The NP/DP-language Distinction as a Scale and Parameters in Minimalism

Hiromune Oda, Ph.D.

University of Connecticut, 2022

The dissertation examines a wide range of phenomena from the perspective of the NP/DP-language distinction, establishing a number of novel typological generalizations, and discussing their relevance to various aspects of the linguistic theory.

Bošković originally proposed a two-way distinction of NP/DP-languages, namely languages with and without definite articles, but Talić (2017) argued that the two-way distinction is not sufficient, with a three-way distinction needed: article-less languages, affixal article languages, and non-affixal article languages. In this dissertation, I first provide additional support for Talić's proposal with two novel typological generalizations: extraction of a conjunct from coordinate structures may be allowed only in affixal article and article-less languages, and what I call compositional indeterminate pronouns may be productive only in languages with affixal articles and without articles. Relatedly to the second generalization, I establish the generalization that large-scale pied-piping is possible in a language only if it has compositional indeterminate pronouns and the projection to be pied-piped is head-final. I also offer deductions of these generalizations, which have consequences for a number of phenomena/mechanisms, e.g. with respect to phases, Agree, labeling, and the syntax of wh-in-situ. Regarding Agree, I argue that it can be deduced from Minimal Search, hence eliminated from the grammar.

Most importantly, I argue that Talić's three-way distinction of NP/DP-languages is not sufficient either; we need a more fine-grained "scale" of NP/DP-language distinction, from canonical to non-canonical DP languages, rather than a two-way or a three-way "cut". I demonstrate that a number of languages (e.g. Italian and Greek) that have been treated as canonical DP-languages by Bošković and Talić exhibit some properties of NP-languages, and argue that the presence/absence of a definite article does not necessarily correspond to the presence/absence of DP in a given language or a construction. I propose that the definite article has an option of not projecting DP, by being base-generated adjoined to another head in the nominal domain, an option which I show

enables us to capture the scale of DP-hood. I also show that the scale approach to the NP/DP-language distinction is an appropriate point of parameterization in minimalism.

The NP/DP-language Distinction as a Scale and Parameters in Minimalism

Hiromune Oda

B.A., The University of Tokyo, 2013

M.A., The University of Tokyo, 2015

M.A., University of Connecticut, 2021

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2022

APPROVAL PAGE

Doctor of Philosophy Dissertation

The NP/DP-language Distinction as a Scale and Parameters in Minimalism

Presented by

Hiromune Oda, B.A., M.A., M.A.

Major Advisor: _____
Željko Bošković

Associate Advisor: _____
Vicki Carstens

Associate Advisor: _____
Ian Roberts

Associate Advisor: _____
Mamoru Saito

University of Connecticut

2022

ACKNOWLEDGEMENTS

I wasn't a good student, and I'm not an ideal person. Still, I managed to write up this dissertation and survive the Ph.D program, solely because I was blessed with people here at UConn Linguistics.

First of all, I would like to express my deepest gratitude to my major advisor, Željko Bošković. Željko has been my ideal advisor; he has always been extremely enthusiastic, inspiring, supportive, accessible, and encouraging. Whenever I talked about my half-baked ideas (or just raw ingredients) in an appointment or in a terribly written term paper, he said "THIS IS EXTREMELY INTEREST-ING (or IMPORTANT)!!!!", and he explained in detail how I could organize, combine, and write them up in the way that they contribute to the field. And whatever I wrote, whether it was a term paper, a conference abstract, a proceedings paper, a generals paper, a journal article, or a disser-tation, he gave me detailed comments on every single point in a flash, which always pointed to a clear direction that I should take. I also enjoyed his comparative syntax seminars, which were volcanic eruptions of his insights and love of syntax. What's amazing is that he kept explaining his ideas and answering questions from the audience with the same amount of heat for three hours (or more), and then he moved onto a meeting with me (and other students), where he never cooled down. I always exhausted all my energy on Tuesdays, especially when I had a meeting with him after his seminar, but my brain was always filled with new knowledge and research topics as well as joy of linguistics. All these have enabled me to come up with this dissertation topic and write it up as the 300-page dissertation, with which Željko helped me literally until the last moment. Even after writing this dissertation, my ideas are still expanding thanks to the stock and influence of all his past seminars and appointments. Now one of my goals as a linguist is to be a volcano like Željko.

I would also like to thank Mamoru Saito, who has been in my advisory committee since my first year. Every time he visited UConn Linguistics, I occupied him with a series of appointments that I asked him for. He must have been jet-lagged after the long travel from Japan, but every single time he kindly and patiently listened to my half-baked ideas, and gave me critical comments from the biggest picture of the theory, which shaped me as a theoretical linguist. There are two major lessons from Mamoru that I loved to listen to and 100% agree with. One is from his seminars and some of the appointments: "the more radical a theory is, the more interesting it is." This actually led me to propose some radical theories in this dissertation, which I hope to be interesting (for him and for other people). The other lesson comes from the UConn Linguistics 50th anniversary event: "UConn people are nice". Yes, this is why I was able to survive and enjoy the Ph.D program here at UConn, and there is no doubt that Mamoru is one of the representatives of such UConn people.

I am also grateful to Ian Roberts, who I/we think is a Husky. I was so lucky to meet him here at UConn and join his seminars and reading group; his insights and knowledge of the linguistic theory, typology, diachronic linguistics, and parameter theories have direct influences on this dissertation, as you can see later (if you ever read it). I also enjoyed discussing Romance linguistics with him, which was essential for my other research projects; in fact, if I had not joined his seminar co-taught with Andrea, I might not have gotten engaged in Romance linguistics. (Well, his assumption that I speak Italian is a good pressure for me to keep learning Italian.) It is my precious memory to have visited his house in Cambridge for a reception after the CamCoS conference and chatted with him and his wife, as well as other guests. I also loved his British joke after my syntax generals defense; "I was expecting too much, but I wasn't disappointed". I hope this dissertation didn't disappoint him (if he ever expected too much).

I am also thankful to Vicki Carstens. She arrived at UConn during the pandemic time, so we didn't have so many opportunities to see each other in person, unfortunately. Nonetheless, she was so kind and encouraging to me. I sent her a prospectus and asked for an appointment around the New Year's Day, which I don't think is a great idea, but she replied to me and had a meeting after the New Year's Day, giving me critical comments on the general picture of my proposals. I actually owe her not only the New Year's Response, but also a Trader Joe's shopping bag that I had left at Andrea's house. I forgot how it ended up at her place, but she kept it and even brought it to the department for me. The bag had some items that were important to feed my son at that time, so she saved his life.

The other faculty at UConn Linguistics were also essential for my Ph.D journey. Stefan Kaufmann, who was my second generals supervisor and a former committee member, deserves special mention. He patiently guided me to crystalize my vague ideas for the semantics generals, waiting for me to take small steps for a long time. And when I encountered a problem that I could't handle by myself, he magically solved it and explained the magic to me in a way that I could understand. Without him, the semantics generals, and a journal article that is based on it, would not have been possible. I still feel deeply bad that I had to confess to him the change of the dissertation committee, but he has nevertheless been supportive to me in all respects, and the work supervised by him is actually cited in this dissertation. I am also thankful to Magdalena Kaufmann. Magda (with Stefan) always encouraged me to do semantics, which I appreciated, and I in fact loved to discuss a lot of interesting data and topics about Japanese and German with her. I am also grateful to Magda and Stefan, as well as Sebastian and Tobias, for welcoming Miyu and Koreaki, who also enjoyed talking with them. Sebastian and Tobias are Kore-kun's role models; they kindly offered one of their toys to him, which Kore-kun has been really grateful for and has played with carefully but also joyfully. He was a lot inspired by their playing the violin and the piano, and he behaved like a good "onii-chan" after he played with them. We breathed an atmosphere of culture and family at their place.

I'm grateful to Jon Gajewski for his support and generosity. At the semantics generals defense, he said "when you submit this to a journal, you should do . . . ", which sounded like it was taken for granted. I tried not to show my emotion at that time, but I was internally really happy, and relieved. He was always supportve as Department Head; especially when I made some mistakes as President of LingClub, he cheered me up and worked hard to resolve the issue for me. It was also a lot of pleasure and honor to work for him when the department hosted the 50th anniversary event, where I confirmed with Mamoru's words that UConn people are nice, as represented by Jon.

I learned a lot from Diane Lillo-Martin and William Snyder. As a syntactician, I was wondering how I could contribute to language acquisition study, especially in mininalism. Discussions with them in their classes in my first year made me continue wondering about it throughout the Ph.D program, and I'm finally addressing issues of language acquisition in this dissertation. I was also fortunate to have a chance to work as an RA for William. Joining a real acquisition study with him encouraged me to think about language acquisition study from a more realistic perspective, which also lets me enjoy seeing how my son acquires Japanese (as well as English at this point) even more.

I cannot proceed without mentioning influences from Adrian Stegovec, who I met when he was a graduate student here. I was astonished every time he gave comments in classes and presentations; every single comment contained significant insights and interesting data that provide an angle that I had never thought about. This inspired me to try to give some comments in as many

classes and presentations as possible inside and outside of UConn, and I realized even more how good Adrian is. But I have kept and will keep trying, which I believe is the only way to get close to his level.

Jonathan Bobaljik and Susi Wurmbrand gave me a lot of lessons too, although I only had one year or so to interact with them here at UConn. Jonathan's comments were always surprisingly critical, but also productive and helpful. I wish I had had more opportunities to talk with him, but I was still lucky that I had one. Susi is my first syntax teacher here at UConn, and I couldn't wait for her class every week. The logical, critical, and encouraging way of her teaching is one of my ideals. I also truly appreciated her words in her mid-semester feedback; "you always have really good things to say - so, just let us know your thoughts and answers a bit more often." At that time I was too shy to speak up (even more so than now), but her words gave me confidence to believe that I can keep doing syntax.

When it comes to teaching, it's impossible not to mention Jon Sprouse. His seminar on experimental syntax made me notice (again) the importance of designing linguistic experiments carefully, as well as fundamental issues of the linguistic theory in the big picture of cognitive science. And I also learned a lot from serving as a TA for his Ling1010. His materials were not only quite accessible to beginners of linguistics, but also very helpful for us graduate students to recognize fundamental issues of linguistics again from the cogitive scientific perspective.

Speaking of Ling1010, I also loved to work as a TA for Harry van der Hulst. Harry's materials centered linguistics in the context of the history of philosophy, which encouraged me to reconsider fundamental issues of linguistics from the philosophical side, and to enjoy philosophy itself. It was also a lot of fun to chat with him in basic Japanese and Dutch in the morning in front of his office; it was a good start of a day to see him showing the big smiles with coffee in his hand.

And yes, coffee. It is now part of my life, and it has become even more so thanks to Andrea Calabrese. It is definitely one of the best memories at UConn to visit his gorgeous house to prepare for Open House and to discuss Italian linguistics, where he continuously fed us with excellent food and coffee. I always loved to listen to his funny stories (which were frequently updated to my surprise). When it comes to linguistics, he was always critical, trying to be a devil for me (and Pietro), but I was/we were very glad when he said "But I don't agree with this!", which was actually a sign of him liking our idea. Moreover, this dissertation would not have been possible without him. His review comments on my prospectus shaped some of the core proposals in chapter 5, and experiences to work for his Ling1030 founded the ground for me to dig into typology. And on top of everything, he gave me (and Pietro) an important thesis to work on in my entire career as a linguist, which had originated from Shige-Yuki Kuroda: Italian and Japanese are the same! It is not an exaggeration to say that I'm working on Italian to prove this Kuroda-Calabrese thesis. I hope I/we can keep him saying "I don't agree with this!" when it comes to Italian linguistics, but "yes, we are the same!" when it comes to Italian-Japanese comparative linguistics.

I would also like to express my gratitude to visiting professors, Jairo Nunes, Norbert Corver, and Guglielmo Cinque. I loved to hear Jairo saying "your points are well taken", and how he discussed Portuguese and some other languages in connection with fundamental issues of the linguistic theory. I also received an important lesson from him; when you have a child, you can never have a good chunk of time for your work, but your concentration and productivity can (and must) somehow improve. At that time I didn't have a child, but now I understand his words from the bottom of my heart. Norbert is an extremely pleasant person to talk with, and I appreciated his humble way of expressing his attitude toward science, which I believe all scientists need to keep

in mind: "I may be wrong, but I think of it this way for now." Dutch linguistics became quite close to me after Norbert demonstrated detailed documentations and analyses of Dutch data from a typological perspective. I, and especially my wife Miyu, are eager to visit Utrecht again, to see him and Nijntje. It was fortunate to have a chance to interact with Guglielmo (well, actually I never called him this way as a polite Japanese who has a tiny bit of sense of the Italian culture, Professor Cinque) before the pandemic. His gentle introduction of astonishingly broad range of typological observations and careful argumentations made me love typology and comparative syntax even more. I also appreciated his casual and funny way of expressing his opinions on linguistics and other stuff. It is my treasure and honor to have welcomed him to a seminar on Italian linguistics at Andrea's office, where we discussed data on a Piedmont dialect from Pietro and my field work.

Special thanks to Tamara Cohen for her invaluable administrative help in various occasions.

UConn Linguistics is definitely a special place for me, but in the first place, it would have been impossible to enter the Ph.D program at UConn without instructions and support by Akira Watanabe and Noriko Imanishi at the University of Tokyo. My interest in comparative syntax, especially comparison of Japanese with Romance, actually originated from Akira's seminars, supervision, and works (well, again, as a polite Japanese, who was advised by him in Japan, I would call him Watanabe-sensei; . . . Mamoru? Sorry, I'm "Americanized" when I talk about and with him in English). To be honest, sometimes I couldn't fully follow his comments on my work at that time, solely because of my poor understanding and knowledge of linguistics. However, as I learned linguistics more, I came to understand what he had wanted to tell me, and I appreciated his supervision and insights all the more. Also, chapter 3 and chapter 4 of this dissertation are based on his works on indeterminate pronouns. So I can safely say that I was led by my two advisors in the US (Željko) and in Japan (Watanabe-sensei) to write this dissertation. He was also frank and friendly to me all the time, which I really appreciated. Imanishi-sensei understood my interests and weakness as a linguist and a person, and even seemed to know about me more than I did. At that time I was the only student of linguistics in the Master's program there, so it was really tough to read all the papers assigned and give presentations on them alone, but that definitely improved my skills of reading papers critically and understanding importance of each work. After I was accepted by UConn Linguistics, she let me read some of Željko's papers in her seminar (or practically appointments), including the 2012 paper on the NP/DP-language distinction (*On NPs and clauses*), which is the point of departure of this dissertation. Probably she already knew at that time that I would write this dissertation. My thanks also go to senior students and alumni of the English Language and Literature Department at the University of Tokyo, who supported and "raised" me with extraordinary kindness to this immature kid: Tomoe Arii, Naomi Harada, Shun'ichiro Inada, Sakumi Inokuma, Yuki Ishihara, Mioko Miyama, Harumasa Miyashita, Chigusa Morita, Motoki Nakajima, Terue Nakato, Chizuru Nakao, Tohru Noguchi, Junya Nomura, and Ayaka Sugawara. Special thanks to Isso Nakamura, who has always inspired and motivated me with his enthusiasm and diligence toward linguistics since we met more than 10 years ago; it looked as if I taught him linguistics, but the truth is that I learned from him more than I could teach him.

I am truly grateful to my cohort, without whom I would not have enjoyed the Ph.D journey this much. Gabriel Martínez Vera is a cool, smart, and pleasant guy from whom I learned a lot as a friend and a linguist. I always respect the way he makes connections with people (and I always wonder how to do it), and my first days at UConn would have been much more difficult without him, as well as Tania Torres. It was a lot of fun to talk with Pasha Koval, who sometimes introduced me different perspectives from mine, which widened my view of the world. I still remember that

always blessed with friends and teachers". Yes, UConn Linguistics is no exception. I can't find words to express my gratitude to my wife, Miyu, who always organized everything for me to work effectively and comfortably. The life here must have been quite tough for her in all the respects, especially at the beginning. She didn't have friends in Connecticut at first when she arrived here (except for Pietro and his family, who helped her from her day 1). She also had a hard time in communicating with people in English at first, and at the same time she had to get used to different culture, food, systems, everything. But she never complained, just behaving like enjoying the new life, especially in front of me. And she worked hard, and everybody in the department welcomed her, for which I'm deeply grateful. Eventually, she even made more friends in the department and Connecticut than me. Once at an Open House, she talked with all prospective students, and she introduced me to them; "this is Hiro, my husband, who is a grad student here". Now I'm an appendix of Miyu at UConn Linguistics. And we welcomed a new family member, Koreaki, in 2018. As Jairo said (and also saw from some of the faculty here), I couldn't spend as much time for work as I wanted after his birth. However, it's nothing, there have been far more rewards than losing my time; every single day, every single moment with Koreaki and Miyu is the treasure of my life. And it was their unconditional love that kept cheering me up and motivating me to work harder. 今まで本当にありがとう。これからもよろしく。

Right before I finish this dissertation, my grandmother, Sakano Oiso, passed away at the age of 91. I dedicate this dissertation to her, without whom this dissertation, and of course I, would not have existed in this world.

Contents

6 Article Drop in Non-canonical DP-languages: P-N Affinity, Uniqueness, and Relevance of Modifiers

Chapter 1

Introduction

1.1 Background: The NP/DP-language distinction

This dissertation is a broad typological investigation of the categorial status of the nominal domain, which will be shown to have consequences, through typological correlations, for a number of phenomena. The point of departure is the so-called NP/DP-language distinction. Bošković (2008b, 2012) argues that there is a structural difference in nominal phrases between languages with definite articles and languages without definite articles. His main argument is based on a number of syntactic and semantic cross-linguistic generalizations that he establishes, some of which are given in (1).[1] All his other generalizations also concern syntax and semantic properties, such as NEG-raising, clitic doubling, the superiority effect of multiple wh-fronting, sequence of Tense, etc. (see Bošković 2012, 2016b for more generalizations, some of which are noted below).[2]

(1) a. Only languages without definite articles may allow adjunct extraction out of a nominal phrase.

     b. Only languages without definite articles may allow Left Branch Extraction of an adjective

---

out of a nominal phrase.[3]

c. Only languages with definite articles allow the majority superlative reading.

d. Head-internal relatives display island-sensitivity in article-less languages, but not in languages with articles.

e. Radical pro-drop is possible only in article-less languages.

f. Second-position clitic systems are found only in article-less languages.

To illustrate, adjunct extraction out of a nominal phrase is disallowed in English, which has definite articles, as shown in (2a), whereas it is allowed in Serbo-Croatian, which lacks definite articles, as seen in (2b). Likewise, as shown in (3a), Left Branch Extraction of an adjective out of a nominal phrase is blocked in English, whereas it is possible in Serbo-Croatian, as shown in (3b).[4]

(2) a. *[From which city]$_i$ did Peter meet [t_i girls]?

b. [Iz   kojeg grada]$_i$ je Ivan sreo [djevojke t_i]?
   from which city   is Ivan met   girls   (Bošković 2008b)

(3) a. *Expensive$_i$ he saw [t_i cars].

b. Skupa$_i$ je vidio [t_i kola].
   expensive is seen   car   (Bošković 2008b)

For those cross-linguistic generalizations, what counts as a definite article is important. The definition of a definite article is given in (4), which is taken from Bošković (2016b).

---

3. See also Uriagereka (1988) and Corver (1990).
4. Regarding Left Branch Extraction of adjectives, Franks (2007) observes that Colloquial Finnish has developed a definite article and disallows LBE, unlike literary Finnish, which does not have a definite article and allows LBE. The history of Ancient Greek is also important here. Based on Taylor's (1990) corpus study, Bošković (2012) observes that the Homeric Greek, the period in which Ancient Greek did not have articles, allowed LBE, while Koine Greek, where Ancient Greek developed definite articles, disallowed it. Bošković (2012) takes these observations (regarding diachronic changes in Finnish and Ancient Greek) as strong evidence for his generalization regarding the correlation between the availability of definite articles and adjective LBE (see also Bošković 2012 for the list of languages that allow adjectival LBE, all of which lack definite articles).

(4) DEFINITION: A *definite article* (i) has the meaning of an iota operator, (ii) obligatorily occurs in a nominal phrase with a definite interpretation, (iii) occurs only once in a nominal phrase, and (iv) has a form distinct from demonstratives.

As Bošković notes, (4ii) follows from (4i) and Chierchia's (1998) proposal regarding type-shifting of NPs. Chierchia proposes that covert type-shifting operations in general are only possible in the absence of a lexical item that has the same function; in other words, presence of a lexical item that works as a type-shifter blocks application of a covert type-shifting operation. Since a definite article is a lexical item that functions as an iota operator (i.e., it is a type-shifter from type $\langle e,t \rangle$ to type e), covert type-shifting from type $\langle e,t \rangle$ to type e is blocked in languages that have definite articles. Thus, a definite article must be present for definite interpretation in languages that have definite articles. On the other hand, in languages that lack definite articles, the iota operator can apply covertly, yielding definite interpretation. This is illustrated in Serbo-Croatian (5), where the subject and the object, both of which are bare nominals, have definite interpretation.

(5) Mačke razbiše prozor.
     cats   broke   window
     '**The** cats broke **the** window.'   (Bošković 2016b)

Note here that, as mentioned above, Bošković's generalizations concern syntactic and semantic phenomena. Bošković (2008b, 2012, 2021a) takes this as indicating that the relevant cross-linguistic differences are not simply phonological, where definite articles would simply be null in languages such as Serbo-Croatian. If languages with no overt definite articles had a phonologically null definite article and the only difference between languages with and without definite articles were phonological (i.e., whether the definite article is overt or covert), the semantics of superlatives and locality of movement, for instance (cf. the generalizations in (1a-d)), would need to be phonological phenomena. In addition, as discussed in Bošković (2012), extraction patterns are completely different between languages with and without definite articles, as partially described in (1a) and (1b). If the difference between languages with and without articles is simply whether

they have overt or covert definite articles, which project their own phrase, all languages would have the same structure in the relevant domains. Then, the locus of parametric variation regarding extraction in these domains would be the locality system, which is currently stated in terms of phases and is part of the computational system of human language. In other words, we would need to posit variation in the computational system of human language, i.e., UG. However, in minimalism, UG is considered to be invariant (see especially Boeckx 2011 for relevant discussion), and all cross-linguistic variation is attributed to lexical properties, i.e., feature specifications of lexical items (see, e.g., Borer 1984, Chomsky 1995b, Baker 2008a,b; note that the difference regarding the presence/absence of DP can be stated in lexical terms, see e.g. Bošković 2021a). Thus, assuming a phonologically null definite article in article-less languages would have a consequence to the the architecture of the grammar which is undesirable under the current generative linguistic theory (see Bošković 2021a for more discussion of this issue).

Another argument against phonologically null definite articles comes from Neg-raising discussed by Bošković and Gajewski (2011). Bošković (2008b) establishes the generalization that languages with definite articles allow Neg-raising and those without definite articles disallow it (see Bošković 2008b for what counts as Neg-raising for the purpose of the generalization). As seen in (6), in English, a strict NPI *until yesterday/tomorrow* requires a clause-mate negation licensor, but it can be licensed by a non-clause-mate negation when the matrix predicate is a Neg-raising predicate such as *believe*. In contrast, in Serbo-Croatian, the counterpart of *believe* does not allow licensing of a strict NPI in the embedded clause, as shown in (7).

(6) a. John didn't leave/*left <u>until yesterday</u>.

  b. *John didn't claim [Mary would leave <u>until tomorrow</u>].

  c. John didn't believe [Mary would leave <u>until tomorrow</u>].   (Bošković 2008b)

(7) a. *Marija ju je posjetila najmanje dvije godine.

      'Mary visited her in at least two years.'

   b. Marija je nije posjetila najmanje dvije godine.

      'Mary has not visited her in at least two years.'

   c. *Ivan ne vjeruje [da ju je Marija posjetila najmanje dvije godine].

      'Ivan does not believe that Mary has visited her in at least two years.'

<div align="right">(Bošković and Gajewski 2011)</div>

Bošković and Gajewski (2011) offer a deduction of this generalization based on Gajewski's (2005) semantic analysis of Neg-raising, in which a Neg-raising predicate such as *believe* in English essentially denotes a definite plural, i.e., it is essentially a combination of a modal base and a definite article, and this definite article is responsible for the availability of Neg-raising (see Gajewski 2005 and Bošković and Gajewski 2011 for technical details; the analysis is based on certain parallelisms in the semantics of Neg-raising predicates and distributive plural definite descriptions). What this means is that a language can have Neg-raising predicates only if it has a definite article; languages that lack definite articles cannot have a Neg-raising predicate. If languages without definite articles had a phonologically null definite article, it would be unclear why those languages cannot use the phonologically null definite article and allow Neg-raising. Note also that Neg-raising is a semantic phenomenon where phonology plays no role (Gajewski 2005 in fact offers a semantic account), and there is no overt definite article involved in Neg-raising in the first place. Thus, merely assuming a phonologically null definite article is of no help to capture the correlation between semantics (i.e, Neg-raising) and the presence/absence of an overt definite article in a given language.

    Bošković (2021a) provides additional arguments that there is no phonologically null article based on the unavailability of anaphoric interpretation of bare nouns in article-less languages discussed by Jenks (2018) and Despić (2019). Jenks (2018) observes that in Mandarin Chinese, a bare noun cannot be used to refer to an antecedent of a donkey sentence, as shown in (8a); for the relevant interpretation, a demonstrative must be used, as shown in (8b). Note that in English, the

definite article is used in the donkey sentence in (9).

(8) a. Mei ge [you yi zhi shuiniu de] nongfu dou hui da **shuiniu**.
   every CL have one CL buffalo REL farmer all will hit buffalo
   'Every farmer that has a buffalo hits buffalo (generally).' (no bound reading)

   b. Mei ge [you yi zhi shuiniu de] nongfu dou hui da **na zhi shuiniu**.
   every CL have one CL buffalo REL farmer all will hit that CL buffalo
   'Every farmer that has [a buffalo]ᵢ hits [that buffalo]ᵢ.'   (Jenks 2018:503)

(9) Every farmer that has a buffaloᵢ hits **the** buffaloᵢ.

If Mandarin Chinese, which lacks overt definite articles, had a phonologically null definite article, the bare noun 'buffalo' in (8a) should be able to refer to its antecedent, just like its English counterpart in (9), which has the definite article (since both would then have a definite article). Likewise, Despić (2019) observes that in article-less languages, a bare mass noun cannot refer to its antecedent, as represented by Turkish (10a). For the anaphoric reference, the mass noun requires a demonstrative, as shown in (10b). This is contrasted with English, where the definite article is used in the same context, as shown in (11).

(10) a. #Ömrüm boyunca üzüm yetiştirdim. **Meyve** herşeyim oldu.
   my.life throughout grape produce fruit my everything
   'I have been producing grapes my whole life. Fruit is everything to me.'

   b. Ömrüm boyunca üzüm yetiştirdim. **Bu meyve** herşeyim oldu.
   my.life throughout grape produce this fruit my everything
   'I have been producing grapes my whole life. This fruit is everything to me.'

   (Turkish, Despić 2019:266)

(11) We have been growing grapes for generations – and you know, we have made millions on **the** fruit.   (Despić 2019:268)

In both cases (i.e., (8) and (10)), if there is a phonologically null definite article, the bare nouns in (8a) and (10a) should behave like their counterparts in English, which have a definite article. Thus, Bošković (2021a) concludes that article-less languages do not have a phonologically null definite article.

Bošković (2008b, 2012) in fact offers deductions of a number of his generalizations by proposing that languages with definite articles have DP in the nominal domain, whereas languages without definite articles lack DP. In particular, to account for (1a) and (1b), Bošković adopts Chomsky's (2000) Phase Impenetrability Condition (PIC), according to which only the edge of a phase is accessible for movement outside of the phase. He also adopts the anti-locality condition argued for in Bošković (1994), Saito and Murasugi (1999), Ishii (1999), Abels (2003), Grohmann (2003), among others: in the formulation of Bošković (2013a), movement has to cross at least one full phrase, not a segment. In addition, Bošković (2013a, 2014) argues that the highest phrase in the extended projection of a lexical head, including N, constitutes a phase. As a result, in languages with definite articles, DP is a phase in the nominal domain (as the highest phrase in the extended projection of N) whereas in languages without definite articles, NP is a phase in the nominal phrase. Consequently, in languages with definite articles, when an NP adjunct is extracted from the nominal phrase, the movement either has to violate the PIC to satisfy the anti-locality condition if it moves directly out of DP, as illustrated in (12a), or has to violate the anti-locality condition to obey the PIC if it moves to Spec,DP before Spell-Out since it crosses just a segment, not a full phrase, as illustrated in (12b). In contrast, in languages without definite articles, the highest projection of a nominal phrase is NP, so that the adjunct undergoing movement violates neither the PIC nor the anti-locality condition, as schematized in (13).

(12) a.

b.

(13)
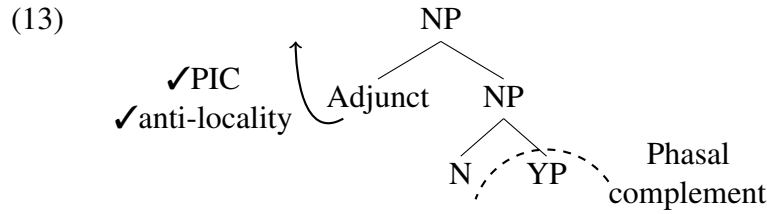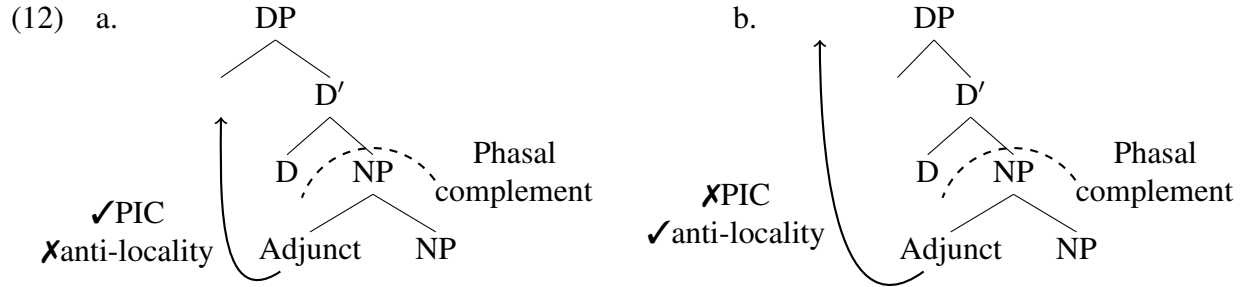
The same holds for LBE of an AP, which Bošković (2013a) assumes is adjoined to NP.[5] As schematized in (14), in languages with definite articles, where DP projects above NP, movement of an adjective out of DP would violate the PIC or the anti-locality condition. On the other hand, in article-less languages, where DP does not project above NP, the movement in question satisfies the PIC and the anti-locality condition, as illustrated in (15).[6]

(14) a.

b.

---
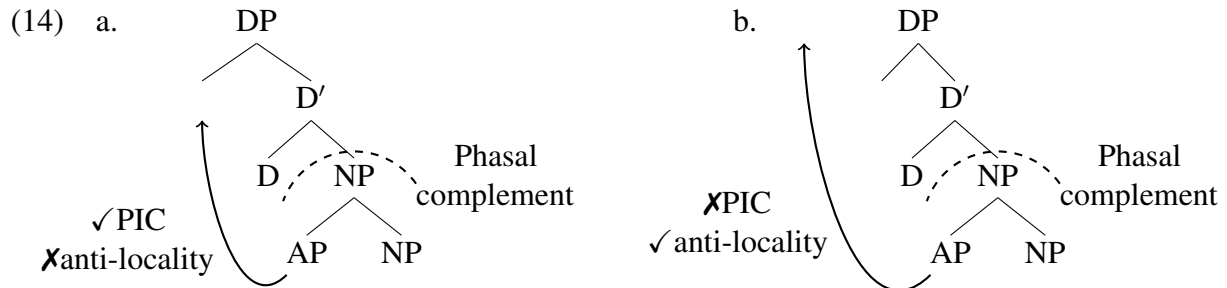
5. Bošković (2005) proposes an alternative analysis of adjective LBE, in which adjectives project AP that dominates NP in languages with definite articles and DP projects above the AP, so that the adjective (AP) cannot move out of the DP to the exclusion of the NP. See also fn. 11.

6. However, if the NP from which LBE and adjunct extraction take place is directly dominated by another NP in Serbo-Croatian, both LBE and adjunct extraction are blocked, which follows given that the higher NP is a phase in such cases, see Bošković (2012) for data and discussion.

(15)
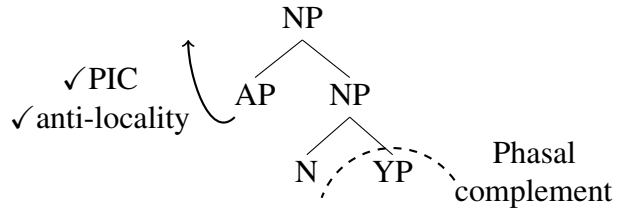
Thus, the presence/absence of DP plays a crucial role in accounting for Bošković's generalizations (see Bošković 2008b, 2012 for other generalizations).

Now, one important aspect of Bošković's generalizations is that they have a two-way language cut, i.e., whether a language has definite articles ("DP-language") or not ("NP-language"). Interestingly, however, it has been noticed that languages that have affixal definite articles (e.g., Icelandic) exhibit some similarities with languages that lack definite articles (e.g., Serbo-Croatian), i.e., they don't always pattern with languages that have non-affixal definite articles (e.g., English). Thus, Reuland (2011) and Despić (2011, 2015) observe that languages that have reflexive possessives either lack definite articles or have suffixal articles, which is exemplified in (16) (see Despić 2011, 2015 for the list of languages).[7,8]

(16)   a. *He loves **himself's** neighbors.

        b. Svaki dečak$_i$ je video **svog$_i$** oca.
            every boy   is seen   self's father
            'Every boy$_i$ saw his$_i$ father.'   (Serbo-Croatian, Despić 2015:209)

        c. Egil$_i$ vantar bókina **sína$_i$/*hans$_i$**.
            Egil needs book   self's/his
            'Egil needs his book.'   (Icelandic, Thraínsson 2007:463)

---

7. Marelj (2011) argues that languages that have reflexive possessives allow adjective LBE, based on a number of article-less languages, but Despić (2011, 2015) shows that this is not the case, since there are languages such as Icelandic that disallow adjective LBE but have reflexive possessives. Note that for the LBE generalization illustrated in (3), affixal article languages do behave like languages with non-affixal articles.

8. To be more precise, Despić's original generalization is that languages that have reflexive possessives either lack definite articles or have postnominal articles. The latter also include Koromfe, whose "definite article" is postnominal but does not look suffixal. It should, however, be added that the "definite article" in Koromfe does not seem to be obligatory in contexts of definite interpretations, so it may not be a real definite article from the viewpoint of Bošković's treatment of definite articles (cf. (4)). I thus put aside Koromfe here and adopt the version of the generalization in the text.

This indicates that a more fine-grained distinction of NP/DP-languages may need to be made.

Not only do affixal article languages behave like article-less languages in some domains, but they also show some properties that neither non-affixal article languages nor article-less languages do. Thus, Bošković (2008a) observes that movement of D-linked wh-phrases and relativization out of multiple wh-islands are allowed in Albanian, Bulgarian, Hebrew, Icelandic, Norwegian, Romanian, and Swedish, which are all affixal article languages.[9] This is contrasted with English, which has non-affixal article, and SC, which lacks articles. (17)-(19) are quoted from Bošković (2008a).

(17) *I saw a book which I wonder who knows who sells.   (English)

(18) *Vidio sam knjigu koju$_i$ se   pitam   ko zna   ko prodaje t$_i$.
    Seen am book which REFL wonder-1sg who knows who sells
    'I saw a book which I wonder who knows who sells.'   (SC)

(19) Vidjah edna kniga, kojato$_i$ se   čudja   koj znae koj prodava t$_i$.
    saw-1sg one book   which-the REFL wonder-1sg who knows who sells
    'I saw a book which I wonder who knows who sells.'   (Bulgarian)

Talić (2015, 2017) also shows that affixal article languages do not always behave like non-affixal article languages, and they behave like article-less languages in some respects. Her main argument concerns Left Branch Extraction of intensifier adverbs out of APs. As (20) shows, languages that have non-affixal definite articles disallow such adverb LBE. Talić shows that this holds for Dutch, German, Spanish, Brazilian Portuguese, French, Italian, Hungarian, and Cypriot Greek. In contrast, languages that lack definite articles allow it, as (21) shows. This also holds for Polish, Russian, Slovenian, and Persian.

(20) ***Terribly$_i$**, he was [**t$_i$** tired.]   (English, Talić 2015:420)

---

9. The generalization crucially concerns extraction from *multiple* wh-islands.

(21) **Strašno$_i$** je bila [**t$_i$** umorna].

terribly is been tired.F.SF

'She was terribly tired.' (BCS, Talić 2015:420)

Crucially, she observes that Bulgarian, Icelandic, and Romanian, which have affixal definite articles, also allow adverb LBE, as represented by Bulgarian (22).

(22) **Užasno$_i$** bjah [**t$_i$** umoren].

terribly was tired

'I was terribly tired.' (Bulgarian, Talić 2015:421)

Thus, Talić (2015, 2017) establishes the following generalization:[10]

(23) *Generalization of adverb LBE out of predicative TAPs*: Languages that allow Adv-extraction out of predicative TAPs either lack definite articles or have affixal definite articles.

To account for (23), Talić (2015, 2017) extends Bošković's analysis of (1a) and (1b) mentioned above to adverb LBE. Talić proposes the Structural Parallelism hypothesis, as formulated in (24).

(24) *Structural Parallelism*

a. If a language allows bare lexical structure without a functional layer in the domain of one lexical category, it may allow bare lexical structure in the domain of other lexical categories (e.g., a language can have both bare NP and bare AP).

b. If a language never allows bare lexical structure, that is, it always requires a functional layer in the domain of one lexical category, it must have a functional layer in the domain of all lexical categories (e.g., such a language will never have bare NP or bare AP).

As stated by (24b), Structural Parallelism generalizes presence of a functional projection in the domain of one lexical category (e.g., nominal domain) to another domain. Talić thus proposes that there is a functional projection in the adjectival domain that corresponds to DP in languages that have non-affixal articles (F$_{adj}$P). F$_{adj}$P as the highest projection in the adjectival domain is a

---

10. Note that this is a one-way correlation.

phase, and hence LBE of an intensifier adverb, which Talić assumes is adjoined to AP, is blocked due to the interplay of the PIC and the anti-locality condition in the same way as adjective LBE is blocked in Bošković's proposal. In contrast, in languages that lack definite articles, $F_{adj}P$ can be absent, since there is no DP layer in those languages and hence the absence of the functional projection is generalized to the adjectival domain, due to (24a). Since the problematic projection is not present, intensifier adverbs can move out of an adjectival domain without violating the PIC or the anti-locality condition in the relevant languages.[11]

Recall now that affixal article languages pattern with article-less languages with respect adverb LBE, although they have definite articles and hence are expected to be DP-languages. Given the Structural Parallelism and the account of adverb LBE by Talić, it is expected that affixal article languages should pattern with article-less languages (i.e., they should behave like NP-languages) in the nominal domain in some respects. Talić shows that this is indeed the case, building on Dubinsky and Tasseva-Kurktchieva's (2014) observation regarding adjunct extraction out of a nominal phrase. As mentioned above, Bošković (2008b, 2012) establishes the generalization that only languages without definite articles may allow adjunct extraction out of a nominal phrase. Thus, the extraction in question is disallowed in English, which is a non-affixal article language, as shown in (25a), whereas it is allowed in Serbo-Croatian, which is an article-less language, as shown in (25b) (see Bošković 2012 for the list of relevant languages). Bošković proposes that DP is projected above NP in languages with definite articles in general, and this DP blocks the extraction in question, in the same was as LBE is blocked in those languages as discussed above. In contrast, in article-less languages, DP does not project above NP, so the extraction in question is allowed.

(25) a. *[From which city]$_i$ did Peter meet [girls t$_i$]?

    b. [Iz   kojeg grada]$_i$ je Ivan sreo [djevojke t$_i$]?

        from which city   is Ivan met girls   (Bošković 2008b)

---

11. Adjective LBE is blocked in affixal article languages even in the absence of the article (Bošković 2005, Despić 2011). Despić (2011) proposes, following Bošković (2005), that AP dominates NP in affixal and non-affixal article languages and DP must project above the AP for the entire phrase to function as an argument (cf. fn. 5). See chapter 5 and chapter 6 for relevant discussion.

Crucially, Dubinsky and Tasseva-Kurktchieva (2014) observe that in Bulgarian, which is an affixal-article language, the extraction in question is disallowed when the definite article is present with a quantifier or a prenominal possessive, as shown in (26a) and (26c), but it is allowed when the article is absent in such environments, as shown in (26b) and (26d).

(26) a. *[Ot koj universitet]$_i$ sreštna-ha nyakolko-**to** studenti t$_i$?
       from which university   met-they   several-the   students
       'From which university did they meet several students?'

    b. [Ot koj universitet]$_i$ sreštna-ha nyakolko studenti t$_i$?
       from which university   met-they   several   students
       'From which university did they meet several students?'

    c. *[Ot koj universitet]$_i$ sreštna-ha nejni-**to** studenti t$_i$?
       from which university   met-they   her-the   students
       'From which university did they meet her students?'

    d. [Ot koj universitet]$_i$ sreštna-ha nejni studenti t$_i$?
       from which university   met-they   her   students
       'From which university did they meet her students?'

                                                        (Dubinsky and Tasseva-Kurktchieva 2014)

Building on Bošković's analysis mentioned above, Dubinsky and Tasseva-Kurktchieva (2014) and Talić (2015, 2017) proposes that DP is present when the definite article is present and it is absent when the definite article is absent in Bulgarian. Talić further argues that affixal article languages in general may omit the DP layer when the definite article is absent (see chapter 2 for more arguments provided by Talić).

As Talić argues, the above observations indicate that the two-way cut of the NP/DP-language distinction originally proposed by Bošković (2008b, 2012) is not sufficient. Under Bošković's two-way cut, whether a language has a definite article or not correlates with whether DP is *always* present or not in the language. However, DP may be absent in the absence of a definite article

in affixal article languages. In other words, affixal article languages sometimes behave like DP-languages, and sometimes like NP-languages, which is in fact confirmed by Reuland and Despić's generalization regarding reflexive possessives and Talić's generalization regarding adverb LBE. In addition, as noted above, Bošković (2008a) observes that affixal article languages show selective island sensitivity, which neither non-affixal article languages nor article-less languages show. Thus, the literature has revealed that we need not just a two-way distinction (whether a language has definite articles or not), but also a three-way distinction (whether a language has non-affixal articles, affixal articles, or no article) of NP/DP-languages.

1.2   Main claims of the dissertation

In this dissertation, I will examine a wide range of phenomena from the perspective of the NP/DP-language distinction, establishing a number of novel typological generalizations, discussing their relevance to various aspects of the generative linguistic theory (on both the empirical and the theoretical level).

I will first establish and discuss two novel cross-linguistic generalizations in which affixal article languages pattern with article-less languages rather than non-affixal article languages, reinforcing Talić's argument that we need an (at least) three-way distinction with respect to the NP/DP-language distinction. I will then propose deductions of those new generalizations, which will be shown to have a number of consequences for the architecture of the grammar.

One of the two generalizations concerns the Coordinate Structure Constraint (CSC). As is well-known, the CSC bans extraction out of a coordinate structure, as shown in (27).

(27) *[Which table]$_i$ will he buy [t$_i$ and the chair]?   (Ross 1967)

However, I will show that there are a number of languages from different language families that allow violations of the CSC. Crucially, the languages that can violate the CSC are either affixal article languages or article-less languages. Thus, we find the three-way distinction of NP/DP-languages with respect to violations of the CSC. In addition, I will argue that the CSC, which was

originally proposed as a single condition by Ross (1967), should actually be separated into two conditions. I will then discuss relevance of phases and labeling to the two separated CSCs, arguing that those two CSCs are deduced from different mechanisms in the grammar.

The other generalization concerns the typology of indefinite pronouns, which has been extensively discussed in the non-generative literature. Haspelmath (1997) shows that languages can be classified into two major groups with respect to the morphological make-up of indefinite pronouns. In one of the two groups, indefinite pronouns are derived from interrogative pronouns, as shown by Mandarin Chinese (28), where the interrogative pronoun *shenme* is used to express the interpretation of 'something'. Haspelmath calls them *interrogative-based* indefinite pronouns.

