it is not blocked by the first wh-word in the highest SpecCP since it has its own A'-SpecCP serving as an escape hatch. This is the same idea as proposed in Richards 1997.



In A-languages, on the other side, there is only one A'-SpecCP, the others are A-SpecCPs. Only the A'-SpecCP counts as an edge, whereas the others do not. They are blocked by the A'-SpecCP. This is due to the assumption that there is an A'-part of the CP and an A-part. For wh-constructions, both are needed (for declaratives, the top-part can be missing). Whatever wants to leave the embedded clause needs to move through the A'-SpecCP part of the CP. In A-languages, there is only one A'-SpecCP. If this position is already occupied with another wh-word (or its trace), there is no chance that anything else could move through it. Hence, all wh-words positioned in embedded A-SpecCPs cannot move out of their CP since the only A'-SpecCP is already occupied by another wh-word.



6.4 Long-Distance Questions

In A-languages, there is an interesting phenomenon found in long-distance questions. They behave fundamentally different than local wh-questions in that they suddenly pattern with A'-languages: They show Superiority and WCO-effects. Take this B/C/S example from Richards 1997 for Superiority in long-distance questions:

- (78) a. Ko si koga tvrdio da je istukao?
 who AUX whom claimed that AUX beaten
 'Who did you claim beat whom?'
 b. * Koga si ko tvrdio da je istukao?
 whom AUX who claimed that AUX beaten
B/C/S (Richards 1997: p. 51; originally from Bošković 1997)

This means, their wh-words are not A-moved but A'-moved in these cases. Hence, their CP resembles the one of A'-languages in that it only has Specifiers with A'-properties. The question is, how do those wh-words suddenly A'-move instead? If we assume that they are in an A-position in local wh-questions then this A-position underlies the regular CP-barrier. It does not, like the A'-SpecCP, act as a Left Edge and hence, wh-words moved into it cannot escape the CP. Wurmbrand (To Appear) mentions that in certain instances of Hyper-ECM, the ACC-DP can move out of the embedded clause into the matrix clause by A-movement. Additionally,

in Hyperraising, DPs also move across the CP boundary via A-movement. This option is not available for wh-movement. I claim that this is due to the same fact as proposed above with wh-islands: CP has an A'-part and an A-part. Wh-constructions need both (possibly due to selection). In declaratives, the A'-part can be absent (as is the case in CCA). The lower part of the CP can be an A-domain in A-wh-languages as well as in CCA-languages. In CCA-languages, then, the A'-part can be absent which pushes the A-SpecCP to be at the edge and allow movement out of it. In wh-constructions, the A'-CP-part must always be present; If a wh-word wants to leave the embedded clause, it must first move to this A'-position and from there it can escape the clause. Thus, if any wh-word wants to move higher up into the matrix clause, their movement needs to be A'-movement, any other derivation would fail.

That long-distance questions in A'-languages show A'-properties is the simple effect of that: A-movement would not be able to reach a higher clause, hence the wh-words must be A'-moved if they want to be moved, therefore long-distance questions only show A'-wh-movement. In the following sections I will present different hypotheses on how this movement could take place. However, what is most important for my claim here, is that A-SpecCPs behave like regular A-positions: They do not let anything escape the CP. If something wants to be raised out of the CP, it has, as always, to be in an A'-left-edge position. I assume that in A-languages, the second SpecCP can change to a regular A'-position and let the wh-words pass. How exactly this might be possible can be found in the hypotheses on wh-movement in the following section.

6.5 Weak Crossover

In Richards 1997 we have seen that A-languages do not show WCO-effects whereas A'-languages do. If we assume that in A-languages, the first wh-word A-moves before it A'-moves to SpecCP, this can be predicted correctly. A detailed approach

on how this derivation looks like will follow in the next section. For the moment, let us assume that in A-languages, each wh-word first adjoins to TP (via Scrambling or Focus-movement). This step of movement is A-movement. From there, it is attracted by C and properly A'-moves into A'-SpecCP. This leaves a trace (or copy, depending on the framework) with A-quality. This trace can be used to bind the lower pronoun and by that, a WCO-effect is avoided. In Van Urk 2015, the same is proposed for English: English A'-moves its wh-words and by that, WCO effects arise.

- (79) * Who_i did her_i friend see t?
English (Van Urk 2015: p. 41)

However, if the wh-word is A-moved first, no WCO-effects occur:

- (80) Which girl_i seemed to her_i friend [t to be smart]?
English (Van Urk 2015: p. 41)

In (80), the wh-phrase *which girl* is A'-moved to SpecCP like each other wh-phrase. However, it first stops in the SpecTP of *seem* and thereby has an intermediate A-stop. It leaves a trace/copy in this A-position which then binds the lower pronoun *her*. Assuming that in A-languages all wh-words first A-adjoin to TP, the same idea can be applied: The wh-word, before being A'-moved to SpecCP, leaves a trace/copy in the A-position in TP which can bind a pronoun and thus avoid WCO.

On the other hand, one could also assume a composite-probe account as in Van Urk 2015: C has both, an A- and A'-feature. Whatever agrees with it inherits those features. Hence, if one wh-word is raised, it must have both properties, A- and A'. Due to its A-feature, it might avoid WCO-effects.

6.6 What Makes All Those Wh-Words Move? Five Hypotheses

I am assuming that wh-movement might take place via multiple Agree (as proposed, amongst others, in Hiraiwa 2001). This means, C probes for wh, finds it, raises it to SpecCP and keeps on probing, finds another one, tucks it in below the highest SpecCP into a new SpecCP and so on. This is the basic assumption on which some of the following hypotheses will be built on. Additionally, I claim that all wh-words must move to CP, at least at LF. This is due to the fact that they need to be interpretable which can only happen if they are located in CP. For A'-languages, applying this assumption is easy: It follows Richards 1997 completely. C has only [A']-features and probes down for wh-words which also only have [A']-features. It finds the first one (due to *Attract Closest*) and A'-raises it to SpecCP. Then it goes on with probing, finds the second one and tucks it in below the highest SpecCP (due to *Shortest Move*) via A'-movement. Therefore, there is a hierarchy between them. What is more interesting, are the A-languages. The main question about them is, why is it that some wh-words get there via A-movement and some via A'-movement. Is this a property inherent to the wh-words, to the C probe or to the SpecCP landing site?

I will now bring forward five hypotheses in order to account for how all those wh-words move into CP. All of those ideas should be able to explain...

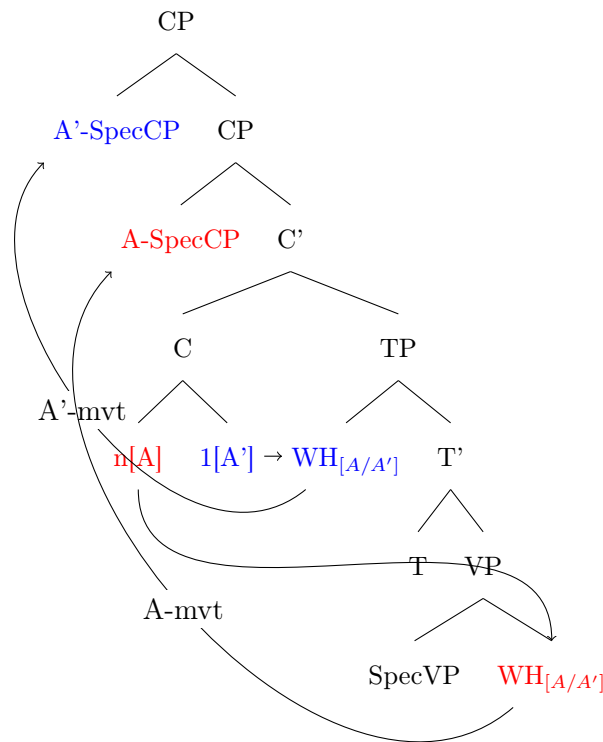
- a. ... why wh-words A'-move in A'-languages but A-move in A-languages
- b. ... that there is Superiority in A'-languages but not in A-languages
- c. ... that long-distance wh-questions in A-languages behave like A'-languages
(in that they show Superiority and all other A'-characteristics)

Concerning c): If we assume that long-distance wh-questions suddenly behave like A'-languages, then we also have to assume that in these cases, all wh-words move to A'-SpecCPs in the matrix clause. This means, the CP environment of A-languages suddenly changes to the CP environment of A'-languages, namely in that it consists

of A'-SpecCPs only.

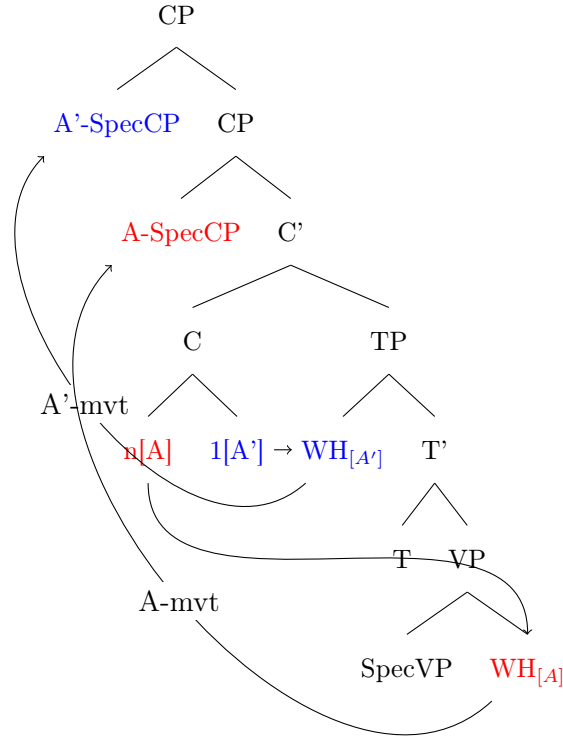
I have to admit that none of my hypotheses is yet able to properly account for all three criteria and that they therefore need to be thought through and for now should only be taken as rough ideas. Out of the five hypotheses presented below, the fifth one seems to be the most promising. The main point of my thesis is not to exactly explain why what moves where but rather to present a model of possible wh-landing-sites inside CP, leaving open how they get there.

Hypothesis 1: In A-languages, all wh-words bear both, an [A] and an [A']-feature. C has a single, unsaturated [A']-probe but multiple unsaturated [A]-probes. The wh-words serve as a goal for them. Therefore, C finds one wh-word and A'-raises it and then finds the other wh-words and A-raises them. This puts one wh-word in A'-SpecCP and the others in A-SpecP. The SpecCPs are not predefined as A' or A but become either one by what is moved into them.



The problem with this idea is that it is not able to explain the Superiority-effects in A-languages correctly. In local wh-questions, there are no Superiority effects. If wh-raising would happen in the way just described, then the [A']-probe on C would have to find and raise the highest wh-word first (due to *Attract Closest*) and then go on looking for the others with its [A]. This would render a hierarchy between the wh-words and therefore Superiority. Since there is no Superiority in A-languages, this hypothesis cannot properly describe the derivation. On the other hand, it could describe long-distance wh-questions properly: If the wh-words themselves bears both, [A] and [A'], they can be A'-attracted by the matrix C and A'-moved there to multiple A'-SpecCPs. This would explain why long-distance questions always behave like in A'-languages.