(28)   Ta yiwei wo xihuan **shenme**.
        he think I   like   what
        'He thinks I like something.'

In the dissertation, I will establish a novel generalization that a particular type of 'interrogative-based' indefinite pronouns are found only in affixal article languages and article-less languages. I will then propose a deduction of this generalization based on parameterization in feature specification and structure of the relevant pronouns. I will also show that this deduction of the generalization sheds new light on the typology of wh-questions, which has been one of the most widely discussed topics in the generative literature. In particular, it will be shown that the deduction allows us to treat multiple wh-fronting found in languages such as Serbo-Croatian and wh-in-situ of the Japanese type in a uniform manner.

In addition, I will discuss relevance of the indefinite pronouns in question for large-scale pied-piping, which is found in languages such as Basque. I will establish a novel cross-linguistic generalization that large-scale pied-piping is possible only if the language has the relevant indefinite pronouns and the projection to be pied-piped is head-final. I will show that this generalization can be deduced from the perspective of the labeling theory. Chomsky (2015) proposes the concept of "weak heads", which are essentially defective with respect to labeling. I will propose a principled

criterion for defining such heads based on their morphological properties. I will then show that this notion interacts with the head-directionality of complementizers and their morphological status to deduce the proposed generalization regarding large-scale pied-piping. I will also show that the proposed conception of weak heads enables us to deduce Agree from Minimal Search, which is a third factor principle external to UG. This means that we can eliminate Agree from UG.

Note that the above discussion of the CSC and indefinite pronouns concerns a three-way distinction of NP/DP-languages. In this dissertation, however, I will further argue that the three-way distinction (languages with non-affixal articles, languages with affixal articles, and languages without articles) is not sufficient either, and that we need a more fine-grained "scale" of NP/DP-language distinction, from canonical DP-languages to non-canonical DP-languages. Specifically, I will demonstrate that a number of languages that have been previously considered as DP-languages in the literature (the languages in question have free stranding (i.e., non-affixal) articles) actually exhibit some properties that NP-languages do, and argue that the presence/absence of a definite article does not necessarily correspond to the presence/absence of DP in a given language or a construction. I will then propose that the definite article has an option of not projecting DP in the nominal domain, since the definite article can be base-generated as adjoined to another head (this is in fact what will enable us to capture the exceptional behavior of certain DP languages noted in this passage).

A word of caution is in order here regarding the relationship between generative linguistics and typology. As mentioned above, I will establish a number of novel typological generalizations, and discuss relevance of those typological observations for the generative linguistic theory. In the non-generative typological literature, however, it has been observed that almost all typological generalizations have some exceptions (i.e., they are statistical universals), based on which non-generative typologists have argued that there is no such thing as UG that all human languages share, because UG should not allow exceptions. As Bošković (2021a) points out, there is actually a misunderstanding here; typological generalizations, whether they have exceptions or not, are descriptive generalizations that are independent of any particular theory, hence those descriptive

generalizations themselves are not UG. What is relevant for UG are the *deductions* of those generalizations. Thus, even if there are some exceptions to a descriptive generalization, those exceptions themselves do not serve as direct counterarguments to UG (though a deduction of the generalization based on UG may need to be reconsidered so that it leaves room for exceptions). In fact, the generalizations established in this dissertation will be described in a theory-neutral way, without appealing to UG; some of my generalizations even have exceptions. The deductions of the generalizations offered in this dissertation will, however, be based on mechanisms of UG, and they will also leave room for exceptions. It should also be added here that all the typological generalizations I establish in this dissertation are implicational hierarchies (i.e., one-way correlations), stated as, e.g., "languages that have a property X have a property Y"; thus, there can be languages that have Y but not X, and those languages do not invalidate the generalizations or their deductions. To put it differently, there can be additional factors that disallow languages that have Y to have X, and those factors may be independent of the deduction of an implicational hierarchy. For more extensive discussion of the relationship between generative linguistics and typology, see Baker and McCloskey (2007), Polinsky and Kluender (2007), and especially Bošković (2021a). What is important here is that typological investigations can not only be compatible with, but also shed new light on, the generative linguistic theory, which this dissertation aims to demonstrate.

1.3   Organization of the dissertation

The dissertation is organized as follows.

In chapter 2, I will discuss the Coordinate Structure Constraint (CSC) from the perspective of the NP/DP-language distinction. I will first show that the CSC can be violated in a number of genetically unrelated languages, as noted above. Crucially, the violations of the CSC involve extraction *of* a conjunct, while extraction *out of* a conjunct is still disallowed in languages where extraction *of* a conjunct is allowed. Based on this, I will argue that the CSC should be separated into two independent conditions: the ban on extraction *of* a conjunct (CSC I) and the ban on

17

extraction *out of* a conjunct (CSC II). I will then establish a novel generalization that languages that allow violations of the CSC I either have affixal definite articles or lack definite articles. I will offer a deduction of this new generalization by extending Bošković's (2005, 2008b, 2012, 2013a) deduction of LBE out of a nominal domain and Talić's (2015, 2017) Structural Parallelism to the coordinate structure. I will also briefly discuss Bošković's (2020b) account of the CSC II, and claim that the CSC I and the CSC II are deduced from different mechanisms: the CSC I is a purely syntactic condition and the CSC II is an interface condition.

In chapter 3, I will discuss the typology of indefinite pronominal systems, which has been extensively discussed in the non-generative literature, from the perspective of the NP/DP-language distinction. As mentioned above, Haspelmath (1997) observes that there are two types of indefinite pronouns: what he calls "generic-noun-based" indefinite pronouns and "interrogative-based" indefinite pronouns; I will focus on the latter. The "interrogative-based" indefinite pronouns are so called because they appear to be composed of an interrogative pronoun and a quantificational affix/particle. However, by paying close attention to the morphology of the relevant pronouns, I will argue that the term "interrogative-based" indefinite pronoun is misleading, and propose a new classification of the relevant pronouns building on Kuroda (1965); specifically, I will propose that there are two types of relevant pronouns, bare indeterminate pronouns and compositional indeterminate pronouns. I will then establish a novel generalization that languages that have productive compositional indeterminate pronouns either have affixal definite articles or lack definite articles. I will offer a deduction of this new generalization based on Saito's (2017) parameterization of bare indeterminate pronouns in Chinese and compositional indeterminate pronouns in Japanese. I will also argue that this deduction of the generalization in question sheds new light on the typology of the syntax of wh-questions; in particular, multiple wh-fronting found in e.g., Slavic languages and wh-in-situ of the Japanese type receive a unified treatment, which is in fact supported by certain parallelisms between these two types of wh-questions. Furthermore, it will be argued that the proposed analysis of indeterminate pronouns can account for various types of wh-in-situ observed in the literature. Thus, the proposed account of indeterminate pronouns enables us to take a fresh

perspective on the investigation of the typology of the syntax of wh-questions.

In chapter 4, I will show that the availability of the relevant indefinite pronominal system interacts with the availability of large-scale pied-piping observed in languages like Basque. Specifically, as mentioned above, I will establish a novel cross-linguistic generalization that large-scale pied-piping is possible in a language only if the language has productive compositional indeterminate pronouns and the projection to be pied-piped is head-final. In order to deduce this new generalization, as noted above, I propose a morpho-syntactic condition on "weak heads" under Chomsky's (2015) labeling framework, in which weak heads are realized as bound morphemes, as well as a criterion for determining weak heads, which generalizes the notion of weak heads to all heads that have unvalued features at the point of External Merge. I then show that this conception of weak heads captures the availability of large-scale pied-piping, tied with a cross-linguistic morphological difference between head-initial and head-final complementizers observed by Inaba (2011). In addition, as noted above, I will demonstrate that this new conception of weak heads allows us to deduce Agree from Minimal Search, which is a third factor principle external to UG. In other words, we can eliminate Agree from the computational system of language and hence minimize UG. I will also show that this deduction can capture the variation in the superiority effects that is found with multiple wh-questions in combination with Epstein et al.'s (2020) theory of Minimal Search and feature valuation. Moreover, I will discuss Bošković's (2008a) generalization that D-linked and relative indeterminate phrases are insensitive to multiple wh-islands in languages with affixal definite articles from the perspective of the labeling framework discussed in this chapter. I will show that this generalization can be captured by the syntactic nature of a head amalgam created by External Pair-Merge of two heads in Epstein et al.'s (2016) sense.

In chapter 5, I will show that some languages that have been treated as non-affixal article languages by Bošković (2008b, 2012) and Talić (2015, 2017) exhibit different properties from both prototypical non-affixal article languages such as English and prototypical affixal article languages such as Bulgarian. Based on a close investigation of such cases, I will propose that we need an even more fine-grained differentiation of the NP/DP languages distinction. In particular, as

mentioned above, the distinction to be made is not a two-way or three-way "cut", but a "scale" from a canonical DP-language to a canonical NP-language. A theoretical issue to be addressed is how to capture those more fine-grained distinctions, or the "scale". I will propose that definite articles, which have been considered to always project DP in the literature, actually need not project in some languages, which I argue is an option allowed in the Bare Phrase Structure Theory; more specifically, definite articles in, e.g., Italian can be base-generated adjoined to a nominal head, without projecting DP. Thus, the scale of NP/DP-language distinction can be captured by different options for realization of definite articles in the structure, which are in fact allowed by the current syntactic theory. I will also discuss the proposed fine-grained scale of NP/DP-languages distinction from the perspective of the so-called emergentist view of parameters, which conforms to the three factor design of language proposed by Chomsky (2005). The scale of NP/DP-languages distinction will be argued to be an appropriate parameter given the locus of parameterization in minimalism and economy considerations of language acquisition.

In chapter 6, I will discuss cases in which definite articles are omitted in contexts where definite interpretation is obtained and hence definite articles would be expected to occur under Bošković's (2016b) definition of definite articles in (4). Specifically, certain PPs and kinship possessums in some non-canonical DP-languages in the NP/DP-language scale argued for in chapter 5 resist definite articles in these domains even though they have definite interpretation. I will propose a syntactic account for the article drop in PPs, and a semantic account for kinship possessums. Regarding PPs, I will propose that P can be the highest functional projection in the nominal domain in the sense of Grimshaw (2000) and Bošković (2014), which is motivated by the categorial feature specification originally proposed by Chomsky (1970). P being the highest functional projection in the nominal domain blocks projection of DP, and the feature responsible for the definite interpretation that is otherwise contained in D becomes part of the feature bundle of P in such cases. As for kinship possessums, I will propose, following Dobrovie-Sorin (2002, 2004) that they are of type $\langle e,e \rangle$, so that the definite article cannot be combined with the possessive phrase due to a type mismatch. It will also be argued that DP is absent in both article drop in PPs and article drop

with kinship possessums, which is consistent with the proposal in chapter 5 that DP can be absent in non-canonical DP-languages. In addition, I will show that the presence of the definite article is forced in these environments when there is an adjective that modifies the head noun in the PPs or the kinship possessum. I will take this as indicating that the presence of an adjective in a nominal phrase plays a crucial role in projection of DP in environments where DP is otherwise omitted. I will then offer syntactic and semantic accounts for those cases.

Chapter 7 concludes the dissertation and addresses some open questions.

Chapter 2

Decomposing and Deducing the Coordinate Structure Constraint

2.1 Introduction

This chapter explores the Coordinate Structure Constraint (CSC) from a cross-linguistic perspective and establishes a novel typological generalization regarding the CSC. The CSC was originally formulated by Ross (1967) as in (1), based on English examples such as (2).

(1)   In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

(2)   a. *[**Which table**]$_i$ will he buy [**t**$_i$ and the chair]?

      b. *[**Which trombone**]$_i$ did [[the nurse polish **t**$_i$] and [the plumber compute the tax]]?

<div align="right">(Ross 1967)</div>

The more or less standard view in the literature is that the CSC is universal across languages. Postal (1998:52) in fact notes that "the CSC is widely regarded as the most problem-free syntactic constraint ever discovered". This constraint, unlike other island constraints formulated by Ross, is also accepted as a universal constraint in non-generative literature (e.g., Haspelmath 2004, 2007,

Kazenin and Testelets 2004, Peterson and VanBik 2004) and is used as one of the criteria for coordinate structures in comparison with subordinate and comitative constructions. As far as I know, there has been no detailed cross-linguistic research that examines the CSC.

Interestingly, however, it has occasionally been observed that there are cases where the CSC is violated in languages other than English. For instance, Bošković (2009c) and Stjepanović (2014) observe that Serbo-Croatian (SC) allows violations of the CSC, as shown in (3).

(3) ?**Knjige$_i$** je Marko [**t$_i$** i   filmove] kupio.

     books is Marko   and movies bought

     'Marko bought books and movies.'

Yatabe (2003) also observes that scrambling out of a coordinate structure is possible in Japanese.

(4) ?**Kyoodai-to$_i$**   kanojo-wa [**t$_i$** Toodai]-ni   akogareteiru.

     Kyoto.University-and she-TOP   Tokyo.University-DAT admire

     'She admires Kyoto University and Tokyo University'   (adapted from Yatabe 2003)

Given the standard view on the CSC, one might simply consider (3) and (4) to be arbitrary exceptions peculiar to these two languages. In this chapter, however, I demonstrate that the CSC can be violated in a wide range of languages, which furthermore share a common property; this means that cases such as (3) and (4) are not arbitrary at all from a more cross-linguistic perspective. Based on this, I then establish a novel cross-linguistic generalization regarding languages that allow violations of the CSC. I then propose a deduction of this new generalization based on a contextual approach to phasehood advocated by Bošković (2013a, 2014) and the Structural Parallelism hypothesis proposed by Talić (2015, 2017).

The cross-linguistic data regarding violations of the CSC will also lead me to argue, in the spirit of Grosu (1973) and Postal (1998) (for more recent discussion see also Bošković in press, Stjepanović 2014), that the CSC, which was originally formulated as a single condition as in (1), should be separated into two different conditions, and that both ConjP and individual conjuncts

are islands independently of each other.[1] This will shed new light on the long-standing debate regarding the nature of the traditional CSC, namely, the question of where in the grammar the CSC applies (narrow syntax, Johnson 2002, Postal 1998, Ross 1967 among others, LF, Fox 2000, Goodall 1987, Kehler 1996, Lin 2001, Munn 1993, Ruys 1993 among others, or PF, Merchant 2001, Kasai and Takahashi 2002 among others; see Kato 2006 for an overview). I will provide a more fine-grained answer to this issue that could not have been provided without the separation of the CSC into two conditions that is argued for here. More specifically, the present chapter will argue that the two different conditions that result from the separation of the traditional CSC (which concern the islandhood of ConjP and the conjuncts respectively) are deduced from different mechanisms in the architecture of the grammar: one is a purely syntactic condition, and the other one is an interface condition.

A word of caution is in order here. As this is a typological work – in order to establish the new generalization, a number of languages need to be discussed – it is impossible to go into a detailed discussion of any individual language discussed here for space reasons. As a result, I will not discuss some cases where violations of the CSC are disallowed due to some additional interfering factors even in languages such as SC and Japanese which in principle allow such violations.[2] This does not invalidate my argument, however, because the goal of this chapter is to make a distinction between languages that in principle allow CSC violations and those that never allow them; in other words, the primary goal of this chapter is to establish and deduce the new generalization regarding languages that *in principle* allow violations of the CSC of the kind in (3) and (4) (we will see that such languages constitute a natural class), which are, crucially, *never* allowed in languages like

---

1. Grosu (1973) argues for the separation of the CSC based on pronominalization and pseudo-clefts, which is very different from our current concerns. In fact, this chapter will present the first case for the separation of the CSC into separate conditions from a broad cross-linguistic perspective.
2. For example, extraction of the first conjunct out of a nominal phrase is banned in Japanese, as shown in (i).

(i) *Sensei-to$_i$ Taro-wa   [$_{NP}$ Hanako-kara-no   [$_{ConjP}$ t$_i$ tomodachi]-e-no tegami]]-o yonda.
   teacher-and Taro-TOP   Hanako-from-GEN   friend-to-GEN   letter-ACC read.PAST
   'Taro read a letter from Hanako to her teacher and friend(s).'

Such cases are not our main concern here, since, as noted in the text, what is important is that languages like Japanese *in principle* allow violations of the CSC as in (4), in contrast to languages like English (at any rate, extraction out of a nominal domain in Japanese is independently severely restricted; cf. Takahashi and Funakoshi 2013, Shiobara 2017, Arano and Oda 2019).

English (which constitute another natural class).

The chapter is organized as follows: in section 2.2, I provide a cross-linguistic survey of the possibility of CSC violations and argue that the CSC has to be separated into two conditions, which indicates that both the whole coordinate structure (i.e., ConjP) as well as the individual conjuncts themselves are islands independently of each other. In section 2.3, I provide a new generalization regarding the property that languages that allow CSC violations share, and propose a phase-based account of this generalization in connection with the structure of the conjunction phrase. In section 2.4, I briefly discuss Bošković's (2020b) approach to one part of the traditional CSC and its consequences for the proposal in this chapter as well as to the status of the CSC. Section 2.5 concludes the chapter.

2.2   Separating the CSC

As mentioned above, the CSC has standardly been considered to apply in all languages. Thus, extraction out of a coordinate structure is not allowed in English, Dutch, Spanish, Italian, and Brazilian Portuguese, as illustrated in (5)-(9).

(5)   English (Indo-European, Germanic)

   * **The wine$_i$**, he bought [**t$_i$** and the cheese].

(6)   Dutch (Indo-European, Germanic)

   * **De wijn$_i$** kocht   Jan [**t$_i$** en   de   kaas].

    the   wine   bought   Jan   and   the   cheese

    'Jan bought the wine and the cheese.'

(7)   Spanish (Indo-European, Italic)

  ?* **El   vino$_i$**, compré   [**t$_i$** y   el   queso].

    the   wine   I.bought   and   the   cheese

    'I bought the wine and the cheese.'

(8) Italian (Indo-European, Italic)

  *Il vino_i, ho   comprato [t_i e   il formaggio].

  the wine I.have bought   and the cheese

  'I bought the wine and the cheese.'

(9) Brazilian Portuguese (Indo-European, Italic)

  *O vinho_i, ele comprou [t_i e   o   queijo].

  the wine   he bought   and the cheese

  'He bought the wine and the cheese.'

However, there are languages that allow a type of extraction out of a coordinate structure, in addition to SC and Japanese, which were noted above. Let us consider (10)-(22).[3]

(10) SC (Indo-European, Slavic)

  ?Knjige_i je Marko [t_i i   filmove] kupio. (= (3))

  books   is Marko   and movies bought

  'Marko bought books and movies.'

(11) Russian (Indo-European, Slavic)

  Kn'ig'i_i Pasha [t_i i   f'il'my] kupil.

  books   Pasha   and movies bought

  'Pasha bought books and movies.'

(12) Polish (Indo-European, Slavic)

  ?Ksiązki_i Jan [t_i i   filmy] kupił.

  books   John   and movies bought

  'John bought books and movies.'

---

3. Regarding the placement of the conjunction in (13), see Oda (2017), who argues that the conjunction cliticizes to the first conjunct in narrow syntax, and is then carried along by the movement of the conjunct. Stjepanović (2014) shows that something similar actually happens in SC (10), the only difference being that the conjunction in SC procliticizes to the second conjunct instead of encliticizing to the first conjunct. See Stjepanović (2014) and Oda (2017) for more detailed discussion.

(13) Japanese (Japonic)

    a. ?**Kyoodai-to$_i$**   kanojo-wa [**t$_i$** Toodai]-ni   akogareteiru. (= (4))

        Kyoto.University-and she-TOP   Tokyo.University-DAT admire

        'She admires Kyoto University and Tokyo University'

    b. ?**Nani-to$_i$** Taro-ga   [**t$_i$** mizu]-o   katta   no?

        what-and Taro-NOM   water-ACC bought Q

        lit. 'What did John buy and water?'

(14) Korean (Koreanic)

    a. ?**Hankuk-kwa$_i$** kuney-nun [**t$_i$** ilpon]-ul   conkyinghanta.

        Korea-and   she-TOP   Japan-ACC admire

        'She admires Korea and Japan.'

    b. ?**Mwuess-kwa$_i$** John-i   [**t$_i$** mwul]-ul   sanni?

        what-and   John-NOM   water-ACC buy.Q

        lit. 'What did John buy and water?'

(15) Old English (Indo-European, Germanic)

    And he **hine$_i$** miclum [**t$_i$** ond his geferan]   mid feo   weorðude

    and he him   greatly   and his companions with money honored

    'And he much honored him and his companions with money'

                                       (*Anglo-Saxon Chronicle* 878:Lightfoot 1999)

(16) Latin (Indo-European, Italic)

a. neminem **sapientiae$_i$** laudem [**t$_i$** et eloquentiae] sine summo studio et

    nobody wisdom reputation and eloquence without greatest effort and

    labore et doctrina consequi posse.

    industry and study obtain can

    'no one can achieve high distinction for wisdom and eloquence without a very great

    amount of zeal and industry and study.' (Cicero, *de Oratore* 2.363)

b. Etrusci campi, qui **Faesulas$_i$** inter [**t$_i$** Arretium-que] iacent

    Etruscan plains which Faesulae between Arretium-and lie

    'the Etruscan plains between Faesulae and Arretium' (Livy, 22.3.3)

(17) Classical Greek (Indo-European, Hellenic)

polémou$_i$ péri [**t$_i$** kaì aspʰaleíās]

war.GEN about safety.GEN

'about war and safety' (Thucydides 5.11.4, Agbayani and Golston 2010:144)

(18) Sanskrit (Indo-European, Indo-Aryan)

[imān ca lokān]$_i$ upa-hváyate [**t$_i$** etāni ca sāmāni]

these.ACC.SG and world.ACC.SG summon.3SG.PRES these.ACC.PL and chant.ACC.PL

'He summons these worlds and these chants.'

(Śathapathabrahmaṇa1.8.1.19, Mitrović 2011:78)

(19) *Gitksan*

Gwi-hl$_i$ gubis Henry [**t$_i$** gan-hl miyup]?

what-CN eat.TRA.PN Henry and-CN rice

'What did Henry eat and rice?' (Davis and Brown 2011:58)

(20) *Nisgha*

Ksax̱ **haxwadakw-hl_i** dii   jabit,   [**t_i** gan-hl   hawil].

only   bow-CN   CNTR make-TRA-3SG.II   and-CN arrow

'He did nothing but make bows and arrows.'   (Tarpent 1987:452)

(21) *Shona*

?**Ndi-Ø-ani_i** wa-vaka-teng-er-a   [**t_i** na-Ø-Tendai]   ma-rokwe?

NI-1a-who   1a.NSE-2.SM-TA-buy-APPL-FV   and-1a-Tendai 6-dress

'Who(m)_i did they buy [t_i and Tendai] dresses?'   (Zentz 2016:137)

(22) *Tümpisa Shoshone*

Niiii **isapaippii-a_i** punikkappiihantii [**t_i** tunga kammuttsi(-a)].

I   coyote-OBJ   saw   and   jackrabbit(-OBJ

'I saw a coyote and a jackrabbit.'   (Dayley 1989:341)

All the examples above involve extraction *of* a conjunct. However, the languages that allow extraction *of* a conjunct listed above still ban another type of extraction out of a coordinate structure. Consider (23)-(27).

(23) *SC*

?***Knjigu_i**, Milan je rekao da   je Jovan [[kupio   **t_i** danas] i   [prodao kompjuter

book   Milan is said   that is Jovan   bought   today   and   sold   computer

juče]].

yesterday

'Milan said that Jovan bought a book today and sold a computer yesterday.'

(24) *Russian*

?***Knigu_i**, Vasja skazal čto Vanja [[kupil **t_i** sjegodnja utrom] i [prodal komp'yuter

book Bill said that John bought today morning and sold computer

včera]].

yesterday

'Bill said that John bought the book this morning and sold a computer yesterday.'

(25) *Polish*

?***Książkę_i** to Piotr powiedział, że Jan [[kupił **t_i** dzisiaj] i [sprzedał komputer

book TOP Peter said that John bought today and sold computer

wczoraj]].

yesterday

'Peter said that John bought the book today and sold the computer yesterday.'

(26) *Japanese*

***Taro-o_i** John-wa [Yamada-kyoozyu-ga [**t_i** home] (&) [Hanako-o shikatta] to]

Taro-ACC John-TOP Yamada-Prof.-NOM praise Hanako-ACC scolded C

itta.

said

'John said that Prof. Yamada praised Taro and scolded Hanako.' (Kato 2005:317)

(27) *Korean*

***Taro-lul_i** John-un [Yamada-kywosu-ka [**t_i** chinchanha-ko] [Hanako-lul

Taro-ACC John-TOP Yamada-Prof.-NOM praise-and Hanako-ACC

pipanhasstako]] malhassta.

criticize said

'John said that Prof. Yamada praised Taro and scolded Hanako.'

The common characteristic of the data in (23)-(27) is that they involve movement from within a conjunct. This means that extraction *out of* a conjunct is banned even in the languages that allow extraction *of* a conjunct.

If the CSC were a single locality condition as Ross (1967) originally formulated it as in (1), it would be mysterious why extraction *of* a conjunct is allowed but extraction *out of* a conjunct is banned in the same languages. This leads us to the conclusion that the CSC should be separated into two conditions, as in (28).

(28)   a.   CSC I: a conjunct may not be extracted out of a coordinate structure.

       b.   CSC II: an element inside a conjunct may not be extracted out of a coordinate structure.

In languages like SC, Russian, Polish, Japanese, and Korean, the CSC I can be violated whereas the CSC II cannot be. In languages like English, Spanish and Italian, neither the CSC I nor the CSC II can be violated.

I take the above cross-linguistic pattern to indicate that the whole coordinate structure (ConjP) and the conjuncts themselves independently ban extraction from inside themselves, and I interpret this state of affairs to mean that both ConjP and individual conjuncts are islands independently of each other.[4] An immediate question then arises: why the islandhood of ConjP but not of a conjunct can be voided in certain languages (namely, in those that allow extraction of conjuncts but not out of conjuncts). This issue will be discussed in the following sections.

2.3   A Phase-based Approach to Violations of the CSC I

2.3.1   What languages allow violations of the CSC I?

Now that we saw that a number of languages allow violations of the CSC I, a question that arises is whether there is a common characteristic among languages that allow such violations.

---

4. In fact, extraction out of a conjunct is even worse than extraction of a conjunct in languages that disallow both according to my informants. This intuition is straightforwardly captured by the current proposal, since extraction out of a conjunct crosses two islands (ConjP and a conjunct), whereas extraction of a conjunct crosses just one (ConjP).

Interestingly, all the languages that allow violations of the CSC (namely, SC, Russian, Polish, Japanese, Korean, Old English, Latin, Classic Greek, Sanskrit, Gitksan, Nisgha, Shona, and Tümpisa Shoshone) lack definite articles. This is reminiscent of Bošković's (2008b, 2012) NP/DP language distinction, according to which languages that do not have definite articles show many properties that languages which have definite articles do not show. Bošković gives a number of properties of this kind, several of which in fact involve extraction (e.g., Left Branch Extraction (LBE) of adjective may be allowed only in languages without articles). It then seems that the following generalization can be made regarding CSC I violations:

(29)   Generalization of CSC I violations (to be revised)

   Only languages that do not have a definite article may allow CSC I violations.

In all the examples in (10)-(22) the first conjunct is extracted out of a coordinate structure. Extraction *of* a conjunct is thus allowed in these languages, in contrast to the languages in (5)-(9), where extraction of a conjunct is disallowed. Again, this is unexpected by Ross's original formulation of the CSC.[5]

   However, the situation is more complex than that; we do not seem to be dealing here simply with a distinction between languages with and without definite articles. Thus, Johannessen (1998) reports that Norwegian, Swedish, and Old Norse allow violations of the CSC I, even though these languages have a definite article.[6,7]

---

5. See below for extraction of non-initial conjuncts.
6. (i) shows that Old Norse had a definite article.

(i) hestr-**inn**
   horse-DEF
   'the horse' (Faarlund 2009:619)

7. Anders Holmberg (p.c.) points out that an intonational pause is not needed between *jag* 'I' and *och* 'and' in (31), unlike its counterpart in English, which requires an intonational break before *and* (indicating an afterthought). This indicates that the part *och hans gamla dragspelsorkester* is not an afterthought (the same holds for languages like Japanese or SC). However, he also points out that when *i går* 'yesterday' appears after the second conjunct as in (i), the sentence becomes ungrammatical without an intonational break before *och* 'and' (if there is an intonational break there, the sentence is grammatical). The same pattern is noted by Johannessen (1998) for Norwegian.

32

(30) **Per$_i$** så jeg [**t$_i$** og Ola].

Per saw I and Ola

'Per, I saw, and Ola.' (Norwegian, Johannessen 1998:216)

(31) *Swedish*

Kalle Jularbo$_i$ hörde jag [**t$_i$** och hans gamla dragspelsorkester].

K.J. heard I and his old accordion.band

'K.J., I heard and his old accordion band.' (Andersson 1982:35)

(32) *Old Norse*

Skegg-Ávaldi$_i$ átti búð saman [**t$_i$** ok Hermundr], son hans.

Beard-Avald.SG had.SG hut together and Hermund.SG son his

'Beard-Avald and his son, Hermund, lived together.' (Nygaard 1917:13)

Icelandic, which has a definite article, also allows a CSC I violation.[8]

(i)*[**Kalle Jularbo**]$_i$ hörde jag [**t$_i$** och hans gamla dragspelsorkester] i går.

K.J. heard I and his old accordion.band in yesterday

'K.J., I heard, and his old accordion band yesterday.'

One possibility to explain this effect could be that the remnant of the coordinate structure after movement of the first conjunct needs to be focalized by the Nuclear Stress Rule, which essentially affects the last element in an intonational phrase. In Spanish, wh-in-situ is limited to the sentence final position (if an adverb follows it, it must be separated by a pause); Reglero (2007) argues that this can be explained by an interaction of the Nuclear Stress Rule with the Focus Stress Rule. In SC, CSC I violations are most acceptable if the remnant of the coordinate structure precedes the verb, which is a focus position in the language. This can be taken to indicate that there may be a requirement to focalize the remnant of CSC I violations on a par with Spanish wh-in-situ. Turning to Swedish (and Icelandic in footnote 8), in these languages, the Nuclear Stress Rule applies, assigning stress to the rightmost element in an intonational phrase (usually the last element in a sentence; see Ambrazaitis 2009 and Myrberg and Riad 2015 for Swedish and Ámason 1985 for Icelandic). It is, then, not implausible that the remnant of the movement involving CSC I violations in these languages may have to be focalized, with the Nuclear Stress Rule applying to it in the same way as in Spanish wh-in-situ. ((i) is acceptable with a pause before the adverb, on a par with wh-in-situ in Spanish, since the domain of Nuclear Stress Rule application is an intonational phrase.)

8. Gísli Rúnar Harðarson (p.c.) points out the same pattern as in Swedish regarding an intonational break and an element after the second conjunct: in (33) an intonational break is not required before *og* 'and', unlike its English counterpart, but when there is a phrase after the second conjunct as in (i), the sentence is ungrammatical unless there is an intonational break before *og*. (See footnote 7 for a possible account.)

(i)***Pétur$_i$** sá ég [**t$_i$** og Maríu í gær].

Pétur.ACC saw I and Mary.ACC in yesterday

'I saw Peter and Mary yesterday.'

(33) *Icelandic*

Pétur_i sá ég [t_i og Maríu].

Pétur.ACC saw I and Mary.ACC

'I saw Peter and Mary.'

Pană Dindelegan (2016) also reports that a violation of the CSC I is attested in Old Romanian, which also had a definite article.[9]

(34) *Old Romanian*

nu puteţi **[lu Dumnezeu]_i** sluji [t_i şi lu Mamon].

not can.PRES.2PL LUI.DAT God serve.INF and LUI.DAT Mammon

'you cannot serve God and Mammon.'

(Coresi, *Tâlcul Evanghelülor* 56ᵛ; Pană Dindelegan 2016:574)

What is then the class of languages that allow CSC I violations? More specifically, what is the difference between the languages with a definite article that allow CSC I violations and those which do not? The answer lies in the nature of the definite articles. Crucially, definite articles in the languages that allow CSC I violations are affixal, whereas those in the languages that do not allow such violations are non-affixal. Thus, the more precise generalization regarding CSC I violations which puts together (29) and the facts noted above is given in (35).[10]

(35) *Generalization of CSC I violations*

Languages with non-affixal articles disallow CSC I violations, whereas languages without definite articles and languages with affixal definite articles may allow them.

Notice now that the classification of the languages that allow violations of CSC I is the same as the one regarding adverb extraction out of an adjectival domain discussed by Talić (2015, 2017),

---

9. (i) shows that Old Romanian had a definite article.

(i) oameni-**i**
men.M-DEF.PL.NOM
'the men' (Coresi, *Evanghelie cu învăţătură* 13:Pană Dindelegan 2016)

10. Note that (35) is a one-way correlation, which means that there can be languages that lack definite articles or have affixal definite articles but do not allow violations of the CSC I due to some additional factors. See footnote 15 on this.

which is repeated here as (36) from Chapter 1.

(36) *Generalization of adverb LBE out of predicative TAPs*

Languages that allow Adv-extraction out of predicative TAPs either lack definite articles or

have affixal definite articles.

Thus, the generalization (35) supports Talić's argument that the two-way distinction of NP/DP

languages is not sufficient, and that we need (at least) a three-way distinction of NP/DP languages.

Below, I will offer a deduction of the generalization (35) based on Talić's deduction of (36).

In section 2.3.2, I summarize Talić's deduction of (36), which appeals to the Structural Parallelism

hypothesis and the contextual approach to phasehood advocated by Bošković (2013a, 2014). In

section 2.3.3, I provide a deduction of (35) by extending Talić's account of (36) to the coordinate

structure.

2.3.2   Talic's (2015,2017) Structural Parallelism

As mentioned above, the generalization in (35) is quite similar to the one regarding adverbial LBE

established by Talić (2015, 2017). Talić shows that an intensifier adverb can be extracted from

a predicative traditional adjective phrase (TAP) only in languages without definite articles and

languages with affixal definite articles. Thus, English, Spanish, Italian and Brazilian Portuguese

do not allow adverb LBE (36)-(39), whereas SC, Russian, Polish, Icelandic and Romanian allow

it (40)-(44). Japanese, which was not discussed by Talić, patterns with BCS, Russian, Polish,

Icelandic, and Romanian in this respect, as shown in (45).

(37) *English*

  *$Terribly_i$ he was [t_i tired].

(38) *Spanish*

　　***Extremadamente**_i (yo) estoy [**t**_i cansado].

　　extremely　　　　　I　am　　　tired

　　cf. (Yo) estoy extremadamente cansado.

　　'I am extremely tired.'　　　　　　　　　　　　　　　(Talić 2015:420)

(39) *Italian*

　　***Estremamente**_i è [**t**_i intelligente].

　　extremely　　　　is　　smart

　　cf. È estremamente intelligente.

　　'He is extremely smart.'　　　　　　　　　　　　　　(Talić 2017:21)

(40) *Brazilian Portuguese*

　　***Terrivelmente**_i eu estou [**t**_i cansado].

　　terribly　　　　　I　am　　　tired

　　cf. Eu estou terrivelmente cansado.

　　'I am extremely tired.'　　　　　　　　　　　　　　　(Talić 2015:420)

(41) *BCS*

　　Strašno_i je bila　[**t**_i umorna].

　　terribly　is　been　　tired.F.SF

　　cf. Je bila strašno umorna.

　　'She was terribly tired.'　　　　　　　　　　　　　　(Talić 2015:420)

(42) *Russian*

　　Ušzasno_i ja byl [**t**_i rad　　tebja videt'].

　　terribly　I　was　　glad.SF　you　see

　　cf. Ja byl Užasno rad tebja videt'.

　　'I was very glad to see you.'　　　　　　　　　　　　(Talić 2015:420)

36

(43) *Polish*

Okropnie$_i$ on był [**t$_i$** zmęczony].

terribly   he was   tired

cf. On był okropnie zmęczony.

'He was terribly.'   (Talić 2015:420)

(44) *Icelandic*

Rosalega$_i$ er hún [**t$_i$** falleg].

extremely is she   beautiful.SG.F

cf. Hún er rosalega falleg.

'She is extremely beautiful.'   (Talić 2015:420)

(45) *Romanian*

Foarte$_i$ sunt [**t$_i$** obositā].

very   am   tired

cf. Sunt foarte obositā.

'I am very tired.'   (Talić 2015:420)

(46) *Japanese*

Totemo$_i$ John-wa [**t$_i$** shinsetsu] da.

very   John-TOP   kind   is

cf. John-wa totemo shinsetsu da.

'John is very kind.'

Based on these data, Talić (2015, 2017) provides the generalization in (47), repeated from (36):

(47) *Generalization of adverb LBE out of predicative TAPs*

Languages that allow Adv-extraction out of predicative TAPs either lack definite articles or
have affixal definite articles.

Notice again that Talić's classification of the languages that allow adverb LBE out of predicative

37

TAPs is the same as the one regarding CSC I violations that I provided in the previous section.

Talić proposes an account of the generalization in (47) based on Bošković's (2013a, 2014) approach to LBE out of traditional noun phrases (TNPs). Bošković (2008b, 2012) established the generalization regarding LBE out of TNPs in (47).[11]

(48) *Generalization of adjective LBE out of TNPs*[12]

Only languages without definite articles may allow adjective LBE, while languages with definite articles never allow it.

(49) *BCS*

Pametni_i su oni [t_i studenti].

smart are they students

'They are smart students.'

(50) *English*

****Smart**_i they are [**t**_i students].

To account for (48), Bošković adopts Chomsky's (2000) Phase Impenetrability Condition (PIC):

---

11. Bošković (2012) gives the following languages as allowing adjective LBE, all of which lack definite articles: SC, Russian, Polish, Czech, Ukrainian, Slovenian, Latin, Mohawk, Southern Tiwa, Gunwinjguan languages, Hindi, Bangla, Angika, and Magah. Franks (2007) observes that Colloquial Finnish has developed a definite article and disallows LBE, unlike literary Finnish, which does not have a definite article and allows LBE. Taylor's (1990) also observes that the occurrences of split NPs/wh-phrases decreased from the Homeric Greek period to the Koine Greek period, which Bošković (2012) takes as indicating that the development of definite articles played an important role in the loss of adjective LBE and hence as evidence for his generalization regarding the correlation between the availability of definite articles and adjective LBE.

Interestingly in this context, extraction of a conjunct is attested in the Homeric Greek period (8th century BC) and the Classical Greek period (5th-4th century BC), where definite articles were not fully developed, but it is not reported in the Koine Greek period (1st century AD), where definite articles were fully developed (Agbayani and Golston 2010, Devine and Stephens 2000). This coincides with the above observation that the development of a definite article led to loss of adjective LBE in Ancient Greek.

12. Note that this is a one-way correlation, which means that there are languages without definite articles that do not allow adjective LBE due to additional factors (see Bošković 2013a on these additional factors). The reader should also bear in mind that the possibility of adjective LBE does not necessarily have to coincide with the possibility of violations of the CSC I in an article-less language, since there can be additional factors that block either phenomenon in a particular language (see footnote 15 on the CSC I). (In fact, adjective LBE is blocked in affixal article languages (Despić 2011, 2015, Talić 2015, 2017); see Despić (2011, 2015), who proposes that AP dominates NP in languages that have definite articles, as Abney (1987) proposes (whether they are affixal or not), hence AP cannot move to the exclusion of NP. See also chapter 5 and 6 on this.) Note also that this holds for Talić's generalization (47) as well. In other words, the issue here is that in all these cases we are dealing with one-way, not two-way, correlations; as a result, the phenomena in question will not necessarily co-occur in a given language.

after Spell-Out (completion of a phase), only the head of the phase and its edge (specifiers and adjuncts) remain accessible for further syntactic operations, as a result of which movement out of a complement of a phase head is blocked after Spell-Out. He also adopts the anti-locality condition argued for in Bošković (1994), Saito and Murasugi (1999), Ishii (1999), Abels (2003), Grohmann (2003), among others: in the formulation given in Bošković (2013a), movement has to cross at least one full phrase, not a segment. In addition, Bošković (2013a, 2014) argues that the highest phrase in the extended projection of a lexical head, including NP, constitutes a phase. Based on a number of cross-linguistic generalizations like the one in (48), Bošković argues that there is a structural difference between languages with and without definite articles where the latter lack DP. As a result, in languages with definite articles, DP is a phase in the TNP (as the highest phrase in the extended projection of N) whereas in languages without definite articles, NP is a phase in the TNP. Consequently, in languages with definite articles, when an AP, which Bošković assumes is adjoined to NP, undergoes LBE, this AP either has to violate the PIC to satisfy the anti-locality condition if it moves directly out of DP, as in (51a), or has to violate the anti-locality condition to obey the PIC if it moves to Spec,DP before Spell-Out since it crosses just a segment, not a full phrase, as in (51b).

(51) a.

b.
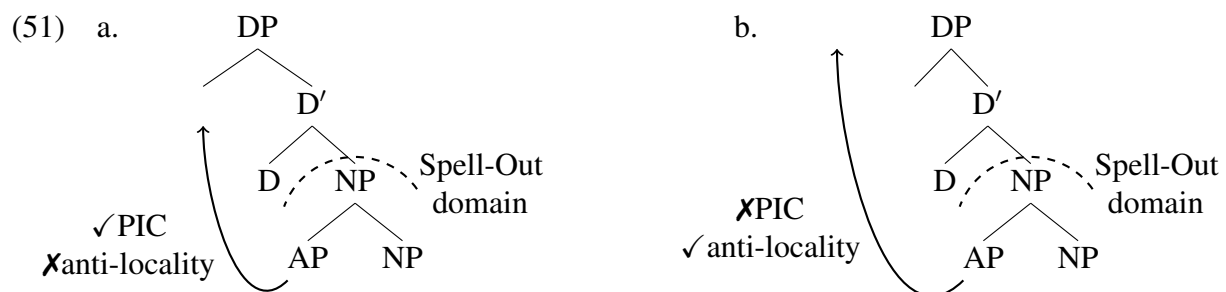
In contrast, in languages without definite articles, the highest projection of a TNP is NP, so that the AP undergoing LBE violates neither the PIC nor the anti-locality condition, as shown in (51).

(52)

$$\text{NP} \quad\quad \overset{\displaystyle\text{NP}}{\overset{\diagup\diagdown}{\text{AP}\quad\text{NP}}}$$

✓ PIC
✓ anti-locality
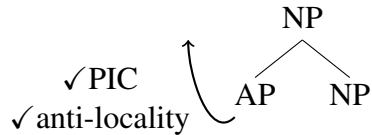
Talić (2015, 2017) extends this idea to adverbial LBE. She proposes the following condition regarding a functional layer above a lexical projection:

(53)   Structural Parallelism (Talić 2015, 2017)

   a.   If a language allows bare lexical structure without a functional layer in the domain of one lexical category, it may allow bare lexical structure in the domain of other lexical categories (e.g., a language can have both bare NP and bare AP).

   b.   If a language never allows bare lexical structure, that is, it always requires a functional layer in the domain of one lexical category, it must have a functional layer in the domain of all lexical categories (e.g., such a language will never have bare NP or bare AP).

According to (53a), given that languages without definite articles may lack DP above NP, those languages may also lack a functional projection above AP. Thus, in languages like BCS, adverb LBE out of a predicative TAP is possible on a par with LBE out of a TNP, as in (54), the underlying assumption being that the adverb in question is AP-adjoined.

(54)

$$\text{AP} \quad\quad \overset{\displaystyle\text{AP}}{\overset{\diagup\diagdown}{\text{AdvP}\quad\text{AP}}}$$

✓ PIC
✓ anti-locality
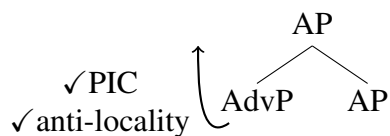
On the other hand, as (53b) states, languages with non-affixal definite articles always have a functional projection above NP or AP. This means that as in the NP domain, there must be a functional projection above AP, as a result of which adverb LBE out of a predicative TAP is not allowed, just like LBE out of a TNP is not, as illustrated in (55).

(55) a. F$_{adj}$P
          |
          F$_{adj}$′
         /   \
      F$_{adj}$ AP ⟍ Spell-Out
                 / \   domain
              AdvP   AP

      ✓PIC
      ✗anti-locality

      b. F$_{adj}$P
          |
          F$_{adj}$′
         /   \
      F$_{adj}$ AP ⟍ Spell-Out
                 / \   domain
              AdvP   AP

      ✗PIC
      ✓anti-locality
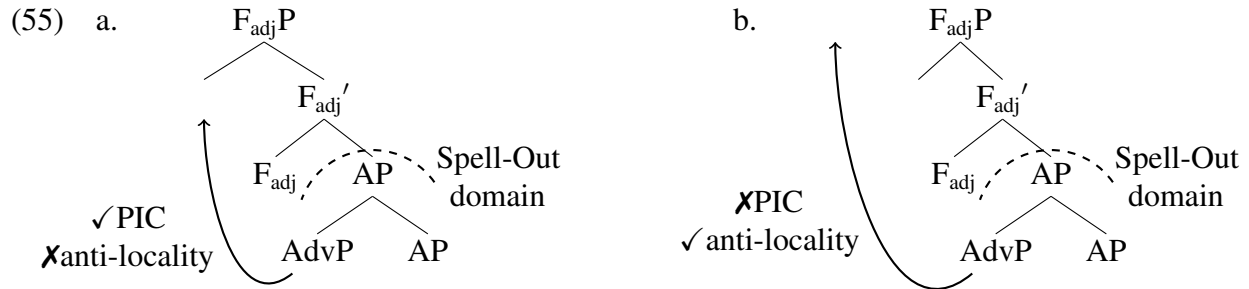
What about affixal-article languages? The data regarding adverbial LBE indicate that affixal article languages may lack a functional layer above a TAP, since these languages pattern with article-less languages in the relevant respect. We may then expect to find similarities between affixal article languages and article-less languages in the nominal domain that would indicate that TNPs in affixal article languages may lack a DP layer (when the article is not present). Talić (2015, 2017) in fact argues that this is indeed the case. More specifically, she shows that in affixal article languages a definite article is not required when a definite article is not semantically motivated: that is, when definiteness/uniqueness is encoded in something other than the definite article. One such case is superlatives, whose uniqueness is standardly assumed to be encoded by the superlative morpheme (*-est* in English). Thus, unlike in English (56), the definite article is optional in Bulgarian as in (57).

(56)   Ivan has ***(the)** best albums by U2.

(57)   *Bulgarian*

   a.   Ivan ima naj-dobri-**te**   albumi ot U2.
        Ivan has SPRL-good-the albums by U2

   b.   Ivan ima naj-dobri   albumi ot U2.
        Ivan has SPRL-good albums by U2

        'Ivan has the best albums by U2.'   (Pancheva and Tomaszewicz 2012:295-296)

There is also a difference in terms of interpretation of superlatives. Pancheva and Tomaszewicz

(2012) observe that English (56) only has the interpretation (59a), whereas Polish (58) allows both (59a) and (59b). They note that Czech, BCS, and Slovenian pattern with Polish in this respect.

(58) *Polish*

Iwan ma naj-lepsze albumy U2.

Ivan has SPRL-better.ACC albums.ACC by

'Ivan has the best albums by U2.' (Pancheva and Tomaszewicz 2012:295)

(59) a. 'Ivan has better albums by U2 than anyone else does.'

b. 'Ivan has better albums by U2 than by any other band.'

Crucially, they also observe that Bulgarian superlatives without a definite article as in (57b) have both the reading (59a) and the reading (59b). Shen (2014) in fact argues that the DP layer is absent in Bulgarian when the definite article is absent, which supports Talić' idea that when a language allows a bare AP, it also allows a bare NP (and vice versa).

In addition, definite articles in affixal article languages can be omitted in an environment where a prototypical interpretation of a definite article is absent (so-called *weak definites*; see Aguilar-Guevara 2014 and Scholten 2010). Thus, in (59)-(61), a definite article is omitted even though it is obligatory in English.

(60) *Icelandic*

a. Hún fór til tannlæknis.

she went to dentist

'She went to the dentist.'

b. Ég tók rútu í skóla-nn.

I took bus in school-the

'I took the bus to school all my life.'

42

c. Hann fór   út   í   búð.

   he   went out in store

   'He went to the store.'   (Talić 2015:432)

(61) *Bulgarian*

   a. (Toj) slusha radio.

      (he) listens radio

      'He is listening to the radio.'

   b. (Tja) otide na zəbolekar.

      (she) went to dentist

      'She went to the dentist'

   c. Cjal   jivot pətuvah s   avtobus.

      whole life   travelled with bus

      'I travelled with the bus all of my life.'   (Talić 2015:432)

(62) *Romanian*

   S-a   dus   la pravalie.

   REFL-has went to store.INDEF

   'He went to the store.'   (Talić 2015:432)

Talić (2015, 2017) takes this as another piece of evidence that the DP layer can be absent in TNPs

of affixal article languages.[13]

   To summarize so far, we have seen that both violations of the CSC I and adverb LBE from pred-

icative TAPs may be allowed only in languages without definite articles and languages with affixal

definite articles. I have then discussed Talić's (2015, 2017) phase-based proposal regarding adverb

---

13. See footnote 12 for an analysis of why LBE out of TNPs is disallowed in affixal article languages. See also Despić
(2015) for a similarity between article-less languages and affixal article languages regarding reflexive possessives,
which was briefly mentioned in chapter 1 and will be discussed in chapter 5; his analysis also appeals to the affixal
nature of D in the latter type of languages.

LBE which is tied to the Structural Parallelism hypothesis, according to which languages that lack the functional layer in a lexical projection (e.g., NP) may lack it in another lexical projection (e.g., AP). I now turn to the deduction of the generalization in (35).

2.3.3 Deduction of the generalization (35): The Structural Parallelism in the Conjunction Phrase

The discussion in section 2.3.2 immediately raises a question regarding CSC I: given that adverb LBE and CSC I violations are both (in principle) allowed in both article-less and affixal article languages, can we explain CSC I violations in line with Talić's (2015, 2017) approach to adverb LBE? In this subsection, I propose a phase-based account of the cross-linguistic variation regarding CSC I violations by extending Talić's Structural Parallelism hypothesis to the coordinate structure.

First, I propose, following Chino and Hiraiwa (2014), Kayne (1994), and Zwart (2005, 2009), that ConjP is universally head-initial, even if the language is otherwise head-final (see especially Zwart 2005, 2009 for a cross-linguistic survey). In addition, I propose, following Kayne (1994) and Stjepanović (2014), that the first conjunct is left-adjoined to ConjP. Thus, the structure of a coordinate structure is (63).

(63)   ConjP
          /\
      XP   ConjP
            /\
        Conj   YP
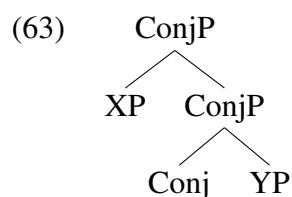
The proposal that conjuncts are left-adjoined can capture some parallelisms between LBE and CSC I violations, on the assumption that adjectives and adverbs are adjoined to NPs and APs respectively (Bošković 2013a, Talić 2015, 2017). First, Stjepanović (2014), who discusses SC, observes that LBE and CSC I violations are allowed and disallowed in the same syntactic environments in SC. Thus, both LBE and CSC I violations are disallowed from a genitive complement of a noun (64) and from a complement of a noun modified by a quantifier (65).

(64) a. *Čije_i je on [djecu   [_{NP} t_i prijateljice]] vidio?   (LBE)

　　　　whose is he kid.ACC   friend.ACC   seen

　　　　'The kids of whose friend did he see?'

　　b. *Marije_i   je on [djecu   [_{ConjP} t_i i   Petra]]   video.   (CSC I violation)

　　　　Mariha.GEN is he kid.ACC   and Petar.GEN seen

　　　　'He saw [Marija and Peter]'s kids.'   (Stjepanović 2014:162)

(65) a. *Čije_i je on upoznao [mnogo djece [t_i majke]]?   (LBE)

　　　　whose is he met   many kids   mother

　　　　'Whose mother did he meet many kids of?'

　　b. *Marije_i   je on upoznao [mnogo djece [t_i i   Petra]].   (CSC I violation)

　　　　Marija.GEN is he met   many djece   and Petar.GEN

　　　　'He met many Marija and Peter's kids.'   (Stjepanović 2014:162)

On the other hand, LBE and CSC I violations are both allowed from an inherently case-marked complement of a noun (65) and from an adjunct (66).

(66) a. Kakvom_i   ga je prijetnja [t_i smrću]   uplašila?   (LBE)

　　　　what.kind.GEN him is threat   death.INSTR scared

　　　　'The threat of what kind of death scared him?'

　　b. Zatvorom   ga je prijetnja [t_i i   ubistvom]   uplašila.   (CSC I violation)

　　　　prison.INSTR him is threat   and murder.INSTR scared

　　　　'The threat of prison and murder scared him.'   (Stjepanović 2014:162)

(67) a. [Zbog   čijih]ᵢ je on to   [tᵢ studenata] uradio?   (LBE)

    because.of whose is he that   students   done

    'Because of whose students did he do it?'

    b. [Zbog   mene]ᵢ je on to   [tᵢ i   njih] uradio.   (CSC I violation)

    because.of me   is he that   and them done

    'He did it because of me and them.'   (Stjepanović 2014:162)

Stjepanović (2014) therefore claims that LBE and CSC I violations are essentially the same phe-
nomena.

The second argument for the similarity between LBE and CSC I violations comes from recon-
struction effects in Japanese scrambling. Recall from section 2 that Japanese allows violations of
the CSC I. In addition, Takahashi and Funakoshi (2013) and Shiobara (2017) show that LBE in
Japanese is possible (but rather restricted). Crucially, Arano and Oda (2019) show that neither LBE
nor movements that involve CSC I violations which do not cross a clause boundary affect scope
and binding, even though clause internal scrambling (i.e., scrambling that does not cross a clause
boundary) can otherwise affect scope and binding. It is well-known that Japanese is a scope-rigid
language, as exemplified by (68a), (68c), and (68e).[14] As shown in (68b), when a scope bearer
('everyone') undergoes clause-internal scrambling and crosses another scope bearer ('someone'),
the scrambled element can take wide scope over the other scope bearer. However, when the same
scope bearer ('everyone') undergoes LBE within a single clause, it cannot take scope over the other
scope bearer, as shown in (68d). Crucially, when the first conjunct is scrambled clause-internally,
it does not affect scope either, just like LBE, as shown in (68f).

---

14. It should be noted that *minna-e-no* 'to-everyone' in (68c) is not a complement of *tegami* 'letter' but a modifier
adjoined to the NP, unlike its English counterpart; see Takahashi and Funakoshi (2013).

(68) Scope (taken from Arano and Oda 2019)

    a.   Dareka-ga   minna-e   tegami-o   kaita.

       someone-NOM everyone-to letter-ACC wrote

       'Someone wrote a letter to everyone.'   ($\exists > \forall$; *$\forall > \exists$)

    b.   **Minna-e$_i$**   dareka-ga   **t$_i$** tegami-o   kaita.   (scrambling)

       everyone-to someone-NOM   letter-ACC wrote

       'To everyone, someone wrote a letter.'   ($\exists > \forall$; $\forall > \exists$)

    c.   Dareka-ga   [$_{NP}$ minna-e-no   [$_{NP}$ tegami]]-o kaita.

       someone-NOM   everyone-to-GEN   letter-ACC wrote

       'Someone wrote [[a letter] to everyone].'   ($\exists > \forall$; *$\forall > \exists$)

    d. ?**Minna-e-no$_i$**   dareka-ga   [$_{NP}$ **t$_i$** [$_{NP}$ tegami]]-o kaita.   (LBE)

       everyone-to-GEN someone-NOM   letter-ACC wrote

       '[To everyone]$_i$ someone wrote [[a letter] t$_i$].'   ($\exists > \forall$; $\forall > \exists$)

    e.   Dareka-ga   [$_{ConjP}$ san-bon-izyoo-no   ronbun-to hon]-o   yonda.

       someone-NOM   three-CL-more.than-GEN paper-and book-ACC read

       'Someone read [[more than three papers] and books].'

                                       ($\exists >$ more than 3; *more than 3 $> \exists$)

    f. ?**[San-bon-izyoo-no**   **ronbun-to]$_i$** dareka-ga   [$_{ConjP}$ **t$_i$** hon]-o   yonda.

       three-CL-more.than-GEN paper-and   someone-NOM   book-ACC read

       lit.'[More than three papers and]$_i$ someone read [t$_i$ books].'   (CSC I violation)

                                    ($\exists >$ more than 3; more than 3 $> \exists$)

Binding tests show the same pattern as scope. (69a), (69c), and (69e) are cases where the anaphor 'each other' is not c-commanded by its antecedent. When the antecedent undergoes clause-internal scrambling, it can bind the anaphor, as in (69b). However, when the antecedent undergoes LBE within a clause, it does not bind the anaphor, as shown in (69d). The same result is observed with extraction of the first conjunct as in (69f).

(69) Binding (adapted from Arano and Oda 2019)

    a. *[Otagai$_i$-no   sensei]-ga   [John-to   Mary]$_i$-o   hihanshita.

       each.other-GEN teacher-NOM   John-and   Mary-ACC   criticized

       'Each other's teachers criticized John and Mary.'

    b. **[John-to   Mary]$_i$-o**   [otagai$_i$-no   sensei]-ga   **t$_i$** hihanshita.   (scrambling)

       John-and   Mary-ACC   each.other-GEN teacher-NOM   criticized

       '[John and Mary]$_i$, each other's teachers criticized t$_i$.'

    c. *Otagai$_i$-ga   [$_{NP}$ [John-to   Mary]$_i$-no [$_{NP}$ tegami]]-o yonda.

       each.other-NOM   John-and   Mary-GEN   letter-ACC   read

       'Each other read John and Mary's letter.'

    d. *[**John-to   Mary**]$_i$**-no**$_1$ otagai$_i$-ga   [$_{NP}$ **t$_1$** [$_{NP}$ tegami]]-o yonda.   (LBE)

       John-and   Mary-GEN   each.other-NOM   letter-ACC   read

       '[John and Mary's]$_i$ each other read [t$_i$ letter].'

    e. *Otagai$_i$-ga   [$_{ConjP}$ karera$_i$-to John]-o   hihanshita.

       each.other-NOM   they-and   John-ACC   criticized

       'Each other criticized them and John.'

    f. ***Karera$_i$-to**$_1$ otagai$_i$-ga   [$_{ConjP}$ **t$_1$** John]-o   hihanshita.   (CSC I violation)

       they-and   each.other-$_{nom}$ John-ACC   criticized

       lit.'[Them-and]$_1$, each other criticized [t$_1$ John].'

The observations regarding scope and binding thus show that LBE and CSC I violations are similar to each other, which can be interpreted as indicating that a conjunct is adjoined to ConjP just like an adjective and an adverb are adjoined to NP and AP, respectively, which gives us a similar configuration in all these cases.

Now, the proposed structure (63) can explain the CSC I violation cases. I assume that ConjP projects a phasal domain (see also Bošković in press and Stjepanović 2014). Following Bošković's

(2013a, 2014) contextual phasehood approach, the highest phrase in the conjunction domain is then the phase. In article-less languages, ConjP is the phase, so that the initial conjunct, which is left-adjoined to the edge of ConjP, can move out of ConjP after Spell-Out without violating the PIC or the anti-locality condition, on a par with adjective LBE out of TNPs and adverb LBE out of TAPs. This is illustrated in (69).[15]

(70)
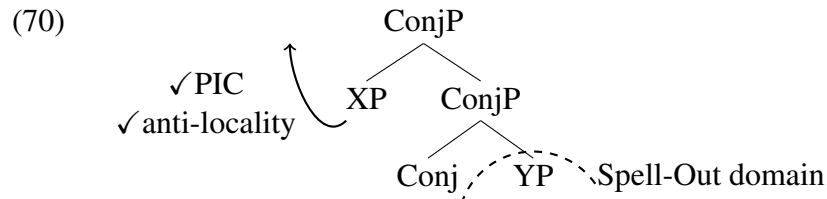
This analysis predicts that the second conjunct cannot move out of a coordinate structure. When the second conjunct (YP) is extracted, it either has to violate the PIC to satisfy the anti-locality condition if it moves directly out of ConjP, or has to violate the anti-locality condition to obey the PIC if it moves to the edge of ConjP before Spell-Out since it would cross just a segment, not a full phrase. This prediction is borne out, as shown in (70).[16]

(71) Japanese

*$\mathbf{Toodai\text{-}ni_i}$	kanojo-wa	[Kyoodai-to	$\mathbf{t_i}$]	akogareteiru.
Tokyo.University-and	she-TOP	Kyoto.University-and		admire

‘She admires Kyoto University and Tokyo University.’   (Oda 2017)

Turning to non-affixal article languages, I propose that Talić's (2015, 2017) Structural Parallelism is extended to ConjP; more specifically, just like these languages require a functional projection above NP and AP, they also require a functional projection above ConjP, which I call FconjP.[17] Given that the highest phrase of an extended projection is a phase (Bošković 2014), it

---

15. Recall that like (47) and (36), (35) is a one-way generalization. It does not mean that all languages without definite articles will allow CSC I violations. In fact, Slovenian and Tamil, which lack definite articles, disallow CSC I violations. It then seems that there is an additional factor concerning CSC I violations. For what this additional factor may be, see Stjepanović (2014) and Oda (2017).

16. The same holds for SC; see Stjepanović (2014). We will, however, see below that an additional problem arises with extraction of lower conjuncts.

17. A candidate for the realization of the head of FconjP may be 'both'. See below for discussion.

follows that this FconjP is a phase instead of ConjP in non-affixal article languages.

At this point, one might wonder if it is appropriate to extend the Structural Parallelism, which is originally stated to apply to lexical projections, to ConjP, which seems to be a functional projection rather than a lexical projection. It is standardly assumed that ConjP "inherits" certain properties of the conjuncts. I thus assume that ConjP "inherits" the nature of lexical projections from its conjuncts, so that ConjP can be considered as a sort of a lexical projection. The intuition behind this idea is that when NPs, APs, or VPs are conjoined, the whole coordinate structure also functions as NP, AP, or VP, respectively. In fact, Zoerner (1995) argues that ConjP lacks inherent categorial features such as $[\pm V]$ and $[\pm N]$, and inherits the relevant feature specifications of its conjunct (see also Biberauer et al. (2014) for a similar idea from a viewpoint of word order restrictions). Thus, based on Zoerner (1995), I suggest that ConjP has an unspecified/unvalued categorial feature, whose value is determined by "feature-sharing" (Pesetsky and Torrego 2007, Bošković 2011a; Chomsky 2013) with its conjuncts. (It should be added here that ConjP does not end up being NP, VP, etc. after the feature sharing, but remains as ConjP with its categorial feature specified (e.g., ConjP$_{[+N]}$, ConjP$_{[+V]}$).) Under this proposal, ConjP can be considered as an unspecified lexical category that takes over categorial status of conjuncts, as a result of which the application of the Structural Parallelism would not go beyond the scope of Talić's original statement.[18]

Notice now that this FconjP structure correctly excludes extraction of the first conjunct in non-affixal article languages. More specifically, when the first conjunct moves out of a coordinate structure, this movement has to violate either the PIC or the anti-locality condition in the same way as LBE. If the first conjunct moves to Spec,FconjP not to violate the PIC, it violates the anti-locality condition since XP crosses a segment and not a full category (72a). If it moves to a higher projection than FconjP to obey the anti-locality condition, it violates the PIC (72b). Thus, the first conjunct cannot move out of the coordinate structure.

---

18. Another possibility is that the Structural Parallelism is not limited to lexical projections but can be extended to functional projections (under the assumption that ConjP is a functional projection). It is worth mentioning here that Bošković (2012) suggests that article-less languages may also lack TP in the clausal spine as the counterpart of lacking DP in the TNP (see also Kang 2014 and Todorović 2016), which is essentially Structural Parallelism holding for functional projections.

(72) a.

```
            FconjP
               |
            Fconj′
         Fconj   ConjP    Spell-Out
   ✓PIC              domain
  *anti-locality  XP   ConjP
                     Conj   YP
```

b.

```
            FconjP
               |
            Fconj′
         Fconj   ConjP    Spell-Out
   *PIC               domain
  ✓anti-locality  XP   ConjP
                     Conj   YP
```
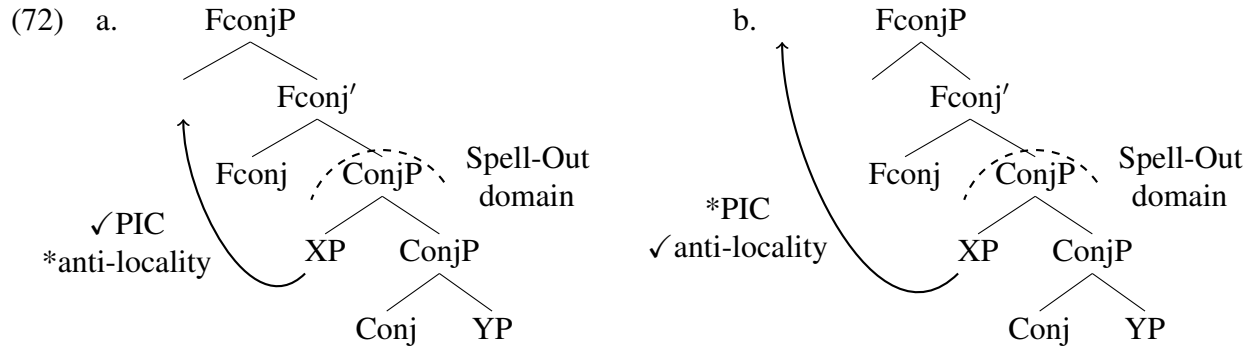
As for the second conjunct, I suggest that the movement of the second conjunct is blocked by Rizzi's (1990, 2004) Relativized Minimality, i.e., we are dealing here with another intervention effect. Following Bošković (2020b) and Johnson (2002), I assume that the first conjunct induces an intervention effect for extraction of the second conjunct. To implement this suggestion, I assume that each conjunct has a coordination feature which is necessary to participate in coordination, and thus this feature induces a Relativized Minimality violation when the second conjunct moves across the first conjunct, although the precise technicality has to be worked out in future research.[19]

Let us now turn to affixal article languages. Recall that affixal article languages allow CSC I violations, just like article-less languages. This can be naturally explained by the current proposal. The Structural Parallelism allows languages that have a bare lexical structure in one domain to have a bare structure in another domain. When this is extended to ConjP, it follows that ConjP may lack FconjP in affixal article languages, since these languages lack a functional projection above TAPs, which allows adverb LBE out of predicative TAPs. Thus, the coordinate structure in these languages has the same structure as in article-less languages, and hence a CSC I violation is allowed both in affixal article languages and article-less languages.

Notice that this analysis does not exclude the possibility that affixal article languages that allow CSC I violations may have FconjP in some circumstances. As mentioned in footnote 17, 'both' can be a candidate for realization of Fconj^0. In fact, in Swedish and Icelandic, when 'both' appears in the coordinate structure, extraction of the first conjunct is disallowed, as shown in (i). This can be interpreted as evidence for the suggestion that 'both' is Fconj^0, which blocks CSC I violations.

---

19. Actually, this analysis also extends to languages that lack FconjP.

(73)   a. *Pétur_i sá   ég **bæði** [t_i og   Maríu].

        Peter saw I   both   and Mary   (Icelandic: Gísli Rúnar Harðarson, p.c.)

   b. *[kalle Jularbo]_i hörde jag **både** [t_i och hans gamla dragspelsorkester]

        K.J.   heard I   both   and his   old   accordion.band

                                 (Swedish: Anders Holmberg, p.c.)

The optional presence of FconjP in these languages makes sense under the current proposal which appeals to the Structural Parallelism. Recall from section 2.3.2 that affixal article languages can optionally have a definite article with superlatives, and when a definite article is present, only one interpretation is possible just like in non-affixal article languages, which indicates that a functional projection relevant for disallowing ambiguity of superlatives is present with the presence of a definite article in affixal article languages. This is quite similar to the case of 'both' in the coordinate structure: when 'both' is present, there is a functional projection relevant for blocking extraction of the first conjunct. Thus, it is not implausible that FconjP can be present in affixal article languages in the presence of 'both' as realization of Fconj^0.

To conclude this section, I have shown that CSC I violations pattern with adverb LBE out of predicative TAPs in that both are allowed in article-less and affixal article languages but not in non-affixal article languages. I have then proposed a phase-based account of CSC I violations following Talić's (2015, 2017) Structural Parallelism and her account of adverb LBE out of predicative TAPs.

2.4   The CSC II and Across-the-board Movement: Bošković's (2020) Labeling Approach

So far, I have discussed the CSC I, one of the two locality conditions of a coordinate structure. In this section, I discuss the other condition, the CSC II, based on Bošković's (2020b) labeling approach in connection to the current proposal for the CSC I.

Recall that I have argued that the traditional Coordinate Structure Constraint should be sepa-

rated into two conditions, the CSC I and the CSC II, and that both traditional ConjP (TConjP) and each conjunct are islands independently of each other. There are two important questions that arise from these arguments. First, what is the nature of the CSC II? Under the current proposal, the CSC I is deduced from independently established syntactic conditions: the PIC and the anti-locality condition. Here one might argue that the CSC II should be treated in the same way as the CSC I, because they are both related to islandhood of a coordinate structure (TConjP and each conjunct) and the islandhood of TConjP (the CSC I) is deduced from the PIC. Notice, however, that there is no principled natural connection (in the relevant respect) between TConjP and the conjuncts apart from both being involved in a coordinate structure. Rather, it is perfectly logically possible that the islandhood of the TConjP and that of the conjunct could come from completely different mechanisms.

The second question concerns the so-called across-the-board (ATB) movement. It is well-known that extraction out of a coordinate structure is possible even in non-affixal article languages like English, when an element is extracted out of each conjunct, as exemplified by (74).

(74)   **Who$_1$** did you see [[friends of **t$_1$**] and [enemies of **t$_1$**]]?

This is surprising given the argument in the present chapter that TConjP and the conjuncts are islands, because who in (74) is extracted out of TConjP, the first conjunct, and the second conjunct, which should lead to a violation of both the CSC I and the CSC II. However, the fact is that the sentence is perfectly grammatical. This raises an issue for the current proposal.

In this context, Bošković (2020b) proposes an interesting account of the CSC II and ATB movement. He argues that the CSC II is essentially a requirement on conjunct labeling and that ATB movement does not violate this requirement (in contrast to extraction out of only one conjunct). He assumes following Bošković (in press) and Oda (2017) that conjuncts are phases. This is derived from Bošković's (2013a, 2014) contextual phasehood approach that I have also adopted here: whatever the category of the conjunct is (e.g., NP, AP, VP), when a Conj head merges with the conjunct, the extended domain of the conjunct is closed, so that the highest phrase of the conjunct becomes a phase. When a phrase undergoes successive-cyclic movement, the phrase has to move

to the edge of phase to avoid the PIC, which in this case means that movement from a conjunct has to proceed through the edge of the conjunct.

Before showing how this deduces CSC II, it should be noted that a question that arises here is whether Bošković's contextual phasehood approach to conjuncts is compatible with the proposal regarding the application of the Structural Parallelism to ConjP suggested in section 2.3.3. Recall from section 2.3.3 that I suggested that ConjP "inherits" the categorial status of conjuncts. The implementation of the category inheritance suggested there is that Conj has an unvalued categorial feature which is valued via feature-sharing with conjuncts. An immediate issue that arises here is that if ConjP inherits categorial status of conjuncts as a lexical category, it might count as an extended projection of the conjuncts, which would prevent the conjuncts from being phases under Bošković's contextual approach to phasehood. I suggest that a specific definition of extended projections and the timing of evaluation matter here. The intuition behind the notion of extended projections is that certain functional categories share the same categorial status with their lexical base (e.g., DP, QP and NP as nominal elements within the TNP). Biberauer et al. (2014) in fact propose that extended projections have the same categorial feature inherently specified such as [+V], [-N]. Based on this proposal, I suggest that whether a phrase counts as an extended domain of a lexical projection depends on whether the head of the phrase has the same inherently specified categorial feature with its sister at the point of merger. If head X, which has a [+N] feature, merges with YP, which also has a [+N] feature, XP is part of the extended domain of the nominal projection to which YP belongs. On the other hand, if a head X merged with the YP does not have a [+N] feature (e.g., a verb which has [+V]), the nominal domain is closed at YP and hence YP becomes a phase under Bošković's implementation of phasehood. Recall now that Conj does not have an inherently specified categorial feature under the proposal in this chapter. This means that when Conj merges with the conjuncts, it does not count as an extended projection of the conjuncts due to the lack of an inherently specified categorial feature.[20] Then, the highest projection of the conjuncts

---

20. The assumption that ConjP is universally head-initial even in otherwise head-final languages (Chino and Hiraiwa 2014, Kayne 1994, Zwart 2005, 2009) could be deduced from this proposal. Whether a phrase is head-final or head-initial depends on the inherent categorial status of the phrase (see, e.g., Biberauer et al. 2014). In German, for example, nominal projections are head-initial but verbal projections are head-final. As for conjunction, whether it is head-initial

serves as a phase and ConjP counts as a distinct phasal domain, as proposed by Bošković, even though ConjP later inherits the categorial status of the conjuncts. This explains the dual status of Conj as a lexical domain under Talić's Structural Parallelism and as a distinct phasal domain from conjuncts (more precisely, as closing the conjunct phasal domain) under Bošković's contextual approach to phasehood.

Turning back to the account of the CSC II, Bošković also adopts Chomsky's (2013) labeling theory. In this theory, when a phrase merges with a head, the head projects, but when a phrase merges with another phrase, either they have to undergo feature-sharing or one of them has to move to a higher position so that the other one can project a label. Crucially, in this theory, when successive-cyclic movement targets a phase edge, the highest node is unlabeled since two phrases are merged together without feature sharing, as illustrated in (75). (All successive-cyclic movement is treated this way in the labeling framework: lack of feature sharing creates an intermediate structure like (75), which then forces movement.)

(75)
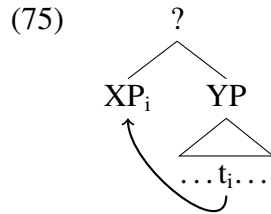
In addition, Bošković assumes a version of the Coordination-of-Likes requirement (CL) (Chomsky 1957, Schachter 1977, Williams 1978, Gazdar 1981, Sag et al. 1985, Bowers 1993, Beavers and Sag 2004, among many others), which "requires conjuncts to be parallel in their categorial status" (Bošković 2020b:136) and which applies derivationally (i.e., when ConjP is formed). Combining all the ingredients, the CSC II is now deduced from the CL. When movement of an element takes place out of only one of the two conjuncts, this movement delabels the conjunct so that the conjuncts are no longer categorially parallel, which results in a CL violation. This is why extraction

---

or head-final cannot be inherently specified, since the categorial status of conjunction is not inherently specified as discussed in the text. If Kayne (1994) is interpreted as indicating that languages are universally head-initial by default, it then follows that ConjP has to be head-initial as a default of UG in any language.

out of a conjunct is banned under Bošković's account.[21]

Let us now look at how Bošković (2020b) accounts for the CSC II and ATB movement. Consider first (76), where an element is extracted from only one conjunct, violating the CSC II.

(76) *Who$_1$ did you see [[enemies of t$_1$] and John]?

In (76), each conjunct is built first and then enters the coordinate structure. Within the first conjunct, *who* undergoes successive-cyclic movement to the edge of DP, which makes the topmost node unlabeled. After this movement, the conjuncts enter the coordination structure, at which point the CL is evaluated. Crucially, there is no parallelism in terms of categorial status between the first and the second conjuncts; the former lacks a label, while the latter is DP, as illustrated in (77).

(77)   [[$_?$ who$_1$ enemies of t$_1$] and [$_{DP}$ John]]

Thus, the CL is violated, which in turn means that the CSC II is violated, and hence the sentence is ungrammatical.

Consider next the ATB movement case (74), which is repeated as (78) here.

(78)   **Who$_1$** did you see [[friends of **t$_1$**] and [enemies of **t$_1$**]]?

Here, successive-cyclic movement occurs in both conjuncts. More precisely, Bošković (2020b) adopts Nunes's (2004) sideward movement analysis of ATB, in which *who* undergoes successive-cyclic movement to the edge of the second conjunct from the complement of *(enemies) of*, and then merges to the complement of *(friends) of*; from there who moves to the edge of the first conjunct. It should be noted here that the copy of *who* at the edge of the second conjunct and the copies of *who* in the first conjunct do not form a chain at this point, since they do not c-command each other. As a result, neither the copy at the edge of the first conjunct nor the one at the edge of the second conjunct counts as a trace (there is no higher copy that c-commands either of these elements). What we then have here is an {XP, YP} structure at the topmost node of each conjunct

---

21. Bošković shows that the ban in question actually holds only for successive-cyclic movement out of conjuncts, since only such movement has a de-labeling effect.

without feature sharing, which makes each conjunct unlabeled. The unlabeled conjuncts then enter the coordinate structure as in (79).

(79)   [[$_?$ who$_1$ friends of t$_1$] and [$_?$ who$_1$ enemies of t$_1$]]

Since both conjuncts are unlabeled, the CL is not violated. In other words, the two conjuncts are parallel in their categorial status in that both of them are unlabeled. Thus, under Bošković's approach to the CSC II, ATB movement is correctly predicted to be licit.

It should be noted here that the reason why Bošković adopts Nunes's sideward movement is to avoid an intervention effect. As is standardly assumed (and also assumed here), the first conjunct is structurally higher than the second conjunct; in other words, the first conjunct asymmetrically c-commands the second conjunct (see Munn 1993). Bošković then argues that if *who* were to move out of the second conjunct without sideward movement in (79), the first conjunct would count as an intervener, which would block extraction of *who* from the second conjunct. Notice, however, that movement of an element inside the second conjunct, which itself is not a conjunct, would not be blocked by the presence of the first conjunct under the Relativized Minimality account of the first conjunct intervention effect adopted in section 3.3. Recall that I suggested in section 3.3 that extraction of the second conjunct over the first conjunct is blocked by Relativized Minimality because both conjuncts have a coordination feature and that this feature induces an RM violation. Under this proposal, the first conjunct should not block movement out of the second conjunct (it only blocks movement *of* the second conjunct), since the element moving from inside the second conjunct itself is not a conjunct and hence does not have the coordination feature that is required to participate in coordination, which would then allow ATB movement even without Nunes's sideward movement. Thus, we do not have to assume Nunes's sideward movement to account for ATB movement under the current proposal.

At this point, the reader may wonder whether Bošković's proposal regarding the CSC II is compatible with the one in this chapter regarding the CSC I. Recall that I proposed in section 3.3 that in non-affixal article languages there is FconjP above ConjP, which serves as a phase, and that the reason why CSC I violations are not allowed in such languages is that FconjP counts

as a phase and when FconjP is completed, movement of the (initial) conjunct either violates the PIC or the anti-locality condition. Given this reasoning, one might argue that my proposal would incorrectly rule out ATB movement, because the complement of FconjP which includes the element undergoing ATB movement would be sent to Spell-Out.

I argue that the problem can be resolved once we consider the nature of $Fconj^0$ in detail. Recall that the categorial status of ConjP is dependent on each conjunct; that is, ConjP inherits the categorial status of each conjunct by categorial feature-sharing. Given Bošković's argument that successive-cyclic movement to the edge of a conjunct delabels the conjunct (before it merges with $Conj^0$), it follows that ConjP is unlabeled in the case of ATB movement, as illustrated in (80), since $Conj^0$ cannot inherit relevant features from unlabeled nodes. (Note that this does not violate the CL as discussed above.)

(80)

$? \Leftarrow$ unlabeled due to each conjunct being unlabeled

first conjunct $\Rightarrow$?

?

who_i XP Conj $? \Leftarrow$ second conjunct

$\ldots t_i \ldots$ who_i YP

$\ldots t_i \ldots$
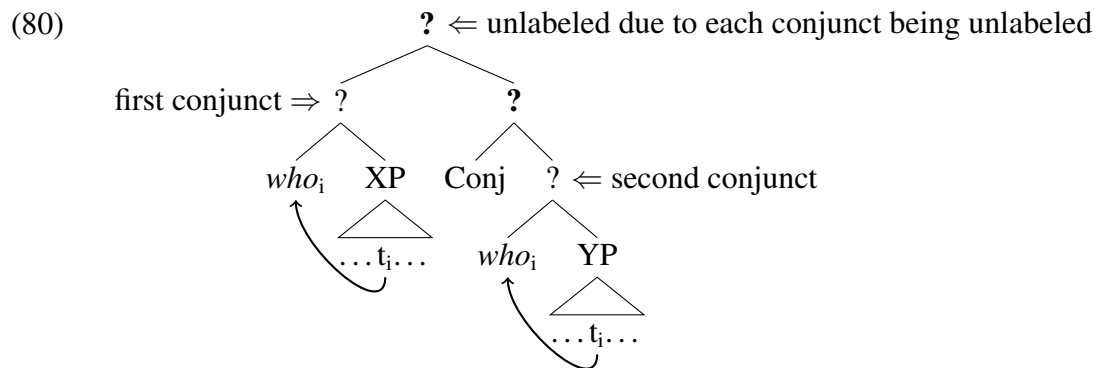
Recall also that FconjP is an extended functional projection of the ConjP domain. Since extended projections of a domain share the same categorial status with the bottom phrase of the same domain, it follows that FconjP has no categorial status either and hence is not labeled in the ATB case. This is illustrated in (81).

(81) **?** ⇐ unlabeled due to "ConjP" being unlabeled

```
                              ?  ⇐ unlabeled due to "ConjP" being unlabeled
                          ╱      ╲
                      Fconj       ?
                              ╱       ╲
        first conjunct ⇒  ?             ?
                        ╱   ╲         ╱   ╲
                    whoᵢ    XP     Conj    ?  ⇐ second conjunct
                          ╱‾‾╲            ╱   ╲
                         …tᵢ…          whoᵢ   YP
                                             ╱‾‾╲
                                            …tᵢ…
```
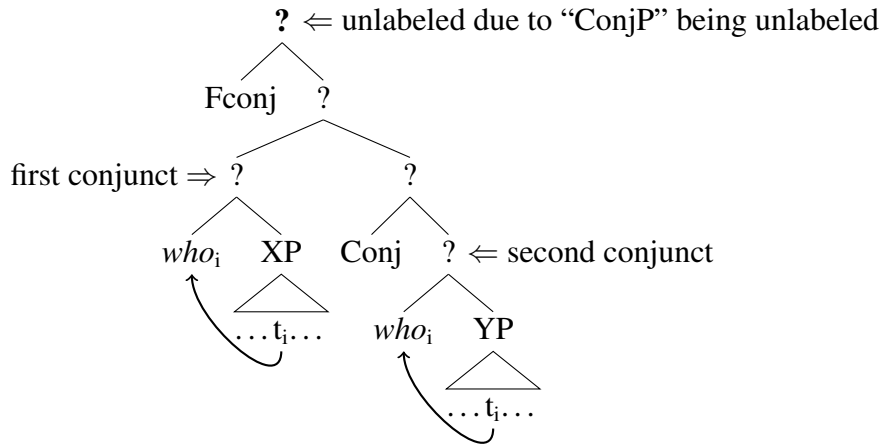
The issue now is whether the unlabeled highest node in (81) counts as a phase. Bošković (2020b) argues that unlabeled syntactic objects cannot be phases: under Chomsky's original proposal, CP, vP, and DP are phases, but unlabeled syntactic objects are none of these. Under the contextual phasehood approach (e.g., Bošković 2014), the highest phrase of an extended projection is a phase, but there is no way to determine whether the unlabeled node is the highest phase of an extended projection or not, since no categorial information relevant for determining an 'extended projection' is provided. Thus, I conclude that the highest node in (81), which is supposed to be FconjP, cannot be a phase. It then follows that ATB movement out of either conjunct is not blocked by the PIC, and thus ATB movement is allowed under the current proposal that assumes FconjP.[22]

We can now also address the question regarding the nature of the CSC I and the CSC II raised at the beginning of this section. The CSC II follows from the CL under Bošković's (2020b) account, which is essentially an interface condition required for interpretation, unlike the CSC I, which is essentially a constraint within narrow syntax.[23] Given this account, ATB is no longer a problem for the view that both TConjP and each conjunct are islands. The "islandhood" of each conjunct is now understood as the categorial parallelism requirement on conjuncts; as long as the requirement is met, extraction out of each conjunct is allowed.

---

22. Notice that this movement can target a projection higher than FconjP, which obeys the anti-locality condition. The reader should bear in mind that the current account of ATB differs from Bošković's since it resolves the intervention effect issue that arises under his account without the need to adopt Nunes's (2004) sideward movement account.

23. It should be noted that the CSC II is not a purely semantic condition though, since phases and the PIC are also involved. There is, however, no semantic component in the CSC I, in contrast to the CSC II.

It is worth noting here that ATB movement ameliorates CSC II violations but not CSC I violations. As (82) shows, an element can be extracted *out of* two conjuncts at the same time, but if two conjuncts are extracted at the same time, the sentence is ungrammatical. This contrast itself can be taken as another argument against unifying the two parts of the traditional CSC, as argued in section 2 of this chapter. It appears that ATB should save the CSC violations in both (82a) and (82b) if the two parts of the traditional CSC were a single condition.

(82)   a.   What$_i$ did [[Mary buy t$_i$] and [John sell t$_i$]]?

        b. *What$_i$ did Mary buy [t$_i$ and t$_i$]?

The impossibility of saving (82b) in contrast to (82a) by ATB movement in fact follows from the present proposal. Recall that extraction out of a single conjunct induces a CL violation, which is remedied by ATB movement out of each conjunct. What ATB movement remedies here is in fact only the CL/labeling problem that arises with CSC II violations. Extraction *of* conjuncts themselves, however, cannot be saved by ATB movement, because the CSC I is a pure syntactic locality condition and there is no CL/labeling problem involved in the first place. In other words, there is no CL/labeling problem to begin with here, so there is nothing that ATB can remedy. (Rather, when the second conjunct crosses the first conjunct, Relativized Minimality is violated; we are then dealing here with a pure syntactic locality violation.)

Another case where the separation of the CSC is relevant is the temporal sequence exception to the CSC (Ross 1967, Postal 1998, Bošković 2020b). As these authors observe, when there is a temporal sequence between the two conjuncts, extraction *out of* the second conjunct is possible, as shown in (83). However, extraction *of* the second conjunct itself is not allowed, even if the two conjuncts in (83b) are interpreted as a temporal sequence. This can also be taken as evidence against the unification of the CSC, since extraction *out of* a conjunct and extraction *of* a conjunct should be saved by the same operation under a uniform CSC.

(83)   a.   What$_i$ did you [[go to the store] and [buy t$_i$]]?

        b. *What$_i$ did you buy [the whisky and t$_i$]?

The current proposal may also enable us to capture the contrast between the CSC II and the CSC I with respect to the (un)availability of the temporal sequence exception. Since the semantic interpretation is relevant, there is room for the CSC II, as an interface condition related to interpretation, to be affected by an interpretational difference, whereas there is none for the CSC I, which is a pure syntactic condition (although I leave developing the technical details of the CSC II exception in this context for future research). Thus, the current proposal can be extended to explain the (un)availability of the two exceptional cases noted above.

2.5   Conclusion

In this chapter, I have discussed the traditional CSC from a cross-linguistic perspective and showed that the CSC I can be violated in a number of languages. In particular, I have established the generalization that the CSC I can only be violated in languages without definite articles and languages with affixal definite articles. I have also argued based on the cross-linguistic data that the traditional CSC has to be separated into two conditions: the CSC I, which bans extraction *of* a conjunct, and the CSC II, which bans extraction *out of* a conjunct. This has led me to conclude that the traditional ConjP and each conjunct are islands independently of each other. The islandhood effect associated with extraction out of them was deduced from different mechanisms. The CSC I is essentially a syntactic constraint that derives from the interaction of the PIC and the anti-locality constraint, whereas the CSC II is essentially an interface condition that derives from the Coordination-of-likes and the interaction of Spell-Out and labeling, which is required for interpretation at the C-I interface according to Chomsky (2013). This conclusion is partially compatible with a widely discussed view in the literature that the traditional CSC is an LF condition, as mentioned in section 1 (see, e.g., Kato 2006). However, it should be noted here that this view in the literature has mainly focused on the CSC II in the current terms, not on the CSC I, and hence did not realize the possibility that different mechanisms can be responsible for the traditional CSC effects. The present chapter has made it possible to investigate the traditional CSC in a more fine-grained manner from

a cross-linguistic perspective.