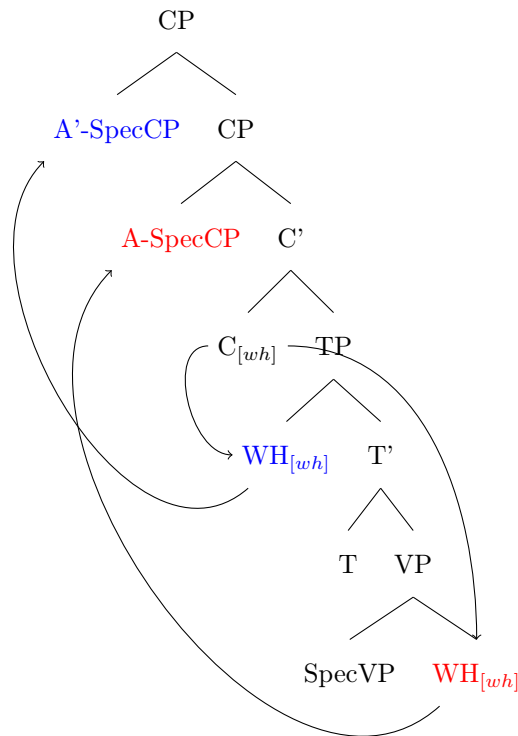
Hypothesis 2: In A-languages, C has only one [A']-probe and multiple [A]-probes. One of the wh-words has [A']-properties whereas the others have [A]-properties. The [A']-probe on C finds the [A']-wh-word and raises it to A'-SpecCP. Then the [A]-probe on C finds the remaining [A]-wh-words and A-moves them to A-SpecCPs below the highest one.



This could explain Superiority: Only one wh-word has A'-properties. So if [A'] on C starts probing, it finds the (only) [A']-wh-word and raises it. Each wh-word can bear the [A']-feature, hence each one could be the one to be raised highest and therefore the first in Superiority. However, the problem of this account is that the A/A' difference cannot be inherent to the wh-words themselves since in long-distance questions, all of them are A'-raised. Therefore, at least in these cases they all need to have [A']-features. If we assume that some of them only have [A]-features, they can never be attracted by an [A']-probe from a matrix clause. Furthermore, what determines which wh-word has [A'] and which one has [A]?

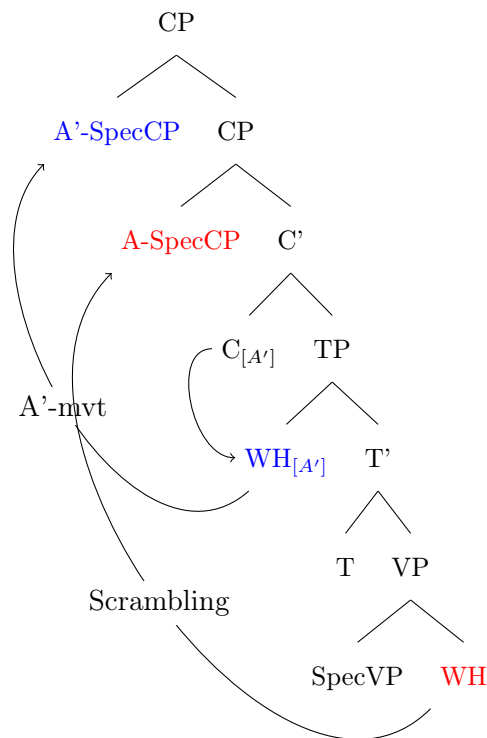
Hypothesis 3: Maybe, however, it is also the case that not the kind of probe or the kind of wh-word defines the movement of the wh-words but the kind of landing site. Let us assume that A-languages have one SpecCP with A'-qualities and one (or more) SpecCP with A-qualities. C attracts all wh-words with a non-specific feature

and moves them up to the SpecCPs. At that point, they are neither A or A'-moved. The feature which attracts them is either both, A and A' or simply undefined, let us call it [wh] for the moment. The first wh-word is attracted and moved to the highest SpecCP which is an A'-position. Due to that, the wh-movement is A'-movement and the wh-word finds itself in an A'-position. The other wh-words are then attracted and moved to CP but there are only A-positions left in CP. So, they are landing in A-positions and therefore their movement is A-movement.



The problem with this idea is that again, it cannot explain Superiority correctly. If all wh-words are attracted by the same feature, the highest one would have to be raised first due to *Attract Closest*. This would render Superiority between the raised wh-words. Additionally, if we assume that the CP structure in A-languages consists of one A'-SpecCP and one or some A-SpecCPs, why is it that in long-distance questions the matrix CP consists of several A'-SpecCPs?

Hypothesis 4: It could also be the case that in A-languages, only the first wh-word is attracted by C and wh-moved whereas all other wh-words move into CP via another kind of operation. Bošković 1997 claims that in A-languages, wh-movement is focus-movement rather than proper movement. Maybe this is actually the case or at least something similar to that is responsible, like scrambling. Let us say, C probes for [A'] only, finds one wh-word and raises it to SpecCP. Then, all other wh-words are moved via Scrambling into the lower A-SpecCPs. Possibly, this could also be true for A'-languages: The first wh-word is properly A'-raised whereas all others are A'-scrambled to lower SpecCPs. This goes along with Richards 1997, who claims that A'-languages allow only A'-scrambling whereas A-languages allow A-scrambling as well (they also allow A'-scrambling).

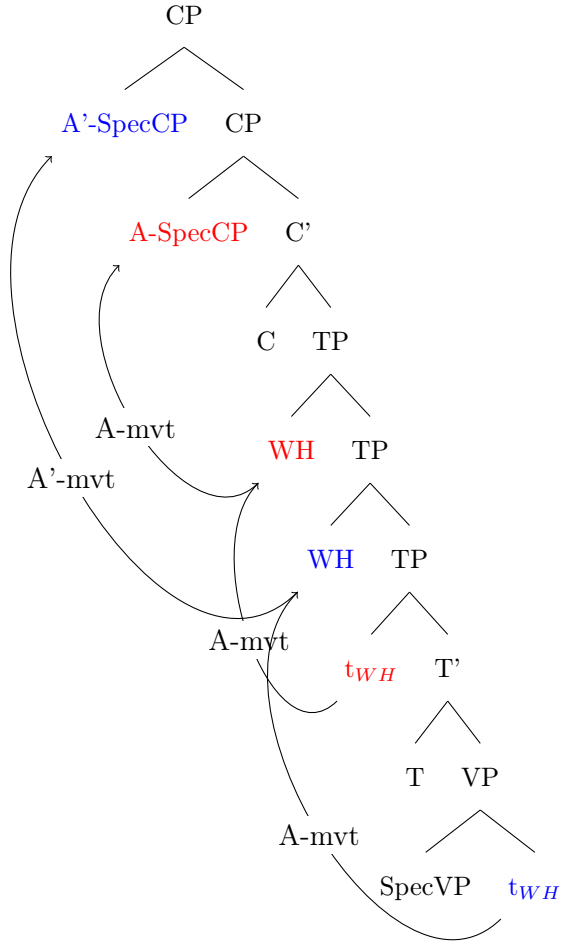


When there is a matrix clause, like in long-distance questions, the wh-words have to undergo A'-scrambling to reach the matrix CP. This is, because the embedded CP is still a barrier for movement to the matrix clause. If all wh-words want to move to

the matrix clause, then one can do this via A'-wh-movement, triggered by a feature on matrix C whereas all others have to be A'-scrambled to reach the highest CP. Naturally, they then show A'-properties like Superiority. Unfortunately, I do not know how to explain the lack of Superiority in local wh-questions in this account. If C searches for one wh-word to saturate its features, it should, due to *Attract Closest*, again find the highest wh-word and not just anyone, which implicates Superiority. On the other hand, this idea could explain why in A'-languages (or, at least very visibly in Bulgarian) there is Superiority between the first and the second wh-word but not between the second and the third. If the second and third wh-words were moved via scrambling, this would describe the lack hierarchy between them since scrambling can happen to one or the other first.

Hypothesis 5: All the previous hypotheses proposed share one problem. The analysis presented in Richards 1997 claims that in A-languages, all wh-words adjoin to TP first via A-movement and then one of them is raised to SpecCP via A'-movement. This means, they all undergo A-movement first. This explains for example the absence of WCO-effects found in A-languages. WCO is a typical phenomenon observed for A'-movement. If an anaphoric constituent moves to an A'-position over its antecedent, it is no longer in the binding domain, hence is not properly bound and thus the sentence becomes ungrammatical. If the same anaphoric constituent A-moves to a position above the antecedent, it remains in the binding domain and properly bound and the sentence stays grammatical. If, as assumed in the accounts presented previously, the first wh-word always A'-moves whereas all others only A-move, this cannot explain the lack of WCO-effects in A-languages. In A-languages, there is no WCO-effect when there is only one wh-word, which means that this wh-word as well has first undergone A-movement. In order to make these facts coherent with the theory proposed here, even the first wh-word in A-languages must first A-move to some position before it can A'-move to A'-SpecCP. There are different options to

build this into my account. All wh-words might adjoin to TP first, as proposed in Richards 1997, and from there move up to CP. This could be done via Scrambling or Focus-movement. Let us assume that all wh-words first A-move via Focus-movement to TP (they end up adjoined to TP). This idea is taken from Bošković 1997 and 2002 who proposes Focus-driven movement of wh-words to TP. After that focus-movement, C picks the highest wh-word and A'-moves it to the highest SpecCP. By that, there is no Superiority. All wh-words are adjoined to TP and hence equidistant from C. (This is the same idea as proposed in Richards 1997.) However, differently to Bošković 1997, 2002 and Richards 1997, I claim that the remaining wh-words do not stay adjoined to TP but are A-moved to the A-SpecCPs. This is, because all wh-words need to be valued in CP. This account can properly explain the differing behaviour in long-distance questions: Focus-movement is clause-bound. If the wh-words first focus-move, they need to remain in their local clause, hence the highest position they can reach from there is the closest CP. The only way to escape the local CP is via regular straight wh-movement to SpecCP without focus-movement inbetween.



In this account, all wh-words have the option to focus-move to TP before they move on to CP in order to have A-properties and by that avoid WCO-violations and Superiority. If the wh-word do not take this option, their movement is pure, direct A'-movement to CP which is the case in long-distance questions.

In my opinion, this account is the most promising one. It adapts Bošković 1997, 2002 and Richards 1997 but improves them in a way that all wh-words move to CP where they can be interpreted as variables whilst still maintaining their A-quality. Additionally, it can explain all A-language-phenomena like Superiority and long-distance questions. One problem that is unsolved in Richards 1997 is that as soon as the highest wh-word moves up to A'-SpecCP, it should have regular A'-qualities and not allow A-phenomena like WCO. Still, it does. This could be solved by an

idea proposed in Van Urk 2015: If an element first A-moves and then A'-moves, then it is able to avoid WCO-effects. By moving to an A-position first, it leaves a trace or copy which then can be used to bind a pronoun. In our case this means that a single wh-word first adjoins to TP via A-movement and then moves on to CP. The trace or copy it leaves in TP binds the pronoun and no WCO effects arise. What could also be the case, loosely based on Van Urk 2015, is that C has both, A- and A'-features. Whatever is attracted by that C inherits both features and by that is is both, A and A'. This could explain the lack of WCO-effects with single wh-words: Let us assume that in A-languages, C has mixed A/A'-features and wh-movement must satisfy these features. If there is only one wh-word, it must satisfy both features. By that, it has both, A- and A'-quality. That in turn means that a single wh-SpecCP would be a mixed position. As such, it may not trigger WCO.

Conclusion: All those hypotheses have problems explaining the observed phenomena like the lack of Superiority in A-languages and the A'-qualities of A-languages in long-distance questions. They all have certain flaws and probably, none of them is applicable. However, Hypothesis five seems to make the most accurate predictions and will therefore be my preferred account. What remains the most important assumption, is that in A-languages, the first wh-word ends in and A'-position in CP whereas the others are A-moved to an (or more) A-positions in CP. Those landing-sites are the main claim of my thesis and not the syntactic operations that lead to them.