As noted above, the present chapter has also established the generalization regarding violations of the CSC I. In particular, it can be violated in languages without definite articles and languages with affixal definite articles. The generalization makes the same language cut as the possibility of adverb LBE out of predicative TAPs, which is also allowed only in languages without definite articles and languages with affixal definite articles. These two types of languages constitute a natural class under Talić's (2015, 2017) Structural Parallelism, according to which a bare structure in one lexical domain is possible if there is a bare structure in another domain. Based on this, a fine-grained structure of the traditional ConjP has been proposed, and the difference between article-less and affixal article languages on the one hand and non-affixal article languages on the other hand has been explained by the absence/presence of a functional projection above ConjP. The discussion of the CSC I in this chapter thus provides additional evidence for Talić's claim that we need (at least) three-way distinction of NP/DP languages.

Chapter 3

Typology of Indefinite Pronouns and Syntax of Wh-questions

3.1 Introduction

In this chapter, I discuss typology of indefinite pronouns, which was originally investigated in detail in the literature on non-generative typology. In particular, I discuss it from the perspective of the NP/DP languages distinction, by paying a close attention to the morphological compositions of indefinite pronouns.

In his seminal work, Haspelmath (1997) conducts a large-scale cross-linguistic survey of indefinite pronouns, and shows that languages can be classified into two major groups with respect to the morphological make-up of indefinite pronouns: in one, indefinite pronouns are derived from a generic noun, such as *something* in English, which is composed of the quantificational element *some* and the generic noun *thing*. In the other group, indefinite pronouns are derived from interrogative pronouns (the term *derive* should be taken "non-technically" here), as shown by Mandarin Chinese (1), where the interrogative pronoun *shenme* is used to express the interpretation of 'something'.

(1) Ta yiwei wo xihuan **shenme**.

     he think I like what

     'He thinks I like something.' (Li 1992:125)

Haspelmath calls the former *generic-noun-based* indefinite pronouns, and the latter *interrogative-based* indefinite pronouns. Haspelmath also raises the question whether there is a typological correlation between the type of indefinite pronouns and other properties of relevant languages, but he leaves it open.[1]

In this chapter, I show that the availability of a certain type of indefinite pronouns actually correlates with Talić's (2015, 2017) three-way distinction of NP/DP-languages, similarly to the one regarding the Coordinate Structure Constraint discussed in chapter 2. As a point of departure, I argue that the term "interrogative-based" pronoun actually does not reflect the nature of the relevant indefinite pronouns correctly. The term is misleading in that it presupposes that the interrogative form is the primitive, i.e., basic form, of the pronouns in question. However, it will be shown that the interrogative form is actually not the primitive, and the pronouns in question should be considered as *indeterminate pronouns* in Kuroda's (1965) sense. In addition, based on their morphological composition, I propose that indeterminate pronouns should be further classified into two types: compositional indeterminate pronouns, which require a quantificational affix/particle for indefinite use, and bare indeterminate pronouns, which can be used as indefinite pronouns with no such particle/affix. I then establish a novel cross-linguistic generalization regarding compositional indeterminate pronouns in the spirit of Talić's (2015, 2017) three-way distinction of NP/DP-languages; namely, this type of indeterminate pronouns are allowed only in languages that have affixal definite articles or lack definite articles. This generalization is also deduced from (a revised version of) Saito's (2017) analysis of indeterminate pronouns in Japanese and Chinese and

---

1. Haspelmath attempts to associate the above division of indefinite pronouns with some other properties of the languages in the two groups (Haspelmath 1997:ch.9). He first hypothesizes that the word order in VP (or head-directionality more generally) could be relevant; in a small scale analysis of languages from Europe, he finds weak correlations between the head-initial order and generic-noun-based indefinite pronouns and between the head-final order and interrogative-based indefinite pronouns. However, after examining more languages, he concludes that there is no correlation between head-directionality and the type of indefinite pronouns.

the presence/absence of DP a lá Talić (2015, 2017).

I also discuss the typology of wh-questions, which has been one of the most widely discussed topics in the generative linguistics. Cheng (1991) proposed the Clausal Typing Hypothesis, in which the presence/absence of a Q-particle correlates with the absence/presence of wh-fronting. However, Bruening (2007) shows, based on a broader typological survey, that the Clausal Typing Hypothesis is problematic. It has, then, remained an open issue what properties correlate with the syntax of wh-questions. In this chapter, it is shown that the deduction of the new generalization regarding indeterminate pronouns offered in this chapter also sheds new light on this issue. Specifically, it allows us to treat multiple-wh-fronting in languages like Slavic and wh-in-situ of the Japanese type in a uniform manner. It is also argued that the proposed analysis of indeterminate pronouns can account for various types of wh-in-situ observed in the literature by accommodating previous works. Thus, the proposed account of indeterminate pronouns enables us to take a fresh perspective on the investigation of the typology of the syntax of wh-questions.

The chapter is organized as follows. In section 3.2, I argue that the term "interrogative-based" indefinite pronoun should be replaced with Kuroda's (1965) indeterminate pronoun and that we need to classify indeterminate pronouns into two types. In section 3.2.1, I propose to separate the pronouns in question into two types, defining compositional indeterminate pronouns and bare indeterminate pronouns based on their morphological make-up. In sections 3.2.2 and 3.2.3, I show that these two types of indefinite pronouns also exhibit semantic and syntactic differences, which further supports the proposed classification. In section 3.2.4, it is argued that the semantic difference can be captured by examining the role of a morpheme used in one of the two types of indefinite pronouns in question, and that the syntactic differences can be best explained by Saito's (2017) Agree-based analysis of the syntax of wh-questions in Japanese and Chinese. In section 3.3.1, I establish a novel cross-linguistic generalization regarding compositional indefinite pronouns. In section 3.3.2, I offer a deduction of the generalization based on a modification of Saito's analysis discussed in section 3.2 and Talić's treatment of the NP/DP-language distinction. In section 3.4, I discuss consequences and extensions of the proposed analysis to certain domains

of the syntax of wh-questions. After reviewing the literature on the typology of wh-questions in section 3.4.1, it is shown in section 3.4.2 that under the proposed analysis multiple wh-fronting languages and wh-in-situ languages of the Japanese type can receive a unified treatment despite the difference with respect to the surface placement of "interrogative pronouns", which distinguishes them from single wh-fronting languages such as English. In section 3.4.3, I also argue that cross-linguistic variation in the exact syntax of wh-in-situ can be captured by the proposed analysis of indeterminate pronouns, by integrating previous accounts of wh-in-situ in the literature. Section 3.5 concludes the chapter.

3.2 Bare vs. compositional indeterminate pronouns

In this section, I argue that "interrogative-based indefinite pronouns" in Haspelmath's (1997) terminology need to be reconsidered from a perspective of morphology, syntax, and semantics. Specifically, I claim, building on Kuroda (1965), that the term "interrogative-based" should be dispensed with and the pronouns in question should rather be considered true *indeterminate pronouns*. In addition, I show that indeterminate pronouns should further be separated into two sub-classes from the viewpoint of morphology, and then demonstrate that these two types also exhibit semantic and syntactic differences. I then offer a formal account of the distinction in question based on Saito's (2017) Agree-based analysis of Japanese indeterminate pronouns.

3.2.1 Morphological difference and new terminology

Haspelmath (1997) observes that there are actually two ways to derive "interrogative-based" indefinite pronouns. In one, "interrogative pronouns" and "indefinite pronouns" are morphologically identical; e.g., Chinese *shenme*, which can mean 'what' or 'something', depending on the syntactic context. When it occurs in a non-interrogative environment, *shenme* means 'something', while it means 'what' when it occurs in an interrogative environment, as shown in (2).

(2) a. Ta yiwei wo xihuan **shenme**.

       he think I   like   what

       'He thinks I like something.'

b. Ta yiwei wo xihuan **shenme**?

       he think I   like   what

       'What does he think I like?'   (Li 1992:125)

In the other way of deriving relevant pronouns, "interrogative pronouns" require a quantificational particle/affix to compose into indefinite pronouns. A representative is Japanese *nani*, which requires the particle *ka* to be attached to it in order to have the interpretation of 'something', as illustrated in (3a). In an interrogative environment where *nani* is supposed to be interpreted as 'what', no particle is added to it, as seen in (3b).

(3) a. Kare-wa watashi-ga **nani-*(ka)**-ga sukida to omotteiru.

       he-TOP I-NOM   what-KA-NOM like   C think

       'He thinks I like something.'

b. Kare-wa watashi-ga **nani-(*ka)**-ga sukida to omotteiru no?

       he-TOP I-NOM   what-KA-NOM like   C think   Q

       'What does he think I like?'

From this perspective, the term "*interrogative-based* indefinite pronoun" is quite misleading. As noted above, Chinese *shenme* is interpreted as an interrogative pronoun (meaning 'what') in an interrogative force context, but as an existential indefinite pronoun (meaning 'something') in a non-interrogative force context (and with no dedicated particle for this usage). Likewise, Japanese *nani* is interpreted as an interrogative pronoun (meaning 'what') with interrogative force, but as an existential indefinite pronoun with the particle *ka*. Thus, the interpretation of the relevant pronouns depends on the morpho-syntactic environment in which they occur. This means that those pronouns do not have an inherent quantificational force, and hence the "interrogative form" is not a primitive

form of the pronouns under discussion. If they were inherently interrogative, the interrogative interpretation would need to be "canceled" somehow in the indefinite usages, and it is unclear how this could be technically implemented.

Actually, this point was already noticed and discussed in the generative literature, as early as Kuroda (1965). Kuroda (1965) calls the relevant pronouns in Japanese *indeterminate pronouns*.[2] In particular, Kuroda (1965:101) states that "[i]t can be said that the role of the indeterminate pronouns [is] very much like that of yet unbounded logical variables", expressing the intuition that the interpretation of the pronouns in question is dependent on the morpho-syntactic context in which they occur. The idea that those "interrogative pronouns" in languages like Chinese and Japanese are not inherently interrogative but are actually indeterminate pronouns in Kuroda's sense has been discussed a great deal and elaborated on in the formal linguistic literature (e.g., Huang 1982, Nishigauchi 1990, Cheng 1991, Shimoyama 2006, among many others). Thus, I conclude that we should dispense with the term "interrogative-based" indefinite pronoun, and redefine the relevant pronouns in such a way that the above observations are correctly reflected.

It should be immediately added here that Kuroda's indeterminate pronouns are not sufficient to define the pronouns in question, either. Recall that Chinese indeterminate pronouns do not require any quantificational particle/affix for the indefinite pronominal usage, unlike those in Japanese, which require a quantificational particle/affix. I takes this as indicating that indeterminate pronouns should further be separated into two types. Specifically, I define the Chinese-type indeterminate pronouns as in (4) and the Japanese-type indeterminate pronouns as in (5).

(4)   DEFINITION 1: A *bare indeterminate pronoun* is a pronoun whose interrogative and indefinite usages have the same form.

(5)   DEFINITION 2: A *compositional indeterminate pronoun* is a pronoun which functions as an interrogative pronoun in the context with interrogative force or as an indefinite pronoun when a specific quantificational particle/affix is added to it.

---

2. Kuroda in fact takes insights on this from the traditional Japanese grammar, where the pronouns in question are called *futeigo* 'indeterminate words'.

Hereafter the terminology such as "interrogative pronoun" and "wh-phrase" that is used in the relevant literature will be adapted to the above terms, and in example sentences I gloss bare and compositional indeterminate pronouns with English interrogative pronouns such as 'who', 'what', etc., only for presentational purposes. It should be kept in mind that neither type of indeterminate pronouns inherently has the interrogative interpretation by themselves.

Below I argue that the distinction between the two types of indeterminate pronouns is also required from semantic and syntactic perspectives. I then offer a formal analysis of the distinction in question.

3.2.2 Semantic difference

Haspelmath (1997) and Bruening (2007) observe a semantic difference between bare indeterminate pronouns and compositional indeterminate pronouns that holds cross-linguistically. Specifically, Haspelmath notes that bare indeterminate pronouns can only be interpreted as non-specific and cannot occur in environments where the specific reading is forced. In Bruening's terms, when used as existential indefinites, bare indeterminate pronouns cannot take wide scope and are typically interpreted as non-specific, whereas compositional indeterminate pronouns can take wide or narrow scope.[3] For instance, Li (1992:127) shows that *shenme* in Chinese cannot take scope over negation, as seen in (6). Bruening points out that the same holds in Passamaquoddy, as seen in (7).

(6) Ta **bu**   xihuan **shenme**.
     he NEG like   what
     'He doesn't like anything.' ($^{ok}\neg > \exists$ / $* \exists > \neg$)

---

3. To be more precise, Bruening states that bare indeterminate pronouns cannot take the widest scope. Thus, when there are three scope possibilities, only the intermediate and the narrowest are possible. I put aside this detail here, because what is important for our current purpose is that bare indeterminate pronouns and compositional indeterminate pronouns behave differently in the same environment.

(7) **Ma=te**   **wen** '-kisi-tomh-a-wiy-il   Piyel-ol.

NEG=EMPH who 3-PERF-beat-DIR-NEG-OBV Piyel-OBV

'No one beat Piyel.' (*'There is someone who didn't beat Piyel.')   (Bruening 2007:161)

In contrast, Nishigauchi (1990:121) observes that the compositional indeterminate pronoun *dare* in Japanese can have a specific or non-specific interpretation in (8).

(8) **Dare-ka**-kara henna   tegami-ga   todoita.

who-KA-from strange letter-NOM arrived

'A strange letter arrived from somebody.'

In addition, when a compositional indeterminate pronoun is used with sentential negation, just as in Chinese (6) and Passamaquoddy (7), it is strongly preferred that it takes scope over negation, as illustrated in Japanese (9).

(9) John-wa **dare-ka**-o   sasowanakatta.

John-TOP who-KA-ACC didn't.invite

'There is somebody who John didn't invite.'

Thus, bare indeterminate pronouns and compositional indeterminate pronouns behave differently with respect to specificity/scope.

I suggest that this difference can be attributed to the morphological make-up of these two types of indeterminate pronouns. Reinhart (1997), Winter (1997), Chung and Ladusaw (2004), López (2012), among others, propose that the specific/wide-scope reading of indefinites obtains by means of choice functions. Interestingly, Hagstrom (1998), Cable (2007, 2010), Yatsushiro (2009), among others, propose that *ka* in Japanese, which is combined with a compositional indeterminate pronoun to compose an existential indefinite pronoun as in (9), is a variable over choice functions (see Cable 2007, 2010 for application of this analysis to Tlingit and Sinhala). Extending this to compositional indeterminate pronouns in general, I suggest that compositional indeterminate pronouns can have the specific reading, or take wide scope, for the existential indefinite usage because

of the presence of the quantificational particle/affix that serves as a choice function variable.[4] On the other hand, bare indeterminate pronouns cannot do so due to the lack of a quantificational particle/affix for the existential indefinite usage. Thus, the difference in semantics between the two types of indeterminate pronouns follows from the morphological difference and the function of the relevant quantificational particles/affixes.

3.2.3 Syntactic difference

Another important difference between bare and compositional indeterminate pronouns concerns the syntax of interrogatives, in particular sensitivity to wh-islands. It has been observed in the literature on Japanese that compositional indeterminate pronouns are "unselectively bound" by a licensor such as interrogative C and a quantificational particle, in the sense that multiple indeterminate pronouns can be associated with one licensor. In (10a), the two compositional indeterminate pronouns, *dare* and *doko*, are interpreted as question words under the scope of the question particle *ka*. The same holds in the case of universal quantification; in (10b), the same compositional indeterminate pronouns are associated with the additive particle *mo* and receive the universal quantificational interpretation.

(10) a. Watashi-wa [dare-ga   doko-ni itta   **ka**] shiranai.

              I-TOP   who-NOM where-to went Q   not.know

              'I don't know who went where.'

    b. Watashi-wa [dare-ga   doko-ni itte   **mo**] kamawanai.

              I-TOP   who-NOM where-to go.INF also not.care

              'I don't care no matter who goes where.'

Nishigauchi (1990) argues, building on Kuroda (1965), that compositional indeterminate pronouns in Japanese are variables that lack quantificational force on their own; they are unselectively bound

---

4. For the non-specific interpretation (or narrow scope) of compositional indeterminate pronouns, I follow Reinhart (1997), Winter (1997), Chung and Ladusaw (2004) in assuming that existential closure for choice functions can apply at any point in the semantic composition, so that the existential quantifier can take scope under another scope bearer. I refer the reader to the above references for technical details.

by their licensors. At the same time, Nishigauchi (1990) notes that the binding relation between compositional indeterminate pronouns and quantificational particles in Japanese is not truly unselective, in that there is a locality restriction on the relation in question. Specifically, when there is more than one quantificational particle that can potentially bind compositional indeterminate pronouns, only the closest one to compositional indeterminate pronouns can be the binder of them. This is illustrated in (11), where there are two potential binders, the additive particle *mo* and the interrogative particle *ka*. The interrogative particle *ka* can be used in wh-questions or yes-no questions; if it does not bind an indeterminate pronoun, it is interpreted as marking a yes-no question. In (11a), only *mo* can bind the compositional indeterminate pronoun *dare*, since it is closer to *dare* than *ka*. As a result, *dare* receives the universal quantificational interpretation, and *ka* is used as a yes-no question marker. Likewise, in (11b), *ka* is closer to *dare* than *mo*, so *dare* can only be bound by *ka*. Thus, *ka* functions as a marker of a wh-question, and the compositional indeterminate pronoun is interpreted as a question word meaning 'who'.

(11) a. Kimi-wa [[dare-ga   kite   **mo**] ikanai **no**?

          you-TOP   who-NOM come also not.go Q

          'Are you not going, whoever may come?'

          NOT 'For which *x*, *x* a person, are you not going even if *x* is coming.'

      b. John-wa   [[dare-ga   kuru **ka**] shitteite **mo**] ikanai.

          John-TOP   who-NOM come Q   know   also not.go

          'John will not go even if he knows who will come.'

          NOT 'For all *x*, *x* a person, John will not go even if he knows whether *x* is coming.'

                                                                (Nishigauchi 1990:148)

The same effect is observed when potential binders are the same element. In (12), where the question particle *ka* occurs in the matrix and the embedded clause, neither of the two compositional indeterminate pronouns can take scope in the matrix clause, because the closest *ka* must bind them. As a result, the sentence is interpreted as a multiple embedded wh-question. Nishigauchi (1990) points out that this is essentially an instance of the wh-island effect.

(12) Tanaka-kun-wa   [<u>dare</u>-ga   <u>nani</u>-o   tabeta **ka**] oboeteimasu **ka**?

Tanaka-DIM-TOP   who-NOM   what-ACC   ate   Q   remember   Q

'Does Tanaka remember who are what?'

NOT 'For which *x*, *x* a person, does Tanaka remember what *x* ate?'

NOT 'For which *y*, *y* a thing, does Tanaka remember who ate *y*?'

Crucially, this "local" unselective binding property of Japanese compositional indeterminate pronouns is contrasted with the "genuine" unselectivity of Chinese bare indeterminate pronouns. As is well-known, Chinese bare indeterminate pronouns do not show wh-island effects, as shown by Huang's (1982) classic example in (13).

(13) Ni   xiang-zhidao [shei   mai-le   shenme]?

you   wonder   who   buy-ASP   what

'What is the thing x such that you wonder who bought x' or

'Who is the person x such that you wonder what x bought'

This can be considered as a "genuine" unselective binding case, in the sense that there is no locality restriction on the binding relation between the bare indeterminate pronouns and the interrogative C, unlike Japanese (12), where the compositional indeterminate pronouns can only take scope in the embedded clause. This indicates that bare indeterminate pronouns and compositional indeterminate pronouns are licensed by different mechanisms.

In the following subsection, I introduce Saito's (2017) covert movement analysis, which I show can capture the syntactic difference between Chinese bare indeterminate pronouns and Japanese compositional indeterminate pronouns discussed in this subsection. I then propose a modification of his analysis, based on a semantic consideration.

3.2.4   Analysis of the syntactic difference

As mentioned above, Nishigauchi (1990) acknowledges the difference between Chinese and Japanese with respect to sensitivity of indeterminate pronouns to wh-islands, and he suggests that the com-

positional indeterminate pronoun in Japanese undergoes covert movement to Spec,CP, which he argues is responsible for the wh-island effect.[5] Building on this, Saito (2017) proposes an Agree-based account of Japanese indeterminate pronouns. He adopts Bošković's (2007b) Agree theory, in which an element that has an unvalued feature undergoes movement to a position from which it can probe down a goal in its c-commanding domain. (14) is an illustration of movement of an external argument DP to Spec,TP under Bošković's theory. (14a) is a configuration where a DP that has an unvalued Case feature is base-generated in Spec,vP. In (14b), T probes down the DP and have its ϕ-features valued. Crucially, the Case feature of the DP is still unvalued at this point. In (14c), then, the DP moves to Spec,TP. It can then probe down T since T is now in the c-command domain of the DP. The Case feature of the DP is then valued as [NOM].[6]

(14) a.
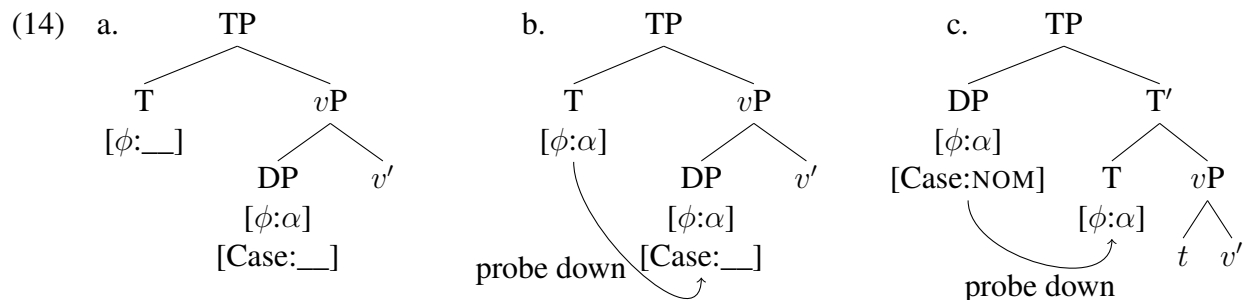
Extending this to Japanese indeterminate pronouns, Saito (2017) proposes that compositional indeterminate pronouns in Japanese have an unvalued operator feature, which he assumes is responsible for the interpretation of wh-items, and that they undergo covert movement to a position where they can probe down (i.e., c-command) a quantificational particle (e.g., Spec,CP in the case of the question particle *ka*).[7] This is illustrated in (15).

---

5. Nishigauchi argues that this movement is driven by a requirement on government of compositional indeterminate pronouns under the GB theory.

6. Bošković (2007b) eliminates the Activation Condition (Chomsky 2000), which requires a goal to bear an unvalued feature to be targeted by Agree. All that is needed is that a probe (i.e., an unvalued feature) always probes down to Agree with its goal in its c-commanding domain. For empirical evidence for this approach, see Villa-García (2015), who discusses multiple complementizer sentences in Spanish, where a DP is base-generated higher than its Case-licensor.

7. This covert movement is implemented as deletion of the higher copies under the copy theory of movement (Chomsky 1995b); see e.g., Bobaljik (1995, 2002), Brody (1995), Groat and O'Neil (1996).

(15) a. Taro-wa [[Hanako-ga nani-o tabeta] ka] sitteiru.

Taro-TOP Hanako-NOM what-ACC ate Q know

'Taro knows what Hanako ate.'

b.

```
        CP                →                    CP
      /    \                                 /    \
    TP      ka                            nani     C′
    /\    [Q(uestion)]              [Op:Q(uestion)]  /  \
  …nani…                                           TP    ka
  [Op:__]                                          /\  [Q(uestion)]
                                                 …__…
```

(Saito 2017:19)
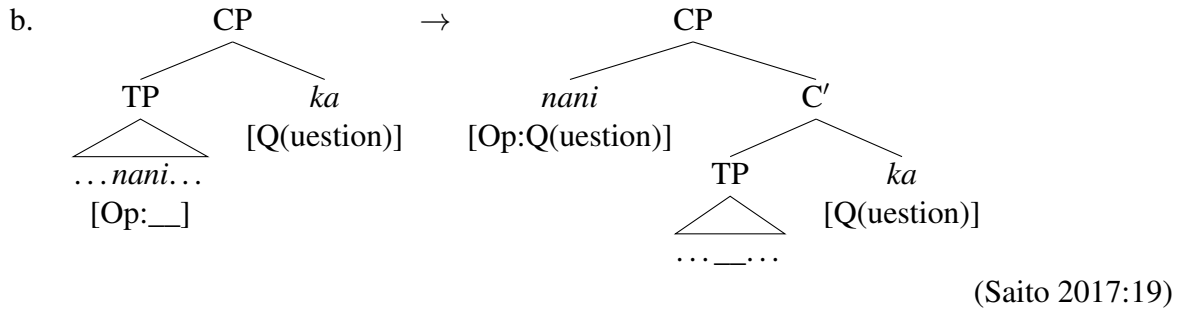
In the case of the additive particle *mo*, the operator feature of the compositional indeterminate pronoun is valued as [Conjunctive], which Saito takes to be universal quantificational by means of conjunction of all individuals in the domain of discourse.

(16) a. [[Dare-ga kaita hon] mo] toshokan-ni aru.

who-NOM wrote book also library-in is

'For every x, x a person, (also) a book that x wrote is in the library.'

b.

```
        FP                →                    FP
      /    \                                 /    \
    NP      mo                            dare     F′
    /\   [Conjunctive]             [Op:Conjunctive]  /  \
  …dare…                                           NP    mo
  [Op:__]                                          /\  [Conjunctive]
                                                 …__…
```

(Saito 2017:19)
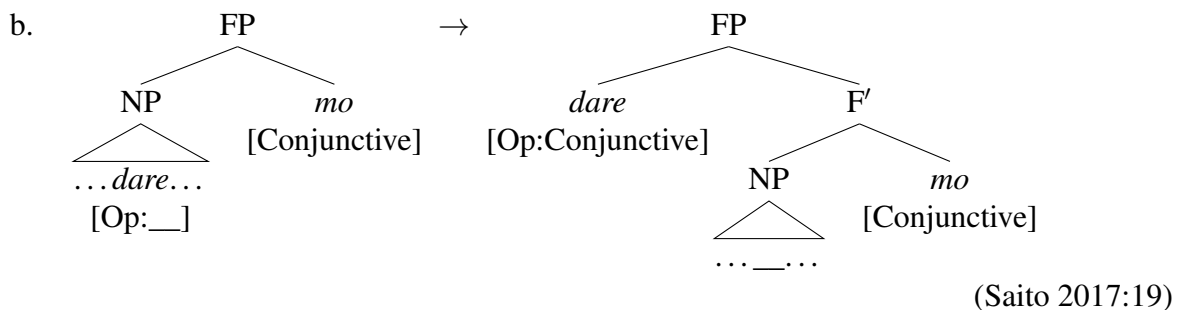
Saito shows that the wh-island effect (and the locality of quantificational particles in general) reduces to Rizzi's (2010) criterial freezing in this system. Rizzi's criterial freezing essentially states that a syntactic object with an unvalued feature F must stay in the position where F is valued. In (17a), there are two potential valuers of the operator feature of the indeterminate pronoun *nani*: *ka* in the matrix CP (CP$_1$) and *ka* in the embedded CP (CP$_2$). The operator feature drives movement of *nani* to a higher position. When *nani* moves to Spec,CP$_2$, its operator feature is valued, so it

75

is frozen in Spec,CP$_2$ (i.e., there is no feature that drives further movement). Thus, *nani* must be interpreted in the embedded clause and the wh-island effect obtains. This is schematized in (17b).

(17) a. [$_{CP1}$ Taro-wa [$_{CP2}$ Hanako-ga   nani-o   tabeta ka] sitteimasu ka]?

          Taro-TOP   Hanako-NOM what-ACC ate   Q   know   Q

        'Does Taro know what Hanako ate?'

b.

```
         CP₁          →              CP₁
        /   \                       /    \
      TP     ka                   TP      ka
     /  \    [Q]                  / \     [Q]
   CP₂                          CP₂
   /  \                         /    \
  TP    ka                    nani    C₂′
 /  \   [Q]                  [Op:Q]   /  \
…nani…                               TP    ka
[Op:__]                             /  \   [Q]
                                   … __ …
```
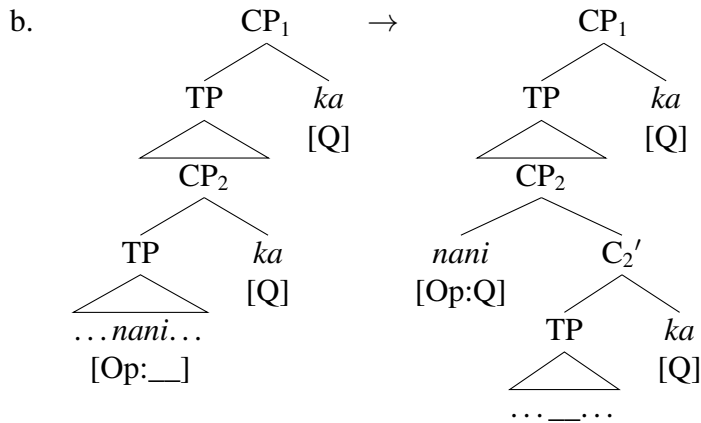
In contrast, Chinese bare indeterminate pronouns are assumed to lack the relevant operator feature. In other words, bare indeterminate pronouns have no syntactic dependency with any elements.[8] Thus, they do not undergo movement, unlike Japanese compositional indeterminate pronouns, and hence no wh-island effect is obtained, as seen in (13).

I argue that Saito's covert movement analysis of Japanese compositional indeterminate pronouns receives independent support from the behavior of local anaphors. Japanese has two types of anaphors: a long-distance subject-oriented anaphor *jibun* 'self', which can refer to a non-clause-mate subject as well as a clause-mate subject, and morphologically more complex anaphors composed of *jishin* 'self' and another pronominal element, such as *kare-jishin* 'himself' and *kanojo-jishin* 'herself', which can only refer to a clausemate nominal. For instance, (18a) shows that *kanojo-jishin* 'herself' in the embedded clause cannot refer to the matrix subject *Mary*. Interestingly, however, when a clause-mate anaphor is contained in an indeterminate phrase that takes scope in the embedded clause, it can refer to the matrix subject, as shown in (18b).[9] The same

---

8. Their interpretation is determined solely by semantic operators at LF.
9. To be more precise, *kanojo-jishin* can in principle have a logophoric interpretation, by which it can refer to the

contrast obtains with another matrix predicate that selects an embedded question, as seen in (19).

(18) a. *Mary$_i$-wa [John-ga   kanojo-jishin$_i$-no ronbun-o   yonda to] shitta.

          Mary-TOP John-NOM her-self-GEN   paper-ACC read   C   found.out

          Lit: 'Mary found out that John had read herself's paper.'

     b. Mary$_i$-wa [John-ga   **dono**-kanojo-jishin$_i$-no ronbun-o   yonda ka] shitteiru.

          Mary-TOP John-NOM which-her-self-GEN   paper-ACC read   Q   know

          Lit: 'Mary knows which herself's paper John had read.'

(19) a. *Mary$_i$-wa [John-ga   kanojo-jishin$_i$-no ronbun-o   yonda ka] tazuneta.

          Mary-TOP John-NOM her-self-GEN   paper-ACC read   Q   asked

          Lit: 'Mary asked that John had read herself's paper.'

     b. Mary$_i$-wa [John-ga   **dono**-kanojo-jishin$_i$-no ronbun-o   yonda ka] tazuneta.

          Mary-TOP John-NOM which-her-self-GEN   paper-ACC read   Q   asked

          Lit: 'Mary asked which herself's paper John had read.'

This is correctly predicted by the covert movement analysis; since the entire indeterminate pronominal phrase that contains the local anaphor covertly moves to the edge of the embedded clause in order to have its operator feature valued in (18b), the matrix subject becomes accessible to the anaphor.

It should be added here that this contrast cannot be captured by any of the previous works that do not assume movement of the entire indeterminate pronominal phrase. For instance, Watanabe (1992a) proposes that a null operator is base-generated with an indeterminate pronoun and only this operator moves to Spec,CP, leaving the indeterminate pronoun in-situ. Maki (1995), Hagstrom (1998), Takahashi (2002), and Cable (2007, 2010) propose that a quantificational particle, rather than an indeterminate pronoun, moves to the surface position (e.g., C). Shimoyama (2006) proposes a semantic account of indeterminate pronouns in which neither an indeterminate pronominal

matrix subject. What is important here is, though, that there is a clear contrast between (18a) and (18b).

phrase nor a quantificational particle moves and the local unselective binding behavior is captured by a Hamblin alternative semantics. Crucially, none of these approaches to Japanese compositional indeterminate pronouns can capture the contrast in (18), since *kanojo-jishin* 'herself' contained in the indeterminate pronominal phrase would not move to the edge of the embedded clause. Thus, Saito's (2017) analysis, which integrates the insight of Nishigauchi's (1990) work, has a better empirical motivation compared with the previous approaches.

Moreover, under Saito's analysis of Chinese bare indeterminate pronouns, it is predicted that a local anaphor in an embedded clause in Chinese should not be able to refer to the matrix subject even when it is contained in an indeterminate phrase, since bare indeterminate pronouns do not have an operator feature that would drive movement, unlike compositional indeterminate pronouns in Japanese. This is indeed borne out. (20a) shows that the local anaphor *ta-ziji* 'herself/himself' in the embedded clause cannot refer to the matrix subject. Crucially, as seen in (20b), when it is contained in the bare indeterminate phrase, it cannot refer to the matrix subject either, unlike the Japanese local anaphor *kanojo-jishin* 'herself' in (18b).

(20) a. Mary$_i$ faxian [Zhangsan$_j$ du-le ta-ziji$_{*i/j}$-de lunwen]
          Mary find.out Zhangsan read-ASP her/himself-GEN paper
          Lit: 'Mary found out that Zhangsan read himself's paper.'

     b. Mary$_i$ zhidao [Zhangsan$_j$ du-le ta-ziji$_{*i/j}$-de **na**-pian lunwen].
          Mary know Zhangsan read-ASP her/himself-GEN which-CL paper
          Lit: 'Mary knows which himself's paper Zhangsan read.'

The contrast between Japanese (18) and Chinese (20) thus supports Saito (2017) parameterization of Japanese compositional indeterminate pronouns and Chinese bare indeterminate pronouns, in which the former have an unvalued operator feature whereas the latter do not.

However, I submit that Saito's analysis needs to be amended from a semantic perspective. As noted above, Japanese compositional indeterminate pronouns acquire their own quantificational force via Agree with a quantificational particle under his proposal. But this seems to give rise to

redundancy in quantification. In the recent semantic literature on "wh-items"/indeterminate pronouns (Kratzer and Shimoyama 2002, Shimoyama 2006, Cable 2007, 2010, among many others), it is standardly assumed that "wh-items"/indeterminate pronouns denote a set of individuals (e.g., people, things, etc.), which Saito also adopts. In this line of analyses, indeterminate pronouns themselves do not have a quantificational force, and how the set of individuals is quantified over is determined by a quantificational particle. For instance, Shimoyama (2006) proposes that *mo* contributes universal quantification over a set of individuals denoted by an indeterminate pronoun, as formulated in (21). An example is given in (22).

(21) For $\llbracket \alpha \rrbracket^g \subseteq D_e$,

$\llbracket \alpha \text{ mo } \rrbracket^g = \{\lambda P \forall x [x \in \llbracket \alpha \rrbracket^g \to P(x) = 1]\}$ (Shimoyama 2006:155)

(22) a. Dono-gakusei-mo   odotta.

which-student-also   danced

'Every student danced.'

b. $\llbracket (22a) \rrbracket = \{\forall x [x \in \{y: \text{student}(y)\} \to \text{dance}(x)]\}$ (Shimoyama 2006)

Thus, the additive particle *mo* has a substantial contribution to the semantic composition of the indeterminate pronoun. In fact, Saito (2017:22) also assumes that *mo* itself has the additive interpretation, and claims that "[t]he analysis allows the particles to be interpreted uniformly whether they appear in the presence or absence of [indeterminate pronouns]". There are indeed some works that attempt to unify the additive *mo* and the universal quantificational *mo* (e.g., Ohno 1989, Kobuchi-Philip 2007; see also Uegaki 2018 for unification of disjunctive *ka*, interrogative *ka*, and existential *ka*). This means that both the indeterminate pronoun and the additive particle would contribute the universal quantification under Saito's proposal. Now, there arises a problem of semantic composition of those two elements. Since the indeterminate pronoun has its own quantificational force, the additivity or universal quantification contributed by the additive particle would be redundant. Actually, the issue is more serious than just redundancy; if both the indeterminate pronoun and the additive particle have universal quantificational force on their own, it would be unclear how

these two elements are semantically composed in the current semantic theory. Although it might be possible to resolve this redundancy, it would be more straightforward to simply assume that indeterminate pronouns do not have their own quantificational force, as has been proposed in the literature.

How can, then, this issue be resolved while maintaining the core insight of Saito's analysis? The reason why Saito considers indeterminate pronouns in Japanese to have quantificational force is that the interpretation of indeterminate pronouns is dependent on overt quantificational particles, in contrast with Chinese indeterminate pronouns, which can be used as indefinites without a quantificational particle/affix. Saito then attributes this dependency to agreement between compositional indeterminate pronouns and quantificational particles; in particular, compositional indeterminate pronouns have an unvalued *interpretable* operator feature, while bare indeterminate pronouns lack it. Note now that the agreement relation need not involve interpretable features. In the literature, it has been standardly assumed since Chomsky (2000) that uninterpretable features need to be valued in order to be deleted before the syntactic structure is sent to the interfaces.[10] Thus, I propose that the operator feature that is to be valued is an *uninterpretable* feature rather than an interpretable feature, *pace* Saito (2017). This modification still maintains the syntactic dependency between compositional indeterminate pronouns and quantificational particles in Japanese, with the semantic redundancy eliminated and the covert phrasal movement analysis intact.

To summarize this section, I have argued that Haspelmath's (1997) "interrogative-based" indefinite pronouns are better understood as indeterminate pronouns in Kuroda's (1965) sense. I have then proposed to define two types of indeterminate pronouns, namely, compositional and bare indeterminate pronouns. It has been shown that compositional indeterminate pronouns and bare indeterminate pronouns show different morphological, semantic, and syntactic behavior. In order to capture those differences between the two types of indeterminate pronouns, I have introduced Saito's (2017) analysis of indeterminate pronouns in Japanese, in which a compositional

---

10. This, of course, does not deny the possibility that interpretable features enter into a syntactic computation being unvalued and later get valued in the course of the derivation, as has actually been proposed by Pesetsky and Torrego (2007) and defended by Bošković (2011a). Important here is that the feature in question need not be an interpretable feature.

indeterminate pronoun (but not a bare indeterminate pronoun) has an unvalued operator feature and this feature drives covert movement of the indeterminate pronoun to a position from which it c-commands and probes down a quantificational particle (e.g., Spec,CP). I have then proposed a modification of the analysis, by which the operator feature is an uninterpretable feature rather than an interpretable feature as proposed by Saito.

In the next section, building on the new classification of indefinite pronouns, I address the question of what property correlates with the availability of certain types of indefinite pronouns, the issue Haspelmath (1997) left open, as mentioned in section 3.1. Specifically, I establish a novel cross-linguistic generalization regarding availability of compositional indeterminate pronouns, on which availability of those indefinite pronouns correlates with the NP/DP-language distinction similarly to the Coordinate Structure Constraint discussed in chapter 2. I then propose a deduction of the generalization based on a modified version of Saito's (2017) account discussed here.

3.3   Indeterminate pronouns and the NP/DP languages distinction

3.3.1   Establishing a novel generalization

To the best of my knowledge, the first (and only) generative work that addresses the issue of potential correlation between typology of indefinite pronouns and other linguistic properties is Watanabe (2004a). Watanabe first divides Haspelmath's (1997) "interrogative-based" indefinite pronouns into the Chinese-type and the Japanese type, which correspond in my terminology to bare indeterminate pronouns and compositional indeterminate pronouns, respectively.[11] Interestingly, Watanabe notes that the productivity of compositional indeterminate pronouns correlates with absence of definite articles. For instance, Japanese and Russian, which lack definite articles, have productive compositional indeterminate pronouns (for space reasons, I present partial

---

11. However, Watanabe does not provide clear definitions of these two types.

paradigms taken from Watanabe 2004a).

(23) **Japanese**

	indeterminate	existential	neg-concord	universal
person	dare	dare-ka	dare-mo	dare-mo
thing	nani	nani-ka	nani-mo	nani-mo-kamo
place	doko	doko-ka	doko-mo	doko-mo
time	itsu	itsu-ka	—	itsu-mo

(24) **Russian**

	indeterminate	existential	neg-concord
person	kto	kto-to	ni-kto
thing	cto	cto-to	ni-cto
place	gde	gde-to	ni-gde
time	kogda	kogda-to	ni-kogda

There is also a striking diachronic change that shows this correlation. Thus, Latin, which lacked definite articles, had productive compositional indeterminate pronouns, whereas most Modern Romance languages, which have acquired definite articles, do not have them.[12]

(25) **Latin**

	indeterminate	existential	polarity	free choice
person	quis	ali-quis	quis-quam	qui-vis
thing	quid	ali-quid	quid-quam	quid-vis
place	ubi	ali-cubi	usquam	ubi-vis
time	quando	ali-quando	umquam	—

---

12. Watanabe (2009) also observes that Old English, which lacked definite articles, had productive compositional indeterminate pronouns, in contrast to Present English, which has the definite article and lost productive indeterminate pronouns.

(26) **Italian**

	indeterminate	existential	neg-concord
person	chi	qualcuno	nessuno
thing	che	qualche cosa, qualcosa	niente, nulla
place	dove	in qualche luogo	in nessun luogo
time	quando	qualque volta	(mai)

While the correlation between articles and compositional indeterminate pronouns appears to be robust, Watanabe acknowledges that Bulgarian, Romanian, and Hungarian have indeterminate pronouns although they have definite articles. Watanabe in fact does not provide a clear descriptive generalization regarding indeterminate pronouns that accommodate these languages. He attempts to offer an analysis in which indeterminate pronouns undergo agreement with quantificational affixes/particles, but his analysis is not fully empirically motivated due to the lack of a descriptive generalization regarding indeterminate pronouns. Very often, establishing a descriptive generalization is a prerequisite for the analysis, and hence it is an issue for Watanabe's work that he fails to establish one.[13]

This being said, there is a possibility that arises from insights of previous works. Notice that Bulgarian and Romanian, two exceptional languages which have compositional indeterminate pronouns but also have definite articles, are languages with affixal definite articles (see also the discussion of Hungarian below). We have seen in the previous chapters that languages with affixal definite articles pattern with languages without definite articles in a number of respects. It may then be that affixal article languages pattern with article-less languages in the domain of indeter-

---

13. Watanabe's analysis faces another problem. Watanabe classifies languages into three in terms of their D system: non-agreeing D languages such as English and Modern Romance (except for Romanian), agreeing D languages such as Japanese, Slavic, Latin, and Romanian, and no D language such as Chinese. For him, compositional indeterminate pronouns are possible only in agreeing D languages, where quantificational particles/affixes as D undergo agreement with compositional indeterminate pronouns. Crucially, Modern Romance languages such as Italian and Spanish are assumed to have non-agreeing D and hence lack productive compositional indeterminate pronouns under his analysis, unlike e.g., Romanian and Bulgarian, which are assumed to have agreeing D and have productive compositional indeterminate pronouns. However, those Romance languages do show agreement between a head noun and a definite article, the latter of which has been standardly analyzed as D, just like Romanian and Bulgarian. It is then not clear how agreement between a noun and a definite article as D would be distinguished from agreement between a compositional indeterminate pronoun and a quantificational particle/affix as D in a principled way.

minate pronouns, too. In order to confirm if this is indeed the case, I have conducted a large-scale cross-linguistic survey of indefinite pronouns, in which I have checked 138 languages that have indeterminate pronouns. Here I focus on compositional indeterminate pronouns.[14] Many of the languages come from Haspelmath (1997) and *The World Atlas of Language Structures (WALS)*, but I have put aside some languages from these two works, due to the lack of accessible sources that are required to confirm the issues under investigation here. Regarding WALS, it should also be noted that one can combine the features "indefinite pronouns" and "definite articles" in the search on WALS online, but the classifications (especially regarding the latter) are occasionally wrong, and the database does not distinguish compositional indeterminate pronouns and bare indeterminate pronouns, so I have checked each language with independent sources.

The result is as follows. Among the 138 languages that have indeterminate pronouns, 80 languages are identified as having productive compositional indeterminate pronouns (the remaining 58 use bare indeterminate pronouns). Among those 80, 66 languages do not have definite articles. The 66 languages are: Ainu, Awa Pit, Badimaya, Bawm, Bengali, Buriat, Cahuilla, Chantyal, Djingili, Old English, Estonian, Evenki, Garo, Georgian, Gitksan, (West) Greenlandic, Hayu, Hunzib, Hupa, Jakaltek, Old Japanese, Present Japanese, Kannada, Ket, Kham, Kodava, Korean, Korku, Latin, Latvian, Lezgian, Limilngan, Lithuanian, Maithili, Malayalam, Manipuri, Meithei, Micmac, Mundai, Muruwari, Nanai, Navajo, Newar, Nez Perce, Ngankikurungkurr, Ngiyambaa, Okinawan, Iron Ossetic, Polish, Huallaga Quechua, Imbabura Quechua, Russian, Serbo-Croatian, Shipibo-Konibo, Shoshone, Sinhala, Takelma, Tamil, Telugu, Tiwi, Udihe, Ukrainian, Warndarang, Yakut, Yup'ik, and Yuwaalaraay. Among the remaining 14 languages, 11 have affixal articles: Assamese, Basque, Bulgarian, Itzaj, Karok, Lillooet, Macedonian, Digor Ossetic, Romanian, Tonkawa, and Wichita. The remaining three languages, which appear to have non-affixal articles, are Hungarian, Yiddish, and Sorbian.

A word of caution is needed here regarding the languages that appear to have non-affixal definite articles. Hungarian definite articles are often considered to be non-affixal, but MacWhinney

---

14. Bare indeterminate pronouns are found in languages like Chinese, which lack definite articles, but also in German, Lakhota, etc., which have definite articles. See footnote 17 for discussion.

(1976:398) notes the following: "[t]he Hungarian definite article is *az* before vowels and *a* before consonants. This is the only morpheme in Hungarian which undergoes a morphophonemic alternation that is dependent upon the shape of the beginning of a root. For this reason, Hungarian linguists have often treated the definite article as a prefix." Given this, Hungarian can be classified as an affixal article language (see also chapter 5 for discussion). For Yiddish, the definite articles do not have a form distinct from demonstratives, the two being differentiated only by stress (Margolis 2011:122). Given Bošković's (2016b) definition of definite articles I adopt here, under which definite articles obligatory occur in a nominal phrase with a definite interpretation and have a distinct form from demonstratives, Yiddish articles may actually not be articles. For Sorbian, Schaarschmidt (1984) reports that the younger generation of speakers, who only use Sorbian in schools, use definite articles considerably less frequently than the older generation of speakers, who learned Sorbian through German. Jentsch (1980) and Lötzsch (1968) also note that definite articles in Sorbian are not obligatory in the context of definite interpretation and that they are not used in some cases where definite articles would be expected in German. These points indicate that Sorbian articles may actually not be (fully grammaticalized) articles. Given this, I propose the following generalization:

(27) *Generalization of compositional indeterminate pronouns*

Languages that have productive compositional indeterminate pronouns either have affixal definite articles or lack definite articles.

Note that this generalization is quite similar to the one regarding adverb LBE established by Talić (2015, 2017) and the one regarding the CSC in the previous chapter, in that affixal article languages and article-less languages pattern together. It should also be emphasized that (27) is a one-way correlation. Thus, there can be affixal article languages and article-less languages that do not have productive compositional indeterminate pronouns. What is important here is that there are no non-affixal article languages that have productive compositional indeterminate pronouns. In the next subsection, I provide a deduction of (27), the gist of which is that DP projects above (compositional) indeterminate pronouns in non-affixal article languages, which prevents quantificational

affix/particles from being attached to indeterminate pronouns, while in affixal article languages and article-less languages, the projection in question may be absent and hence quantificational affixes/particles can be attached to them.

3.3.2 Deduction of the generalization

Let us start from the structure of indeterminate pronouns. Kuroda (1965) proposes that Japanese compositional indeterminate pronouns consist of PRO(noun) and IND(terminate); essentially, PRO specifies the domain of quantification (e.g., person, thing), and IND marks the entire phrase as a compositional indeterminate pronoun. Regarding the categorial status of indeterminate pronouns, Huang (1982) proposes that they are generally NPs (except for 'how' and 'why', see below for discussion). Building on these two works, I propose that indeterminate pronouns in general are NPs which consist of Root that specifies the domain (e.g., person, thing), and N (or n; I use the label N hereafter only for presentational purposes). In addition, I suggest that this N is the locus of the parametric variation in the presence/absence of an unvalued uninterpretable operator feature discussed in section 3.2.4.[15] If this N bears an unvalued uninterpretable operator feature, the entire NP is a compositional indeterminate pronoun of the Japanese type, as schematized in (28). This operator feature is valued as [∀], [∃], etc. by a quantificational particle/affix for indefinite (and similar) usages, or as [Q(uestion)] in interrogatives after the entire NP moves to a position from which it c-commands a goal (i.e., interrogative C). On the other hand, if this feature is absent on N, we obtain a bare indeterminate pronoun of the Chinese type, as schematized in (28). Since there is no operator feature that requires valuation, the entire NP does not require a quantificational particle/affix and does not undergo movement in interrogatives (the interpretation of bare indeterminate pronouns is determined solely by semantic operators at LF, as mentioned in footnote 8).

---

15. This parameterization is consistent with the so-called Borer-Chomsky Conjecture (Borer 1984, Chomsky 1995b; Baker 2008a,b), in which all parametric variation is reduced to different specifications of formal features in the lexicon. See chapter 5 for more discussion.

(28) a. Compositional indeterminate   b. Bare indeterminate

$$\text{NP}_{\text{uOp[ ]}}$$
$$\text{N}_{\text{uOp[ ]}} \quad \text{Root}$$
$$\text{PERSON,THING,}\ldots$$

$$\text{NP}$$
$$\text{N} \quad \text{Root}$$
$$\text{PERSON,THING,}\ldots$$
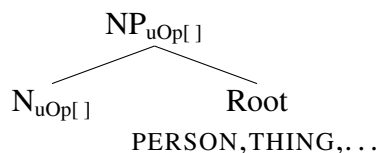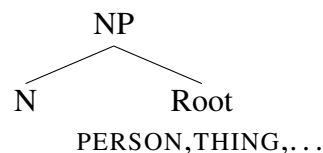
Turning to the distinction between non-affixal article languages on the one hand and affixal article languages and article-less languages on the other hand, I follow Talić (2015, 2017) in proposing that DP must project above NP in non-affixal article languages, whereas it can be absent in affixal article languages and article-less languages. Thus, in non-affixal article languages DP must project above indeterminate pronouns, which are NPs, whereas it can be absent in affixal article languages and article-less languages. In addition, I propose that this DP (i.e., D) bears a valued interpretable operator feature $iOp_{[Q]}$, which gives the value to the operator feature of N and marks the indeterminate pronoun as an "interrogative pronoun" in the traditional sense. The structure of "interrogative pronouns" in non-affixal article languages is schematized in (29).[16] Thus, indeterminate pronouns (but not "interrogative pronouns") are in a sense primitive in all languages, but in non-affixal article languages the D in question makes them "interrogative pronouns".

(29) a.
$$\text{DP}_{\text{iOp[Q]}}$$
$$\text{D}_{\text{iOp[Q]}} \quad \text{NP}_{\text{uOp[ ]}}$$
$$\text{N}_{\text{uOp[ ]}} \quad \text{Root}$$

b.
$$\text{DP}_{\text{iOp[Q]}}$$
$$\text{D}_{\text{iOp[Q]}} \quad \text{NP}_{\text{uOp[Q]}}$$
$$\text{valuation} \quad \text{N}_{\text{uOp[Q]}} \quad \text{Root}$$
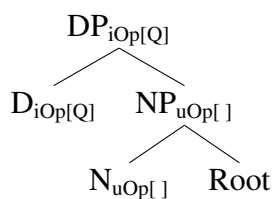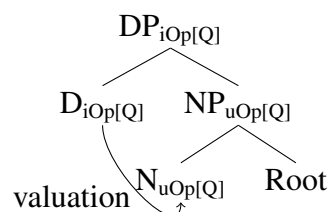
Notice now that the value of the operator feature of indeterminate NPs in non-affixal article languages is practically always [Q] because of the obligatory presence of the relevant D. This means that the operator feature of indeterminate NPs in those languages cannot have other values such as [∀], [∃], which would be given by quantificational particles/affixes. It thus follows that non-affixal

16. Here the structure is presented in the conventional fashion, but to be precise, the NP node in (29) is actually a set of features of the indeterminate pronoun projected under the Bare Phrase Structure Theory (Chomsky 1995a). Thus, the projected features of the indeterminate pronoun, including the operator feature, c-commands and probes (the features of) D.

article languages do not allow productive compositional indeterminate pronouns. In contrast, in article-less languages and affixal article languages, this D may be absent, so that if indeterminate NPs have the operator feature, it can be valued by quantificational particles/affixes, just like in Japanese. Thus, the generalization (27) is deduced from the operator feature analysis of indeterminate pronouns a lá Saito (2017) and the presence/absence of D that has a valued operator feature in the spirit of Talić (2015, 2017) (regarding affixal article languages). (Note again that (27) is a one-way correlation. The deduction here leaves room for the possibility that languages that lack definite articles and languages that have affixal definite articles may lack productive compositional indeterminate pronouns. The present proposal explains why productive compositional indeterminate pronouns are *in principle* allowed in those languages, whereas they are *never* allowed in non-affixal article languages.)

At this point, it should be noted that the DP under discussion only projects above indeterminate NPs; if it were to project above any NP, we would expect any noun to be able to be used as a question word, contrary to the fact. This restriction actually makes sense once we consider what counts as an extended projection in a lexical domain. Grimshaw (2000) and Biberauer et al. (2014) suggest that functional projections in the extended projection of a lexical category share some relevant feature(s) with the lexical category: e.g., [+N] in the case of the nominal domain. In the case of indeterminate pronouns, the DP in question shares an operator feature with the indeterminate NP. We can, then, maintain that the presence of the operator feature is the criterion for the DP in question being an extended projection of indeterminate NPs. Thus, the DP can only project above an indeterminate pronoun because of the presence of the operator feature on both elements; it cannot project above NPs that lack the operator feature.[17]

I would like to add here that the decomposition of indeterminate pronouns as N and Root is supported by their morphological compositions in comparison with demonstratives in some lan-

---

17. This reasoning leaves room for some non-affixal article languages to have bare indeterminate pronouns, given that bare indeterminate pronouns do not have an unvalued operator feature on which projection of the DP in question is contingent. As mentioned in footnote 14, Lakhota is indeed a language that allows bare indeterminate pronouns and has non-affixal articles. Interestingly, Williamson (1984) observes that Lakhota wh-questions, just like those in Chinese, are not subject to any island constraints, including wh-islands, as shown in (i).

guages. In Japanese, for instance, certain indeterminate pronouns share a root that specifies the domain with demonstratives, as shown in (30).[18] This pattern is also observed in Tamil, as in (31).

(30) Japanese: Kuno (1973), Martin (1975)

	proximal (*ko*)	medial (*so*)	distal (*a*)	indeterminate (*do*)
individual (*re*)	ko-re	so-re	a-re	do-re
place (*ko*)	ko-ko	so-ko	aso-ko	do-ko
adnominal (*no*)	ko-no	so-no	a-no	do-no
manner (*nna*)	ko-nna	so-nna	a-nna	do-nna
adverbial (*u*)	ko-u	so-u	a-a	do-u

(i) [tuwa takuwe cheya hą   ki] Marie inųǧa   he?
    who why   cry   DUR C Marie you.ask Q
    'Who did you ask Mary why (he) was crying?'   (Williamson 1984:269)

Indeterminate NPs in Lakhota can thus be analyzed as lacking the operator feature just like those in Chinese, and hence the D with the operator feature does not project above the indeterminate NP despite Lakhota being a non-affixal article language.

German, which appears to be a non-affixal article language, also allows bare indeterminate pronouns. It should, however, be noted that the environments where a bare indeterminate pronoun in German can occur are quite limited. Haspelmath (1997) observes that bare indeterminate pronouns in German must "cliticize" onto the finite verb, and they cannot precede the verb, as illustrated in (ii). I take this as indicating that the apparent bare indeterminate in (i) should not be analyzed in the same way as such elements in Chinese.

(ii) a. Da   kommt **wer**.
        there comes who
        'Someone is coming.'

     b. Jemand/***wer** kommt da.
        someone/who   comes there
        'Someone is coming.'

I suggest that indeterminate pronouns in German actually have an unvalued operator feature (which is usually valued by the valued operator feature of the D discussed in the text), and the apparent bare indeterminate in (i) can be analyzed as a sort of incorporation of indeterminate NP onto the verb. Baker (1988:285) claims that "[noun incorporation] is only possible if no [DP] is generated above the NP". In chapter 6, I will argue that German actually allows omission of DP in some very limited environments. In particular, German has strong and weak definite articles in Schwarz's (2009) sense; I will analyze the latter as not projecting DP (rather, it is head-adjoined to P). It is, then, not unexpected that German allows indeterminate NPs without the D that bears an operator feature discussed in the text in some very limited environments as incorporation onto the verb (note that nouns that can be incorporated are "bare" nouns that accompany no modifier or determiner, which can be analyzed as consisting of Root and N, just like indeterminate NPs). Regarding the unvalued operator feature, Baker proposes that incorporated NPs become "invisible" for Case assignment and do not need Case. Incorporation is standardly analyzed as involving head-adjunction, which is implemented by Pair-Merge in the current syntactic theory (Chomsky 2004, 2015; Epstein et al. 2016). Interestingly, Chomsky (2015) suggests that one of the heads in the ordered-pair created by Pair-Merge is "invisible" for Minimal Search, and unvalued features of the invisible head can be left unvalued (see chapter 4 for more discussion on this). The indeterminate NP in (iia) can, then, be analyzed as becoming invisible for Minimal Search due to the incorporation onto the verb, hence its operator feature can be left unvalued.

18. This is called the *ko-so-a-do* paradigm in the traditional Japanese grammar.

(31) Tamil: Asher (1985), Dixon (2003)

	proximal (*i*)	distal (*a*)	indeterminate (*e*)
nominal (*nta*)	i-nta	a-nta	e-nta
place (*ngke*)	i-ngke	a-ngke	e-ngke
time (*ppa*)	i-ppa	a-ppa	e-ppa
quantity (*ttane*)	i-ttane	a-ttane	e-ttane
manner (*ppati*)	i-ppati	a-ppati	e-ppati

From the perspective of the current proposal, the indeterminate pronouns and the demonstratives in these languages with the same root can be analyzed as having different feature specifications of N; the former have an unvalued operator feature, whereas the latter have a feature that encodes deixis (here I call this feature Dem(onstrative) for ease of exposition). This is illustrated in (32).

(32)   a.   Japanese

$ko_{\text{Dem[PROX]}}/so_{\text{Dem[MED]}}/a_{\text{Dem[DIST]}}/do_{\text{uOp[\_]}}$ *re/ko/no/nna/u*

   b.   Tamil

$i_{\text{Dem[PROX]}}/a_{\text{Dem[DIST]}}/e_{\text{uOp[\_]}}$ *nta/ngke/ppa/ttane/ppati*
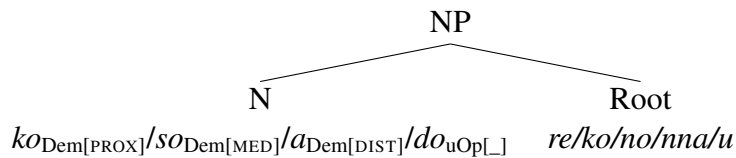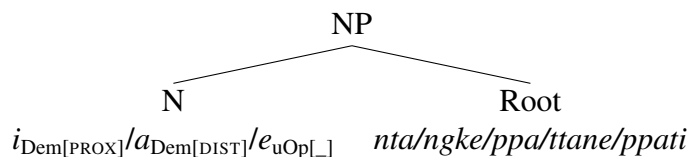
This analysis could actually be extended to English, where certain "interrogative pronouns" appear to share the root with demonstratives: e.g., *wh-at* vs. *th-at*, *wh-ere* vs. *h-ere* vs. *th-ere*, *wh-en* vs. *th-en* (cf. Kayne 2004, Nishiyama 2013). Crucially, however, the D that bears a valued operator feature must project above the indeterminate NP in English (English being a non-affixal article language), so the indeterminate pronouns in English necessarily function as "interrogative pronouns" in the traditional sense.[19]

---

19. It should be mentioned that there are "interrogative pronouns" that are also used as relative pronouns (e.g., English *who, which, where*, etc.). A possible explanation of this could be that the operator feature of D that projects above indeterminate NPs can have the value [Rel(ative)] instead of [Q]. The idea that both relative pronouns and "interrogative pronouns" have an operator feature is not new; see, e.g., Abels (2012), Bošković (2008a), Haegeman (2012), Starke (2001).

The proposal that indeterminate pronouns are generally NPs is supported by the observation that even locative, temporal, and manner indeterminate pronouns behave like NPs in languages like Japanese. It is well-known that nominal phrases in Japanese accompany case particles, as seen in (33a). Interestingly, those case particles can be attached to locative, temporal, and manner indeterminate pronouns, as shown in (33b)-(33d).[20]

(33) a. Nichiyoubi-**ga** ii.

Sunday-NOM good

'Sunday is good (for me).'

b. **Doko-ga** ii desu ka?

where-NOM good COP Q

'Where's good (for you)?'

c. **Itsu-ga** ii desu ka?

when-NOM good COP Q

'When's good (for you)?'

It should also be noted that English, which does not have productive indeterminate pronouns, still uses *where* for the locative indefinite pronouns (i.e., *somewhere*, *everywhere*, *anywhere*, *nowhere*). Cheng (1991) suggests that these are lexical compounds and hence sort of exceptional, but here I suggest the possibility of a different account from the perspective of the current proposal. In chapter 6, I will argue that P can function as the highest functional projection in the nominal domain (see also Grimshaw 2000, Bošković 2013a, Zanon 2020). There I will show that there are cases in which a definite article is omitted in locative PPs in some languages, which I will take as indicating that D can in principle be absent in locative PPs. Notice now that the indefinite pronouns in English in question are locative. Under the current proposal, locative indeterminate pronouns are inherently NPs and PP projects above them for adverbial usages (see also discussion in the text below). It is then possible that DP which is supposed to project above the locative indeterminate NP can be omitted because of the presence of locative P, and hence quantificational elements such as *some*, *every*, *any*, and *no* can merge above the indeterminate NP (and below PP). Thus, the current proposal has a potential to offer a principled explanation for this exceptional behavior of the locative *where* in the English indefinite pronominal system. (Note that other wh-items such as *who*, *what* cannot be used as compositional indeterminate pronouns, since they are not locative and hence the relevant P cannot merge above them.)
20. It is worth noting that the locative indeterminate pronouns in Japanese also require postpositions for an adverbial usage, as seen in (i). This can be taken as another piece of evidence that *doko* is NP rather than PP.

(i) Doko-*(**de**) ohiru-o tabemashita ka?
where-LOC lunch-ACC ate Q
'Where did you have lunch?'

d. **Dou-ga**   ii   desu ka?

how-NOM good COP Q

Lit. 'How's good (for you)?'

Note that *where* and *when* can also be used as subjects in colloquial English, as seen in the translation of (33b) and (33c). These data can be taken as indicating that these apparently adverbial "interrogative pronouns" in English are actually inherently NPs, and their adverbial usage can be analyzed as involving a null P that projects above the indeterminate NP (see, e.g., Huang 1982).

Interestingly, Huang (1982) argues that the reason indeterminate pronoun (which is conventionally translated as 'why') is inherently a PP. His argument is based on island sensitivity; the reason indeterminate pronoun *weishenme*, unlike other indeterminate pronouns, shows island sensitivity in Chinese.[21] This is illustrated in (34), where *weishenme* cannot be interpreted in the matrix clause, unlike *shei*. Huang essentially claims that the reason indeterminate pronoun being a PP causes an ECP problem (see also Lasnik and Saito 1984 for discussion). Nishigauchi (1990) also reaches the conclusion that the reason indeterminate pronoun *naze* in Japanese is a PP rather than an NP based on a similar argument.[22]

(34) ni   ziang-zhidao [shei weishenme mai-le   shu]?

you wonder   who why   buy-ASP book

'Who is the person x such that you wonder why x bought books?'

NOT 'What is the reason y such that you wonder who bought books for y?'

Interestingly for our current context, reason indeterminate pronouns cannot be used as indefinite pronouns cross-linguistically; they are generally used only as "interrogative pronouns" in the conventional sense. For instance, the reason indeterminate pronoun *naze* in Japanese cannot be productively combined with quantificational particles and used as indefinite pronouns. Thus, there is

---

21. Note that *wei* in *weishenme* can be used as a preposition meaning 'for' (*weishenme* is *wei* 'for' + *shenme* 'what').
22. It has been argued in the literature that in some languages 'why' is base-generated in Spec,CP unlike other indeterminate/"interrogative" pronouns (e.g., Rizzi 1990, 2001, Ko 2005, Stepanov and Tsai 2008, Yoshida et al. 2015), which points to the different status of 'why' compared with other indeterminate/"interrogative" pronouns. Relatedly, in French, which allows wh-in-situ in the matrix question, *pourquoi* 'why' is the only indeterminate/"interrogative" pronoun that cannot be used in wh-in-situ questions (Aoun 1986, Rizzi 1990, Bošković 2000).

no form of a universal quantifier, a free choice item, or an NPI/NCI composed of *naze*, in contrast with other indeterminate pronouns, as illustrated in (35). See Haspelmath (1997) for more cross-linguistic data.[23]

(35) a. *naze-mo

        why-also

        Intended: 'for every reason'

     b. *naze-demo

        why-even

        Intended: 'for any reason'

     c. *naze-mo

        why-also

        Intended: 'for no reason'

I suggest that the general impossibility of using reason indeterminate pronouns as indefinite pronouns is due to their PP status. Given Huang's (1982) and Nishigauchi's (1990) proposals, reason indeterminate pronouns can be analyzed as PP composed of P and Root under the current system. This is contrasted with other indeterminate pronouns including locative and temporal, which are NP composed of N + Root and merged with P for the adverbial usage as discussed above. I suggest that the P of the reason indeterminate pronouns bears an unvalued operator feature, and that this feature drives movement of the PP across languages, which is responsible for the island sensitivity of reason indeterminate pronouns in interrogatives (i.e., they need to undergo movement in narrow syntax even in languages like Chinese where other indeterminate pronouns do not undergo movement).[24] The structure of reason indeterminate pronouns is schematized in (36).

---

23. Željko Bošković (p.c.) points out that *zašto* 'why' in Serbo-Croatian has the same morphological composition as Chinese *weishenme*; *za* 'for' + *što* 'what'. Interestingly, although Serbo-Croatian has a productive compositional indeterminate pronominal system, *zašto* can only be used as an NPI (composed with *i*: *i-zašto*) and as an NCI (composed with *ni*: *ni-zašto*, note that these two usages are morphologically related), unlike other indeterminate pronouns. See also footnote 25 for the limited indefinite use of *naze* 'why' in Japanese.

24. As is well-known, extraction of adjuncts out of islands is generally worse than extraction of arguments. Here I put aside this argument-adjunct asymmetry (for an attempt to capture the asymmetry based on features, see, e.g., Starke
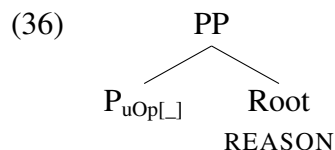
(36)

```
            PP
           /\
  P_uOp[_]    Root
            REASON
```

In chapter 6, I will argue that P can function as the highest functional projection in the nominal domain (see also Grimshaw 2000, Bošković 2013a, Zanon 2020). My suggestion here is that in the presence of P, no quantificational particle can project its own projection (say, QP) above reason indeterminate pronouns, since P "closes" the extended projection of the nominal domain as its highest projection. Thus, whether the language has a bare or compositional indeterminate pronominal system, it follows that the reason indeterminate pronoun cannot accompany a quantificational particle such as 'also', 'even', and hence they cannot be used as indefinite pronouns.[25,26]

2001, Friedmann et al. 2009, Haegeman 2012).

25. An exception to this is existential *ka*, which can be combined with *naze* to compose *naze-ka* 'for some reason' (this *ka* is syntactically different from the interrogative particle *ka*, which occurs sentence-finally; but see Uegaki 2018 for an argument for unification of these two in terms of semantics). It is worth noting here that existential *ka* behaves differently from *mo* and *demo*. In particular, *mo* and *demo* can be detached from their corresponding indeterminate pronouns to have a universal, negative polarity, or free choice interpretation, as shown in (1a)-(1c), while *ka* cannot be detached from an indeterminate pronoun with an existential meaning, as shown in (1d).

(i) a. [[Dare-ga   kaita] hon]-**mo**   yonda.
        who-NOM wrote book-also read.PAST
        'For every x, x a person, I read the book that x wrote.'

   b. [[Dare-ga   kaita] hon]-**mo**   yomanakatta.
        who-NOM wrote book-also not.read.PAST
        'For no x, x a person, I read the book that x wrote.'

   c. [[Dare-ga   kaita] hon]-**demo** yomu.
        who-NOM wrote book-even   read.PRES
        'For any x, x a person, I read the book that x wrote.'

   d. *[[Dare-ga   kaita] hon]-**ka** yonda.
        who-NOM wrote book-or read.PAST
        'For some x, x a person, I read the book that x wrote.'

It is then not unreasonable to conjecture that *mo* and *demo* on the one hand and *ka* on the other have different syntactic status. Specifically, I suggest that *mo* and *demo* project their own projection above indeterminate pronouns as an extended projection in the nominal domain (say, QP), whereas *ka* is simply adjoined to indeterminate pronouns. In the case of *naze*, which is inherently PP, neither *mo* nor *demo* can project above it because P is the highest projection in the nominal domain and no further nominal projection is allowed, whereas *ka* is adjoined to PP, so the issue of the highest functional projection in the nominal domain does not arise (note that the agreement relation between *ka* and *naze* can be established since the former c-commands the latter).

26. Manner indeterminate pronouns have also often been treated as inherently PPs in the literature. For instance, Huang (1982) analyzes the manner indeterminate pronoun *zenme* in Chinese as a PP, not an NP. Interestingly, *zenme* cannot be used as an indefinite pronoun, which can support its PP status just as in the case of *weishenme*. On the other hand, the manner indeterminate pronoun *dou* in Japanese can be used as an indefinite pronoun. In fact, there is cross-linguistic variation regarding whether manner indeterminate pronouns can be used as indefinite pronouns (see

To summarize this section, I have established the novel cross-linguistic generalization that languages that have productive compositional indeterminate pronouns either lack definite articles or have affixal definite articles. I have then proposed a deduction of this generalization: indeterminate pronouns are universally NPs (except for reason indeterminate pronouns); DP with a valued operator feature projects above indeterminate NPs in non-affixal article languages and this DP necessarily marks indeterminate pronouns as "interrogative pronouns" in the traditional sense, whereas it may be absent in article-less and affixal article languages, so that the operator feature of the indeterminate NP can be valued by a quantificational particle/affix.

3.4 Typology of the syntax of wh-questions

The typology of wh-questions has been one of the most widely discussed issues in the syntactic theory. In particular, what property correlates with wh-fronting and wh-in-situ has been one of the central questions.[27] Well-known in the literature is Cheng's (1991) Clausal Typing Hypothesis, according to which presence/absence of a Q-particle for interrogatives correlates with absence/presence of obligatory wh-fronting. This claim has, however, been challenged by Bruening (2007), who shows based on typological surveys that there is no such correlation. Bruening further claims that there is no correlation between the syntax of wh-questions and types of indefinite pronouns. As far as I know, the issue of correlations between the syntax of wh-questions and other linguistic properties has been left open.

In this section, I discuss this issue from the perspective of the classification of indefinite pronouns argued for in this chapter. After reviewing Cheng's (1991)'s Clausal Typing Hypothesis and Bruening's (2007) arguments against it, I introduce Bošković's (2020a) cross-linguistic gen-

---

Haspelmath 1997). It then seems that whether a manner indeterminate pronoun is inherently an NP or a PP is a point of parameterization/cross-linguistic variation. This is not implausible under the suggestion in the text given the Borer-Chomsky-Conjecture, which attributes parametric variation to feature specifications in the lexicon. That is, given that the lexical categories are defined by combinations of categorial features (Chomsky 1970), some languages use [+N,-V] (i.e., N) for the locus of the unvalued operator feature in the manner indeterminate pronoun, and others use [-N,-V] (i.e., P).

27. I use the terms such as *wh-fronting* and *wh-in-situ* only as descriptive terms of placement of indeterminate pronouns following the convention, i.e., I use it only for presentational purposes.

eralization regarding a correlation between multiple wh-fronting and compositional indeterminate pronouns, whose deduction Bošković himself leaves open. I will show that Bošković's generalization can actually be deduced from the proposal regarding compositional indeterminate pronouns presented in this chapter. Moreover, I point out that, under this deduction, wh-in-situ of the Japanese type is expected to behave like multiple wh-fronting despite the surface difference in placement of indeterminate pronouns, and show that this is indeed borne out. Finally, I discuss different types of wh-in-situ observed in the literature, and argue that the current proposal on feature specifications of indeterminate pronouns can capture the varied behavior of wh-in-situ languages.

3.4.1 Clausal Typing Hypothesis and arguments against it

Cheng (1991) proposes the Clausal Typing Hypothesis based on a number of languages, which is given in (37).

(37) Clausal Typing Hypothesis (Cheng 1991:29)

Every clause needs to be typed. In the case of typing a wh-question, either a wh-particle in C^0 is used or else fronting of a wh-word to the Spec of C^0 is used, thereby typing a clause through C^0 by Spec-head agreement.

Essentially, Cheng's claim is that wh-in-situ languages should have a Q-particle in wh-questions and wh-fronting languages should have no such particle in wh-questions. For instance, English, a wh-fronting language, does not have a Q-particle, whereas Japanese, a wh-in-situ language, has a Q-particle. However, Bruening (2007) argues against the Clausal Typing Hypothesis based on broader typological surveys. Bruening first quotes Ultan's (1978) cross-linguistic survey, in which 62 languages are examined. The data are summarized in (38).[28]

---

28. In Ultan's survey, languages that have wh-fronting and a Q-particle are Agta, Albanian, Syrian Arabic, Basque, Burmese, Chontal, Fanti, Finnish, French, Louisiana Frenche, Scottish Gaelic, Gbeya, Grebo, Guarani, Gunwinggu, Hebrew, Hungarian, Irish, Jaqaru, Klamath, Lithuanian, Malagasy, Malay, Ojibwa, Piro, Russian, Squamish, Tagalog, Twi, and Zapotec. Languages that have wh-in-situ and lack a Q-particle are Amharic, Neo-Aramaic, Gujarati, Kurku, and Tongan.

In fact, Slavic languages, which are (multiple) wh-fronting languages, have the Q-particle *li*, which can occur in wh-questions.

(38) Numbers of languages in Ultan's (1978) survey

	wh-in-situ	wh-fronting
particle	14	30
no particle	5	13

Bruening also reports an even larger scale of language samples, based on Matthew Dryer's database (Dryer 2004). The result of the survey is given in (39), quoted from Bruening (2007).

(39) Numbers of languages from Dryer (2004)

	wh-in-situ	wh-fronting
Q-particle	258	123
No Q-particle	143	53
Total	401	176
Percentage Q-particle	64	70

Crucially, there are languages which have both wh-fronting and a Q-particle, and languages that have wh-in-situ but lack a Q-particle. As Bruening points out, this is obviously inconsistent with Cheng's Clausal Typing Hypothesis. Bruening thus concludes that the Clausal Typing Hypothesis is untenable.

Bruening further claims that there is no correlation between types of wh-questions and properties of "interrogative pronouns". He first discusses Cole and Hermon's (1998) proposal that wh-questions universally involve a question operator and its corresponding variable. Cole and Hermon parameterize the relation between the operator and the variable; in wh-fronting languages like English, the operator and the variable are combined as a single lexical item (such as *who*, *what*), whereas in wh-in-situ languages like Chinese, the operator and the variable are two distinct lexical items, the latter corresponding to bare indeterminate pronouns in the current terminology.[29] As Bruening (2007) points out, Cole and Hermon's proposal makes two typological predictions.

---

29. Cole and Hermon's account is thus somewhat similar to the current proposal on indeterminate pronouns, in that the "interrogative pronoun" is not a primitive form and can be composed of an element that specifies the domain of quantification and an operator that is responsible for the interrogative force (in the case of, e.g., English).

First, all wh-in-situ languages should be able to use "interrogative pronouns" as indefinite pronouns (i.e., all wh-in-situ languages should have indeterminate pronouns), since the operator is a distinct lexical item and hence "interrogative pronouns" lack inherent quantificational force. Second, no wh-fronting languages should allow "interrogative pronouns' to be used as indefinite pronouns, which are understood as bare indeterminate pronouns in the current terminology, since "interrogative pronouns" necessarily involve the operator, which forces them to function as questions words under Cole and Hermon's proposal. Crucially, Bruening shows that neither of these predictions is borne out. Regarding the first prediction, Cole and Hermon (1998) themselves acknowledge that Turkish is a counterexample, since it is a wh-in-situ language but does not have indeterminate pronouns.[30] As for the second prediction, Bruening points out that Passamaquoddy and German, which are wh-fronting languages, allow bare indeterminate pronouns.[31]

(40)   a.   Kesq yaq   pemacqim-a-htit   otuhk-ol, on **keq** (')-nutom-oni-ya.

              while QUOT drag-DIR-3.PL.CONJ deer-OBV then what 3-hear-N-3.PL

              'While they were dragging the deer, they heard something.'

                                                    (Passamaquoddy: Newell 1974:5)

       b.   Es hat **wer** geklingelt.

              it   has who rung

              'Somebody has rung the bell'   (German: Postma 1994:188)

Bruening thus concludes that there is no correlation between the type of wh-questions (i.e., wh-fronting or wh-in-situ) and indeterminate pronouns.

---

30. See also Hiraiwa (2009), who shows that Gur languages have wh-in-situ and generic-noun-based indefinite pronouns.

31. See footnote 17 for discussion of German bare indeterminate pronouns under the current proposal. Slavic languages, which have productive compositional indeterminate pronouns and multiple wh-fronting, also appear to allow bare indeterminate pronouns in some limited environments (Izvorski 1996, Bošković 2002b). This can be analyzed in the same way as bare indeterminate pronominal cases in German discussed in footnote 17.

3.4.2 Multiple wh-fronting, indeterminate pronouns, and Japanese wh-in-situ

Note that Bruening's (2007) discussion concerns a rather simplistic cut in the syntax of wh-questions, namely, wh-fronting and wh-in-situ, since the primary purpose of his work is to show that previous theories such as Cheng's (1991) Clausal Typing Hypothesis and Cole and Hermon's (1998) proposal are typologically falsified. However, Bošković (2020a) observes an interesting typological correlation, by examining a particular type of wh-fronting. It is well-known that there are languages such as Slavic languages that force all indeterminate phrases to be fronted, which is called multiple wh-fronting. Interestingly, Bošković points out that Basque, Bulgarian, Czech, Latin, Hungarian, Macedonian, Mohawk, Polish, Romanian, Russian, Serbo-Croatian, Slovenian, Ukrainian, and Yiddish, all of which have multiple wh-fronting, have productive compositional indeterminate pronouns in the current terminology. Thus, he establishes the following descriptive generalization (adapted to the current terminology);[32]

(41) Languages that have multiple wh-fronting have productive compositional indeterminate pronouns.

This generalization is further confirmed by Estonian, Georgian, Latvian, Lezgian, Lithuanian, Ossetic, Tlingit, and Quechua.[33]

Bošković leaves open deduction of this generalization. Here I argue that it can actually be deduced from the current proposal regarding compositional indeterminate pronouns. Recall that under my proposal compositional indeterminate pronouns have an unvalued operator feature, which

---

32. This correlation is actually hinted at by Cheng (1991). However, she neither distinguishes bare and compositional indeterminate pronouns nor establishes a clear descriptive generalization.

33. Bošković (2020a, 2021a) also notes that compositional indeterminate pronouns in multiple wh-fronting languages do not allow the quantificational affix to be detached from them, which Bošković calls a sub-wh system (this is contrasted with Japanese, where quantificational particles can be detached from indeterminate pronouns; see footnote 25). Thus, Bošković refines (41) as (i).

(i) If a language has multiple wh-fronting, it has a sub-wh-system.   (Bošković 2021a:23)

This generalization has an apparent counterexample. Cable (2007, 2010) shows that Tlingit has multiple wh-fronting, but allows quantificational particles to be detached from compositional indeterminate pronouns. I do not discuss this issue here, because what matters for our purpose is that multiple wh-fronting languages have productive compositional indeterminate pronouns.

captures the dependency between compositional indeterminate pronouns and quantificational particles/affixes. In the absence of quantificational particles/affixes, the operator feature drives movement, since it needs to move to a position from which it can c-command and probe down a goal under Bošković's (2007b) refined theory of Agree. Under this proposal, multiple wh-fronting can and in fact needs to be analyzed as being driven by the unvalued operator feature of each compositional indeterminate pronoun in the relevant languages. Thus, it follows that if a language has multiple wh-fronting, it has productive compositional indeterminate pronouns.

Note now that the generalization (41) is a one-way correlation; namely, not all languages that have compositional indeterminate pronouns are expected to have multiple wh-fronting. In fact, Japanese has productive compositional indeterminate pronouns, but has wh-in-situ. Recall also that the current proposal that deduces the generalization is based on Saito's (2017) analysis of compositional indeterminate pronouns in Japanese, in which compositional indeterminate pronouns in Japanese undergo covert movement to a licensing position in interrogatives. This raises an interesting expectation regarding languages that have compositional indeterminate pronouns but do not have multiple wh-fronting. Crucially, the licensing mechanism of compositional indeterminate pronouns in Japanese is exactly the same as that in multiple wh-fronting languages under the current proposal; that is, the unvalued operator feature drives movement to a licensing position, whether it is overt or covert. It is, then, expected that multiple wh-fronting in the above languages and wh-in-situ in Japanese (and more generally languages with compositional indeterminate pronouns) should exhibit similar properties despite the surface fronting vs. in-situ difference.[34]

Two sets of data show that this is indeed borne out. The first concerns the superiority effect. Rudin (1998) observes that multiple wh-fronting in Polish, Serbo-Croatian, and Czech does not exhibit the superiority effect (see also Bošković 2002b for Russian), as exemplified by Serbo-Croatian (42).

---

34. Later I will offer an account of the overt vs. covert distinction in the movement of indeterminate pronouns, the gist of which is that the realization of a copy at PF is contingent on a focus feature on indeterminate pronouns.

(42) a. **Ko** je **što kome** dao?

who has what to.whom given

b. **Ko** je **kome što** dao?

who has to.whom what given

c. **Što** je **ko kome** dao?

what has who to.whom given

d. **Što** je **kome ko** dao?

what has to.whom who given

e. **Kome** je **ko što** dao?

to.whom has who what given

f. **kome** je **što ko** dao?

to.whom has what who given

'Who gave what to whom?'   (Rudin 1998:473)

Interestingly, Nishigauchi (1990) observes that Japanese does not show superiority effects.

(43) a. **Dare-ga dare-ni nani-o** agemashita ka?

who-NOM who-DAT what-ACC gave   Q

b. **Dare-ga nani-o dare-ni** agemashita ka?

who-NOM what-ACC who-DAT gave   Q

c. **Nani-o dare-ga dare-ni** agemashita ka?

what-ACC who-NOM who-DAT gave   Q

d. **Nani-o dare-ni dare-ga** agemashita ka?

what-ACC who-DAT who-NOM gave   Q

e. **Dare-ni dare-ga nani-o** agemashita ka?

who-DAT who-NOM what-ACC gave   Q

f. **Dare-ni nani-o dare-ga** agemashita ka?

who-DAT what-ACC who-NOM gave Q

'Who gave what to whom?' (Adapted from Nishigauchi 1990)

Thus, wh-in-situ in Japanese behaves like multiple wh-fronting in Polish, Serbo-Croatian, and Czech with respect to the absence of the superiority effect, which may be expected if these languages have the same underlying syntax of wh-questions.[35]

It should be added here that there are multiple wh-fronting languages which do show the superiority effect, as exemplified by Bulgarian (44).

(44) a. **Koj kogo** vižda?

who whom sees

'Who sees whom?'

b. *__Kogo koj__ vižda?

whom who sees (Bulgarian: Rudin 1998:472-473)

One might thus argue that the parallelism between multiple wh-fronting and Japanese-type wh-in-situ does not hold. However, there is an important typological difference between languages that show the superiority effect and those that do not. Bošković (2008b, 2012) observes that Czech, Hungarian, Mohawk, Polish, Russian, Serbo-Croatian, Slovenian, and Ukrainian exhibit no superiority effect, whereas Basque, Bulgarian, Macedonian, Romanian, and Yiddish show the superiority effect. Crucially, the latter languages have definite articles. Thus, Bošković proposes the following generalization:

(45) Multiple wh-fronting languages without articles do not display superiority effects (in cases like (42)).

Interestingly, Bošković (1997a) shows that even in those multiple wh-fronting languages with def-

---

35. It should be added here that there are some cases where the superiority effect is obtained in Serbo-Croatian and Japanese; see Bošković (2002b) and Takahashi (1993) respectively. In those cases, there are additional factors such as an overt Q-particle *li* in Serbo-Croatian and indeterminate pronouns in different clauses in Japanese. I put aside these complications here (but see chapter 4 on Serbo-Croatian *li*).

inite articles, the superiority effect only applies to the highest indeterminate phrase. Thus, when there are three indeterminate phrases in Bulgarian, the second and third indeterminate phrases show no superiority effect, as illustrated in (46).

(46) a. Koj **kogo kakvo** e pital?
          who whom what   is asked
          'Who asked whom what?'

      b. Koj **kakvo kogo** e pital?
          who what   whom is asked   (Bulgarian: Bošković 1997a:239)

Bošković (2002b) proposes that in multiple wh-fronting languages that show the superiority effect, the highest indeterminate phrase prior to wh-movement undergoes movement to Spec,CP triggered by a [+wh] feature of C, which is responsible for the superiority effect, whereas in multiple wh-fronting languages that do not show the superiority effect, C does not have this [+wh] feature and indeterminate phrases move to a lower position in the C domain (the same essentially applies to the non-initial indeterminate phrases in Bulgarian (46), hence the lack of the superiority effect). Bošković (2008b) suggests that the D feature is required for the relevant movement to Spec,CP, and this D feature is only present in languages that have definite articles, given the argument often made in the literature that there is a parallelism between DP and CP (see, e.g., Haegeman 2010).[36] At any rate, the superiority effect in the relevant languages is not an inherent property of multiple wh-fronting *per se* but obtains because of some additional factor (see chapter 4 for more discussion on this issue). Thus, I conclude that cases such as (44) do not serve as counterexamples to the above discussion, and that the parallelism between multiple wh-fronting languages and Japanese-type wh-in-situ languages holds.[37]

---

36. Note that this does not mean that the D feature projects DP in the indeterminate phrases; it can rather be naturally assumed under the Bare Phrase Structure Theory that it is part of the sets of features that constitute the indeterminate pronoun and the relevant C head. For related discussion, see chapter 5, where I discuss possibilities for the D-feature not projecting DP.

37. A possible expectation that may arise from this discussion is that wh-in-situ languages that have productive indeterminate pronouns and have affixal articles would show the superiority effect for the highest indeterminate phrase, since such languages are a wh-in-situ version of Bulgarian. Unfortunately, I have not found such a language so far (note that affixal article languages are relatively rare, and detailed syntactic data of the kind we are discussing here are

The second set of data that shows similarity between multiple wh-fronting and Japanese-type wh-in-situ concerns available interpretations in multiple wh-questions. In single wh-fronting languages such as English, multiple wh-questions only allow a pair-list answer. Thus, a legitimate answer to (47) would be something like *Mary bought a book, John bought bread, Sue bought wine, ...,* but not simply *Mary bought a book*.

(47)   Who bought what?   (^{ok}pair-list/*single-pair)

In contrast, as discussed in Bošković (2001a), Japanese multiple wh-questions allow both a pair-list answer and a single-pair answer. Thus, (48) can be answered with *Mary-ga hon-o katta yo* 'Mary bought a book'.