6.7 SP and PL Reading

Citko and Grohmann 2000 and Grohmann 2000 claim that the distinction between two types of languages (with respect to wh-fronting) is whether they allow Single Pair (SP) or Pair List (PL) readings. An SP reading means that the answer to a question is a single proposition. A PL reading means that the answer is a set of propositions:

- (81) Q: Who bought what?
- a. SP: Mary bought a sweater.
 - b. PL: Mary bought a sweater, Jane bought shoes, Anne bought a skirt,...
- (Citko and Grohmann 2000: p. 2)

Languages differ in whether they allow SP or PL answers. English, for example, does not allow SP readings in multiple wh-questions whereas Japanese does:

- (82) *Scenario: John is in a store and off in the distance sees somebody buying an article of clothing, but he does not see who it is and neither does he see exactly what is being bought. He goes to the shop assistant and asks:*
- a. *English:* #Who bought what?
 - b. *Japanese:* Dare-ga nani-o katta no?
'Who bought what?'
 - c. **Answer:** Mary bought a sweater.
- (Citko and Grohmann 2000: p. 2-3)

This means, English disallows SP readings whereas Japanese allows them. With respect to that, Bulgarian patterns with English whereas B/C/S patterns with Japanese:

- (83) *Scenario from above*
- a. *Bulgarian:* #Koj kakvo e kupil?
'Who bought what?'
 - b. *B/C/S:* Ko je šta kupio?
'Who bought what?'
- (Citko and Grohmann 2000: p. 3)

This renders the following pattern:

	Japanese, SC, Hindi, Polish	Bulgarian, Romanian, (German?)
SP	✓	x
PL	✓	✓

Citko and Grohmann 2000 suggest that overt syntactic movement forces PL readings whereas *wh-in-situ* languages are the only ones which allow SP readings. They assume that The Question Particle *Q* is responsible for this distinction. If it has scope over both *wh*-words, it allows SP reading, if it has scope over only one or no *wh*-word, it forces PL reading. Citko and Grohmann 2000 assume that *Q* is base-generated below SpecCP (which is the target-position for any moved *wh*-word). Hence, only in languages where no *wh*-word moves, *Q* can retain scope over both of them. As soon as one or both of the *wh*-words move, *Q* does not have scope over both any more and the SP-reading is lost.²

Really interesting about this, even though analysed very differently in Citko and Grohmann 2000, is the division of the languages: It is the same distribution as in Richards 1997. The A-languages (B/C/S, Japanese, Hindi, Polish) pattern together in that they allow SP-readings and PL-readings whereas the A'-languages (Bulgarian, Romanian) pattern together in that they only allow PL-readings.³ Trying to find an exact explanation of why A'-languages do not allow SP readings whereas A-languages do would go beyond the scope of this thesis. However, I think it is interesting to mention that this seems to be the case.

6.8 Summary

In this section, I presented a new idea of landing-sites for *wh*-words. The assumption is based on a combination of proposals from Richards 1997 and Wurmbrand (To Appear). I adopt the claim brought forward in Richards 1997, that concerning multiple *wh*-questions, languages divide into two types: A'-languages which A'-move

²Citko and Grohmann 2000 assume that *Q* marks the phrase interrogative and therefore, LF movement of the *wh*-words in *wh-in-situ* languages is not longer necessary. That is how all *wh*-words can remain in their base-positions in *wh-in-situ* languages even at LF.

³According to Citko and Grohmann 2000, German patterns with Bulgarian and Romanian here. This is not what is predicted in my or Richards 1997 analysis. German seems to behave very differently concerning multiple-*wh*-questions in general. This strange behaviour is examined in detail in the last chapter of this thesis.

all of their wh-words and A-languages in which certain instances of wh-movement are A-movement. For A'-languages, I completely adopt the account presented in Richards 1997, namely that there are multiple A'-SpecCPs to which all of the wh-words are moved. Contrary to Richards 1997, however, I claim that in A-languages, all wh-words move to CP as well - however, this movement is still (partially) A-movement. This is based on the idea that certain languages allow A-positions inside CP. (See, amongst others, Van Urk 2015 and Wurmbrand (To Appear)). I propose that in A-languages, one wh-word (the highest) moves to an A'-SpecCP whereas all other wh-words move to A-SpecCPs. This assumption is able to give an explanation for the lack of Superiority effects, the fact that wh-islands block wh-movement, the different behaviour of long-distance questions as well as the absence of WCO-effects in A-languages. Eventually, I bring forward five hypotheses of how to analyse the movement operations observed. The most promising hypothesis is based very closely on the assumption made in Richards 1997 that all wh-words A-move before any of them can A'-move. He claims that all wh-words first adjoin to TP from where one then A'-moves up into CP. I propose a slightly expanded version of this idea. I claim that all wh-words first A-move via Focus-movement (similar to Bošković 1997, 2002) and adjoin to TP. From there, one wh-word is A'-moved to the highest SpecCP, having A'-properties.⁴ It might be the case, based on Van Urk 2015, that this highest SpecCP has both, A- and A'-properties. I leave this open to further investigation. Then, the other wh-words are A-moved to lower SpecCPs, having A-properties. The last assumption differs from Richards 1997 who claims that the remaining wh-words stay adjoined to TP. My approach renders a structure where all wh-words are inside CP, where they can be interpreted. I assume that in wh-constructions, CP always needs to have a A'-SpecCP which blocks the lower A-SpecCPs from being Left Edges. Hence everything that wants to escape the CP needs to move

⁴I can and will not make a final claim on the question whether this is the case in all A-languages. There might be instances of languages where no wh-word moves to CP but all of them remain adjoined to TP. A similar thing is assumed in Richards 1997 for Hungarian.

through this A'-SpecCP. This explains why wh-islands block A-wh-movement and the different behaviour of long-distance wh-questions. Additionally, focus-movement is local, hence anything that is first focus-moved cannot be raised out of its local CP. This as well gives an explanation why long-distance wh-questions behave like A'-languages - in order to escape a clause, wh-movement cannot be focus-movement first. Based on Van Urk 2015, I propose that C has both, A- and A'-features which attract the wh-words via multiple Agree.

In the last part I present some work of Citko and Grohmann 2000 and Grohmann 2000 on SP and PL reading. Interestingly, languages seem to pattern similarly to the A'/A-distinction when it comes to SP and PL readings. I do not make a final claim of how the two phenomena could be related but leave this observation open for further discussion. The observation, however, seems to be interesting.

The main idea of this chapter is analysing all wh-words of multiple wh-constructions inside CP. This idea is new and combines two accounts from Richards 1997 and Wurmbrand (To Appear). The approach brought forward here seems to be able to explain some of the observed behaviours of A-languages and opens up a connection between CCA and multiple wh-constructions: In CCA, there seems to be an A-position inside CP as well. This A-position might be the same as the one in A-wh-languages. Using my idea of A-wh-movement into CP could tie together CCA phenomena and A-wh-movement: If they both need an A-position in CP, then maybe the same languages exhibiting CCA will also allow A-wh-movement. This typological connection will be discussed in the following chapter.

7 Typological predictions

In the above chapter I brought forward the assumption that in some languages, multiple wh-movement is partially A-movement and that this A-movement targets an A-position in CP. This is not a completely new idea and mostly taken from Richards 1997 who proposes that wh-movement is A-movement in a set of languages as well as from Wurmbrand (To Appear) who introduces an A-position in CP. What is new is bringing those two ideas together by assuming that an A-position in CP (as suggested amongst others in Van Urk 2015 and Wurmbrand (To Appear)) can serve as landing position for those A-moved wh-words. Wurmbrand (To Appear) suggests that CCA like Hyper-ECM or Hyperraising involve an A-position in CP. If one assumes that this is the same A-position as the one targeted by A-moved wh-words, then there should be a connection between languages allowing CCA and languages A-moving their wh-words. If this A-position is inherent to the structure of one class of languages, then it might be that this class of languages allows both, CCA and A-wh-movement. This is a very strong implicational constraint and probably cannot be held up this way. Therefore, I do not propose that all languages that allow CCA also must allow A-wh-movement or the other way round. A claim like this would require a better established theory on A-movement to A-SpecCP as well as a bigger sample of languages examined. Since this task goes beyond the scope of this thesis, I looked only at a couple of languages which seem to exhibit a correlation between CCA and A-wh-movement. From this small set of languages no typological universal can be drawn but the fact that CCA and A-wh-movement seem to somehow correlate typologically is at least an argument for the idea of having an A-position in CP.

Eventually, I will not be able to make a strong prediction like

(84) *Temporary Generalisation:*

Languages that exhibit CCA also exhibit A-wh-movement (and the other way around)

It will turn out that (84) is too strong and cannot be held up since there will be languages allowing CCA but not A-wh-movement. Hence, any generalisation made must be unidirectional instead of bidirectional. In the end, I will suggest the following final generalisation which is built on the examined language data:

(85) *Final Generalisation:*

Whenever a language A-moves its wh-words, it allows CCA (but not the other way around)

This is due to the fact that (according to my findings) there are not only languages allowing both, A-wh-movement and CCA or neither, but curcially also languages which do not allow A-wh-movement but do allow CCA. On the other hand, I did not find languages which exhibit this pattern the other way around, namely exhibiting A-wh-movement but not CCA. A proper discussion and possible explanation of the results found will be presented in the end of this chapter.

In order to strengthen my hypotheses, I picked a small set of languages and investigated if they showed CCA-phenomena and which type of movement their wh-words undergo. Then I checked if the set of languages that shows CCA-phenomena is identical with the set of languages that A-moves its wh-words. And to some extent, they seem to be congruent. I examined those languages by (some) of the criteria proposed in Richards 1997, which are the following:

	A-languages	A'-languages
<i>Superiority</i>	x	✓
<i>wh-movement out of islands</i>	x	✓
WCO-effects	x	✓

I will first look at languages showing both, A-wh-movement and CCA. Then I will look at those allowing neither. Next, I will present an unexpected class of languages, namely those which do not show A-wh-movement but CCA and finally, I will look at the possible case of such languages which do allow A-wh-movement but not CCA. I did not find an instance of this last class in my small set of examined languages and will be the crucial reason for my unidirectional generalisation. The languages taken into account are Turkish, Japanese, Greek, Hungarian, Korean, Brazilian Portuguese, Mandarin Chinese, Romanian, Bulgarian and English.¹ They will pattern in the following way²:

✓ A-wh-mvt ✓ CCA	✓ A-wh-mvt x CCA	x A-wh-mvt ✓ CCA	x A-wh-mvt x CCA
Turkish		Korean	
Japanese		Braz. Portuguese	English
Greek		Romanian	Bulgarian?
Hungarian		Chinese	

7.1 ✓ A-wh-movement, ✓ CCA

7.1.1 Turkish

Wh-movement Turkish behaves like an A-language concerning wh-movement. It shows no Superiority between the wh-words:

¹I would have liked to look at B/C/S as well but there seems to be no literature on Hyper-ECM or Hyperraising so far. Therefore, I will leave this open to further discussion.

²The possibility for A-Scrambling has been examined for only a few of these languages. However, A-scrambling seems to be, besides Superiority, a very effective criterion to distinguish A-languages from A'-languages. Hence, if this data is used for further work, it should be reexamined with regard to A-scrambling.

- (86) a. Kim Kim-e ne-yi sat-mış?
 who who-DAT what-ACC sell-REP
 'Who has sold what to whom?'
 b. Kim-e kim ne-yi sat-mış?
 c. Ne-yi kim kim-e sat-mış?
 Turkish (Özsoy 1996: p. 4)

Neither can wh-words escape out of wh-islands:

- (87) * Nasıl [t [t [hangi problem-i çöz]-eceğ]-i]-ni
 how which problem-ACC solve-NOM-3-POSS-ACC
 düşün-üyor-sun?
 think-PROG-2.SG
 Intended: 'How do you think which problem you will solve t?'
 Turkish (Özsoy 1996: p. 8)

Given these facts, I claim that Turkish is an A-language when it comes to wh-movement. This means, it moves one wh-word to an A'-SpecCP and the others to A-SpecCPs. Most importantly, this means it has an A-position in CP.

CCA Since Turkish allows an A-position in CP in the environment of wh-movement, it might also allow CCA using this A-position. This is, indeed, the case, as already described in Wurmbrand (To Appear). It shows instances of Hyper-ECM:

- (88) Pelin [dün **Mert-i** sınav-a gir-di diye]
 Pelin-NOM [yesterday **Mert-ACC** exam-DAT enter-PAST C]
 bil-iyor.
 know-PRES
 'Pelin thinks that yesterday, Mert took an exam.'
 Turkish (N. G. Şener and S. Şener 2011: p. 5)

For Turkish, there seems to be a connection between CCA and A-wh-movement. This could mean that both, CCA and wh-movement use the same A-position in CP.

7.1.2 Japanese

Japanese is another good example for my claim, since it has both, A-wh-movement and CCA.

Wh-movement Japanese does not show Superiority effects between its wh-words, which classifies it, according to Richards 1997, as an A-language.³

- (89) a. John-ga nani-o naze katta no?
John-NOM what-ACC why bought Q
'What did John buy why?'
b. ? John-ga naze nani-o katta no?
John-NOM why what-ACC bought Q
Japanese (Richards 1997: p. 31)

Additionally, wh-words in Japanese cannot escape wh-islands:

- (90) ?? John-wa [Mary-ga nani-o katta-ka dooka] siritagatte-iru no?
John-TOP Mary-NOM what-ACC bought whether know-want Q
'What does John want to know whether Mary bought?'
Japanese (Watanabe 1992)

This, as well, classifies it as an A-language concerning wh-movement.

CCA Like already mentioned in Wurmbrand (To Appear), Japanese exhibits CCA, namely Hyper-ECM:

- (91) John-ga [mada **Mary-wo** kodomo-da to] omotta.
John-NOM [still **Mary-ACC** child-COP COMP] thought
'John thought that Mary was still a child'
Japanese (Hiraiwa 2001: p. 72)

³In other literature, amongst others in Jeong 2003, it is claimed that this example is ungrammatical. A wh-adjunct, according to Jeong 2003, cannot precede a wh-argument. The same sentence as below is marked ungrammatical there. I do not know a Japanese speaker to verify the data, hence I will adapt the one presented in Richards 1997 with a '?' instead of a '*'.

Taking these facts together, namely the CCA-phenomenon in Japanese as well as its apparent A-wh-movement, Japanese as well seems to have an A-position in CP which can be used for both, CCA and wh-movement.

7.1.3 Greek

Greek also exhibits both, A-wh-movement and CCA.

Wh-movement Although Greek is not mentioned in Richards 1997 and there is no extensive literature on Greek multiple wh-questions, it seems as if it would behave like an A-language. At least, it does not show Superiority effects and wh-islands seem not to be a problem.⁴

- (92) a. Tin tenia, pjos tin ide pou?
the movie who.NOM it watched.3.SG where
'Who did watch the movie where?'
b. Tin tenia, pou tin ide pjos?
the movie where it watched.3.SG who.NOM
Greek (p.c. Christos Christopoulos; examples adapted from Sinopoulou 2008: p. 225)
- (93) a. Pjiu edose ti o Janis?
who.GEN gave.3.SG what.ACC the Janis.NOM
'Whom did John give what?'

⁴Christos Christopoulos mentioned that (92) and (93) are fine for him but another Greek speaker he asked would mark them with '???'. Additionally, in some other cases, there seems to be a preferred order of the wh-words:

- (1) a. Pote agorase ti o Janis?
When bought.3.SG what.ACC the Janis.NOM
'When did John buy what?'
b. Ti agorase pote o Janis?
What.ACC bought.3.SG when the Janis.NOM
Greek (p.c. Christos Christopoulos; examples adapted from Sinopoulou 2008: (8), p.225)

- b. Ti edose pjiu o Janis?
 what.ACC gave.3.SG who.GEN the Janis.NOM
Greek (p.c. Christos Christopoulos; examples adapted from Sinopoulou 2008: p. 225)

For island effects, I only have an example with one wh-word. However, wh-movement is not possible out of an island (see (94a)). Neither is focus-movement (like in (94b)) which is claimed to be A-movement. Since those two behave similarly, it is likely that they are both A-movement.

- (94) a. * Ti_i ides ton andra pou diavaze t_i?
 what.ACC saw.2.SG the man.ACC that was-reading.3.SG
 'What_i did you see the man that was reading t_i?'
Greek (Sinopoulou 2008: (14a), p.227)
- b. * *Tin efimerida* ida ton andra pou diavaze
 the newspaper.ACC saw.1.SG the man.ACC that was-reading.3.SG
 t_i.
 'I saw the man that was reading *the newspaper*.'
Greek (Sinopoulou 2008: p. 227)

Given these data, I assume Greek to be an A-language concerning wh-movement (in terms of Richards 1997).

CCA Greek allows both, Hyper-ECM and Hyperraising.

- (95) *Hyper-ECM*

- a. *θoro* [ton yani pos ine eksipnos]
 consider-1SG [ACC John COMP be-SG smart]
 'I consider John to be smart'
Greek (Joseph 1976: p.241)

- (96) *Hyperraising*

- a. Ta pedhia arxisan na trexoun.
 the children-NOM started-3.PL SUBJ run-3.PL

'The children started to run.'

Greek (Alexiadou and Anagnostopoulou 1999: p. 21)

(96a) is analysed as Raising out of an finite clause in Alexiadou and Anagnostopoulou 1999. They claim that the subjunctive marker *na* forms a CP and that these cases are no instance of control, hence Hyperraising.

Putting the A-properties of Greek wh-movement together with the fact that it allows CCA, supports my hypothesis that an A-position in CP can be targeted by both, wh-movement and CCA.

7.1.4 Hungarian

Wh-movement Richards 1997 claims that Hungarian is an A-language.⁵ Supportive of this claim is the fact that there are no WCO-effects in (local) Hungarian questions.

- (97) a. Ki szereti az anyját?
who loves the mother-his-ACC
'Who_i loves his_i mother?'
- b. Kit szeret az anyja?
who-ACC loves the mother-his
'Whom_i does his_i mother love?'

Hungarian (Richards 1997: p. 33; originally from Kiss 1987 and Brody 1995)

Neither does Hungarian show Superiority between the wh-words.⁶

⁵However, it behaves slightly different to other A-languages like B/C/S or Japanese since it A-moves all of its wh-words and does not A'-move one of them. This means, in B/C/S, according to the analysis proposed in Richards 1997, all wh-words are adjoined to TP (via A-movement) and then one of them is A'-moved to SpecCP. Richards 1997 claims that in Hungarian, on the other hand, all of the wh-words stay adjoined to TP. I will not investigate this difference in great depth but assume Hungarian to be an A-language due to the A-quality of its wh-movement.

⁶Horvath 1998 mentions that Hungarian does allow wh-movement out of islands which is contrary with its characterization as an A-language; Only A'-languages allow wh-movement out of

- (98) a. Mit ki rendelt?
 what-ACC who-NOM ordered
 'Who ordered what?'
 b. Ki mit rendelt?
 who-NOM what-ACC ordered
Hungarian Horvath 1998: p. 41)

Taking these data together, I adopt Richards 1997 claim that Hungarian is an A-language.

CCA Hungarian does have CCA, namely, it allows Hyperraising.

- (99) A fénymásoló el kell, hogy tűnjön.
 the copier(NOM) dis- needs that appear.SBJ.3.SG
 'The photocopier needs to disappear.'
Hungarian (Den Dikken et al. 2017: p. 15)

By that, Hungarian as well falls into the class of languages which allow both, A-wh-movement and CCA.

7.2 x A-wh-movement, x CCA

7.2.1 Bulgarian

Wh-movement Bulgarian is the main example for an A'-language in Richards 1997. It shows Superiority effects as well as WCO-effects and wh-movement out of islands is possible.

- (100) Superiority

islands.

- (1) a. Kit_i kinek_j akart Mari hogy bemutassunk t_i t_j?
 who-ACC who-DAT wanted Mary-NOM that introduce-SUBJ-1.PL
 'Who did Mary want us to introduce to whom?'
Hungarian (Horvath 1998: p. 41)

- a. Koj kogo vižda?
 who whom sees
 'Who sees whom?'

- b. * Kogo koj vižda?
 whom who sees

Bulgarian (Richards 1997: p. 30; originally from Rudin 1988)

(101) Islands

- a. ? Koja ot tezi knigi se čudiš koj znae koj prodava?
 which of these books REFL wonder-2SG who knows who sells
 'Which of these books do you wonder who knows who sells?'

Bulgarian (Rudin 1988: p. 457)

(102) WCO effects

- a. Koj_i običa majka si_i?
 who_i loves mother his_i
 'Who_i loves his mother_i?'
- b. * Kogo_i običa majka si_i?
 who_i loves mother his_i
 'Whom_i does his_i mother loves?'

Bulgarian (Richards 1997: p. 32)

CCA There are no CCA phenomena in Bulgarian. ECM is not possible, neither into non-finite clauses (introduced by the particle 'da') nor into finite clauses.⁷

- (103) Njama koj / *kogo da otide.
 isn't who-NOM / *whom-ACC to go
 'There isn't anyone to go.'

Bulgarian (Rudin 1986: p. 193)

- (104) Mislja, če tja / *ja e umna.
 think-1.SG that she-NOM / *her-ACC is smart
 'I think she is smart.'

Bulgarian (p.c. Marchela Oleinikova)

⁷Rudin 1986 claims that raising neither is possible, but does not give an example.

Hence, Bulgarian exhibits neither A-wh-movement nor CCA. This can also be predicted by my claim above that A-positions in CP are language-specific and can be targeted by both, wh-movement and CCA. If there is no A-position in CP, none of them is possible.

7.2.2 English

Wh-movement English is, according to Richards 1997, an A'-language. It shows Superiority effects and WCO-effects.⁸

(105) Superiority

- a. Who bought what?
- b. * What did who buy?

English (Richards 1997: p. 46)

(106) WCO

- a. * Whom_i does his_i mother love?

English (Richards 1997, p. 46)

CCA There are no CCA-phenomena in English. Neither Hyper-ECM nor Hyper-raising are possible.

(107) *Hyper-ECM*

- a. * I want that her is happy.
- b. * I want her that is happy.

English (p.c. Sean Anstiss)

(108) *Hyperraising*

- a. * She seems that t is happy.

English (p.c. Sean Anstiss)

⁸However, it does not allow wh-movement out of wh-islands. This is not mentioned in Richards 1997 and therefore, I will as well stick to classifying English as an A'-language.

Therefore, English as well supports my claim that if there is no A-position in CP, it has neither wh-movement nor CCA.

7.3 × A-wh-movement, ✓ CCA

Interestingly, there are several languages which do not A-move their wh-words but do exhibit CCA phenomena. This is not directly predicted by a generalisation which states that either languages have an A-position in CP or they do not. The languages in this class have an A-position in CCA contexts but they do not have it in wh-contexts.