(48)   Dare-ga   nani-o   katta   no?   (^{ok}pair-list/^{ok}single-pair)
       who-NOM   what-ACC   bought   Q

       'Who bought what?'   (Bošković 2001a:2, attributed to Mamoru Saito)

Bošković (2001a) notes that German patterns with English, whereas Chinese and Hindi pattern with Japanese. He points out that the difference correlates with whether the "interrogative"/indeterminate pronoun (overtly) moves to Spec,CP; the indeterminate pronouns in English and German move to Spec,CP, whereas those in Japanese, Chinese, and Hindi do not. He argues that this is further supported by French, which allows wh-fronting and wh-in-situ. In (49a), where the "interrogative" pronoun *qu(oi)* 'what' is fronted, only the pair-list answer is possible. In contrast, the single-pair answer is allowed in (49b), where all the "interrogative" pronouns stay in-situ. Bošković thus concludes that movement to Spec,CP makes the single-pair answer unavailable.

(49)   a.   Qu'a-t-il   donné à qui?   (^{ok}pair-list/*single-pair)
            what-has-T-he   given   to whom

            'Whad did he give to whom?'

---

not always available). I leave investigation of this prediction for future research.

b. Il a donné quoi à qui? (^{ok}pair-list/^{ok}single-pair)

he has given what to whom

'What did he give to whom?' (Bošković 2001a:3)

Crucially, Bošković (2002b) points out that Serbo-Croatian also allows both pair-list and single-pair answers, on a par with Japanese as well as Chinese, Hindi and wh-in-situ in French, despite the overt wh-fronting. (50) thus allows a single-pair answer such as 'Mary bought a book'. See also Stepanov (1998) for Russian and Citko and Grohmann (2001) for Polish, which behave like Serbo-Croatian in the relevant respect. Bošković argues that the single-pair answer is available in these languages because indeterminate pronouns in these languages do not move to Spec,CP.[38]

(50) Ko je šta kupio? (^{ok}pair-list/^{ok}single-pair)

who is what bought

'Who bought what?' (Bošković 2001a:3)

What is important in the current context is that the similarity between multiple wh-fronting in these languages and Japanese with respect to the availability of the single-pair answer is not unexpected under the current proposal that they involve the same underlying syntax, despite the surface placement of indeterminate phrases.

The parallelism between Japanese and multiple wh-fronting languages of the Serbo-Croatian type is also found when the word order of two indeterminate phrases in the above examples is changed. Hagstrom (1998) observes that when the lower indeterminate phrase is overtly moved over the higher indeterminate phrase, the sentence itself is grammatical (that is, there is no superiority effect), but it only has a single-pair reading.

(51) a. **Dare-ga** kinoo **nani-o** katta no? (^{ok}pair-list/^{ok}single-pair)

who-NOM yesterday what-ACC bought Q

'Who bought what yesterday?'

---

38. Bulgarian patterns with English in that it only allows a pair-list answer. Bošković (2002b) attributes this to movement of an indeterminate phrase to Spec,CP, similarly to the superiority effect discussed above.

b. **Nani-o$_i$**   kinoo   **dare-ga**   t$_i$ katta   no?   (*pair-list/^{ok}single-pair)

what-ACC yesterday who-NOM   bought Q   (Hagstrom 1998:74)

Interestingly, Bošković (2001a) observes the same pattern in Serbo-Croatian; in (52b), where the object indeterminate phrase is moved over the subject indeterminate phrase, only the single-pair answer is possible.

(52)   a.   **Ko**   je **šta**   kupio?   (^{ok}pair-list/^{ok}single-pair)

who   is   what   bought

'Who bought what?'

b.   **Šta$_i$**   je **ko**   kupio?   (*pair-list/^{ok}single-pair)

what   is   who   bought   (Bošković 2001a:12)

Thus, despite the surface difference regarding the positions of indeterminate pronouns, Japanese wh-in-situ and multiple wh-fronting (of the Serbo-Croatian type) show similar behavior in the syntax and semantics. I take this as supportive evidence for my argument that multiple wh-fronting languages and wh-in-situ of the Japanese type have the same underlying licensing mechanism, i.e., movement (overt or covert, see below) of indeterminate pronouns driven by the operator feature that is also responsible for the productive compositional indeterminate pronominal system. The only difference between the two types of languages is PF realization of copies, i.e., whether the highest copy or the lowest copy of the chain is pronounced.

A remaining question is what is responsible for the difference in PF realization of indeterminate pronouns. A number of authors have proposed that (multiple) wh-fronting is driven by focus (e.g., Bošković 1999, 2002b, Horvath 1996, É Kiss 1995, Lambova 2001, Rochemont 1986, Stepanov 1998, Stjepanović 1999, Watanabe 2002). In particular, Watanabe (2002) proposes that overt wh-fronting in general is triggered by an uninterpretable focus feature. He examines loss of overt wh-fronting in the history of the Japanese language, and observes that in Old Japanese (in the Nara period) an indeterminate phrase is overtly fronted with the Q-particle *ka* attached to (the phrase that contains) the indeterminate phrase. This is shown in (53a), where the indeterminate phrase + the Q-

particle *izuku-yu-ka* 'from where' is located higher than the nominative subject *imo-ga* 'my wife'. In the later periods the Q-particle was located at the sentence-final position and correspondingly indeterminate phrases stayed in-situ, which is essentially what we find in Present Japanese. The Q-particle *ka* was also used to mark a focused constituent in polar questions, which appears higher than a nominative subject just like indeterminate phrases, as seen in (53b).

(53) a. Kado tate-te to-mo sashi-taru-wo **izuku-yu-ka** <u>imo-ga</u>

gate close-CONJ door-ALSO shut-PAST-ACC where-through-KA wife-NOM

iriki-te yume-ni mie-tsuru?

enter-CONJ dream-LOC appear-PERF

'From where did my wife come and appear in my dream, despite the fact that I closed the gate and shut the door?' (Man'youshuu #3117, Watanabe 2002:182)

b. …**[Hatsuse-no kawa-ha ura na-mi]-ka** <u>fune-no</u> yori-ko-nu?

Hatsuse-GEN river-TOP shore absent-ness-KA boat-NOM approach-come-NEG

'Is it because Hatsuse River has no shore that no boat comes near?'

(Man'youshuu #3225, Watanabe 2002:183)

Watanabe thus suggests that *ka* has a focus feature, which drives movement to Spec,FocP in the left periphery (Rizzi 1997).[39] Watanabe (2004b) also points out that in Imbabura Quechua the focus marker *taj*, which expresses exclusivity, is used in wh-questions. This focus marker is obligatory in wh-questions, and wh-in-situ is not allowed in this language.[40]

---

39. Aldridge (2009, 2018) claims that the relevant movement targets a position lower than Spec,CP (see also Bonan 2019, Bošković 2021b, Dadan 2019 and references therein). The choice between these two proposals does not matter for our discussion here.

40. It is worth adding here that in Ancash Quechua, the presence/absence of *taq*, which is the cognate of Imbabura Quechua *taj*, correlates with wh-fronting/wh-in-situ, as shown in (1).

(i) a. **May-man-taq**_i [José munan [María t_i aywanan-ta]]?
   where-to-Q Jose wants Maria will.go-ACC
   'Where does Jose want Maria to go?'

   b. [José munan [María **may-man** aywanan-ta]]?
    Jose wants Maria where-to will.go-ACC
    'Where does Jose want Maria to go?' (Cole and Hermon 1994:240)

(54) a. {**Ima-ta-taj**} ya-ngui [Juan {*****Ima-ta-taj**} randi-shka]-ta?

what-ACC-Q think-2PL Juan what-ACC-Q bought-C-ACC

'What do you think Juan has bought?' (Imbabura Quechua: Cole 1982:21)

b. Chay-ta-**taj** muna-ni

that-ACC-TAJ want-1

'I want that very one.' (Cole 1982:167)

Watanabe (2002) also maintains that this focus feature should be uninterpretable, because it would be redundant to add an interpretable focus feature to an indeterminate phrase, which itself is inherently interpreted as focus. This argument can be elaborated in a Hamblin alternative semantics, in which "interrogative pronouns"/indeterminate pronouns themselves are assumed to universally denote a set of individuals whether they undergo overt movement or not (see, e.g., Beck 2006). In this framework, focus also yields a set of alternatives; for instance, in a sentence *[MARY]$_F$ likes John*, there are alternatives of individuals in addition to *Mary*. If an "interrogative pronoun"/indeterminate pronoun to be fronted is assigned an interpretable focus feature, it would have different interpretations depending on whether it undergoes overt movement or not (a set of individuals in the case of wh-in-situ and a set of sets of individuals in the case of wh-fronting), which is counterintuitive with respect to their interpretation and complicates the general picture of semantics of indeterminate pronouns. In fact, I have shown above that multiple wh-fronting (of the Serbo-Croatian type) and wh-in-situ (of the Japanese type) can receive the same interpretation, and that compositional indeterminate pronouns, which are naturally assumed to have the same semantic interpretations across the same type of indefinite pronouns (e.g., existential, universal, free choice), are found in (multiple) wh-fronting languages and wh-in-situ languages. Thus, I concur with Watanabe (2002) that wh-fronting is derived by an unvalued uninterpretable focus feature.

It is worth noting here that under the current proposal, wh-fronting in languages that have non-affixal definite articles is solely triggered by an unvalued uninterpretable focus feature. Recall that I have proposed that DP obligatorily projects above the indeterminate NP in non-affixal article lan-

guages. In those languages, the operator feature of the indeterminate NP is valued by the valued operator feature of the D head that obligatorily projects above the indeterminate NP. Thus, the entire phrase (DP) lacks the unvalued operator feature that would otherwise drive (covert) movement under Bošković's (2007b) theory of Agree adopted here. It is, then, expected that "interrogative pronouns" that do not undergo overt movement and hence lack an unvalued focus feature in those languages (such as wh-in-situ in multiple wh-questions) should not show movement behavior. This is indeed supported by parasitic gaps. It is well-known that English wh-fronting licenses parasitic gaps, as illustrated in (55), where *e* represents a parasitic gap.

(55)   Which articles$_i$ did John file t$_i$ without reading e_i?   (Engdahl 1983:5)

Crucially, in multiple wh-questions, a lower wh-phrase that stays in-situ does not license a parasitic gap, as shown in (56).

(56) *I forgot who filed which articles$_i$ without reading e_i?   (Engdahl 1983:14)

The contrast between (55) and (56) indicates that the in-situ wh-phrase in (56) does not undergo movement in narrow syntax.

It should be immediately added here that the failure to license the parasitic gap in (56) cannot be attributed to the surface position of the in-situ wh-phrase. Bošković (2002b) observes that in Romanian, a multiple wh-fronting language, parasitic gaps are licensed by an indeterminate phrase that appears to stay in-situ when there is another homophonous indeterminate phrase, as shown in (57), where there are two instances of *ce* 'what' and the lower one is pronounced in the base position.[41]

(57)   Ce   precede   ce$_i$ fără   să   influenţeze e_i?
        what precedes what   without SUBJ.PART influences
        Lit. 'What precedes what without influencing?'   (Bošković 2002b:374)

If an in-situ "interrogative pronoun"/indeterminate pronoun in general could not license a parasitic

---

41. Note that paragistic gaps are not universally available. Bošković (2002b) notes that Bulgarian and Serbo-Croatian lack parasitic gaps in general, so the test in the text is not applicable.

gap, it would be mysterious why the parasitic gap in (57) is licensed by the in-situ indeterminate pronoun *ce*. Notice that the current proposal can straightforwardly account for the contrast between (56) and (57). In (56), the in-situ wh-phrase has no unvalued feature that would trigger movement, since the operator feature of the indeterminate NP is valued by that of the D head above the indeterminate NP, and this D does not have an unvalued focus feature (this is contrasted with the fronted wh-phrase in (55), whose operator feature is valued but which has an unvalued focus feature that triggers overt movement). On the other hand, the in-situ indeterminate pronoun in (57) has an unvalued operator feature (note that Romanian has productive compositional indeterminate pronouns and hence lacks D above indeterminate NPs), so that it undergoes movement to the C domain in narrow syntax (and the lowest copy is pronounced because the sequence of the homophonous indeterminate pronouns, i.e., *ce ce* is banned by a PF constraint; see Bošković 2002b for discussion). Thus, I conclude that English wh-in-situ indeed does not undergo movement in narrow syntax, which is correctly captured by the current proposal regarding the structure of wh-phrases and overt wh-fronting.

It is worth adding French here, which lacks compositional indeterminate pronouns and allows both wh-fronting and wh-in-situ. Bošković (2002b) observes that in French, overt wh-fronting licenses parasitic gaps, whereas wh-in-situ does not, as illustrated in (58).

(58) a. Qu'$_i$-a-t-il   lu   t$_i$ sans   classer e_i?

         what-has-T-he read   without filing

         'What has he read without filing?'

      b. *Il a   lu   quoi$_i$ sans   classer e_i?

         he has read what without filing

         'What has he read without filing?'   (French, Bošković 2002b:376)

Since French lacks productive compositional indeterminate pronouns and has non-affixal definite articles, "interrogative pronouns" in French should be analyzed as having the same structure as those in English; D with the valued operator feature projects DP above the indeterminate NP. Wh-

fronting is then triggered by a focus feature as in English, and wh-in-situ does not move in narrow syntax due to the lack of a feature that would trigger movement. Thus, the failure to license the parasitic gap in (58b) can be capture by the current proposal.[42]

How about wh-in-situ languages, then? Given the current proposal that compositional indeterminate pronouns in Japanese undergo covert movement due to their operator feature, it is predicted that Japanese wh-in-situ should also license parasitic gaps just like Romanian wh-in-situ. Unfortunately, this is difficult to test, since Japanese is a radical *pro*-drop language and hence it is not immediately clear whether the relevant gap is a parasitic gap or *pro*. In fact, this issue has been controversial in the literature. Takahashi (2006) claims that apparent parasitic gaps in Japanese are ellipsis sites, while Abe (2011) argues that Japanese indeed allows real parasitic gaps. Hirayama (2018), on the other hand, proposes that what look like parasitic gaps in Japanese are best analyzed as *pro*. Discussing this issue is well beyond the scope of this dissertation, and it remains to be investigated whether the above prediction will be borne out.

Still, there is another prediction regarding parasitic gaps with wh-in-situ. Specifically, since bare indeterminate pronouns lack the operator feature that would drive movement under the current proposal, it is predicted that bare indeterminate pronouns cannot license parasitic gaps. This is borne out in Chinese. Lin (2005) observes that Chinese wh-in-situ does not license parasitic gaps despite the availability of object *pro*, as shown in (59). This is contrasted with (60), where the indeterminate phrase is left-dislocated and the parasitic gap is licensed.

(59) a. *Laowang [zai hujian e_i zhiiqian] jiu   kaichu-le shei$_i$?

        Laowang at meet   before   already fire-ASP who

        'Who did Laowang fire before meeting?'

   b. *Laowang [zai do-guo e_i zhihou] jiu   diudiao-le sheme wenjian$_i$?

        Laowang at read-EXP   after   then throw-ASP what   document

        'Which document did Laowang throw away right after reading?'   (Lin 2005:299)

---

42. Malay behaves similarly to French in the relevant respect. See footnote 43 for discussion.

(60) a. Shei$_i$ Laowang [zai hujian e_i zhiiqian] jiu   kaichu-le?

       who Laowang at meet   before   already fire-ASP

       'Who did Laowang fire before meeting?'

b. Sheme wenjian$_i$ Laowang [zai do-guo e_i zhihou] jiu   diudiao-le?

       what   document Laowang at read-EXP   after   then throw-ASP

       'Which document did Laowang throw away right after reading?'   (Lin 2005:300)

One may wonder if the left-dislocation in (60) involves movement, since it is well-known that Chinese left-dislocation does not show island effects, as seen in (61), hence the left-dislocated element in (59) could be base-generated there. However, Lin points out that left-dislocation of an indeterminate phrase exhibits island sensitivity, as seen in (62). This indicates that the indeterminate phrases in (60) have undergone A′-movement, licensing the parasitic gaps.

(61) Weiyu$_i$ Laowang yu-guo   [$_{island}$ xihuan t$_i$ de]   ren.

      tuna   Laowang meet-EXP   like   MOD person

      'Tuna, Laowang has met persons who like [it].'   (Lin 2005:300)

(62) *Shenme you$_i$ Laowang yu-guo   [$_{island}$ xihuan t$_i$ de]   ren

      what   fish Laowang meet-EXP   like   MOD person

      'What fish is it such that Laowang met persons who like it?'   (Lin 2005:300)

Lin further notes that left-dislocation of non-indeterminate phrases also license parasitic gaps, as shown in (63). Interestingly, when there is a parasitic gap, left-dislocation of non-indeterminate phrases shows island-sensitivity as seen in (64), which indicates that the non-indeterminate phrases are not base-generated in the surface position but moved from the complement of the matrix verb.

(63) a. Xiaoli$_i$ Laowang [zai hujian e_i zhiiqian] jiu   kaichu-le.

       Xiaoli Laowang at meet   before   already fire-ASP

       'Xiaoli, Laowang fire before meeting.'

b.　Zhefen wenjian$_i$　Laowang [zai do-guo　　e_i zhihou] jiu　diudiao-le.

　　this　　document Laowang　at　read-EXP　　after　　then throw-ASP

　　'This document, Laowang throw away right after reading.'　　　(Lin 2005:301)

(64)　a. *Xiaoli$_i$　Laowang [zai huijian　e_i　zhiqian] jiu　　　tingdao [$_{island}$ Zhangsan

　　　Xiaoli　Laowang　at　meet　　　before　already hear　　　　　Zhangsan

　　　kaichu-le　t$_i$]　de　　xiaoxi.

　　　throw-ASP　　MOD　news

　　　'Xiaoli, Laowang heard the news before meeting [him] that Zhangsan fired [him].'

　　b. *Zhefen wenjian$_i$　Laowang [zai du-guo　　e_i zhihou] jiu　tingdao [$_{island}$ Zhangsan

　　　this　　document Laowang　at　read-EXP　　after　　then hear　　　　Zhangsan

　　　diudiao-le　t$_i$]　de　　xiaoxi

　　　throw-ASP　　MOD　news

　　　'This document, Laowang heard the news right after reading [it] that Zhangsan threw

　　　[it] away.'　　　　　　　　　　　　　　　　　　　　　　　(Lin 2005:301)

Thus, Lin concludes that the gaps in (60) and (63) are licensed by A′-movement (i.e., left-dislocation) and that Chinese has parasitic gaps, which cannot be licensed by in-situ indeterminate pronouns as in (59). This is predicted by the current proposal, since bare indeterminate pronouns do not undergo movement in narrow syntax due to the lack of the operator feature, unlike compositional indeterminate pronouns.[43]

---

43. As mentioned in footnote 42, Malay is a language where wh-fronting is optional like French. Bošković (2002b) observes that, just as in French, wh-fronting licenses a parasitic gap, whereas wh-in-situ does not, as seen in (i).

(i)　a.　Buku yang mana$_i$ kamu aturkan t$_i$ tanpa　　baca　　e_i?
　　　　book　that which　you　filed　　　without reading
　　　　'Which book did you file without reading?'

　　b. *Kamu aturkan buku yang mana$_i$ tanpa　　baca　　e_i?
　　　　you　filed　　book that　which without reading
　　　　'Which book did you file without reading?'　　　　　　　(Bošković 2002b:376)

Cole and Hermon (1998:240, fn.26) mention that for indefinite usages Malay indeterminate pronouns need to be reduplicated or accompany the morpheme *pun*, and consider them to be similar to those in Japanese (though they do not provide examples; see also Haspelmath 1997). One might then argue that Malay serves as a counterexample to

3.4.3 Different types of wh-in-situ

In section 3.2, we saw that indeterminate pronouns in Japanese and Chinese show different behavior in a number of respects. Specifically, indeterminate pronouns in Japanese need to be accompanied by a quantificational particle (i.e., compositional indeterminate pronouns), whereas those in Chinese need not (i.e., bare indeterminate pronouns). In addition, Japanese wh-in-situ is subject to wh-islands, whereas Chinese wh-in-situ is not. Saito (2017) argues that Japanese compositional indeterminate pronouns have an unvalued operator feature that needs to Agree with a licenser (interrogative C or a quantificational particle) and undergoes covert movement in interrogatives, whereas Chinese bare indeterminate pronouns lack it and hence do not move (and are licensed by unselective binding), a difference which I have shown is confirmed by anaphor binding effects. Thus, the parameter behind Japanese-type compositional indeterminate pronouns and Chinese-type bare indeterminate pronouns is the presence/absence of the operator feature on indeterminate pronouns (more specifically, on N).

In this subsection, I discuss other types of wh-in-situ, which look similar to that of Japanese or Chinese at a first glance. I show that in Sinhala, which appears to have a compositional indeterminate pronominal system like Japanese and has received uniform analyses of the syntax of wh-in-situ with Japanese in the literature, wh-in-situ actually behaves differently from that of Japanese. I also demonstrate that Vietnamese wh-in-situ behaves differently from that of Chinese even though Vietnamese has bare indeterminate pronouns just like Chinese. I argue that the behavior of indeterminate pronouns in these languages can be naturally captured under the current parameterization of indeterminate pronouns.

---

the current proposal. Crucially, however, Cole and Hermon (1998) observe that wh-in-situ in Malay does not exhibit island sensitivity, just like wh-in-situ in Chinese. We can then conclude that indeterminate pronouns in Malay do not have an unvalued operator feature, and hence do not move in narrow syntax at all in the case of wh-in-situ. For the apparent compositional indeterminate system, see section 3.4.3 for a relevant discussion of Sinhala, which can be extended to Malay.

Sinhala

The syntax of wh-questions in Sinhala has received a fair amount of attention in the syntactic literature (e.g., Cable 2007, 2010, Hagstrom 1998, Kishimoto 1992, 2005, Sumangala 1992). Sinhala wh-questions involve an in-situ indeterminate phrase and a Q-particle *də*.

(65) Chitra **monəwa də** gatte?

     Chitra what Q bought.E

     'What did Chitra buy?'   (Kishimoto 2005:3)

Sinhala also has compositional indeterminate pronouns, as illustrated in (66) (cf. Kishimoto 1992, 2005).

(66) a. monəwa-hari

         what-HARI

         'something'

     b. monəwa-t

         what-T

         'everything, anything'

Given the presence of compositional indeterminate pronouns in Sinhala, one might expect that wh-in-situ in this language would behave like that in Japanese. In fact, Sinhala wh-in-situ can occur inside an island similarly to Japanese wh-in-situ, as illustrated in (67) and (68).

(67) [adjunct island Chitra **monəwa** kanə kotə] də Ranjit pudumə unee?

              Chitra what ate time Q Ranjit surprised became.E

     'Ranjit was surprised when Chitra ate what?'   (Sinhala, Kishimoto 2005:30)

(68) [adjunct island Mary-ga **nani-o** tabeta toki]-ni John-wa odoroita no?

              Mary-NOM what-ACC ate time-at John-TOP got.surprised Q

     'John was surprised because Mary ate what?'   (Japanese)

Previous works have indeed proposed unified analyses of wh-in-situ in Sinhala and Japanese. For instance, Hagstrom (1998) and Kishimoto (2005) propose that the Q-particle but not the indeterminate pronoun moves to the licensing C, covertly in Sinhala and overtly in Japanese. In the default case like (65), the Q-particle is base-generated with the indeterminate pronoun, whereas in the cases such as (67) where the indeterminate pronoun is embedded inside an island, the Q-particle is base-generated at the edge of the island.

However, in section 3.3.1 I argued that wh-in-situ in Japanese actually moves covertly, contra Hagstrom's claim. Specifically, if a local anaphor, which itself cannot refer to the matrix subject as seen in (69a), is contained in an indeterminate phrase, it can refer to the matrix subject, as shown in (69b) (these examples are repeated from (18)).

(69) a. *Mary$_i$-wa [John-ga   kanojo-jishin$_i$-no ronbun-o   yonda to] shitta.
         Mary-TOP John-NOM her-self-GEN   paper-ACC read   C   found.out
         Lit: 'Mary found out that John had read herself's paper.'

      b. Mary$_i$-wa [John-ga   **dono**-kanojo-jishin$_i$-no ronbun-o   yonda ka] shitteiru.
         Mary-TOP John-NOM which-her-self-GEN   paper-ACC read   Q   know
         Lit: 'Mary knows which herself's paper John had read.'

I argued that this contrast shows that the indeterminate phrase covertly moves to the edge of the embedded interrogative CP. If the indeterminate phrase were to stay in-situ in narrow syntax as Hagstrom (1998) proposes, the contrast would be mysterious.[44]

At any rate, there is a crucial difference between Sinhala and Japanese regarding island sensitivity of wh-in-situ. As we saw in section 3.3.1, Japanese wh-in-situ is sensitive to wh-islands, as shown in (70), repeated from (12).

---

44. The behavior of Sinhala in this respect remains to be verified, which I leave for future work.

(70) Tanaka-kun-wa [dare-ga nani-o tabeta **ka**] oboeteimasu **ka**?

Tanaka-DIM-TOP who-NOM what-ACC ate Q remember Q

'Does Tanaka remember who are what?'

NOT 'For which *x*, *x* a person, does Tanaka remember what *x* ate?'

NOT 'For which *y*, *y* a thing, does Tanaka remember who ate *y*?'

In contrast, Sinhala wh-in-situ is not subject to wh-islands, as shown in (71).

(71) Ranjit [Chitra monəwa kieuwa də-næddə kiyəla] də danne?

Ranjit Chitra what read.A whether that Q know.E

'For which x, x a thing, Ranjit knows whether Chitra read x?'   (Kishimoto 2005:30)

Saito (2017) argues that the sensitivity of Japanese wh-in-situ to wh-islands in (70) is due to covert movement of the wh-phrase, the view I have also taken here. We are, then, led to conclude that wh-in-situ in Sinhala does not undergo covert movement, unlike that in Japanese.

Kishimoto (2005), who attempts to unify wh-in-situ in Japanese and Sinhala, acknowledges this contrast between these two languages, and suggests in his footnote 27 that 'whether' blocks binding between Q and an indeterminate pronoun in Japanese, whereas it does not in Sinhala some-how. If we unify the syntax of wh-in-situ in Sinhala and Japanese under Saito's (2017) proposal, Kishimoto's suggestion could be restated in a way that 'whether' does not constitute a wh-island in Sinhala. However, there is strong evidence against this. Crucially, Kishimoto himself shows that when the Q-particle occurs inside a wh-island, the sentence is unacceptable, as seen in (72).

(72) ?*Ranjit [Chitra monəwa **də** kieuwa də-næddə kiyəla] danne?

Ranjit Chitra what Q read.A whether that know.E

'Ranjit knows whether Chitra read what?'   (Kishimoto 2005:29)

(72) shows that wh-questions in Sinhala are indeed subject to wh-islands, which indicates that some element associated with the indeterminate phrase undergoes movement. But what is moving in Sinhala wh-question? Given that indeterminate pronouns themselves do not move as argued

above, we are led to conclude that the Q-particle undergoes covert movement from the surface position to the licensing C domain, as Hagstrom (1998) and Kishimoto (2005) propose.[45] What is important here is that Sinhala wh-questios involve a different syntactic mechanism than Japanese wh-questions (or Chinese wh-questions).

What does this mean for the current proposal? Since we have concluded that indeterminate pronouns in Sinhala do not undergo (covert) movement, we are led to analyze them as lacking an unvalued uninterpretable operator feature, unlike those in Japanese; otherwise, the operator feature would trigger (covert) movement of the indeterminate pronouns. Instead, the Q-particle, which is analyzed as moving, should have the relevant feature. Recall now that Sinhala has compositional indeterminate pronouns, just like Japanese (note also that Sinhala lacks definite articles just like Japanese). In my account, compositional indeterminate pronouns undergo agreement with a quantificational particle to value their operator feature. But we have concluded that Sinhala indeterminate pronouns lack the relevant feature. Does this mean that the above account of indeterminate pronouns needs to be modified or even abandoned?

The answer is negative. Note that the operator feature of compositional indeterminate pronouns of the Japanese type is an unvalued *uninterpretable* feature, so its role is solely to ensure a syntactic dependency between a compositional indeterminate pronoun and its licenser; it does not contribute to the interpretation of the indeterminate pronoun. Rather, the quantificational force of indeterminate pronouns is determined by a quantificational element in a given language (in other words, indeterminate pronouns universally denote a set of alternatives in the sense of a Hamblin

---

45. Yang (2021) proposes a slightly but significantly different Q-movement analysis of Sinhala wh-in-situ. Specifically, she argues that the apparent wh-island effect in (72) is not due to the movement of the Q-particle *per se*, but due to the placement of realization of copies of the Q-particle. One of her striking arguments is that even if the embedded clause is a non-island, the Q-particle cannot be located inside the embedded clause for long-distance questions, as seen in (i).

(i) a. *[Ranjit monəwa **də** gatta   kiyəla] kiuwe?
       Ranjit what   Q   bought that   said.E

   b. [Ranjit monəwa gatta   kiyəla] **də** kiuwe?
       Ranjit what   bought that   Q   said.E
      'What did you say that Ranjit bought?'   (Yang 2021:28)

See Yang (2021) for more extensive discussion. What is important for the current purpose is that Sinhala wh-in-situ is analyzed as involving a different licensing mechanism than Japanese wh-in-situ.

alternative semantics). It is, then, logically possible that indeterminate pronouns that lack the operator feature in question can be dependent on a quantificational particle with respect to their quantificational force. Those indeterminate pronouns are simply not syntactically dependent on any element, and their interpretation is purely determined in the semantics. In fact, even bare indeterminate pronouns in, e.g., Chinese, which are analyzed as lacking the feature in question, can co-occur with a quantificational element, which then determines the interpretation of the bare indeterminate pronouns. (73) shows that *dou* 'all' and *ye* 'also' provide the universal quantificational force for the bare indeterminate pronouns in Chinese.

(73) Ta <u>shenme</u> **dou/ye** xihuan.
      he what   all/also like
      'He likes everything.'   (Chinese, Li 1992:148)

Turning back to Sinhala, the apparent compositional indeterminate pronouns in Sinhala can be analyzed in the same way as bare indeterminate pronouns in Chinese in the relevant respect. The indeterminate pronouns in Sinhala lack the operator feature and hence do not need syntactic licensing, just as those in Chinese, with the quantificational force of the relevant pronouns determined by the quantificational particles (the only difference being that Sinhala indeterminate pronouns are accompanied by a particle for the indefinite interpretation unlike those in Chinese).

More generally, my proposal regarding indeterminate pronouns is a one-way correlation; indeterminate pronouns that have an unvalued uninterpretable operator feature in languages where DP need not project above them must be compositional indeterminate pronouns, but indeterminate pronouns that lack the feature can be either compositional indeterminate pronouns or bare indeterminate pronouns. Indeterminate pronouns in Sinhala thus fall under one of the types that the current proposal predicts.

Vietnamese

Another wh-in-situ language that is of particular interest in the current context is Vietnamese. As shown in (74), indeterminate pronouns in Vietnamese do not undergo overt wh-fronting.

(74) a. Tân mua **gì**?

Tan buy what

'What does Tan buy?'

b. *__Gì__ Tân mua t$_i$?

what Tan buy

'What does Tan buy?' (Bruening and Tran 2006:320)

Vietnamese also allows bare indeterminate pronouns, as seen in (75).

(75) Tân không gặp **ai**.

Tan NEG meet who

'Tan does/did not meet anyone.' (Tran 2009:141)

In addition, Vietnamese wh-questions are not sensitive to wh-islands when a sentence final particle is added, as illustrated in (76).

(76) Anh muốn biết [**ai** ăn **cái gì**] vậy/thế?

you want know who eat what PRT

'For which y, you want to know for which x, x ate y?' or

'For which x, you want to know for which y, x ate y?' (Tran 2009:211)

Thus, at a first glance, Vietnamese seems to be similar to Chinese, where bare indeterminate pronouns are allowed and wh-in-situ is not sensitive to wh-islands.

Interestingly, however, when a sentence final particle is absent, Vietnamese wh-in-situ shows island sensitivity. Tran (2009:210) notes that "[i]f [(77)] is uttered out of the blue, with a neutral intonation, the embedded reading is the most prominent one".[46]

---

46. Tran also mentions that one of the indeterminate pronouns can be interpreted in the matrix clause when a heavy

(77) Anh ấy muốn biết [**ai** ăn **cái gì**]?

he want know who eat what

'He wants to know for which person x, for which thing y, x ate y.' (Tran 2009:211)

In the absence of a sentence final particle, wh-in-situ in Vietnamese is also sensitive to other islands such as adjunct island, as shown in (78). This is contrasted with Japanese, where wh-in-situ can be embedded in islands other than wh-islands as seen above.

(78) Tân sẽ thua cuộc [adjunct island vì **ai** làm hu' xe cua anh ta]

Tan ASP lose event because who make damage vehicle belong he

*(thế)?

PRT

'Tan will lose the race because who damaged his car?' (Bruening and Tran 2006:327)

Based on these observations, Bruening and Tran (2006) propose that wh-in-situ in Vietnamese in the absence of a sentence final particle undergoes covert movement, whereas in the presence of a sentence final particle wh-in-situ is licensed by unselective binding. Their proposal can be straightforwardly accommodated under my proposal on the feature specification of indeterminate pronouns. Vietnamese is a language which allows two options for feature specification of indeterminate pronouns; having an unvalued operator feature or lacking it. In the case of wh-questions, when there is an unselective binder in the numeration, which is a sentence final particle, the latter option is chosen, and when it is not chosen in the numeration, indeterminate pronouns have an unvalued operator feature for licensing.

If indeterminate pronouns in Vietnamese can have an unvalued operator feature, it is expected that they can behave as compositional indeterminate pronouns (note that Vietnamese is an article-less language). This is indeed borne out. As shown in (79), an existential quantifier with the specific reading (taking scope over negation) is composed of an indeterminate pronoun and the

---

stress is put on it. This is similar to Japanese wh-in-situ in a wh-island, which can also take scope in the matrix clause if it has a heavy stress (see, e.g., Ishihara 2003).

demonstrative *đó* (see section 3.2.2 for relevance of a particle/affix for the wide scope interpretation). Note that this option is not available in Chinese, which is analyzed here as lacking the operator feature on indeterminate pronouns altogether. Recall that Vietnamese also allows bare indeterminate pronouns as repeated in (80), which takes narrow scope relative to the negation. In this case, the indeterminate pronoun lacks the operator feature, just like those in Chinese.

(79) Tân không gặp **ai** **đó**. ($\exists > \neg$)

      Tan NEG   meet who DEM

      'Tan does/did not meet someone.' (Tran 2009:159)

(80) Tân không gặp **ai**. ($\neg > \exists$)

      Tan NEG   meet who

      'Tan does/did not meet anyone.' (Tran 2009:141)

Thus, the dual nature of wh-in-situ in Vietnamese, i.e., the presence/absence of covert movement, can be captured by the current proposal on the feature specifications of indeterminate pronouns.

3.5 Conclusion of the chapter

In this chapter, I have discussed morpho-syntactic properties of indefinite pronouns from a typological perspective. I have first divided "interrogative-based" indefinite pronouns in Haspelmath's (1997) sense into two types from their morphological composition; compositional indeterminate pronouns and bare indeterminate pronouns. This distinction is supported by their semantic as well as syntactic properties. I have also introduced a revised version of Saito's (2017) analysis of the two types of indeterminate pronouns. I have then established the novel descriptive generalization that languages that have compositional indeterminate pronouns either lack definite articles or have affixal definite articles. As a deduction of this generalization, I have proposed that compositional indeterminate pronouns are NPs and have an unvalued uninterpretable operator feature; in article-less languages and affixal article languages a quantificational particle <u>can</u> project above composi-

tional indeterminate pronouns and provide a quantificational force, whereas in non-affixal article languages a D head that bears a valued operator feature <u>necessarily</u> projects above compositional indeterminate pronouns and marks them as question words.

I have then discussed consequences and extensions of this proposal for various aspects of the syntax of wh-questions. I have shown that multiple wh-fronting and wh-in-situ of the Japanese type show similar behavior in several respects regardless of the apparently different placement of indeterminate pronouns, and argued that this is naturally explained by the current proposal. I have also argued that the difference between multiple wh-fronting languages and single wh-fronting languages such as English with respect to the behavior of wh-in-situ follows from the current proposal. In addition, it has been demonstrated that the proposed account can capture various types of wh-in-situ via feature specifications of indeterminate pronouns in a given language.

Chapter 4

Large-scale Pied-piping, Weak Heads, and Deduction of Agree from Minimal Search

4.1 Introduction

In chapter 3, I discussed the typology of indefinite pronouns and the syntax of wh-questions. In particular, I proposed a new classification of indefinite pronouns. The relevant definitions are given in (1) and (2), which are exemplified in (3) and (4), respectively.

(1) DEFINITION 1: A *bare indeterminate pronoun* is a pronoun whose interrogative and indefinite usages have the same form.

(2) DEFINITION 2: A *compositional indeterminate pronoun* is a pronoun which functions as a pronoun in the context with interrogative force or as an indefinite pronoun when a specific quantificational particle/affix is added to it.

(3) Chinese

    a.   Ta yiwei wo xihuan **shenme**.

          he think I   like   what

          'He thinks I like something.'

b. Ta yiwei wo xihuan **shenme**?

he think I like what

'What does he think I like?' (Li 1992:125)

(4) Japanese

a. Kare-wa watashi-ga **nani-*(ka)**-ga sukida to omotteiru.

he-TOP I-NOM what-KA-NOM like C think

'He thinks I like something.'

b. Kare-wa watashi-ga **nani-(*ka)**-ga sukida to omotteiru no?

he-TOP I-NOM what-KA-NOM like C think Q

'What does he think I like?'

I showed in chapter 3 that bare indeterminate pronouns and compositional indeterminate pronouns exhibit different syntactic and semantic behavior, and argued that the differences can be captured by a parameterization of the feature specification of the two types of indeterminate pronouns. Specifically, compositional indeterminate pronouns have an unvalued uninterpretable operator feature, whereas bare indeterminate pronouns lack it. In addition, I argued that the proposed parameterization of the relevant indefinite pronouns can be extended to capture the typology of the syntax of wh-questions.

In this chapter, I discuss yet another domain in which compositional indeterminate pronouns correlate with a particular type of the syntax of wh-questions. Specifically, I introduce Watanabe's (2004b) observation that the availability of productive compositional indeterminate pronouns in my terminology is a prerequisite for large-scale pied-piping. I then discuss another factor that is relevant for the availability of large-scale pied-piping. In particular, I show that the SOV word order, or more precisely head-finality of the projection to be pied-piped, is crucial for large-scale pied-piping. I thus establish a novel cross-linguistic generalization regarding large-scale pied-piping. In order to deduce this new generalization, I propose a morpho-syntactic condition on

"weak heads" under Chomsky's (2015) labeling framework, in which weak heads are realized as bound morphemes, as well as a criterion for determining weak heads, which generalizes the notion of weak heads to all heads that have unvalued features at the point of External Merge. I then show that this conception of weak heads captures the availability of large-scale pied-piping, tied with a cross-linguistic morphological difference between head-initial and head-final complementizers observed by Inaba (2011).

In addition, I demonstrate that this new conception of weak heads allows, and in fact requires, us to deduce Agree from Minimal Search, which is a third factor principle external to UG. In other words, we can eliminate Agree from the computational system of language and hence minimize UG. I also show that this deduction can capture the variation in the superiority effects that is found with multiple wh-questions in combination with Epstein et al.'s (2020) path-based theory of Minimal Search and feature valuation.

Finally, I discuss Bošković's (2008a) generalization that D-linked and relative indeterminate phrases are insensitive to multiple wh-islands in languages with affixal definite articles, from the perspective of the labeling framework discussed here. I propose that this generalization can be captured by the syntactic nature of a head amalgam created by External Pair-Merge of two heads in Epstein et al.'s (2016) sense, once we consider the timing of the visibility of one of the two heads for Minimal Search in this configuration.

This chapter is organized as follows. In section 4.2 I establish a new cross-linguistic generalization regarding large-scale pied-piping. In section 4.3, I offer a deduction of the generalization by proposing a new morpho-syntactic condition and a criterion regarding weak heads in Chomsky's (2015) sense, as well as accommodating Inaba's (2011) observation that head-final complementizers are generally affixal. In section 4.4, I argue for a deduction of Agree from Minimal Search based on the proposed criterion for weak heads, and account for variation in the superiority effects with wh-questions based on Epstein et al.'s (2020) theory of Minimal Search. In section 4.5, I discuss Bošković's (2008a) generalization regarding insensitivity of D-linked/relative indeterminate phrases to multiple wh-islands in affixal article languages from the perspective of the labeling the-

ory, providing a deduction of his generalization based on the visibility of weak heads, which also resolves some technical issues in Bošković's original proposal. Section 4.6 concludes the chapter.

4.2 Large-scale pied-piping: A novel generalization

In some languages, an entire clause containing an indeterminate pronoun rather than the indeterminate pronoun itself can undergo movement to the licensing position of indeterminate pronouns in interrogatives. This phenomenon is called *large-scale pied-piping*. This is illustrated in Basque (5) and Imbabura Quechua (6).

(5) [**Nor** joango d-ela]$_i$ esan du   Jon-ek   t$_i$?

     who go   AUX-C say   AUX John-ERG

     'Who has John said will go?'   (Basque: Ortiz de Urbina 1989:248)

(6) [**Ima-ta**   wawa miku-chun-taj]$_i$ Maria t$_i$ muna-n?

     what-ACC child eat-FIN-Q   Maria   want-TNS.AGR

     'What does Mary want (that) the child eat?'   (Imbabura Quechua: Hermon 1984:152)

Large-scale pied-piping can also move an entire island that contains an indeterminate pronoun to a licensing position of an indeterminate pronoun without inducing island effects. In (7) and (8), the relative clause, which constitutes a complex NP island, is pied-piped by the indeterminate pronoun, without the island effect being obtained.

(7) [[**Nork**   idatzi zuen] liburua] irakurri du   Peruk?

     who.ERG write AUX book   read   AUX Peter.ERG

     'Who did Peter read the book that (he) wrote?'   (Basque: Ortiz de Urbina 1989:249)

(8) [[**Ima**-ta   randi-shka] runa]-ta-taj riku-rka-ngui?

     what-ACC buy-C   man-ACC-Q see-PAST-2

     'What did you see the man who bought (it)?'   (Imbabura Quechua: Cole 1982:24)

Interestingly, Watanabe (2004b) points out that Basque and Imbabura Quechua have productive compositional indeterminate pronouns in my terminology.[1] Watanabe thus suggests the following generalization (adapted to the current terminology):

(9)   Languages that allow large-scale pied-piping have productive compositional indeterminate pronouns.

As Watanabe suggests, Old Japanese is a language of this type. He points out that Old Japanese had compositional indeterminate pronouns, as shown in (10).

(10)   a.   *Universal Quantification*

...**itsu-mo**   **itsu-mo**   hito-no   yurusa-mu koto-wo-shi   mata-mu.

when-also when-also person-NOM accept-will word-ACC-PRT wait.for-will

'I will always wait for the woman to accept me.'

(Man'youshuu #2770, Watanabe 2004b:76)

b.   *Negative Polarity*

...**tare**-to   ihu hito-**mo**   kimi-ni-ha   masa-ji.

who-QUOT say person-also you-than-TOP superior-NEG

'Nobody would be nicer than you.'   (Man'youshuu #2628, Watanabe 2004b:77)

Aldridge (2009) shows that an indeterminate pronoun in Old Japanese can be embedded in an island, which is fronted to a licensing position of indeterminate pronouns in interrogatives, as seen in (11). This can be analyzed as a case of large-scale pied-piping on a par with (7) and (8).

(11)   Kono toki-fa   [adjunct island **ika-ni**   si-tutu]-**ka**   na-ga   yo-fa   wataru?
this   time-TOP   how-DAT do-while-KA you-NOM world-TOP pass

'At this time, you pass through this world doing what?'

(Man'youshuu #892, Aldridge 2009:560)

The generalization (9) is also confirmed by Latin, which had productive compositional inde-

---

1. Watanabe (2004b) also claims that languages that have head-internal relative clauses have compositional indeterminate pronouns. However, Hiraiwa (2009) shows that Gur languages, which do not have compositional indeterminate pronouns, have head-internal relative clauses.

terminate pronouns. Danckaert (2012) shows that Latin allowed large-scale pied-piping of adjunct clauses by indeterminate pronouns, as seen in (12).

(12) a. [adjunct island **Qu-am** utilitat-em aut **qu-em** fructu-m petent-es]

which-ACC use-ACC or which-ACC benefit-ACC searching-NOM

sci-re cup-imus illa, quae occult-a nobis sunt?

know-PR-INF desire-PR.1PL those.ACC which.NOM hidden-NOM us.DAT be.PR.3PL

'With which goal or benefit do we desire to know those things which are hidden for us?'

(Cicero. Fin. 3.37, Danckaert 2012:173)

b. Tu uero [adjunct island [**qu-ibus** re-bus gest-is] (&) [**qu-o**

you.NOM PRT which-ABL deeds-ABL done-ABL which-ABL

host-e superat-o]] contion-em aduoca-re aus-us es?

enemy-ABL defeated-ABL assembly-ACC convoke-PR.INF dared-NOM be.PR.2SG

'Which deeds have you been accomplished, which enemy has been defeated so that you

dared to convoke the assembly?' (Cicero. Ver. 3.185, Danckaert 2012:173)

Bengali, which has compositional indeterminate pronouns, also supports the generalization. Bengali is regularly an SOV language, but an object can be moved rightward. This rightward movement is generally a marked option, but if a complement clause follows the matrix verb, it is not marked at all (Bayer 1996). The two possible positions of a complement clause are shown in (13).[2]

(13) a. chele-Ta jan-e na [baba aS-be]

boy-CF know-3 not father come-FUT.3

'The boy doesn't know that (his) father will come.'

b. chele-Ta [baba aS-be] jan-e na

boy-CF father come-FUT.3 know-3 not (Bayer 1996:254)

---

2. The complement clause in (13) can also be moved to the leftmost position of the matrix clause, but this is irrelevant for our purposes.

Interestingly, when an embedded clause contains an indeterminate pronoun that is to take a scope in the matrix clause, the entire embedded clause needs to be located in the preverbal position, as shown in (14).[3]

(14) a. tumi [ke baRi kor-be]   bhab-cho?

   you   who house make-FUT.3 think-2

   'Who do you think will build a house?'

   b. *tumi bhab-cho [ke   baRi   kor-be]?

   you   think-2   who house make-FUT.3   (Bayer 1996:273)

In addition, an indeterminate pronoun can be contained in an adjunct clause and be interpreted in the matrix clause without inducing the adjunct island effect, as shown in (15).

(15) [tumi kothaY ge-le] tomar ma   khuSi hO-be?

   you   where   go-CPT your   mother happy become-FUT.2

   'Your mother will be happy if you go where?'   (Bayer 1996:283)

Simpson and Bhattacharya (2003) suggest that these data can be captured if Bengali allows large-scale pied-piping, just like Basque and Imbabura Quechua (Bayer 1996 also points out that Bengali is similar to Japanese in the relevant respects, see below for Japanese). If this is on the right track, Bengali falls under the generalization (9).

---

3. It is also possible to move the indeterminate pronoun out of the embedded clause in the postverbal position, as seen in (i).

(i) jon **ke** bollo [t$_i$ cole gache]
   John who said   left gone
   'Who did John say left?'   (Simpson and Bhattacharya 2003:133)

This is parallel to Basque and Imbabura Quechua, where an indeterminate pronoun can be extracted out of an embedded clause without pied-piping the clause, as shown in (ii) and (iii).

(ii) **Nor** uste duzu [ikusi duela   Peruk t$_i$]?
   who think AUX   seen has.that Peter
   'Who do you think Jon wrote?'   (Basque: Ortiz de Urbina 1993:194)

(iii) **Ima-ta-taj$_i$** ya-ngui   [Juan t$_i$ randi-shka]-ta?
   what-ACC-Q think-2PL Juan   bought-C-ACC
   'What do you think Juan has bought?'   (Imbabura Quechua: Cole 1982:21)

The generalization under discussion is further supported by Japanese, which has productive compositional indeterminate pronouns and has been argued to have covert large-scale pied-piping by a number of authors (e.g., Nishigauchi 1990, Richards 2000, Morita 2009). As shown in (16), wh-questions in Japanese are not subject to the complex NP island effect, similarly to Basque (7) and Imbabura Quechua (8).[4]

(16) Mary-wa [[**dare**-ga kaita] hon]-o yonda no?

     Mary-TOP who-NOM wrote book-ACC read Q

     'Who did Mary read the book that (he) wrote?'

Although large-scale pied-piping voids most of the island effects, it does not void the wh-island effect. Thus, Basque does not allow pied-piping of a wh-island, as shown in (17). Crucially, as Richards (2000) points out, Japanese also disallows pied-piping of a wh-island, similarly to Basque, as seen in (18).

(17) *[wh-island Nor etorriko d-en] galdetu duzu?

        who come AUX-Q asked AUX

     'Who have you asked whether t has come?' (Ortiz de Urbina 1993:197)

(18) *John-wa [wh-island Mary-ga nani-o katta kadooka] shiritagatteru no?

     John-TOP Mary-NOM what-ACC bought whether want.to.know Q

     'What does John want to know whether Mary bought t?' (Richards 2000:195)

Note that this is contrasted with Chinese (19), where a bare indeterminate pronoun embedded in a wh-island can be interpreted in the matrix clause (see chapter 3 for discussion on the nature of Chinese bare indeterminate pronouns).

---

4. The reason indeterminate pronoun *naze* 'why' in Japanese cannot pied-pipe an island, unlike other indeterminate pronouns (Nishigauchi 1990). Thus, *naze* 'why' cannot be embedded inside an island, as shown in (i).

(i)*Mary-wa [[John-ga **naze** kaita] hon]-o yonda no?
    Mary-TOP John-NOM why wrote book-ACC read Q
    Lit. 'Mary read the book that John wrote why?'

See footnote 20 for an analysis.

(19) Ni xiang-zhidao [shei mai-le   shenme]?

you wonder   who buy-ASP what

'What is the thing x such that you wonder who bought x' or

'Who is the person x such that you wonder what x bought'

Morita (2009) provides a number of arguments for existence of covert large-scale pied-piping in Japanese. A striking data among them is the possibility of binding of a local anaphor in an embedded clause by a matrix subject. As shown in (20a), repeated from chapter 3, a local anaphor *kanojo-jishin* "herself" in the embedded clause cannot be bound by the matrix subject. Crucially, when the local anaphor is contained in an indeterminate phrase that takes scope in the matrix clause, it can be bound by the matrix subject, as shown in (20b).[5]

(20) a. *Mary$_i$-wa [John-ga   kanojo-jishin$_i$-no ronbun-o   yonda to] shitta.

Mary-TOP John-NOM her-self-GEN   paper-ACC read   C found.out

Lit: 'Mary found out that John had read herself's paper.'

   b. Mary$_i$-wa [John-ga   **dono**-kanojo-jishin$_i$-no ronbun-o   yonda to] shitta   no?

Mary-TOP John-NOM which-her-self-GEN   paper-ACC read   C found.out Q

Lit. 'Which herself's paper did Mary found out that John had read?'

This contrast cannot be explained if the object indeterminate phrase containing the anaphor in (20b) stays in-situ. The acceptability of (20b) indicates that the indeterminate phrase undergoes covert movement to a position from which it is accessible to the matrix subject, just as the indeterminate phrases in Basque and Quechua undergo overt movement to the edge of the embedded clause.

It should be added here that Heck (2008, 2009) proposes a generalization regarding pied-piping, which is given as (21).

(21) *The Edge Generalization*

If α pied-pipes β, then α must be at the edge of β.

---

5. Morita (2009) uses *jibun-jishin* 'oneself' for the relevant test, but the judgment is less clear because this element can easily have a logophoric interpretation, by which a local anaphor can refer to a non-clausemate nominal. Here I use *kanojo-jishin* 'herself' to make the contrast sharper.

This generalization was originally meant to apply to overt edges. As seen in Basque (5) and Imbabura Quechua (6), the indeterminate phrase that pied-pipes the embedded clause is located at the edge of the embedded clause. We can even extend this to covert edges under the copy theory of movement; namely, a copy of α needs to be at the edge of β in narrow syntax, whether this copy or the one at the bottom of the chain is realized at PF. In Basque, Imbabura Quechua, and Latin, the copy at the edge of β is pronounced, whereas in Bengali and Japanese the lowest copy is pronounced. Thus, the Edge Generalization can be generalized more under the copy theory of movement, by which covert pied-piping falls under this generalization.

At this point, I would like to note that (9) is a one-way correlation; that is, not all languages that have compositional indeterminate pronouns allow large-scale pied-piping (e.g., Slavic languages, which have productive compositional indeterminate pronouns, do not allow large-scale pied-piping). A question that arises is, then, what property is relevant for the presence/absence of large-scale pied-piping among languages with compositional indeterminate pronouns. Notice that the Edge Generalization (21), even if it is about overt edges, does not contribute to this issue, since compositional indeterminate pronouns in Slavic can be located at the edge of a potential pied-pipee. In fact, Slavic languages are multiple wh-fronting languages just like Basque, in which all indeterminate pronouns overtly move to the edge of the C-domain that could potentially be pied-piped. Likewise, the difference between overt vs. covert wh-movement does not give us a right cut, since Basque and Latin are multiple-fronting languages, while Japanese is a wh-in-situ language. The presence/absence of affixal articles is not relevant either, since Basque is an affixal article language and Latin is an article-less language. What can, then, be a factor that differentiates languages that allow large-scale pied-piping and those that do not?

I suggest that it is the canonical word order that matters here. Crucially, the languages that have large-scale pied-piping (i.e., Basque, Imbabura Quechua, Latin, Bengali, Old and Present Japanese) are canonically SOV languages. In contrast, those that do not have it (e.g., Slavic) are all SVO languages. Thus, I revise Watanabe's generalization (9) as (22).[6]

---

6. Note that this is also a one-way correlation; that is, there can be languages that have indeterminate pronouns and SOV word order but do not allow large-scale pied-piping. What is important here is that no language that canonically

(22) Languages that allow large-scale pied-piping have productive compositional indeterminate pronouns and the SOV canonical word order.

The SOV order is captured by head-finality in theoretical syntax. Crucially, the clauses that are pied-piped in the relevant languages are also head-final. Thus, the generalization (22) can be further refined in the following way:

(23) Large-scale pied-piping is possible in a language only if the language has productive compositional indeterminate pronouns and the projection to be pied-piped is head-final.

The next question is how this new generalization can be deduced. In particular, why does the head-finality matter for the availability of large-scale pied-piping? I address this issue in the next section.

4.3   Deduction of the new generalization

In this section, I offer a deduction of the generalization (22), based on the proposal from chapter 3 that compositional indeterminate pronouns have an unvalued uninterpretable operator feature. I propose that a feature-percolation analysis of large-scale pied-piping, which was the standard analysis in the GB theory but has been claimed to have no theoretical status in minimalism, can be revived and implemented under Chomsky's (2015) labeling framework. I then introduce Inaba's (2011) typological work, which shows that the head-directionality correlates with the morpho-phonological status of complementizers. Based on this, I propose that there is a correlation between the syntactic and morpho-phonological status of heads, which I argue is responsible for the large-scale pied-piping being available only in head-final clauses.

4.3.1   Revival of feature percolation under the labeling theory

The more or less standard line of analysis of large-scale pied-piping in the literature is a feature-percolation analysis. Under this analysis, an operator feature of an indeterminate pronoun at the

has the SVO order allows large-scale pied-piping even if it has compositional indeterminate pronouns.

specifier of XP to be pied-piped is *percolated* into XP, as a result of which XP moves to the licensing position of indeterminate pronouns (e.g., Aissen 1996, Cowper 1987, Grimshaw 2000, Moritz and Valois 1994, Nishigauchi 1990, Ortiz de Urbina 1989, Webelhuth 1992, among many others; see Heck 2008 for more references). This is schematized in (24); the operator feature of an indeterminate pronoun in Spec,XP percolates to the XP (24a) and the entire XP moves to Spec,CP instead of the indeterminate pronoun (24b), where the operator feature is valued as [Q].

(24) a.

b.
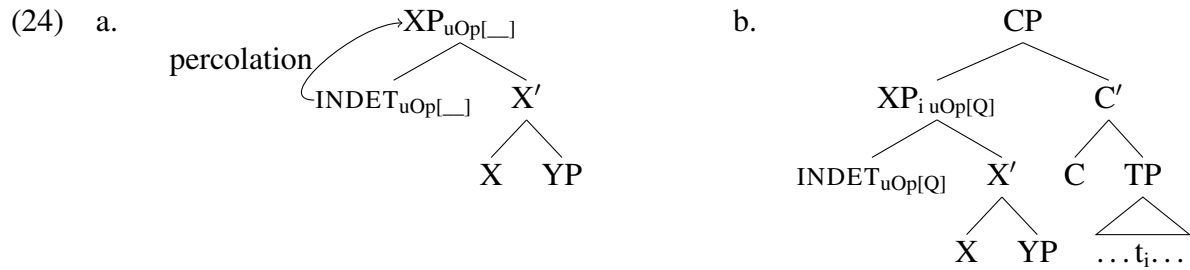
The feature percolation illustrated in (24) has been argued to have no theoretical status in the current minimalism, and in fact it is difficult to implement it under the current framework; see Heck (2008, 2009) for extensive discussion of this issue. Cable (2007, 2010) also argues that pied-piping by means of feature percolation as in (24) should be eliminated from the syntactic theory. In this subsection, however, I show that feature percolation, hence pied-piping, can actually be implemented under Chomsky's (2015) labeling framework.

Before proceeding, I would like to clarify properties of large-scale pied-piping in the context of the current proposal regarding compositional indeterminate pronouns. In chapter 3, I proposed that compositional indeterminate pronouns have an unvalued uninterpretable operator feature, which drive movement until it is valued by probing down a goal (Bošković 2007b). In large-scale pied-piping cases like (5), repeated here as (25), the compositional indeterminate pronoun *nor* moves to the edge of the embedded CP because of the unvalued operator feature.

(25) [$_{CP}$ **Nor** joango d-ela]$_i$ esan du Jon-ek t$_i$?
      who go AUX-C say AUX John-ERG

      'Who has John said will go?' (Basque: Ortiz de Urbina 1989:248)

In (26), I list three important aspects of large-scale pied-piping that need to be captured.

(26)   a.   First, the indeterminate pronoun *nor* lands in the embedded Spec,CP; it does not directly move to the matrix Spec,CP. Instead, the entire embedded CP moves to the matrix Spec,CP. This means that this embedded CP has the relevant unvalued operator feature.

     b.   Second, the operator feature of the indeterminate pronoun should not be valued at the embedded Spec,CP, because receiving a value in the embedded clause means that the indeterminate pronoun would take scope in the embedded CP (just as in indirect questions).

     c.   Third, the operator feature of the indeterminate pronoun should be valued once the embedded CP moves to the matrix Spec,CP and has its operator feature valued. Otherwise, the operator feature of the indeterminate pronoun would remain unvalued at the end of the derivation, resulting in ill-formedness.

In order to address (26a) and (26b), I propose that a head that has an unvalued uninterpretable operator feature, which I will simply call F for ease of exposition, merges with the head of the projection to be pied-piped, e.g., the embedded CP in (25).[7] I suggest that this merger is a head-to-head adjunction as base-generation, which is implemented as External Pair-Merge in the current syntactic theory as discussed by Epstein et al.'s (2016). Then, External Pair-Merge of F with C creates a $\langle C, F \rangle$ amalgam as a complex head. Essentially, External Pair-Merge of F with C adds an unvalued operator feature to C, so that the entire CP is marked as having the operator feature, which captures (26a) (see below for a more precise implementation). At the same time, the operator feature of the $\langle C, F \rangle$ amalgam is unvalued, so the operator feature of the indeterminate pronoun in the specifier of $\langle C, F \rangle$ is not valued in this position. This captures (26b).

Turning to (26c), I suggest that this F is a weak head in Chomsky's (2015) sense, which requires feature sharing with a head of a phrase in its specifier. In Chomsky's labeling framework,

7. It is worth mentioning that Watanabe (1992b) proposes that an operator that functions as WH is merged at the edge of a pied-pipee, which is similar to what I propose here (see also Tsai 1994, 1999). See, however, footnote 21 for discussion of an advantage of the current proposal over the Watanabe-Tsai analyses.

labeling of a syntactic object is not an inherent property of the syntactic object or part of Merge, unlike in earlier syntactic theories (see also Collins 2002 and Seely 2006), but it is determined by the operation *Minimal Search*, which searches the syntactic structure from the highest node. For instance, when a head X and a phrase YP are merged, X projects and provides the label for the configuration, because it is the first head that Minimal Search finds. However, Chomsky (2015) proposes that T in English is a "weak" head, in the sense that it cannot provide a label on its own. In order for the configuration {T, vP} to be labeled, a DP first moves to the specifier of T. Then, Minimal Search finds two heads, namely, D and T. In this configuration, minimal search finds the same feature(s) on the two heads (in this case ϕ-features), and D and T undergo *feature-sharing*, by which the unvalued ϕ-features of T are valued. As a result, the prominent shared features, in this case ϕ-features, project and provide the label for the topmost node as $\langle \phi, \phi \rangle$, and {T, vP} is labeled as TP because T is strengthened by feature-sharing. This is schematized in (27).[8]

(27)   a.
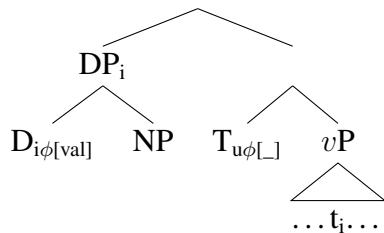
b.
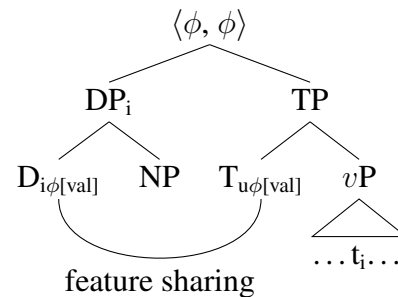
I propose that External Pair-Merge of a weak head H to a non-weak head H' makes the resulting amalgam a complex weak head (see also section 4.3.2 for more discussion on this). Then, the $\langle$C, F$\rangle$ amalgam, which contains the weak head F, is a complex weak head that requires a syntactic object with an operator feature in its specifier position. The relevant derivation is given in (28). In (28a), F Externally Pair-Merges with C, which yields the $\langle$C, F$\rangle$ amalgam. This amalgam is then Externally Set-Merged with TP. Then, the indeterminate pronoun, which is NP (see chapter 3 on this), undergoes movement to the specifier position of the $\langle$C, F$\rangle$ amalgam, which is driven by

---

8. Chomsky argues that this deduces the EPP effect of English finite clauses; namely, T requires a specifier for the labeling reason. See also footnote 15 for more discussion.

137

the unvalued operator feature of N. As shown in (28b), the indeterminate NP and the ⟨C, F⟩ amalgam undergo feature sharing for the unvalued uninterpretable operator features, because Minimal Search finds N and the ⟨C, F⟩ amalgam as the first heads (note that ⟨C, F⟩ amalgam is a complex head which counts as a single unit). The shared features project and provide the label for the highest node as ⟨uOp_[_], uOp_[_]⟩, and ⟨C, F⟩ projects above TP and the ⟨C, F⟩ amalgam, because it is strengthened by feature sharing, on a par with T in (27). Notice now that the entire clause is marked as having the unvalued operator features, which triggers movement of this clause. Thus, the current proposal captures the property of large-scale pied-piping in (26a).

(28)   a.
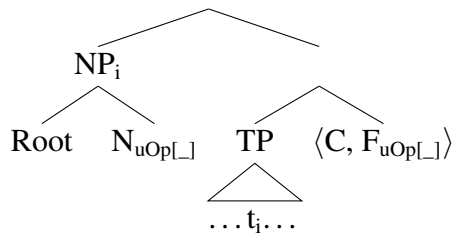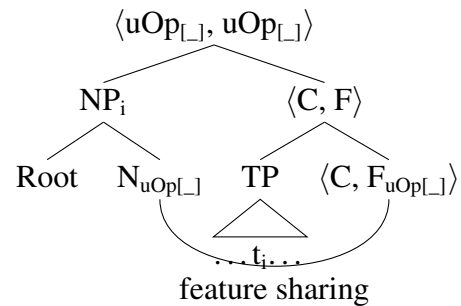
When this embedded clause, which is now labeled as ⟨uOp_[_], uOp_[_]⟩, moves to the specifier of interrogative C, Minimal Search finds N, ⟨C, F⟩, and C.[9] Here I assume that interrogative C has a valued interpretable operator feature iOp_[Q]. Thus, N and ⟨C, F⟩ share the operator feature with C and the unvalued features of N and ⟨C, F⟩ are valued as [Q] at the same time, as illustrated in (29). Notice now that this captures (26c); namely, the operator feature of the compositional indeterminate pronoun is valued when the operator feature of the embedded CP (i.e., the ⟨C, F⟩ amalgam) is valued.

---

9. Under Epstein et al.'s (2020) path-based theory of Minimal Search, which I discuss in section 4.4, N, ⟨C, F⟩, and C are the first heads that Minimal Search finds, and hence they undergo feature sharing. See section 4.4 for details of Epstein et al.'s theory.

(29)

$$\langle Op_{[Q]}, Op_{[Q]} \rangle_j$$

$$\langle uOp_{[Q]}, uOp_{[Q]} \rangle_j \qquad CP$$

$$NP_i \qquad \langle C, F \rangle \qquad TP \quad C_{iOp[Q]}$$

$$Root \quad N_{uOp[Q]} \qquad TP \quad \langle C, F_{uOp[Q]} \rangle \qquad \ldots t_j \ldots$$

$$\ldots t_i \ldots$$

feature sharing and valuation
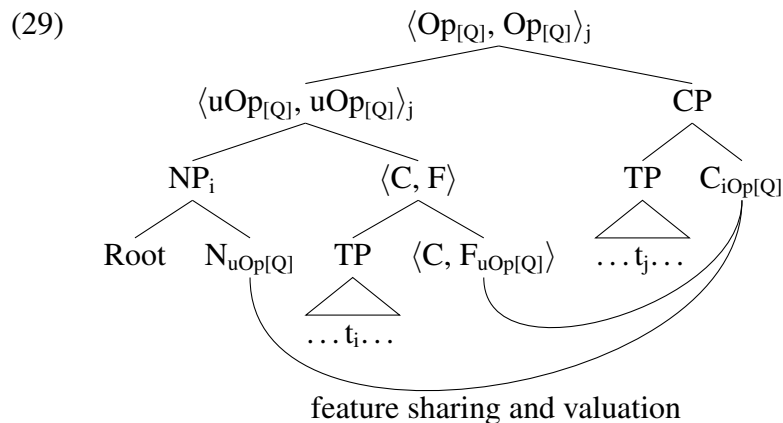
Thus, the current proposal captures the three properties of large-scale pied-piping stated in (26). At the same time, this analysis also captures the Edge Generalization in (21); a compositional indeterminate pronoun is necessarily located in the specifier of the (complex) head of the pied-pipee in order to undergo feature sharing of an operator feature.[10]

It should be added here that this proposal captures both overt large-scale pied-piping found in, e.g., Basque and covert large-scale pied-piping found in, e.g., Japanese. The relevant operations for pied-piping illustrated in (28) take place in narrow syntax, and which copies of the indeterminate pronoun and the pied-pipee are pronounced at PF depends on whether those phrases have an unvalued uninterpretable focus feature or not (Watanabe 2002; see chapter 3). We can hypothesize that in languages in which a pied-pipee moves to the matrix Spec,CP, F has both an unvalued uninterpretable operator feature and an unvalued uninterpretable focus feature. This can be supported by Old Japanese, where *ka* is used in wh-questions that involve wh-fronting as well as with focused phrases, as shown in (30), and Imbabura Quechua, where the particle *taj* found in large-scale pied-piping is also used in focused phrases (that express exclusivity), as shown in (31).

(30) a. Kono toki-fa [ika-ni si-tutu]-**ka** na-ga yo-fa wataru?
this time-TOP how-DAT do-while-KA you-NOM world-TOP pass
'At this time, you pass through this world doing what?'

(Man'youshuu #892, Aldridge 2009)

---

10. For similar ideas of determination of the nature of a projection by feature sharing in other configurations, see, e.g., Baker and Stewart (1999), Moro (2000), Pereltsvaig (2008), Citko (2011).

b. ...[Hatsuse-no   kawa-ha   ura   na-mi]-**ka**   fune-no   yori-ko-nu?

   Hatsuse-GEN river-TOP shore absent-ness-KA boat-NOM approach-come-NEG

   'Is it because Hatsuse River has no shore that no boat comes near?'

<div align="right">(Man'youshuu #3225, Watanabe 2002:183)</div>

(31) a. [Ima-ta   Juan randi-shka]-ta-**taj**$_i$ ya-ngui t$_i$?

       what-ACC Juan bought-C-ACC-TAJ think-2

       'What do you think that Juan bought?' (Cole 1982:21)

   b. Chay-ta-**taj**   muna-ni

      that-ACC-TAJ want-1

      'I want that very one.' (Cole 1982:167)

In such languages, both the operator feature and the focus feature project after feature-sharing, and both features are valued in the matrix Spec,CP.

Note now that the proposed system of large-scale pied-piping correctly predicts that large-scale pied-piping is only possible in languages with productive compositional indeterminate pronouns. In chapter 3, I proposed, following Saito (2017), that compositional indeterminate pronouns bear an unvalued uninterpretable operator feature, whereas bare indeterminate pronouns lack it. Bare indeterminate pronouns thus lack a driving force of movement to the specifier of the ⟨C, F⟩ amalgam, as schematized in (32a). In addition, the above proposal crucially involves feature sharing of operator features. Since bare indeterminate pronouns lack the relevant feature, they cannot undergo feature sharing with the ⟨C, F⟩ amalgam, even if they are located in the specifier of the amalgam, as illustrated in (32b). Thus, a bare indeterminate pronoun would not pied-pipe the entire embedded CP to the matrix Spec,CP.

(32) a.

$$(CP)$$
$$(\langle C, F \rangle)$$
$$\langle C, F \rangle \quad TP$$
*movement $\cdots NP_{Indet} \cdots$

b.

$$(CP)$$
$$NP \quad (\langle C, F \rangle)$$
$$N \quad Root \quad \langle C, F_{uOp[\_]} \rangle \quad TP$$
*feature sharing
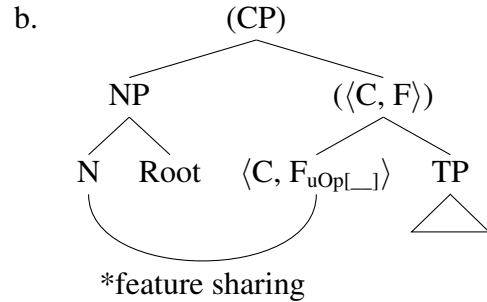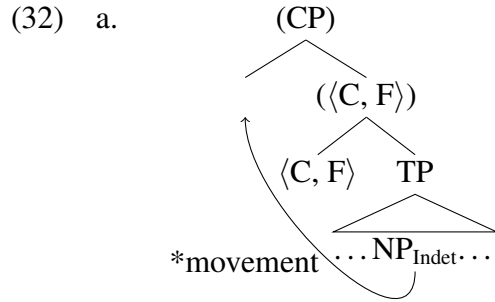
Furthermore, the proposed mechanism of large-scale pied-piping is not possible in non-affixal article languages, where DP that bears a valued interpretable operator feature projects above indeterminate NPs and therefore productive compositional indeterminate pronouns are not allowed (see chapter 3). Even if F Externally Pair-Merges with C, its unvalued operator feature would be valued by the valued operator feature of the D in question via feature sharing. The entire clause would then be labeled as $\langle Op_{[Q]}, Op_{[Q]} \rangle$, which means that the clause would be interpreted as an indirect question.

(33)

$$\langle Op_{[Q]}, Op_{[Q]} \rangle$$
$$DP \quad \langle C, F \rangle$$
$$D_{iOp[Q]} \quad NP \quad \langle C, F_{uOp[Q]} \rangle \quad TP$$
valuation
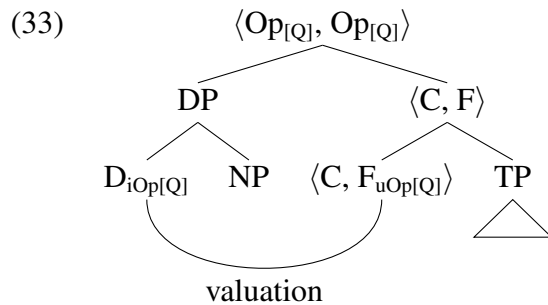
Thus, the proposed analysis of large-scale pied-piping captures the generalization that large-scale pied-piping is allowed only in languages that have productive compositional indeterminate pronouns.

4.3.2 Weak heads as bound morphemes and head-final complementizers

A remaining question is how to explain the second part of the generalization (9), namely, the relevance of head-finality for large-scale pied-piping. Here I propose that weakness of heads interacts with the morpho-syntactic nature of complementizers in a given language, which correlates with

head-finality. The gist of the proposal is that weak heads in Chomsky's (2015) sense are phonologically weak as well, and that only head-final complementizers can be weak heads because they are phonologically weak, in contrast to head-initial complementizers.

Under the traditional Head Parameter, the choice of parameter value (i.e., head-initial or head-final) is arbitrary, and has nothing to do with the morpho-syntax of heads. However, Inaba (2009, 2011) observes an interesting correlation between headedness and morpho-syntax of heads in certain domains. Inaba (2009) points out that initial subordinators are often "general" complementizers in the sense that they can be used to introduce not only complement clauses of verbs but also relative clauses, adverbial clauses, and so forth. In addition, Inaba (2011) notes that they are generally independent words, i.e., free morphemes.[11] A prototypical example is English *that*, as shown in (34). The same observation can also be made for Twi, a Kwa language, as shown in (35).

(34) a. I think [that John is sleeping].

b. the people [that I have never met]

c. He must be crazy [that he should go out now].

(35) a. na ama nim [sɛ kofi yɛɛ adwuma no].
PAST Ama know C Kofi did work the
'Ama knows that Kofi had done the work.'

b. kofi yɛɛ adwuma no [sɛ yaw bɛpɛ n'asɛm].
Kofi did work the C Yaw FUT.like his.manner
'Kofi did the work so that Yaw would like him.'

c. [sɛ kofi yɛ adwuma no â] metua no ka.
C Kofi do work the COND I.FUT.pay him salary
'If/When Kofi does the work, I will pay him.' (Lord 1993)

---

11. It should be noted that this is a statistical universal, i.e., (strong) tendency. See Schachter and Otanes (1972), Kroeger (1993), Richards (1999) for the head-initial complementizer *-ng* in Tagalog, which is suffixed/cliticized onto its preceding word.

In contrast, Inaba (2009) points out that final subordinators usually cannot be used as general complementizers in the above sense, and that they are better analyzed as nominalizers/case-markers or as citation markers.[12] Thus, Uzbek uses a nominalizer to introduce a clausal complement, and Kanuri uses a case-marker for the relevant function.

(36) Men [bu ɔdam-miŋ ǰoǰa-ni   oǧirla-<u>gan</u>-i-ni]   bilaman.

      I   this man-GEN chicken-OBJ steal-NOMN-3SG.POSS-OBJ know.1SG

      'I know that this man stole the chicken.'   (Noonan 2007:96)

(37) [Sá'vá-'nyí íshín-<u>rò</u>]   tɔmăŋɔ́nà.

      friend-my comes-DAT thought.1SG.PERF

      'I thought my friend would come.'   (Noonan 2007:57)

Japanese *to* and Bengali *bole*, which are used as complementizers, are citation markers in the sense that they are used to introduce a complement clause for quotative verbs such as 'say' and 'think', which quote speech or thought of the subject. Thus, these items cannot be used to introduce a complement clause of non-quotative verbs like 'know' and 'see', unlike English *that*. This means that these items are not real counterparts of English *that*, which can be used for the complement clause of quotative and non-quotative verbs (as well as for other usages).

(38) Boku-wa [kare-ga kuru <u>to</u>] omou/*shiranai.

      I-TOP   he-NOM come C   think/not.know

      'I think/don't know that he will come.'   (Japanese, Inaba 2009)

(39) [ram kolkata-y   jacche <u>bole</u>] bhablam/*dekhlam.

      Ram Calcutta-LOC goes   C   I.thought/I.saw

      'I thought/saw that Ram is going to Culcutta.'   (Bengali, Bayer 2001:15)

Interestingly, Inaba (2011) notes that final subordinators are usually particles or suffixes, that is,

---

12. Note that this is a statistical universal, i.e., (strong) tendency; Inaba (2009) acknowledges that the head-final complementizer *(e)n* in Basque behaves like a general complementizer.

bound morphemes. He also points out that the same pattern has actually been observed in other domains in the literature. For instance, Andrews (2007) claims that initial relativizers such as English *that* are generally genuine clausal complementizers, whereas final relativizers typically appear as part of the verbal morphology of the relative clause (or they have no overt form at all). Inaba takes this as indicating that clause-final functional elements are generally affixal.[13]

Turning to weak heads, in Chomsky (2015), they are weak only in a syntactic sense, that is, they cannot provide a label on their own. However, it is logically possible that syntactic properties can be reflected in the PF realization. It is also worth noting here that Chomsky suggests that labels are required for syntactic objects to be interpreted at the C-I and A-P interfaces. We can, then, hypothesize that weak heads are weak not just in the syntactic sense, but also in some interface sense. Now, let us look at weak heads proposed in the literature. Chomsky (2015) proposes that R(oot) and English T are weak heads, as seen above. Saito (2018) proposes that case suffixes in Japanese, which he analyzes as K(ase), are also weak heads. Interestingly, these elements are not free morphemes; roots and affixes.[14] The F head I have proposed above can be also considered to be a phonologically null affix that is adjoined to the head of a projection to be pied-piped. Thus, I propose the following morpho-syntactic condition:[15]

---

13. Inaba's generalization regarding the morphological nature of complementizers is a statistical universal (i.e., strong tendency; cf. footnote 11). See footnote 19 on a note regarding the generalization in (23) in this respect.

14. These elements are not prosodic words (i.e., free) on their own; e.g., *cant-a-re* 'sing' in Italian, where none of the root *cant*, the thematic vowel *a*, and the infinitive suffix *re* are prosodic words on their own. As for English, verbs such as *sing*, which are often assumed to be roots and free morphemes, are analyzed as consisting of a Root and a verbalizer *v* in the Distributed Morphology framework. Accordingly, *sing* is a realization of Root + *v*, not Root itself, and hence Root itself is not a prosodic word on its own (cf. *song*, which is a realization of the relevant Root and *n*).

15. Note that (40) is a one-way correlation; namely, not all non-prosodic morphemes are weak heads. Chomsky (2015) in fact suggests that Italian T, which is a bound morpheme, is a strong head (i.e., a head that can provide a label on its own) in contrast with English T.

However, it may actually be not impossible to strengthen (40), and hypothesize that weak heads are realized as non-prosodic morphemes and non-prosodic morphemes are weak heads in the relevant sense. As mentioned in footnote 8, Chomsky attributes the EPP effect in English to the weakness of T, i.e., T in English requires an overt specifier for a labeling reason. In contrast, Italian T does not require an overt specifier, and hence Chomsky suggests that Italian T can provide a label on its own, i.e., T in Italian is strong. There are actually two cases in which T in Italian does not have an overt specifier: a null subject and a post-verbal subject. A null subject in Italian has been standardly analyzed as *pro*, which is considered to be a phonologically null counterpart of an overt pronoun. Thus, we can assume that *pro* is actually present in narrow syntax and located in Spec,TP in Italian, *pace* Chomsky, who suggests that *pro* is simply absence of a nominal element in narrow syntax. As for post-verbal subjects, it is observed that they are focalized (see, e.g., Belletti 2001, 2004). It is worth noting here that in chapter 3 (see also section 4.3.1) I proposed following Watanabe (2002) that realization of copies of indeterminate pronouns in a chain is dependent on a focus feature on them; if they have a focus feature, the highest copy is pronounced (i.e., wh-fronting), but if they lack it, the lowest copy

(40) Weak heads cannot be realized as free morphemes, i.e., they are morpho-phonologically
      weak.

The intuition here is clear; syntactically weak elements are morpho-phonologically weak as well. This is also compatible with the Distributed Morphology framework (e.g., Halle and Marantz 1993a), under which relations among syntactic objects established in narrow syntax are reflected in morpho-phonology.

The next question is what counts as a weak head (i.e., a head that cannot provide a label on its own). Note that the above mentioned weak heads, T in English, K in Japanese, and F, contain an unvalued feature. T has unvalued ϕ-features, K has an unvalued Case-feature, and F has an unvalued operator feature. As for R, it is natural to assume that it has an unvalued categorial feature uCat$_{[\,]}$, which is valued by a categorizer such as v, n.[16] Based on this, I propose the criterion for weak heads in (41).[17]

(41) Weak heads are heads that bear an unvalued feature when they enter into the syntactic deriva-
      tion (i.e., at the point of External Merge).

Note that the above formulation allows an amalgam of heads that is created by External Pair-Merge to also count as a weak head. Thus, the ⟨C, F⟩ amalgam in (28) is a weak head under this formulation of weak heads, since F has an unvalued operator feature and the amalgam also has this feature at the point of External (Pair-)Merge.[18]

---

is pronounced (i.e., wh-in-situ). One possible extension of this to post-verbal subjects in Italian would, then, be that the subject DP always moves to Spec,TP in Italian in narrow syntax just as in English but the lowest copy is pronounced when the DP has a focus feature. Stjepanović (1999, 2003) in fact proposes that interaction of the Nuclear Stress Rule and focus requires the lower copy of the subject to be pronounced at PF in Italian as well as in Serbo-Croatian. Italian T can, then, be analyzed as a weak head, which requires a valuer of its unvalued ϕ-features in its specifier position, just like English T. Thus, it is not implausible to strengthen (40) as a two-way correlation. A full investigation of this issue is left for future research. (See also footnote 18 regarding an issue of Case.)

16. See also chapter 2, where I proposed that the head of the coordinate structure, Conj, has an unvalued categorial feature which is valued by its conjuncts.

17. Given (41), Italian T, which Chomsky (2015) suggests is a strong head, should actually be a weak head, since it has unvalued ϕ-features. See footnote 15 for discussion.

18. An issue arises regarding treatment of Case in English in connection with the morpho-syntactic nature of D in this language. English D has been more or less standardly assumed to bear an unvalued Case feature, the view I have also adopted here. Given (41), then, English D should not be a prosodic word on its own. One might then wonder if English would be considered to be an affixal article language, although English behaves differently from proto-typical affixal article languages such as Bulgarian in a number of respects discussed in this dissertation (see chapter 2 on the coordinate structure and chapter 3 on the compositional indeterminate pronouns).

Notice now that this theory of weak heads enables us to capture the relevance of head-finality for large-scale pied-piping, in combination with Inaba's (2011) observation. According to Inaba, head-initial complementizers are generally independent words, i.e., free morphemes. However, if F Externally Pair-Merges with a head-initial C, (40) requires the ⟨C, F⟩ amalgam to be realized as a bound morpheme. Under the Distributed Morphology framework, in which morpho-phonological forms of lexical items are determined based on syntactic information (i.e., Vocabulary Insertion takes place after the syntactic computation), there is no vocabulary item for head-initial complementizers that can be realized as a bound morpheme. This can actually be restated under the lexicalist view, where the form of syntactic objects is already specified in the numeration. For instance, English *that*, which itself is a free morpheme, enters into narrow syntactic computation, where it is Externally Pair-Merged with F and marked as a weak head as per (41). At PF, however, the condition (40) is not met since *that* is a free morpheme. Thus, both the lexicalist and non-lexicalist views are compatible with the deduction discussed here. In contrast, head-final complementizers are usually bound morphemes, so that the ⟨C, F⟩ amalgam in head-final clauses can

One possibility to differentiate English from proty-typical affixal article languages is that what matters here may be how "weak" phonologically the relevant element is, i.e., the relevant phonological weakness may not simply be a binary distinction. In fact, English *the* can be stressed (i.e., [ði:]) and used as a prosodic word, while Bulgarian definite articles can never be prosodic words. Also, "pickiness" regarding the morphological host of the definite article may be relevant. English D (when weak) can be hosted by any element following it, including a noun, an adjective, and an adverb, but Bulgarian D cannot be hosted by an adverb (see, e.g., Halpern 1995, Franks and King 2000, Franks 2001 for Bulgarian; see also chapter 5 for discussion of Italian definite articles from a syntactic perspective).

Another possibility could be that the English D, being a weak head, still needs to always be present in the nominal domain and projects DP (after strengthening), unlike Ds in languages like Bulgarian where DP can be omitted, as Talić (2015, 2017) argues (see also chapter 5, where I propose that there are options for introducing definite articles in the syntactic structure other than D projecting DP under the Bare Phrase Structure Theory). In other words, there is a one-way correlation between the morphological and syntactic nature of D; Ds that need not project DP are necessarily weak heads that are not prosodic words on their own, but not all weak D heads can be those that need not project DP. (An issue of phasehood of DP also arises under this treatment of English D, since D would not project DP until its Case feature is valued via Minimal Search/feature sharing. See Appendix for discussion of this issue.)

Yet another possibility may be that Case is exempt from the criterion for weak heads in (41). An intuition behind this is that Case is unique among formal features. It exists for licensing of nominals, as the traditional Case Filter essentially states, and it is the only obligatory formal feature in the nominal domain across languages, i.e., there is no variation with respect to the presence/absence of Case, and it is always present in any nominal phrase in any language, unlike other formal features, whose presence/absence can vary within and across languages (see chapter 3 and chapter 5 for relevant discussion). In addition, under the standard assumptions, there is no interpretable counterpart of Case, unlike e.g., ϕ-features, and there is no exact counterpart of it in the clausal spine (it is standardly assumed that it is valued against T, v, etc., but those heads do not bear an interpretable Case feature). If the unvalued formal features relevant for the criterion for weak heads in (41) are limited to those that can vary within and across languages and have relevance for interpretation, the Case feature, which is universally obligatory and exists only for licensing of nominals, may be exempt from the criterion in question.

satisfy the PF requirement (40), and the derivation for large-scale pied-piping as in (28) is possible in such cases. Thus, the relevance of head-finality for large-scale pied-piping in the generalization (23) can be captured by the interaction of the morpho-syntactic requirement on weak heads and available morphological forms of complementizers that depend on whether they are head-initial or head-final.[19]

The analysis can be straightforwardly extended to large-scale pied-piping of islands. Let us look at Old Japanese (42), repeated from (11).

(42) Kono toki-fa [adjunct island ika-ni   si-tutu]-ka na-ga   yo-fa   wataru?
      this time-TOP   how-DAT do-while-KA you-NOM world-TOP pass
      'At this time, you pass through this world doing what?'

(Man'youshuu #892, Aldridge 2009:560)

Notice that *tutu* 'while', which heads the adjunct island, is a bound morpheme affixed to the verb *si* 'do'. Thus, if F (*ka* here) Externally Pair-Merges with *tutu* (which I assume to be C for ease of exposition), the ⟨C, F⟩ amalgam satisfies the condition (41). This amalgam undergoes feature sharing with the compositional indeterminate pronoun *ika(-ni)* and the adjunct clause is pied-piped, in the same way as complement clauses are pied-piped.

Relative clauses involve a more complex derivation. Let us consider Imbabura Quechua (43), repeated from (8).

---

19. There is actually room for exceptions to the generalization (23). Inaba's generalization regarding the correlation of head-directionality and the morpho-phonological status is a statistical universal, which means that there can be a language that has an affixal head-initial complementizer (though such cases are rare; see footnote 11 on Tagalog). Under the deduction of (23) offered here, if a complementizer Externally Pair-Merges with F, which bears the unvalued operator feature, the amalgam counts as a weak head as per (41), which in turn requires that the amalgam be a bound morpheme, as stated in (40). To put it differently, if the complementizer is a bound morpheme, it can Externally Pair-Merge with F and satisfy the condition (40), whether it is head-initial or head-final. It is, then, predicted that large-scale pied-piping should be in principle possible in languages that have a head-initial affixal non-interrogative complementizer and productive compositional pronouns, since such a head-initial complementizer meets the condition (40) if F Externally Pair-Merges with it. I leave investigation of this prediction for future research (note that large-scale pied-piping is a rather rare phenomenon, and only a small set of languages have been tested in this respect in the literature.)