7.3.1 Korean

Wh-movement Korean does show Superiority effects, at least between wh-adjuncts and wh-arguments. A wh-adjunct cannot precede a wh-argument:

- (109) a. Mwues-ul wae ne-nun sa-ess-ni?
 what-ACC why you-TOP buy-PAST-Q
 ‘Why did you buy what?’
- b. * Wae mwues-ul ne-nun sa-ess-ni?
 why what-ACC you-TOP buy-PAST-Q
 Korean (Jeong 2003: p. 131)

Additionally, wh-islands do not block wh-movement in Korean:

- (110) Mary-nun [Obama-ka nwukwu-ul manna-ss-ul-ttay]
 Mary-TOP-NOM [Obama who-ACC meet-PAST-ADN-when]
 natana-ss-ni?
 appear-PAST-Q
 ‘Who did Mary appear when Obama met?’
 Korean (B. Kim and Goodall 2016: p. 2)

Considering these facts, Korean belongs to the A’-language class concerning wh-

movement.⁹

CCA Hyper-ECM in Korean is possible but restricted to certain types of subjects, namely the major subject. The major subject is, according to Yoon 2007, characterized by the following:

- a. Preference for generic/habitual versus episodic interpretation of the sentential predicate
- b. Preference for the lexical predicate within the sentential predicate to be an individual-level predicate
- c. Preference for the major subject to be more salient than grammatical subject

(Yoon 2007: p. 626)

Hence, a major subject like the generic 'monkeys' in (111) below can undergo Hyper-ECM.

- (111) Cheli-nun wonswungi-lul/ka banana-lul cal meknunta-ko
Cheli-TOP monkey-ACC/NOM banana-ACC well eat-COMP
sayngkakhanta.
thinks

'Cheli thinks monkeys love to eat banana.'

Korean (Yoon 2007: p. 630)

In these instances, Korean shows CCA-effects and must therefore have an A-position in CP at least in those cases.

⁹Contrary to that, Richards 1997 classifies Korean as an A-language concerning wh-movement. He claims that since Korean does show wh-island effects in multiple wh-questions but does not give an example for that. (See Richards 1997, p.25). In general, he patterns Korean with Japanese, hence with the A-languages. Since the data I found in other literature argues against that and I do not have a Korean consultant, I cannot check which data is wrong and which is right. Therefore, this section should be treated with caution concerning the validity of the grammatical judgements.

7.3.2 Brazilian Portuguese

Wh-movement Brazilian Portuguese seems to behave like an A'-language concerning wh-movement.¹⁰ It shows Superiority effects between the wh-words:

- (112) a. Quem bateu em quem?
 who hit to whom
 'Who hit whom?'
 b. * Em quem bateu quem?
 to whom hit who
 Brazilian Portuguese (p.c. Ingrid Cisneiro Facchim)

Additionally, it exhibits wh-island effects.

- (113) a. * O que você me perguntou quem pode fazer?
 to whom you me asked who can do
 Intended: 'What_i have you asked me who can do t_i?'
 Brazilian Portuguese (p.c. Ingrid Cisneiro Facchim)

CCA Brazilian Portuguese, on the other hand, allows Hyper-Raising.

- (114) Os meninos parecem que fizeram a tarefa.
 the boys seem.3.PL that did.3.PL the homework
 'The boys seem to have done their homework.'
 Brazilian Portuguese (Nunes 2009: p. 5)

Considering these data, BP seems to only allow an A-position in CP in cases of CCA, not in cases of wh-movement.

7.3.3 Mandarin Chinese

Wh-movement Mandarin Chinese belongs, according to Richards 1997, to the A'-languages concerning wh-movement. This means, it A'-moves all its wh-words. Since Chinese is a wh-in-situ language, there is no data on Superiority of its wh-words. Its word-order is too rigid to show anything on Superiority (Richards 1997,

¹⁰This hypothesis is based on my own collected data and thus I cannot guarantee its validity completely.

p.31). However, Chinese exhibits other phenomena typical for A'-languages. On the one hand, it allows wh-movement out of islands:

- (115) Judou xiang-zhidao shei mai-le sheme.
 Joudou want-know who buy-ASP what
 'Judou wonders who bought what.'
 a. 'For which y, y a thing such that Judou wonders who bought y.'
 b. 'For which x, x a person such that Judou wonders what x bought.'
Chinese (L. L. S. Cheng 1997: p. 195)

On the other hand, it lacks A-scrambling and by that cannot fix WCO effects.

- (116) * Meigeren_i tade mama ei_i.
 everyone his mother love
 'His_i mother loves everyone_i.'
Chinese (Richards 1997: p. 29)

CCA Contrary to its behaviour with wh-movement, Chinese shows instances of Hyperraising.

- (117) Ta keneng [hui qu nar]
 he possible [will go there]
 'It is possible that he will go there.'
Chinese (Ura 1994: p. 302)

Ura 1994 claims that, since the embedded clause is finite, this is an instance of Hyperraising. The predicate 'keneng' (*possible*) does not assign a Theta-role and therefore, the subject must come from the embedded clause (and thematically also belongs there). (See also Ademola-Adeoye 2010, p.74).

7.3.4 Romanian

Wh-movement Romanian, as well as Chinese, belongs to the A'-languages. This means, it exhibits Superiority:

- a. Cine ce a spus?
 who what has said
 'Who said what?'
- b. * Ce cine a spus?
 what who has said
Romanian (Rudin 1988: p. 474)

In addition, Romanian wh-islands do not block wh-movement.

- (118) Pentru care clauza_i vrei să afli cine nu a decis încă
 for which paragraph want-2.SG to learn who not has decided yet
 ce va vota t_i?
 what will-3.SG vote
 'For which paragraph_i do you want to learn who has not decided yet what
 he will vote t_i?'
Romanian (Rudin 1988: p. 458)

CCA Romanian, however, does exhibit CCA. It allows Hyper-ECM:

- (119) Am vrut [pe cineva să citească cartea]
 I-have wanted [to-ACC somebody SUBJUNCTIVE read book.the]
 'I have wanted somebody to read the book'
Romanian (Rivero 1991: p. 276)

Hence, Romanian also belongs to the class allowing CCA but not A-wh-movement.

7.4 ✓ A-wh-movement, x CCA

Interestingly, I did not find any language exhibiting this pattern. I have to admit, though, that B/C/S might fall into this group. It behaves like an A-language concerning wh-movement and might not have CCA phenomena. Unfortunately, I am not sure about the latter since I did not find any literature on B/C/S Hyper-ECM, Hyperraising or LDA. I did, however, neither find information about regular ECM and Raising. The only thing that might be interesting is some data I gathered

on ECM; It seems to be possible only with small clauses (120a) and not with full complement clauses (120b)):

- (120) a. Smatram Ivana pametnim.
 consider-1.SG Ivan-ACC smart-INSTR.
 'I consider Ivan to be smart.'
 B/C/S (Lasnik 1998: p. 2)
- b. * Smatram da je Ivana pametan/pametnim.
 consider-1.SG that is Ivan-ACC smart-NOM/INSTR
 Intended: 'I consider that Ivan-ACC is smart.'
 B/C/S (p.c. Violeta Blagojevic; Example adapted from Lasnik 1998: p. 2)

Despite this, I did not gather any data on B/C/S Hyperraising, so there might be instances of it which I do not know about and therefore cannot classify B/C/S into any of the language groups named here.

Interestingly, if we assume that B/C/S does not fall into this group, is that the combination 'no CCA but A-wh-movement' is not attested. This means, there are languages exhibiting CCA but not A-wh-movement but there are no such languages, exhibiting A-wh-movement but not CCA. This is a good reason to believe that maybe, A-wh-movement and CCA do target the same A-position in CP but that they are not dependent on each other in both directions. Therefore, it might be the case that there is a very general A-position in CP for A-wh-movement which can also be used for CCA. The other way around, there might be a A-position for CCA in CP which is rather restricted or not as general and hence cannot be used for wh-movement. I will come back to this in the discussion which is about to follow.

7.5 Discussion

In this section, I examined different languages with respect to their behaviour of wh-movement and CCA. An interesting yet not expected pattern seems to arise. As predicted, there are such languages exhibiting both, A-wh-movement and CCA.

Then, there are languages allowing neither of the two, which was foreseeable as well. Contrary to my expectations, however, there is a large class of languages which A'-moves its wh-words but still shows CCA. The contrary group, namely languages A-moving their wh-words but not exhibiting CCA, is nonexistent in my small amount of collected data.¹¹

✓ A-wh-mvt ✓ CCA	✓ A-wh-mvt x CCA	x A-wh-mvt ✓ CCA	x A-wh-mvt x CCA
Turkish		Korean	
Japanese		Braz. Portuguese	English
Greek		Romanian	Bulgarian?
Hungarian		Chinese	

The number of languages looked at does not allow to make a universal statement about the cooccurrence of the two phenomena, still it seems to show a tendency. This tendency leads me to the proposal that there might be a unidirectional correlation between A-wh-movement and CCA.

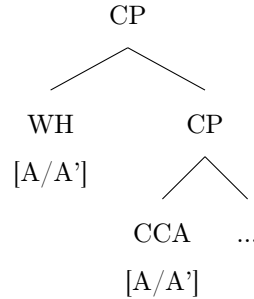
- (121) Whenever a language A-moves its wh-words, it allows CCA (but not the other way around)

This statement subsumes the typological data introduced above. It excludes the non-observed case which is a language with A-wh-movement but no CCA. Again, the amount of languages examined does not allow to make this a universal generalisation. These observations support my framework. If a language A-moves its wh-words, it automatically has an A-position in its CP. This position can be targeted and used by CCA. This implication does not hold the other way around. Connecting these unidirectional implications could result in an analysis like the following¹²: The CP-domain of languages consists of the highest CP, which is A' or A. This CP is targeted

¹¹Leaving B/C/S completely out of consideration.

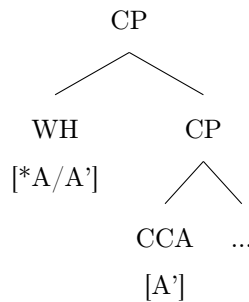
¹²This idea was strongly influenced by personal conversation with Susi Wurmbrand.

by wh-movement. Below, there is a CP for CCA which as well can either be A or A'. I do not want to specify here what nature these CPs have or if they are in fact only different parts of the same CP or stocked specifiers.



Assuming that both CPs can have either property (A or A'), the following restrictions must be made in order to explain the typological observations correctly. One needs find a rule which allows each combination of A/A'-properties except for Wh-CP being A whilst the CCA-CP being A'. This can be done with the following framework:

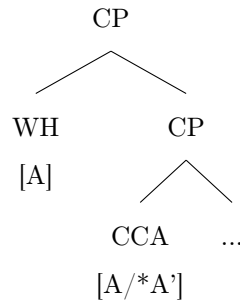
When the CCA-CP has A'-properties, everything above must have A'-properties as well since a higher CP cannot be more restrictive than a lower one. When the CCA-CP is A, there is no CCA. If the CCA-CP is A', the CP above must be A' as well and therefore will not allow A-wh-movement.



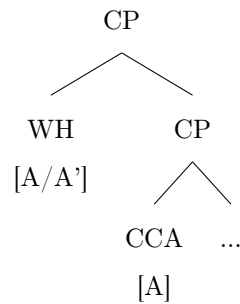
This results in the generalisation in (122) which is supported by the typological observations made above:

(122) If there is no CCA, there is no A-wh-movement.

Contrary, if the highest CP has A-properties, it forces the CCA-CP to have A-properties as well. This is due to the fact the an A'-position cannot be embedded into an A-position in the same clause.



However, if the CCA-CP is A, nothing can be said about the highest CP - it could either have A'- or A-properties.



These two claims result in the generalisation in (123), which describes the typological data correctly:

(123) If there is A-wh-movement, there is CCA.

Assuming this structure, the unidirectional implication between A-wh-movement and CCA, repeated in (124), can be explained correctly.

(124) *Final Generalisation*

Whenever a language A-moves its wh-words, it allows CCA (but not the other way around).

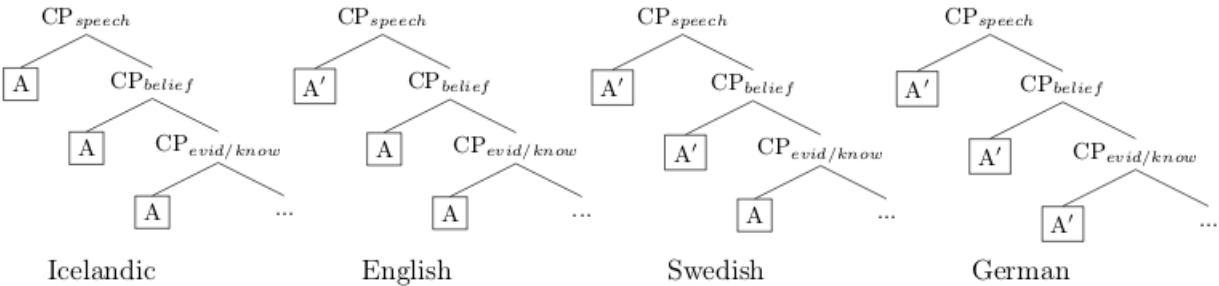
Thereby, it is not the same A-position in CP that is used by wh-movement and CCA together; Two separat CPs need to be assumed, each accountable for only one operation. However, they influence each other and thereby allow to make typological

predictions.

Additionally, this unites the account of Wurmbrand (To Appear) of an ECM-hierarchy and the connection between A-wh-movement and CCA. Wurmbrand (To Appear) brings forward that if an A-position is possible within CP then a complement clause allowing ECM does not necessarily have to be smaller than a CP (contrary to former assumptions) but might also be a CP containing an A-position which can receive case from the matrix clause. We remember from above that regular ECM increases along the *ICH*: If a language allows ECM for *Attitude* complements, it must allow it for *Irrealis* and *Tenseless* as well (Table taken from Christopoulos and Wurmbrand 2020).

<i>regular ECM</i>	say	believe	consider
Icelandic	✓	✓	✓
English	*	✓	✓
Swedish	*	*	✓/?
German, Dutch	*	*	*

Considering this, there are language-specific optional A-positions in CP on different points of the *ICH*, this results in the following assumption: The CP domain differs typologically, depending on which matrix verbs allow an A-position in CP. However, as soon as there is an A-position in CP, it can be used for A-phenomena such as ECM (structures from Wurmbrand and Lohninger 2019, p.15).



This can be combined with my proposal of a wh-CP above all CCA-CPs. I assume

that the wh-CP is higher than all CCA-CPs and they influence each other as proposed above. This includes A-wh-movement into the account brought forward by Wurmbrand (To Appear) and Wurmbrand and Lohninger 2019. Interestingly, the same hierarchy as shown for ECM above holds for Shifted Indexicals: They as well underlie the ICH typologically. Hence, they could be included into the account too. Additionally, this idea explains why CCA seems to need a smaller CP structure than wh-movement. In chapter five, I claimed that in wh-constructions, there must always be an A'-part of the CP above the A-part. In declaratives, this A'-part can be absent: For CCA-languages, this means that whatever is in A-SpecCP is then on the left-edge and can escape the clause via A-movement. This does not work for wh-constructions: They always need an A'-SpecCP to be the highest one and serve as left-edge. This means, that CCA-constructions are smaller than wh-constructions, which is predicted by the analysis brought forward here.

8 Open Issues

Having proposed a theory of multiple wh-movement to CP via A-movement, several questions remain unanswered. I leave open a lot of issues for further research, amongst them the crucial questions:

- What happens on the surface? With the assumption in mind that on LF, everything moves, then why do there occur so many differences at PF? Why are there three types of movement-behaviours (all-fronting, single-fronting, in-situ) and what triggers them? Do the LF properties examined in this thesis influence the PF behaviour? Or, the other way around, does PF-behaviour give the reason why the classes A-languages and A'-languages are not as clear-cut as proposed? This leads directly to the next question:
- Is it possible to divide languages into only two categories, A- or A'-languages? Could it be possible that the distinction between them is not discrete but they build the two ends of a scale on which languages find their place? How could this be tested? What would be the criteria or implicational hierarchies?
- Which hypothesis of the five brought forward in chapter 6.5 makes the best predictions? What could a profoundly integral syntactic theory of A-wh-movement look like?
- What if we look at a larger set of languages? Do the observations and implications met in chapter 7.5 still hold?
- How could the whole idea fit into a split CP structure as proposed in Rizzi 1997?

- How do the languages examined in the typological section behave regarding A-Scrambling? How much has Scrambling to do with wh-movement at LF and at PF?
- What about B/C/S? Is there an instance of CCA? This should be tested with consultants.
- What about those wh-in-situ languages which do not move their wh-words at LF but show inverse linking of wh-words? Do they behave similarly?
- How does D-linking fit into my theory?
- Is there a difference between questions with two wh-words and three wh-words? Does the third wh-word always have to have the same A/A'-properties as the second one?
- How can one explain languages that do not allow multiple wh-questions at all (like Italian)?

Finally, I will now discuss German which behaves very difficultly when it comes to wh-questions. Since it is my mother tongue, I could not resist looking at it in more detail, with the result of being more confused than ever before.

8.1 The Everlasting Problem: German

German seems to pose a puzzle for a lot of analyses illustrated in this thesis. In Richards 1997, it is put together with A-languages like B/C/S since it behaves, to some degree, like an A-language. In Citko and Grohmann 2000, on the other hand, German patterns with Bulgarian when it comes to SP and PL readings. Finally, Grohmann 2000 claims that German does not allow multiple wh-questions at all. Since there seems to be a lot of disagreement in literature about what to do with German, I did not include it into the typological discussion above. I will neither

make a final claim about how to classify German or where it fits in the pattern but only list a number of reasons, why it is complicated to be categorized.

8.1.1 German as an A-language

Richards 1997 claims that German is an A-language and patterns with B/C/S, Japanese, etc. This is due to the fact that to some degree, German multiple wh-questions do behave like A-movement. To begin with, there is no Superiority between the wh-words in local questions whereas there is in long-distance questions.

(125) *Local questions*

- a. Wer hat was gekauft?
who has what bought
'Who bought what?'

- b. Was hat wer gekauft?
what has who bought

German (Richards 1997: p. 46; originally from S.-S. Kim and Sternefeld 1997)

(126) *Long-distance questions*

- a. Wer t glaubt, dass Hans wen gesehen hat?
who believes that John whom seen has
'Who believes that John saw who?'

- b. ?? Wen glaubt wer, dass Hans t gesehen hat?
whom believes who that John seen has
'Whom does who believe that John saw?'

German (Richards 1997: p. 46; originally from S.-S. Kim and Sternefeld 1997)

Next, German allows local A-scrambling which fixes WCO violations. On the other hand, it does not show WCO effects with wh-movement.

(127) *A-scrambling fixing WCO*

- a. *... dass seine_i Kinder jeden_i lieb haben
that his children everyone love

'... that his children love everyone.'

- b. ... dass jeden_i seine_i Kinder lieb haben
that everyone his children love

German (Richards 1997: p. 45)

(128) *WCO with wh-movement*

- a. ? Wen_i liebt seine_i Mutter?
Whom loves his mother
'Who does his mother love?'

German (Richards 1997: p. 46)

On the other hand, German does show WCO-effects in long-distance wh-movement.

- (129) * Wen_i glaubt seine_i Mutter, dass jeder t_i liebt?
whom believes his mother that everyone loves

'Who does his_i mother think that everyone_i loves?'

German (Richards 1997: p. 46)

Finally, not mentioned in Richards 1997, islands block wh-movement in German.

- (130) * Wen weißt du nicht, wann Brutus t erstochen hat?
whom know you not when Brutus stabbed has

'For which person don't you know when Brutus stabbed this person?'

German (D'Avis 1996: p. 92)

Considering all these data, German fits perfectly well into the class of A-languages. However, it does not show instances of CCA and would therefore fall into my class of unattested languages.

8.1.2 SP and PL Readings

Citko and Grohmann 2000 divide languages by the criterion of whether they allow both, SP and PL answers to multiple wh-questions or only PL answers. Interestingly, those languages which allow both, SP and PL answers are the same languages that A-move their wh-words whereas those languages that A'-move their wh-words are the same like those which only allow PL answers. This means, there seems to be

the same split between languages. Unfortunately, there is one exception, namely German. Due to Citko and Grohmann 2000, German patterns with A'-languages in these cases in only allowing PL answers and no SP answers. Citko and Grohmann 2000 claim that German never allows SP readings but always forces a PL reading. This is the same as Bulgarian. B/C/S or Japanese for example allow both, SP and PL readings.

(131) **PL reading**

- a. *Scenario: John sees Jim finishing off his daily sales. He sees a bunch of people that he knows walk away and Jim stacking his left-over merchandise, so he asks him:*
- b. Wer hat was gekauft?
 who has what bought
 'Who bought what?'
German (Citko and Grohmann 2000: p. 4)
- c. Koj kakvo e kupil?
 who what has bought
 'Who bought what?'
Bulgarian (p.c. Aline Panajotov)

(132) **No SP reading**

- a. *Scenario: John is in a store and off in the distance sees somebody buying an article of clothing, but he does not see who it is and neither does he see exactly what is being bought. He goes to the shop assistant and asks:*
- b. # Wer hat was gekauft?
 who has what bought
 'Who bought what?'
German (Scenario from Citko and Grohmann 2000: p. 2; German judgement added)
- c. # Koj kakvo e kupil?
 who what has bought
 'Who bought what?'

Bulgarian (Citko and Grohmann 2000: p. 3)

Putting these together, German clearly patterns with Bulgarian and not with the other A-languages like B/C/S or Japanese. This makes it difficult to unify it with Richards 1997 or my account.

8.1.3 No Multiple Wh-Questions At All

Finally, Grohmann 2000 claims that German does not allow multiple wh-questions at all. There are languages which do not allow multiple wh-questions (like Italian) and Grohmann 2000 proposes that German patterns with those. For Italian, this looks like the following, due to Rizzi 1978 and Calabrese 1984:

- (133) a. * Chi ha scritto che cosa?
 who has written what
 'Who wrote what?'

- b. * Chi è partito quando?
 who is left when
 'Who left when?'

Italian (Grohmann 2000: p. 271; originally from Calabrese 1984)

Similar to Italian, according to Grohmann 2000, multiple wh-questions are no proper wh-questions in German. This assumption is based on the difference between (134) and (135) below.

- (134) *Context: A customer comes to a newsstand and just sees three customers leaving in different directions, each fiddling with their purchases. After asking the salesman whether he sold anything interesting, the salesman replies: „I can't believe it. I sold a Dutch porno magazine, a Japanese comic and a Bosnian newspaper.”*

- a. Wer hat was gekauft?
 who has what bought

'Who bought what?'

German (Grohmann 2000: p. 268)

(135) *Context: A jeweler comes home to his wife and says excited: „I had a great day. I sold a platinum watch, a gold necklace and a brilliant ring.”*

a. # Wer hat was gekauft?
who has what bought

'Who bought what?'

German (Grohmann 2000: p. 269)

The answer to (135) is not felicitous. This is because in (135) no set of individuals has been introduced to the context before. Grohmann 2000 claims that for multiple wh-questions in German, the discourse must provide non-empty sets of referents for both wh-elements. This means a felicitous multiple wh-question is only possible if the set of individuals which the wh-word refers to is common ground between the hearer and speaker. This is the case in (134) but it is not in (135). From that follows, according to Grohmann 2000, that all wh-words in German multiple wh-questions must be D(iscourse)-linked. D-linking means that a wh-word is linked to the discourse. It does not behave like a regular wh-word anymore but moves higher up than SpecCP or gains scope via another operation (see Pesetsky 1987 and Pesetsky 2000). Grohmann 2000 assumes that D-linked wh-words move to TopP and not CP and therefore behave completely different than properly wh-raised words in other languages. He claims that all German wh-words are not "regular" wh-words but D-linked in some kind. Therefore, they move to TopP. This additionally explains the lack of Superiority in German; There is not Superiority between D-linked wh-words, since they do not undergo the regular wh-movement process. Even in Bulgarian and English, which usually obey strict Superiority, D-linked wh-words do not underlie Superiority.

(136) a. Koe momiče koja kniga e kupil?
which girl which book has bought
'Which girl bought which book?'

- b. Koja kniga koe momiče e kupilo?
 which book which girl has bought
Bulgarian (Krapova 2002: p. 2)
- (137) a. Which person bought which book?
 b. Which book did which person buy?
English (Pesetsky 2000: p. 16)

Hence, if all wh-words in German are D-linked, there are no 'regular' multiple wh-questions. Since D-linked wh-words cannot be compared to 'regular' wh-words (as they target different positions), German cannot be integrated into a discussion about multiple wh-movement.

8.1.4 Superiority

The claim brought forward in Richards 1997 is that German does not show Superiority effects and thus patterns with A'-languages. However, this is not entirely correct. As soon as the contexts become more specific or adverbs are involved, German in fact does show Superiority effects. Wiltschko 1997 reveals the following data:

- (138) a. Wer hat denn schon oft was gesehen?
 who has PRT already often what seen
 'Who has often seen what?'
- b. *? Was hat denn schon oft wer gesehen?
 what has PRT already often who seen
German (Wiltschko 1997: p. 107)

It seems that Superiority arises as soon as the wh-words are embedded between two adverbs or at least interfere with them. Wiltschko 1997 claims, contrary to Grohmann 2000 above, that there are non-D-linked wh-words in German. According to her, those are introduced by *was für...* ('what kind of...'). These non-D-linked wh-words show clear Superiority-effects, even without adverbs:

- (139) a. Was für Raucher bevorzugen denn was für Zigaretten?
 what for smokers prefer PRT what for cigarettes
 'What kind of smokers prefer what kind of cigarettes?'

- b. * Was für Zigaretten bevorzugen denn was für Raucher?
 what for cigarettes prefer PRT what for smokers
German (Wiltschko 1997: p. 115)

Those *was für...* wh-constructions interestingly also show WCO effects:

- (140) * [Was für einen Jungen]_i wird sein_i Bruder t besuchen?
 [What for a boy] will his brother visit
 Int.: 'What kind of boy will his brother visit?'
German (Wiltschko 1997: p. 115)

Contrary to that, Richards 1997 claims that German avoids WCO-effects because he classifies it as an A-language:

- (141) ? Wen_i liebt seine_i Mutter?
 whom loves his mother
 'Who does his mother love?'
German (Richards 1997: p. 46)

We see, it is not as easy as proposed in Richards 1997. German definitely shows some A'-behaviour in the cases brought forward in Wiltschko 1997.

Contrary to Grohmann 2000, Wiltschko 1997 claims that not all wh-words are D-linked in German but that 'regular' wh-words definitely exist. She ties this to the difference observed in Pesetsky 1987 for English: There, inherently D-linked wh-words cannot be aggressively non-D-linked.

- (142) a. What the hell book did you read that in?
 b. * Which the hell book did you read that in?
English (Pesetsky 1987: p. 111)

In German, the same behaviour arises:

- (143) a. Was zum Teufel für Bücher hast du gelesen?
 what to-the devil for books have you read
 'What the hell book did you read?'
 b. * Welche zum Teufel Bücher hast du gelesen?
 which to-the devil books have you read

Int.: 'Which the hell book did you read?'

German (Wiltschko 1997: p. 117)

Hence, according to Wiltschko 1997, the difference between D-linking and non-D-linking also exists in German. She claims that D-linked wh-words have been scrambled and therefore are very high in the structure where they do not show Superiority any more. Non-D-linked wh-words, however, did not undergo scrambling, remain inside the relevant syntactic structure and do show Superiority. This can be tested with adverbs: When both wh-words appear before the adverb, they have been scrambled. If, however, one of them occurs after an adverb, it cannot be scrambled, hence not D-linked and should therefore show Superiority effects. This is indeed the case, as observed in Wiltschko 1997:

(144) *D-linked = Scrambled*

a. Wer hat was [denn schon oft gesehen?
who has what [PRT already often seen
'Who has often seen what?'

b. Was hat wer [denn schon oft gesehen?
what has who [PRT already often seen

(145) *Non-D-linked = Not Scrambled*

a. Wer hat [denn schon oft was gesehen?
who has [PRT already often what seen
'Who has often seen what?'

b. *? Was hat [denn schon oft wer gesehen?
what has [PRT already often who seen

German (Wiltschko 1997: p. 128)

There should be a more thorough investigation of which wh-words are scrambled and which are not in German. A proper analysis and thus integration into my framework is limited by the scope of this thesis but will hopefully be pursued in subsequent work on the topic.

The arguments introduced in this chapter show clearly that German exhibits be-

haviour that makes it difficult to fit it into the framework presented in this thesis. Thus, I decided not to include German into my typological investigations. Another thing these works show is that German seems to not pattern completely with one of the classes; A- or A'-languages. This additionally raises the suspicion that a discrete two-way distinction of languages like the one proposed in Richards 1997 might not be enough to cover all typological configurations.

9 Conclusion

In this thesis, I examined multiple wh-constructions. My main attempt was to present a new syntactic analysis for multiple wh-questions. In order to do so, I adopted an account by Richards 1997 who divides languages into two classes concerning their LF: Those which A-move their wh-words and those which A'-move them. Although Richard's account contains some problems and a two-folded division of languages does not seem to be sufficient, I adopt the idea for the sake of the argument. The crucial claim I add to Richard's proposal is the following: Recent work on CCA (mostly Wurmbrand (To Appear)) proposes an A-position inside CP in order to explain A-relations across CP-borders. I use this idea to assume an A-position inside CP to serve as landing-site for wh-movement. This suggestion is new and has some advantages over former approaches. Mainly, it localizes all wh-words inside CP, no matter whether they are A- or A'-moved. Hence, the assumption that at LF, all wh-words must move into CP in order to be interpretable, can be held up for both, A- and A'-languages. I present an exact syntactic analysis for A-wh-movement languages which, despite facing some minor issues, seems to make the right predictions concerning Superiority, wh-islands, long-distance questions and weak crossover effects.

The second part of the thesis deals with typological predictions of the framework presented. Assuming an A-position in CP for both, wh-movement and CCA-phenomena leads to the suspicion that there might be correlation between the two. Examining them typologically shows that this might indeed be the case. Despite the fact that the set of investigated languages is very small, there seems to be a unidirectional im-

plication between A-wh-movement and CCA: Whenever a language A-moves its wh-words, it allows CCA but not the other way around. This means, A-wh-movement seems to entail CCA. I try to give a syntactic interpretation of this generalisation by assuming that the part of CP responsible for wh-movement is higher than the one responsible for CCA and thus conditions its quality.

In the end, I present several works on German and show that it does not properly fit into one of the classes proposed in Richards 1997. It seems to have properties of both, A- and A'-languages, which raises the suspicion that a discrete split between them might not be enough to cover all possible configurations. I leave this direction of thought open to further research as well as several other issues listed in the last chapter.

To conclude, I think my thesis combines two different linguistic observations, namely A-wh-movement and CCA and their syntactic explanations. It depicts a typological correlation between two syntactically independent phenomena and presents an analysis which can explain these connections. It leaves open several questions, amongst them the most striking one of what happens at the surface whilst the LF does all that.

Bibliography

- Ademola-Adeoye, Feyisayo Fehintola (2010). “A Cross-linguistic Analysis of Finite Raising”. PhD thesis. University of KwaZulu-Natal (cit. on p. 99).
- Alexiadou, Artemis and Elena Anagnostopoulou (1999). “Raising without infinitives and the nature of agreement”. In: *Proceedings of WCCFL 18*, pp. 14–26 (cit. on p. 92).
- Anand, Pranav (2006). “De de se”. Thesis. MIT (cit. on p. 54).
- Anand, Pranav and Andrew Nevins (2004). “Shifty operators in changing contexts”. In: *Proceedings from Semantics and Linguistic Theory XIV*. Ed. by R. Young. Vol. 14. Cornell University, CLC Publications, pp. 20–37 (cit. on p. 54).
- Aoun, Joseph and Yen-hui Audrey Li (1993). “Wh-elements in situ: Syntax or LF?”. In: *Linguistic Inquiry* 24.2, pp. 199–238 (cit. on p. 36).
- Baker, Carl L. (1970). “Notes on the description of English questions: The role of an abstract question morpheme”. In: *Foundations of language*, pp. 197–219 (cit. on p. 18).
- Beck, Sigrid (1996). “Quantified structures as barriers for LF movement”. In: *Natural language semantics* 4.1, pp. 1–56 (cit. on p. 10).
- Bondarenko, Tatiana (2017). “ECM in Buryat and the optionality of movement”. In: *Proceedings of the 12th Workshop on Altaic Formal Linguistics (WAFL 12)*. MIT Working Papers in Linguistics. 83 (cit. on pp. 48, 54).
- Bondaruk, Anna (2004). *PRO and control in English, Irish and Polish*. Lublin: Wydawnictwo KUL (cit. on p. 50).

- Bošković, Željko (1997). “Superiority effects with multiple wh-fronting in Serbo-Croatian”. In: *Lingua* 102.1, pp. 1–20 (cit. on pp. 5, 14, 28, 29, 63, 65, 68, 76, 78, 79, 83).
- (2002). “On multiple wh-fronting”. In: *Linguistic Inquiry* 33.3, pp. 351–383 (cit. on pp. 5, 6, 14, 60, 78, 79, 83).
- Brody, Michael (1995). “Hungarian focus and bare checking theory”. In: *Arbeitspapiere des Sonderforschungsbereichs* 340 (cit. on p. 92).
- Calabrese, Andrea (1984). “Multiple questions and focus in Italian”. In: *Sentential complementation*, pp. 67–74 (cit. on p. 112).
- Chang, Lisa (1997). “Wh-in-situ phenomena in French”. PhD thesis. University of British Columbia (cit. on pp. 17, 25).
- Cheng, Lisa Lai Shen (1997). *On the typology of wh-questions*. Taylor & Francis (cit. on pp. 16–19, 33, 99).
- Cheng, Lisa Lai-Shen et al. (2003). “Wh-in-situ”. In: *Glott International* 7.4, pp. 103–109 (cit. on p. 25).
- Chomsky, Noam (1992). “A minimalist program for linguistic theory”. In: *MIT occasional papers in linguistics* 1, pp. 167–217 (cit. on pp. 20, 21).
- (1995). “The minimalist program”. In: *Cambridge et al* (cit. on pp. 9, 20).
- (2000). “Minimalist inquiries: The framework”. In: *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, pp. 89–155 (cit. on p. 21).
- (2008). “On phases”. In: *Current Studies in Linguistics Series* 45, p. 133 (cit. on p. 21).
- Christopoulos, Christos and Susi Wurmbrand (2020). “Germanic infinitives”. In: *Cambridge Handbook of Germanic Linguistics*. Ed. by Richard Page and Michael Putnam. Cambridge: Cambridge University Press (cit. on p. 105).
- Citko, Barbara (1998). “On multiple WH movement in Slavic”. In: *Proceedings of the Annual Workshop on Formal Approaches to Slavic Linguistics. The Connecticut Meeting 1997*, pp. 97–113 (cit. on pp. 15, 60).

- Citko, Barbara and Kleanthes Grohmann (2000). “The non-uniqueness of multiple Wh-fronting: German= Bulgarian”. In: *Proceedings of the Formal Approaches to Slavic Linguistics: The Bloomington Meeting* (cit. on pp. 80–82, 84, 108, 110–112).
- Comorovski, Ileana (1986). “Multiple wh movement in Romanian”. In: *Linguistic Inquiry*, pp. 171–177 (cit. on p. 5).
- D’Avis, Franz-Josef (1996). “On wh-islands in German”. In: *On Extraction and Extraposition in German* 11, p. 89 (cit. on pp. 34, 110).
- Deal, Amy Rose (2017). “Covert hyperraising to object”. In: *Proceedings of NELS 47*. Ed. by Andrew Lamont and Katerina Tetzloff (cit. on p. 48).
- Den Dikken, Marcel et al. (2017). “Predication in the syntax of hyperraising and copy raising”. In: *Acta Linguistica Academica. An International Journal of Linguistics (Until 2016 Acta Linguistica Hungarica)* 64.1, pp. 3–43 (cit. on p. 93).
- Dotlačil, Jakub (2007). “Why clitics cannot climb out of CP: A discourse approach”. In: *Proceedings of FASL 15*. Michigan Slavic Publications, pp. 76–93 (cit. on p. 50).
- Ferreira, Marcelo (2009). “Null subjects and finite control in Brazilian Portuguese”. In: *Minimalist essays on Brazilian Portuguese syntax* 17, p. 50 (cit. on p. 44).
- Fong, Suzana (2019). “Proper movement through Spec-CP: An argument from hyperraising in Mongolian”. In: *Glossa: a journal of general linguistics* 4.1 (cit. on pp. 44, 46–49).
- Grohmann, Kleanthes (2000). “Prolific peripheries: A radical view from the left”. PhD thesis. University of Maryland, College Park (cit. on pp. 80, 84, 108, 112–115).
- Hiraiwa, Ken (2001). “Multiple agree and the defective intervention constraint in Japanese”. In: *MIT working papers in linguistics* 40.40, pp. 67–80 (cit. on pp. 21, 48, 71, 89).
- Horn, Stephen Wright (2008). “Syntax, semantics, and pragmatics of accusative-quotative constructions in Japanese”. PhD thesis. The Ohio State University (cit. on p. 45).

- Horvath, Julia (1998). “Multiple wh-phrases and the wh-scope-marker strategy in Hungarian interrogatives”. In: *Acta Linguistica Hungarica* 45.1-2, pp. 31–60 (cit. on pp. 92, 93).
- Huang, C-T James (1998). *Logical relations in Chinese and the theory of grammar*. Taylor & Francis (cit. on pp. 16, 17).
- İşsever, Selçuk (2009). “A syntactic account of wh-in-situ in Turkish”. In: *Essays on Turkish Linguistics*, pp. 103–112 (cit. on p. 6).
- Jeong, Youngmi (2003). “Deriving anti-superiority effects”. In: *Multiple Wh-Fronting*. Amsterdam: John Benjamins, pp. 131–140 (cit. on pp. 89, 96).
- Joseph, Brian D (1976). “Raising in Modern Greek: A copying process”. In: *Harvard studies in syntax and semantics* 2, pp. 241–281 (cit. on p. 91).
- Kim, Boyoung and Grant Goodall (2016). “Islands and non-islands in native and heritage Korean”. In: *Frontiers in psychology* 7, p. 134 (cit. on p. 96).
- Kim, Shin-Sook and Wolfgang Sternefeld (1997). “Superiority vs. Cross Over”. In: *Workshop on the Minimal Link Condition, Potsdam* (cit. on pp. 29, 109).
- Kiss, Katalin (1987). “Configurationality in Hungarian”. In: *Budapest: Akadémia Kiadó* (cit. on p. 92).
- Krapova, Iliyana (2002). “On D-linking and multiple wh-fronting in Bulgarian”. In: *Proceedings of FASSBL* 4 (cit. on pp. 11, 19, 114).
- Krapova, Iliyana and Guglielmo Cinque (2005). “On the order of wh-phrases in Bulgarian multiple wh-fronting”. In: *Working Papers in Linguistics*, 15, 2005, pp. 171-197. (Cit. on p. 66).
- Lasnik, Howard (1998). “Exceptional case marking: Perspectives old and new”. In: *Proceedings of Formal Approaches to Slavic Linguistics*. Vol. 6, pp. 187–211 (cit. on p. 101).
- Mahajan, Anoop Kumar (1990). “The A/A-bar distinction and movement theory”. PhD thesis. Massachusetts Institute of Technology (cit. on pp. 36, 37).
- Marušič, Franc (2005). “On non-simultaneous phases”. Thesis. Stony Brook University (cit. on p. 50).

- May, Robert and Alan Bale (2017). “Inverse linking”. In: *The Wiley Blackwell Companion to Syntax, Second Edition*, pp. 1–29 (cit. on pp. 17, 25).
- Messick, Troy (2016). “The morphosyntax of self-ascription: A cross-linguistic study”. Thesis (cit. on p. 55).
- Nissenbaum, Jonathan W (2000). “Investigations of covert phrase movement”. PhD thesis. Massachusetts Institute of Technology (cit. on p. 21).
- Nunes, Jairo (2009). “Brazilian Portuguese under minimalist lenses”. In: *Minimalist Essays on Brazilian Portuguese Syntax* 142, p. 1 (cit. on pp. 43, 98).
- Özsoy, A Sumru (1996). “A’-dependencies in Turkish”. In: *VI. Turkish Linguistics Conference* (cit. on p. 88).
- Pesetsky, David (1987). “Wh-in-situ: Movement and unselective binding”. In: *The representation of (in) definiteness* 98, pp. 98–129 (cit. on pp. 16, 18, 113, 115).
- (2000). *Phrasal movement and its kin*. Vol. 37. MIT press (cit. on pp. 9, 10, 32, 63, 113, 114).
- Podobryaev, Alexander (2014). “Persons, Imposters, and Monsters”. Thesis. MIT (cit. on p. 55).
- Pollard, Carl J and Eun Jung Yoo (1998). “A unified theory of scope for quantifiers and wh-phrases”. In: *Journal of Linguistics* 34.2, pp. 415–445 (cit. on pp. 17, 25).
- Richards, Norvin (1997). “What moves where when in which language?” PhD thesis. Massachusetts Institute of Technology (cit. on pp. 1, 2, 5, 6, 8, 25–41, 60–69, 71, 76–79, 82–86, 89–95, 97–99, 108–110, 112, 114, 115, 117–119).
- Rivero, Maria Luisa (1991). “Exceptional Case Marking effects in Rumanian subjunctive complements”. In: *New Analyses in Romance Linguistics, Benjamins, Amsterdam*, pp. 273–298 (cit. on p. 100).
- Rizzi, Luigi (1978). “Violations of the wh-island condition in Italian and subjacency”. In: *unpublished paper* (cit. on p. 112).
- (1997). “The fine structure of the left periphery”. In: *Elements of grammar*. Springer, pp. 281–337 (cit. on pp. 15, 107).

- Ross, John Robert (1967). “Constraints on variables in syntax.” PhD thesis. Massachusetts Institute of Technology (cit. on p. 33).
- Rudin, Catherine (1986). *Aspects of Bulgarian syntax: Complementizers and WH constructions*. Slavica Publishers (cit. on p. 94).
- (1988). “On multiple questions and multiple wh fronting”. In: *Natural Language & Linguistic Theory* 6.4, pp. 445–501 (cit. on pp. 5, 11–13, 15, 28, 33, 34, 60, 63, 65, 66, 94, 100).
- Saito, Mamoru (1992). “Long distance scrambling in Japanese”. In: *Journal of East Asian Linguistics* 1.1, pp. 69–118 (cit. on pp. 35, 38).
- Şener, Nilüfer Gültekin and Serkan Şener (2011). “Null subjects and indexicality in Turkish and Uyghur”. In: *Proceedings of the 7th Workshop on Altaic Formal Linguistics (WAFL7)*, pp. 269–284 (cit. on p. 88).
- Şener, Serkan (2011). “Cross clausal licensing of Accusative case on subjects of CPs in Turkish”. In: *Proceedings of NELS*. Vol. 39 (cit. on pp. 46, 48).
- Shklovsky, Kirill and Yasutada Sudo (2014). “The syntax of monsters”. In: *Linguistic Inquiry* 45.3, pp. 381–402 (cit. on pp. 48–50, 54, 55).
- Sinopoulou, Ourania (2008). “Multiple questions and apparent wh-in situ: evidence from Greek”. In: *Proceedings of ConSOLE XV* 223, p. 246 (cit. on pp. 90, 91).
- Sudo, Yasutada (2012). “On the semantics of phi features on pronouns”. Thesis (cit. on p. 54).
- Sundaresan, Sandhya (2012). “Context and (co)reference in the syntax and its interfaces”. Thesis. University of Tromsø/University of Stuttgart (cit. on p. 54).
- (2018). “An alternative model of indexical shift: Variation and selection without context-overwriting”. University of Leipzig (cit. on p. 55).
- Tanaka, Hidekazu (2001). “Conditions on logical form derivations and representations.” In: (cit. on p. 10).
- (2004). “On the categorial status of raising complements”. In: *York Papers in Linguistics*. Ed. by Johnny Butler et al. York, UK: University of York, Department of Language and Linguistic Science, pp. 213–222 (cit. on p. 57).

- Turek, Mariola (2012). “Multiple spell-out approach to wh-fronting”. PhD thesis. University of York (cit. on pp. 19–24).
- Ura, Hiroyuki (1994). “Varieties of raising and their implications for the theory of Case and agreement”. In: *MIT Occasional Papers in Linguistics* 7 (cit. on p. 99).
- Uriagereka, Juan (1999). *Multiple spell-out*. Ed. by S. Epstein & N. Hornstein (cit. on p. 21).
- Van Urk, Coppe (2015). “A uniform syntax for phrasal movement: A case study of Dinka Bor”. PhD thesis. MIT (cit. on pp. 41, 58, 70, 80, 83–85).
- Watanabe, Akira (1992). “Subjacency and S-structure movement of wh-in-situ”. In: *Journal of east Asian linguistics* 1.3, pp. 255–291 (cit. on pp. 34, 89).
- Wiltschko, Martina (1997). “D-Linking, scrambling und superiority in German”. In: *GAGL: Groninger Arbeiten zur germanistischen Linguistik* 41, pp. 107–142 (cit. on pp. 114–116).
- Wurmbrand, Susi ((To Appear)). “Cross-clausal A-dependencies”. In: *Papers from the 54th Regional Meeting of the Chicago Linguistic Society (CLS 54)*. Ed. by Laura Stigliano Eszter Ronai and Yenan Sun. Chicago, Ill.: Chicago Linguistic Society (cit. on pp. 1, 2, 40–43, 45, 47–49, 52–59, 61, 62, 64, 68, 82–85, 88, 89, 105, 106, 118).
- (2001). *Infinitives: Restructuring and clause structure*. Berlin/ New York: Mouton de Gruyter (cit. on p. 50).
 - (2014). “Restructuring across the world”. In: *Complex visibles out there*. Ed. by Ludmila Veselovská and Markéta Janebová. Palacký University, pp. 275–294 (cit. on p. 50).
 - (2015). “Restructuring cross-linguistically”. In: *Proc. of NELS 45*. Ed. by Thuy Bui and Deniz Özyıldız, pp. 227–240 (cit. on p. 50).
- Wurmbrand, Susi and Magdalena Lohninger (2019). “An implicational universal in complementation— Theoretical insights and empirical progress”. In: *Propositional Arguments in Cross- Linguistic Research: Theoretical and Empirical Issues*. Berlin: Mouton de Gruyter (cit. on pp. 23, 51, 52, 105, 106).

- Yoon, James H (2007). “Raising of major arguments in Korean and Japanese”. In:
Natural Language & Linguistic Theory 25.3, pp. 615–653 (cit. on pp. 45, 97).
- Zeller, Jochen (2006). “Raising out of finite CP in Nguni: The case of fanele”. In:
Southern African linguistics and applied language studies 24.3, pp. 255–275 (cit.
on pp. 48, 49).