(43)   [[Ima-ta   randi-shka] runa]-ta-taj   riku-rka-ngui?

       what-ACC buy-C   man-ACC-Q see-PAST-2

       'What did you see the man who bought (it)?'   (Imbabura Quechua: Cole 1982:24)

In (43), the relative complementizer *shka* is affixed to the verb *randi* 'buy'. Thus, if F Externally Pair-Merges with *shka*, the resulting amalgam ⟨C, F⟩ satisfies the condition (41). Note here that what is pied-piped is the entire nominal phrase, not just the relative clause. Under the current proposal, this means that the head of the nominal phrase, which I assume to be the K(ase) head *-ta*, also has an operator feature, and this operator feature is shared with the relative clause as well as the compositional indeterminate pronoun embedded in the relative clause. The presence of the operator feature on K is implemented by External Pair-Merge of F with K, which observes the condition (41), since the K head *ta* is a bound morpheme (here F is realized as *taj*). As for the position of the relative clause, I assume, following Cinque (2013), that relative clauses are hosted in the specifier of a functional projection αP that is located between N and K. The relative clause then moves to the specifier of the ⟨K, F⟩ amalgam, and they undergo feature sharing of the operator features. The relevant structures are given in (44).[20]

20. A question that arises regarding αP is how its label is exactly determined under the labeling framework adopted here. Specifically, the αP level in (44) would actually be an {XP, YP} structure (i.e., {NP, αP}), so this level should be labeled by feature sharing. The question is, then, what feature would be shared here. One possibility that I would like to suggest here is that α has the categorial feature [+N], given that it is part of the extended projections in the nominal domain (cf. Grimshaw 2000, Bošković 2014). This in turn means that F also bears the [+N] feature in the case under discussion. F and α would then undergo feature sharing of this [+N] feature, and the αP node in (44) would actually be ⟨+N, +N⟩ (cf. Baker and Stewart 1999, Citko 2011).

    One advantage of this suggestion is that it can be extended to capture the observation noted in footnote 4 that the reason indeterminate pronoun *naze* 'why' in Japanese cannot pied-pipe an island, including relative clauses. In chapter 3, I suggested that reason indeterminate pronouns are PPs rather than NPs. Under the feature-based classification of traditional lexical categories (Chomsky 1970), P is [-N, -V]. Given that F is [+N] as suggested above, the reason indeterminate pronouns cannot undergo feature sharing of the categorial feature, hence they cannot pied-pipe an island unlike other indeterminate pronouns, which are analyzed as NPs (see also chapter 6 for discussion of the relation between P and N).

(44)   a.
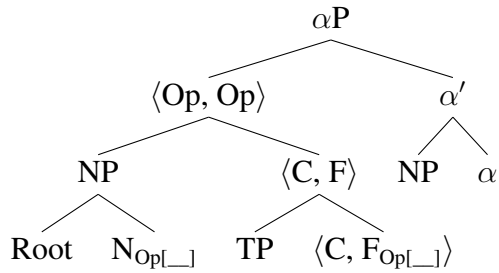
b.
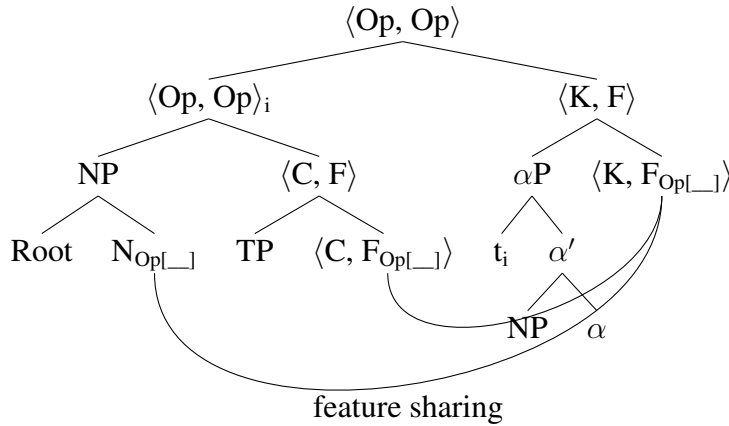
Notice now that the entire nominal phrase is pied-piped by the compositional indeterminate pro-noun in a "roll-up" manner; the compositional indeterminate pronoun pied-pipes the relative clause, and the relative clause pied-pipes the nominal phrase. Actually, Heck (2008, 2009) argues that this is a general property of pied-piping, as stated in (45), where a canonical position amounts to the edge of a pied-pipee.

(45)   *Generalization on recursive pied-piping*

If α can pied-pipe β, and β is in a canonical position to pied-pipe γ, then α can also pied-pipe γ.

This generalization follows from the current proposal. When α pied-pipes β, α is at the edge of β and they undergo feature sharing of an operator feature. This β then undergoes feature sharing of an operator feature with γ and pied-pipes γ when β is at the edge of γ. This is the configuration of pied-piping of γ by α via β. Thus, the current proposal captures the general property of recursive

pied-piping.[21,22]

It should be added here that this explanation of the generalization (23) leaves room for availability of pied-piping on a smaller scale in some domains in head-initial languages. For instance, possessor indeterminate pronouns in general are located at the edge of the nominal phrase, and they are accompanied by a possessor marking such as genitive case, which is generally a bound morpheme. When F is Externally Pair-Merged with the head H that hosts a possessor in its specifier, and the $\langle$F, H$\rangle$ amalgam is forced to be realized as a bound morpheme, which indeed satisfies the requirement (40). The possessor and F undergo feature-sharing of operator features, and the shared operator features drive movement of the entire phrase (or $\langle$Op, Op$\rangle$) to a licensing position of indeterminate pronouns. This is even possible in languages that have non-affixal definite articles and lack productive compositional indeterminate pronouns. Let us consider English (46) as an instance.

(46)   [**Whose** book] did you read?

Whose can be analyzed as consisting of *who* + the genitive marking $/z/$. I assume that *who* is located in Spec,DP and $/z/$ is the head of the DP (see also chapter 6 on an analysis of the genitive

---

21. The current proposal can also explain the impossibility of indeterminate pronouns being interpreted in the matrix clause in (i), where the indeterminate pronouns are embedded in a wh-island that is embedded in a relative clause.

(i) Taro-wa   [[[Hanako-ga   doko-de   nani-o   katta   ka] shitteiru] hito]-ni   atta no?
     Taro-TOP   Hanako-NOM where-LOC what-ACC bought Q   know   person-DAT saw   Q
     'Did Taro see a person who knows where Hanako bought what?'
     NOT 'For which x, x a thing, did Taro saw a person who knows where Hanako bought x?
     NOT 'For which x, x a place, did Taro saw a person who knows what Hanako bought in x?'

Under the current proposal, the operator feature of the indeterminate pronouns is given the value [Q] by *ka* and interpreted in the most embedded clause. Thus, neither *doko* 'where' nor *nani* 'what' can undergo movement to the edge of αP in order to pied-pipe the relative clause. It is worth noting here that the observation in (i) cannot be accounted for by Watanabe's (1992b) proposal, in which an operator that functions as WH can be merged at the edge of a pied-pipee regardless of the position of the indeterminate pronoun. Tsai (1994, 1999) offers a proposal similar to Watanabe's, where an operator, which is responsible for the interpretation of wh-questions, can be merged at the edge of DP, i.e., the relative clause in the case of (i). Their proposals incorrectly predict that the indeterminate pronouns in (i) could be interpreted in the matrix clause, because the relevant operator can merge at the edge of the relative clause and pied-pipe it in (i). Thus, the current proposal is favored over those by Watanabe and Tsai.

22. The current proposal has a potential to offer a principled explanation of head-finality. Kayne (1994) proposes that the default word order is universally head-initial, and the head-final order is derived by roll-up movement of the complement of relevant heads. Biberauer et al. (2014) propose that this roll-up movement is triggered by EPP. Recall that Chomsky (2015) deduces EPP of English T from the labeling requirement of a weak head. It may then not be implausible to hypothesize that the roll-up movement that derives the head-finality is triggered by weak heads, which are realized as suffixes. I leave full investigation of this topic for future research.

marker in English). As discussed in chapter 3, English is a non-affixal article language where DP projects above indeterminate NPs, and this DP has a valued interpretable operator feature $iOp_{[Q]}$, which values the operator feature of indeterminate pronouns and marks them as necessarily questions words. Thus, the entire DP that corresponds to *who* has a valued operator feature. In order for *who* to pied-pipe the possessum *book*, which lacks the operator feature, the head F, which carries the unvalued operator feature, Externally Pair-Merges with D and creates the ⟨F, D⟩, which corresponds to the genitive $/z/$. Since the genitive marker is a bound morpheme, the ⟨F, D⟩ amalgam satisfies the morpho-phonological condition on weak heads (40). When *who* is merged in the specifier of the ⟨F, D⟩ amalgam, they undergo feature sharing of the operator features, with the operator feature of the ⟨F, D⟩ amalgam valued as [Q]. In this way, the possessem *book* is pied-piped by *who*. This is illustrated in (47).

(47)

$$⟨Op_{[Q]}, Op_{[Q]}⟩$$

DP ⟨F, D⟩

$D_{iOp[Q]}$ NP ⟨$F_{uOp[Q]}$, D⟩ NP
  $who_{uOp[Q]}$ book
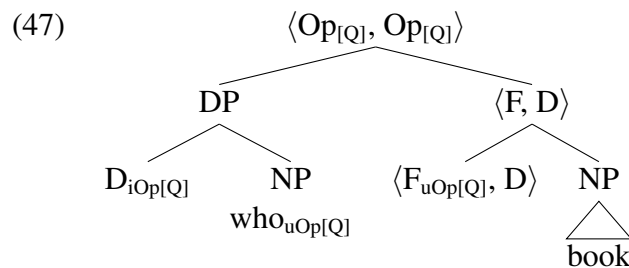
In the cases where the entire phrase *whose book* overtly moves as in (46), both D and F have an unvalued focus feature, which projects with the operator feature.[23]

Ps/PPs can also often be pied-piped by an indeterminate pronoun rather than being stranded, as exemplified by English (48).

(48) [**To whom**] did you give the book?

Bošković (2016b) suggests that Ps in such cases are functional elements rather than lexical elements (see also chapter 6). My suggestion here is that Ps in such cases can be analyzed as bound morphemes that are morphologically dependent on the following nominal phrase.In such cases, F Externally Pair-Merges with P, and the ⟨F, P⟩ amalgam counts as a weak head. The operator feature

---

23. Conversely, when *whose book* stays in-situ as in *Who read whose book?*, the focus feature is absent and only operator features project (see Bošković 2007b).

of the ⟨F, P⟩ amalgam is then valued as [Q] via feature sharing and the shared operator features project as illustrated in (49), which is the configuration of pied-piping.

(49)
$$\langle Op_{[Q]}, Op_{[Q]} \rangle$$

⟨F$_{uOp[Q]}$, P⟩ ⟨Op$_{[Q]}$, Op$_{[Q]}$⟩

D$_{iOp[Q]}$ NP
whom$_{uOp[Q]}$
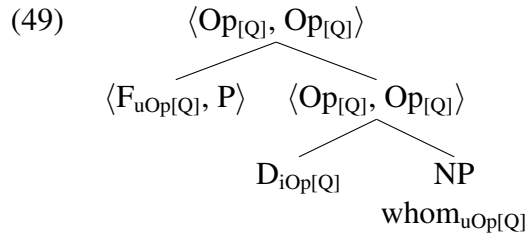
Crucially, when this ⟨F, P⟩ amalgam is sent to PF, it is realized as a bound morpheme, which observes the requirement (40) (overt fronting of *to whom* is trigged by an unvalued focus feature, which I omit in (49)). Thus, the proposed deduction of the generalization (23) can be extended to other pied-piping cases.

To summarize this section, I have proposed that large-scale pied-piping can be captured by a feature-percolation analysis implemented under Chomsky's (2015) labeling theory, in which a head F that has an unvalued operator feature Externally Pair-Merges with the head of the projection to be pied-piped, and F undergoes feature sharing with the head of a compositional indeterminate pronoun that also bears an unvalued operator feature. I have then proposed the criterion of weak heads, by which weak heads are heads that bear an unvalued feature, and a morphological condition of weak heads, in which weak heads are realized as bound morphemes. I have argued that the generalization (23) is deduced from these two properties of weak heads in combination with the feature specification of compositional indeterminate pronouns proposed in chapter 3 and Inaba's (2011) generalization that head-initial complementizers are free morphemes but head-final ones are bound morphemes.

4.4   Deduction of Agree from Minimal Search

In this section, I argue that the formulation of weak heads in (41) allows us to deduce Agree from Minimal Search, a possibility actually hinted at by Chomsky (2013, 2015).

Recall that Chomsky (2015) proposes that weak heads need to be "strengthened" by feature-

sharing. This is technically implemented in the traditional spec-head agreement configuration by Minimal Search, as seen above. Recall also that in chapter 3 I adopted Bošković's (2007b) theory of Agree, in which movement is always driven by an unvalued feature of the moving element. To recapitulate Bošković's theory, (50) is repeated from chapter 3. When a DP is base-generated in Spec,vP and has an unvalued Case feature as in (50a), T first probes down the DP and has its ϕ-features valued, as shown in (50b). The DP then moves to Spec,TP, from where it can probe down T, as a result of which the Case feature of D is valued as [NOM], as seen in (50c).

(50) a., b., c.
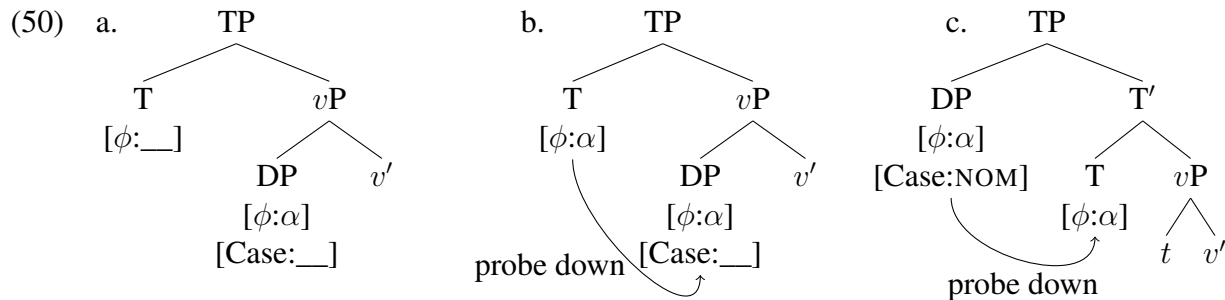
Interestingly for the current context, Saito (2018), who adopts Chomsky's (2015) labeling theory, quotes Noam Chomsky and Hisa Kitahara, who point out that in a configuration in which DP is Spec,TP, Case on D cannot be valued via a probe-goal relation, because D does not c-command T. In Bošković's original system, this issue would actually not arise, because the feature bundle of D including Case projects as DP from where the Case feature c-commands T under the Bare Phrase Structure Theory (Chomsky 1995a); to put it differently, when D projects DP, this DP is essentially like D in terms of feature specifications. Crucially, however, under the current system, in which heads with unvalued features are weak heads and weak heads cannot project on their own, the Case feature on D cannot project up to the conventional DP node and hence cannot c-command the goal.

Saito suggests, following Hisa Kitahara, that Case is valued via Minimal Search, just as ϕ on T is valued. When a DP moves to Spec,TP, Minimal Search finds D and T, which have an unvalued Case feature and unvalued ϕ-features, respectively, and those features are valued against their valuer, as illustrated in (51).

(51) a.

```
            DP
          /    \
        /        \
      DP          ...
     /  \        /  \
    D    NP     T    vP
[Case:__]    [φ:__]
  [φ:α]
```

b.

```
                TP
              /    \
            DP      TP
           /  \    /  \
          D    NP T    vP
        [φ:α]   [φ:α]
      [Case:NOM]
```

valuation by Minimal Search
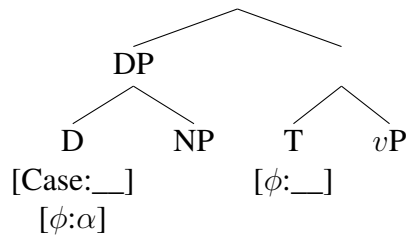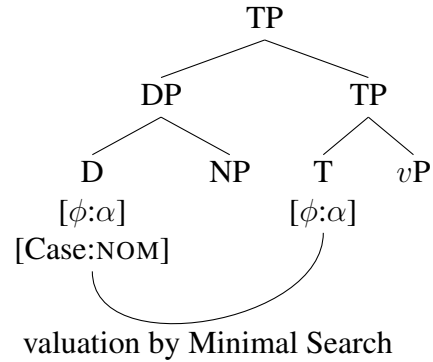
Notice that the current conception of weak heads (see (41)) allows and requires us to generalize this idea to all weak heads, namely, all unvalued features. Whenever Minimal Search finds a weak head whose feature cannot be valued by a head/feature in the search domain, a syntactic object that immediately dominates the weak head needs to move to a position where the weak head can find a valuer and provide a label after feature sharing/strengthening, as discussed above (when a weak head finds a valuer in the pre-movement structure, Minimal Search finds the relevant heads and valuation of the features takes place). This essentially derives the mechanism of Agree in Bošković's system, in which unvalued features always probe down and movement is driven by an unvalued feature of the moving element, which moves so that the relevant feature can probe down its goal (note that Agree also takes place without movement when a probe finds a goal in the pre-movement structure). Crucially, however, under the current system, valuation of all features is done by Minimal Search, which is a third factor principle outside UG according to Chomsky (2013). This means that we can eliminate the operation Agree from the computational system of language, hence minimize UG.[24]

This deduction of Agree also captures two conflicting ideas regarding movement driven by agreement proposed in the literature. On the one hand, as mentioned above, the current argument integrates Bošković's insight that movement is driven by a formal inadequacy of the moving element, which dates back to Move and feature-checking in early minimalism and Case Filter in the

---

24. Under the current system, it would need to be assumed that apparent "upward" valuation, i.e., valuation of a probe that is c-commanded by its goal, in general involves covert movement and pronunciation of a lower copy; e.g., ϕ-agreement between T and a post-verbal subject in Italian is analyzed as involving movement of the subject to Spec,TP and pronunciation of the lower copy. See Stjepanović (2003) and footnote 15 for relevant discussion.

GB Theory. Essentially, under this line of movement theories, the pre-movement structure involves a problem; if movement does not take place, the derivation would crash (e.g., an unvalued feature remaining unvalued). Resolving a problem with the base (i.e., pre-movement) structure thus motivates movement. From this perspective, as Bošković (2021c) points out, Chomsky's (2013) version of labeling theory, which does not assume weak heads, can be considered to be of this type. In this version of labeling theory, movement of an external argument from Spec,vP to Spec,TP takes place in order to label the {DP, vP} structure by leaving a copy in Spec,vP that is ignored for labeling (the {DP, TP} structure is labeled by feature-sharing, but T itself does not require this movement). In other words, if the movement in question does not take place, the labeling problem in the base (i.e., pre-movement) structure would remain, which results in ill-formedness. In fact, unifying Bošković's (2007b) Agree theory and Chomsky's (2013) labeling theory from the perspective discussed here, Bošković (2021c) proposes that an unvalued feature on the moving element causes a labeling problem in the base structure, which then drives movement. On the other hand, movement has also occasionally been assumed to be driven by a formal requirement of the target of movement. Part of Chomsky's (2015) version of labeling theory belongs to this type, since the weak head T requires movement of DP to its (traditional) specifier position for feature-sharing. Crucially, under this version of labeling theory, in this particular case movement is solely driven by weakness of the target, not by a problem in the base-structure.[25] Attract in Chomsky (1995b) and the EPP-driven movement theory in Chomsky (2000, 2001) are also of this type. The current proposal allows both types of movement, but crucially, the motivation for movement is uniform; presence of an unvalued feature (i.e., a weak head). Movement of XP can take place if the target of the movement is a weak head, even if the XP does not have an unvalued feature. Conversely, movement of XP can be motivated solely by an unvalued feature of the XP, even if the target has no unvalued feature. Valuation of Case in Japanese discussed by Saito (2018) falls under this. Saito assumes that T in Japanese does not have unvalued ϕ-features unlike T in English. Under

---

25. As mentioned in footnote 15, Chomsky (2015) assumes that T in Italian is strong and hence does not require a DP in its specifier position. He actually leaves open the structure and the derivation in which an external argument does not move to Spec,TP, which would involve a {DP, vP} structure that should be unlabelable. See footnote 15 for an argument that the issue can be resolved in the system proposed in this dissertation.

the current system, an external argument that bears an unvalued Case feature moves to Spec,TP in order for the Case feature to be valued by Minimal Search, and T itself has no reason to attract the external argument. There is even a third possibility that movement is driven by both the moving element and the target, if both have an unvalued feature. One such case is movement of an external argument from Spec,vP to Spec,TP in English (cf. (51)); D has an unvalued Case feature, which triggers movement of the external argument, and T has unvalued ϕ-features that attract the external argument. In all these three cases, a weak head requires feature valuation by Minimal Search, for which the traditional spec-head configuration is created.

The current proposal that movement can be driven by either the target or the moving element, or both, can account for the cross-linguistic variation regarding the superiority effect in multiple wh-fronting languages discussed in chapter 3. Bošković (2008b, 2012) establishes the generalization that multiple wh-fronting languages without definite articles do not display superiority effects. Thus, Bulgarian, which has definite articles, shows the superiority effect as seen in (52), whereas Serbo-Croatian, which lacks definite articles, does not show the relevant effect, as shown in (53).

(52) a. **Koj kogo** vižda?
who whom sees

b. *__**Kogo koj**__ vižda?
whom who sees
'Who sees whom?'   (Bulgarian, Rudin 1998:472-473)

(53) a. **Ko koga** voli?
who whom loves

b. **Koga ko** voli?
whom who loves
'Who loves whom?'   (Serbo-Croatian, Bošković 2002b:353)

In addition, Bošković (1997a) observes that even in Bulgarian, when there are more than two

compositional indeterminate pronouns, the non-initial ones do not show the superiority effect, as shown in (54).[26]

(54) a. Koj **kogo kakvo** e pital?

who whom what is asked

'Who asked whom what?'

b. Koj **kakvo kogo** e pital?

who what   whom is asked   (Bulgarian: Bošković 1997a:239)

Bošković (2002b) proposes that the highest compositional indeterminate pronoun in Bulgarian undergoes movement to Spec,CP triggered by a [+wh] feature of C, which is responsible for the superiority effect, while the non-initial compositional indeterminate pronouns in Bulgarian (see below for details) and all compositional indeterminate pronouns in Serbo-Croatian move to a lower position. Bošković (2008b) suggests that in languages with definite articles, C has a D feature, which drives movement of an indeterminate pronoun to Spec,CP.

It should be noted first that this explanation is difficult to implement in frameworks such as Bošković's (2007b), in which movement is solely driven by an unvalued feature of the moving element. Crucially, in this framework, the target of movement has no relevance for the movement. Therefore, even if C has a D feature as Bošković (2008b) assumes, it simply probes down a DP in its c-commanding domain and has its D feature valued against it (and the DP need not in principle be an indeterminate pronoun). The highest compositional indeterminate pronoun then moves to Spec,CP due to its operator feature, independently of the D feature of the C head (note that the D feature of DP is naturally assumed to be inherently valued and hence does not trigger movement). Thus, the D feature on C cannot drive movement of the highest compositional indeterminate pro-

---

26. Note that *kogo* 'whom' and *kakvo* 'what' show the superiority effect in the absence of *koj* 'who', as shown in (i).

(i) a. **Kogo kakvo** e pital   Ivan?
whom what   is asked Ivan
'Who did Ivan ask what?'

b.***Kakvo kogo** e pital   Ivan?
what   who is asked Ivan   (Bošković 1997a:239)

noun to Spec,CP, contra Bošković's (2008b) account. It is then unclear why C "attracts" the highest compositional indeterminate pronoun in Bulgarian under Bošković's (2007b) system.

On the other hand, the current proposal can capture the movement in question, if we assume, based on Bošković (2008b), that the D feature of the C head is an unvalued feature and hence this C head is a weak head. Under this assumption, the C head requires a syntactic object that can value its D feature to be in its specifier position, just like Chomsky's (2015) weak T in English. At the same time, the highest compositional indeterminate pronoun also has an unvalued operator feature, so that it also needs to move to a position where Minimal Search can find a head that can give a value to the operator feature. Thus, both the moving element and the target of the movement have a reason to undergo/attract movement in this case.

A question that remains is how the non-initial compositional indeterminate pronouns in Bulgarian and all compositional indeterminate pronouns in Serbo-Croatian undergo movement to a position where the operator feature is valued, without inducing the superiority effect. I argue that Epstein et al.'s (2020) implementation of Minimal Search in multiple-specifier configurations can be extended to account for the lack of the superiority effect in the cases in question. Epstein et al. (2020) discuss what counts as the first heads that Minimal Search finds in multiple specifier configurations. Epstein et al. propose that Minimal Search finds a target via the "shortest" possible path, adopting Chomsky's (1995b) idea that a shorter path is selected over a longer one. The relevant definitions are given in (55).[27]

---

27. This is essentially similar to calculation of a version of Economy of Derivation in early minimalism. Bošković (1997a), building on an early draft of Collins (1994), notes that there are two views of Economy of Derivation. One is a global view, in which a derivation α is more economical than (hence preferred to) a derivation β iff α crosses fewer nodes than β, whether some of the nodes crossed by α and β are the same or not. The other is a local view, in which a derivation α is more economical than (hence preferred to) a derivation β iff the set of nodes crossed by α is a proper subset of the set of nodes crossed by β. Note now that (55) is essentially the local view of Economy of Derivation. Interestingly, Bošković (1997a) argues that the local view is empirically motivated. He provides the following sentences ((ic) is taken from Oka 1993):

(i) a. What$_i$ did [people from where] try to buy t$_i$?

    b. cf. *What$_i$ did who try to buy t$_i$?

    c. ?What$_i$ did you persuade [friends of whom] to buy t$_i$?

    d. cf. ?*What$_i$ did you persuade whom to buy t$_i$?

In (ia), the two wh-phrases *what* and *where* do not c-command each other in their base positions, hence the path of *what* to Spec,CP is not in the subset-superset relation with that of *where* to Spec,CP (the same holds for the two wh-

(55) a. The *path* of α is the set of all SOs of which α is a term.

b. X is a *term* of Y iff X $\in$ Y or X $\in$ Z, Z a term of Y.

c. The path of α is shorter than the path of β iff the path of α is a proper subset of that of β.

Let us first see how this proposal works in a single specifier configuration (56), which schematizes the traditional TP in English.

(56)
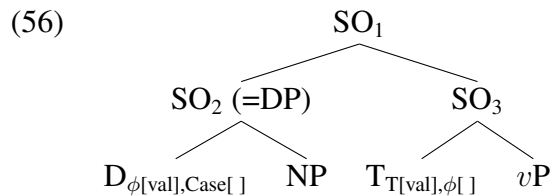
Here the path of D is {SO$_1$, SO$_2$}, and the path of T is {SO$_1$, SO$_3$}, so neither the path of X nor the path of Y is a proper subset of the other. Thus, both X and Y count as the first heads that Minimal Search finds, and they can undergo feature sharing and valuation of the unvalued features.[28]

Let us now look at a multiple specifier configuration. Epstein et al. (2020) use a multiple nominative construction in Japanese as an example.

(57) Bunmeikoku-ga   dansei-ga heikin-jumyou-ga   mijikai.
      civilized.country-NOM male-NOM average-life.span-NOM short.PRES
      'It is in civilized countries that male's average life span is short'

The relevant structure of (57) is schematized in (58).[29]

(58)

```
                          SO₁
              ┌────────────┴────────────┐
            SO₂                         SO₃
         ┌───┴───┐          ┌────────────┴────────────┐
   N₁ Case[ ]   XP        SO₄                         SO₅
                       ┌────┴────┐          ┌──────────┴──────────┐
                 N₂ Case[ ]     YP        SO₆                    SO₇
                                       ┌───┴───┐            ┌─────┴─────┐
                                 N₃ Case[ ]   ZP        T_T[val]       vP
```

In this configuration, the path of N_1 is {SO_1, SO_2}, the path of N_2 is {SO_1, SO_3, SO_4}, the path of N_3 is {SO_1, SO_3, SO_5, SO_6}, and the path of T is {SO_1, SO_3, SO_5, SO_7}. Notice that none of the paths of these heads are a proper subset of the others. Thus, Miminal Search finds N1, $N_{2, 3}$, and T as the first heads at the same time, and valuation of the Case feature of each N takes place.

Epstein et al. further discuss availability of multiple specifiers in different languages. As is well-known, English disallows the multiple nominative construction unlike Japanese. Since the path-based Minimal Search should work universally across languages as a third factor principle, the construction in question would be expected to be possible in English. In order to address this issue, Epstein et al. (2020:6) propose a principle of unique identification as in (59).

(59)   "Multiple-specifier" configurations appear iff Minimal Search finds one and only one valu-
        ing head per agreement-relation; that is, for each unvalued feature *uF-valuee*, there is one
        and only one valued feature *vF-valuer*.

In the case of the Japanese multiple nominative construction, each N has an unvalued Case feature, and there is one and only one valuer for it, namely, (the Tense feature of) T. T is assumed to lack unvalued ϕ-features, so there is no feature that needs to be valued on T. On the other hand, English T has unvalued ϕ-features. Thus, although the Case features of the Ns has one and only one valuer (i.e., the Tense feature of T), Minimal Search finds three valuers for the ϕ-features of T, i.e., N_1, N_2, and N_3. Thus, an English multiple nominative construction would violate the uniqueness condition (59), resulting in ill-formedness of the construction due to the failure to label.

Now, I argue that Epstein et al.'s theory of Minimal Search can be straightforwardly extended

to capture the variation regarding superiority effects discussed above. The gist of my proposal here is that the superiority effect arises if hosting more than one syntactic object causes a labeling problem. Let us first consider the superiority effect in languages like Bulgarian. I assume, following Bošković (2008b) that C has an unvalued D feature in languages with definite articles such as Bulgarian. The C head in those languages cannot then host more than one specifier, because C has an unvalued D feature, which requires one and only one valuer as per (59). Thus, only one compositional indeterminate pronoun moves to Spec,CP. A question that arises now is why the compositional indeterminate pronoun that moves to this position needs to be the highest one. I propose that Minimal Search looks for a syntactic object that has the shortest path among candidates for Internal Merge to the target. To illustrate this, let us look at the schematic structure in (60).

(60)

SO_1
$C_{uD[\ ]}$ SO_2
T SO_3
SO_4 SO_5
'who' v SO_6
V SO_7
'what'
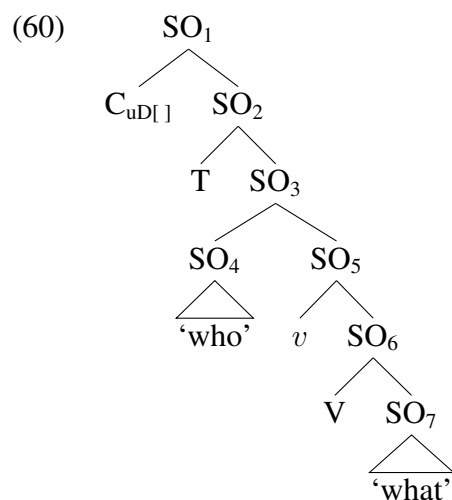
Since what moves is the maximal projection level (i.e., NPs), Minimal Search looks at SO_4 and SO_7. The path of SO_4 (i.e., 'who') is $\{SO_1, SO_2, SO_3\}$, and the path of SO_7 (i.e., 'what') is $\{SO_1, SO_2, SO_3, SO_5, SO_6\}$. Thus, the path of SO_4 is a proper subset of the path of SO_7, and hence it is "shorter". We can therefore conclude that the subject indeterminate pronoun is higher than the object one. Minimal Search then chooses the higher syntactic object for movement (Internal Merge) to the target position, Spec,CP. Thus, the path-based definition of Minimal Search correctly picks the highest element in the structure for Internal Merge, which captures the superiority effect

in multiple wh-fronting in languages like Bulgarian.

Turning to multiple wh-fronting in languages like Serbo-Croatian which do not show the superiority effect, a number of authors have proposed that those indeterminate pronouns that do not show the superiority effect move to a position lower than Spec,CP (see, e.g., Bošković 1997b, 2002b, 2008b, Lambova 2001, Stjepanović 1999, among others). In addition, as discussed in chapter 3, it has been argued that multiple wh-fronting essentially involves focus movement (e.g., Bošković 1999, 2002b, Horvath 1996, Izvorski 1995, É Kiss 1995, Lambova 2001, Rochemont 1986, Stepanov 1998, Stjepanović 1999). Based on these, I suggest that compositional indeterminate pronouns in the cases in question move to Spec,FocP, a position lower than Spec,CP in the left periphery (cf. Rizzi 1997). I also assume that Foc has a valued interpretable Focus feature, and that this feature can value an uninterpretable focus feature and an uninterpretable operator feature (the latter as [Q], just as T values Case as [Nom]). Crucially, Foc does not have an unvalued feature, and hence is a strong head under the current system, unlike C in Bulgarian (or multiple wh-fronting languages with definite articles in general), which has an unvalued D feature and hence is a weak head. Interestingly, Bošković (2001a,b, 2002b) suggests that interrogative C in Bulgarian is a PF affix, whereas that in Serbo-Croatian is not. This is motivated by the observation that Bulgarian C needs to be adjacent to a verbal element (i.e., a verbal element needs to move to C as a host of C), whereas Serbo-Croatian C (which is Foc here) does not, as shown in (61). This is straightforwardly captured by the current proposal, since Bulgarian C is a weak head that bears an unvalued D feature and hence must be realized as a bound morpheme as per (40), whereas Serbo-Croatian C (or Foc) is a strong head that does not bear the D feature and hence is not subject to the condition (40).[30]

(61)   a. *Kakvo (C) toj dade na Petko?

          what   C   he gave to Petko

          'What did he give to Petko?'   (Bulgarian)

---

30. English interrogative C can also be analyzed as an affix just like Bulgarian C, since it requires *do*-support in matrix questions (i.e., it requires a host for affixation; see in fact Bošković 2000). Note that English multiple wh-questions are subject to the superiority effect, and hence can be analyzed in the same way as multiple wh-fronting in Bulgarian.

b. Kakvo dade   toj na Petko?

what   gave+C he to Petko   (Bulgarian)

c. Šta   (C) on dade Ivanu?

what   C   he gave Ivan

'What did he give to Ivan?'   (Serbo-Croatian)

(Bošković 2001a:4)

Let us now consider the derivation of multiple wh-fronting without the superiority effect. When there are two candidates for Internal Merge, the higher one is found by Minimal Search (for the reason discussed above), and it moves to Spec,FocP. Then, the non-initial indeterminate pronoun undergoes Internal Merge to Spec,FocP. Interestingly, the relative height of the indeterminate pronouns in Spec,FocP does not matter for labeling. Whether the non-initial indeterminate pronoun moves to the outer or inner specifier of Foc, each indeterminate pronoun has one and only one valuer for the operator feature and the focus feature (i.e., Foc) and Foc does not have an unvalued feature to be valued, just like the multiple nominative construction in Japanese in (58).[31] Thus, the superiority effect can be analyzed as a matter of labeling of the structure after Internal Merge of relevant syntactic objects.

This analysis can be extended to the non-initial indeterminate phrases in Bulgarian, which do not show the superiority effect on a par with indeterminate pronouns in Serbo-Croatian (cf. (54)). Interestingly, Lambova (2001) proposes that the non-initial indeterminate phrases in Bulgarian move to Spec,FocP (while the highest indeterminate phrase moves to Spec,CP).[32] We can then

---

31. Bošković (2002b) observes that multiple wh-fronting in Serbo-Croatian shows the superiority effect when the overt Q-particle *li* is present, as seen in (i).

(i) a. Ko **li** koga voli?
       who C what loves
       'Who loves what?'

   b.*Koga **li** ko   voli?
       what C who loves

This can be captured if we analyze *li* as a weak head that has some unvalued feature. In fact, *li* is a second position clitic, which is morphologically dependent on the first element in the sentence (see Bošković 2001b for detailed discussion of the property of *li*). The superiority effect shown in (i) can then be accounted for on a par with that in Bulgarian.

32. To be more precise, Lambova calls the relevant projection ΔP, but practically it is equivalent to FocP in the current

---

163

maintain that just as in Serbo-Croatian, each of the non-initial indeterminate phrases in Bulgarian that are located in Spec,FocP finds one and only one valuer of the operator feature, i.e., Foc. Crucially, Foc in Bulgarian can be assumed to be a strong head (i.e., it does not bear an unvalued feature) just like Foc in Serbo-Croatian, so that the superiority effect does not arise for the reason discussed above for Serbo-Croatian.[33]

To summarize this section, I have argued that it is possible to deduce Agree from Minimal Search by assuming that weak heads are heads that have an unvalued feature at the point of External Merge, capturing two conflicting views on agreement in a uniform manner. I have then proposed that this deduction enables us to account for the variation regarding presence/absence of superiority effects in multiple wh-fronting by adopting Epstein et al.'s (2020) path-based calculation of Minimal Search.

4.5   On selective island sensitivity in affixal article languages

In this section, I argue that the above proposal regarding weak heads can be extended to insensitivity of particular indeterminate phrases to wh-islands in a class of languages. Bošković (2008a) observes that movement of D-linked wh-phrases and relativization out of multiple wh-islands are allowed in Albanian, Bulgarian, Hebrew, Icelandic, Norwegian, Romanian, and Swedish, which are all affixal article languages.[34] This is contrasted with English, which has non-affixal article,

---

proposal.

33. Movement of the non-initial indeterminate pronouns to Spec,FocP in Bulgarian violates the Cycle, since it takes place after movement of the highest indeterminate pronoun. One possibility to circumvent this issue is to define the Cycle by phases; e.g., movements within a single phase count as cyclic regardless of the order of the movements (Chomsky 2000, 2001 in fact suggests the possibility that movement to an inner specifier, which violates the Cycle, may be allowed within one phase). Under Bošković's (2007b) contextual approach to phasehood adopted in this dissertation, the highest phrase in the left periphery, i.e., CP, counts as a phase in the C-domain. CP and FocP are then within one phase, so that the movements in question take place within a "phase cycle" and hence do not violate the cycle.

It is worth adding here that there are actually proposals that violate the Cycle in the traditional sense. For instance, Richards (1997, 2001) proposes tucking-in, where non-initial indeterminate pronouns target a position lower than the highest one. Chomsky (2008) proposes that an external argument moves to Spec,TP after C enters the structure and ϕ-features are inherited from C to T. Extending this to the case under discussion, we can propose that non-initial indeterminate pronouns move to Spec,FocP after C enters the structure and, say, a (valued) focus feature is inherited from C to Foc.

34. The generalization crucially concerns extraction from *multiple* wh-islands.

and SC, which lacks articles. (62)-(64) are quoted from Bošković (2008a).

(62) *I saw a book which I wonder who knows who sells.   (English)

(63) *Vidio sam knjigu koju$_i$ se   pitam   ko zna   ko prodaje t$_i$.

     Seen am   book   which REFL wonder-1sg who knows who sells

     'I saw a book which I wonder who knows who sells.'   (SC)

(64) Vidjah edna kniga, kojato$_i$ se   čudja   koj znae koj prodava t$_i$.

     saw-1sg one book   which-the REFL wonder-1sg who knows who sells

     'I saw a book which I wonder who knows who sells.'   (Bulgarian)

Bošković (2008a) thus establishes the following generalization, where selective wh-island insensitivity means the possibility of extraction of D-linked/relative wh-phrases out of multiple wh-islands;

(65) Selective wh-island insensitivity is a property of languages with affixal articles.

<div align="right">(Bošković 2008a:263)</div>

Bošković attempts to explain (65) based on his deduction of freezing effects from Chomsky's (2000) Activation Condition, which requires that a goal have an unvalued feature in order to participate in Agree. As an instance of a freezing effect, he discusses wh-islands in English.

(66) *What$_i$ do you wonder [$_{CP}$ t$_i$ C [$_{IP}$ John bought t$_i$ (when)]]   (Bošković 2008a:256)

He assumes that wh-phrases have a valued wh-feature and an unvalued operator feature, and the latter is checked/valued by an interrogative C in its specifier position (and the unvalued wh-feature of interrogative C is valued). In (66), the embedded interrogative C, which is a probe, Agrees with *what*, which is a goal, and values the operator feature of *what* (at the same time, the wh-feature of the C is valued). Then, *what* moves to Spec,CP to satisfy EPP of this embedded interrogative C. Crucially, when the matrix interrogative C is merged and tries to probe down a goal, *what* is inactive, since its unvalued feature (i.e., the operator feature) has already been valued. Thus, *what*

is "frozen" in Spec,CP of the embedded clause, and cannot move to Spec,CP of the matrix clause. (This explanation can be adapted to Bošković's 2007b Agree system, in which an unvalued feature drives movement, as Bošković 2008a himself notes. See also Saito 2017, who indeed explains wh-island effects in Bošković's 2007b system, as discussed in chapter 3.)

In addition, Bošković (2008a) proposes that there are special D heads for D-linking and relativization which have an unvalued operator feature and project above wh-phrases. He also suggests that in affixal article languages these D heads share operator features with compositional indeterminate pronouns in my terminology due to their affixal nature, as a result of which the full D-linked/relative indeterminate phrases have two operator features. This suggestion is motivated by the observation that in some affixal article languages such as Albanian and Bulgarian, D-linked and relative indeterminate phrases, but no other indeterminate phrases, are accompanied by affixal articles (see *koja-to* 'which-the' in (64)). He then argues that only one of the two operator features of those indeterminate phrases is valued by an embedded C, so the indeterminate phrases remain active even after they Agree with the embedded C. Consequently, the indeterminate phrases in the specifier of the embedded interrogative CP can Agree with the matrix interrogative C.

It should be noted that Bošković leaves open how the feature sharing mechanism by affixation is technically implemented. In particular, it remains to be explained why the affixal nature of the relevant articles, which is a morphological property, influences the syntactic property, i.e., feature sharing. In addition, it is actually not clear how (64) is captured. In his proposal, D-linked/relative indeterminate phrases in affixal article languages have *two* operator features because of feature sharing, one from compositional indeterminate pronouns and the other from D-linking/relative D. Let us consider how this would account for (64), repeated here as (67), with relevant bracketing and traces added.

(67) Vidjah edna kniga, [$_{CP1}$ kojato$_i$ C$_1$ [se čudja   [$_{CP2}$ t$_i$ koj C$_2$ [znae
     saw-1sg one book   which-the   REFL wonder-1sg   who   knows

     [$_{CP3}$ t$_i$ koj C$_3$ [prodava t$_i$]]]]]].
            who   sells

     'I saw a book which I wonder who knows who sells.'   (Bulgarian)

In the bottom-up structure building and derivation, *kojato* first Agrees with C$_3$, and one of its two operator features is valued. It then moves to Spec,CP$_3$, from where it Agrees with C$_2$.[35] Notice that at this point, the remaining operator feature of *kojato* would be valued, and hence no unvalued feature would be left for Agree with C$_1$. It then seems to be expected that *kojato* would be frozen in Spec,CP$_2$. In order to resolve this issue, Bošković suggests that one of the two operator features can remain unchecked in the configuration of the relevant indeterminate phrases. He assumes that D-linked/relative D hosts an indeterminate phrase in its specifier position. Since the two operator features originate in different positions (one in the head of the entire D-linked/relative indeterminate phrase and the other in its specifier), valuation of one of them can be "delayed" until the indeterminate phrase reaches the final landing site. Here again, it is not clear how this delay of feature valuation till the end of the derivation would be technically implemented.

I argue that the current proposal regarding Pair-Merge and Minimal Search can implement Bošković's ideas in a rather straightforward way, but without appealing to feature sharing. First, I follow Bošković in assuming that the D-linked/relative D in affixal article languages has an unvalued operator feature, from which it follows under the current proposal that this D is a weak head and is realized as a bound morpheme. In addition, I propose that this D is Externally Pair-Merged with the N head of compositional indeterminate pronouns, by which a ⟨D, N⟩ amalgam is created. This is schematized in (68) (note that since the ⟨D, N⟩ amalgam is a weak head, it does not project until it undergoes feature sharing, hence the lack of the label on the highest node in (68)).

---

35. Here I ignore intermediate successive-cyclic movement steps such as the edge of vP.

(68)

$$\langle D_{Op[\ ]}, \widehat{N_{Op[\ ]}} \rangle \quad Root$$
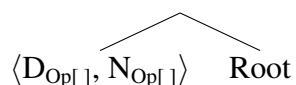
As mentioned above, Chomsky (2015) assumes that one of the heads in an ⟨H, H'⟩ amalgam is invisible for Minimal Search. Recall also that valuation of a feature is done by Minimal Search, which means that syntactic objects that are invisible for Minimal Search do not participate in valuation of features. Thus, in (68), one of the two heads in the amalgam is invisible for Minimal Search, and hence its operator feature remains unvalued when the entire phrase reaches the first embedded Spec,CP. At a first glance, this seems to implement Bošković's suggestion that valuation of one of the two operator features is delayed. Crucially, however, the invisible head remains invisible throughout the derivation under Chomsky's (2015) assumption; the unvalued feature in the invisible head would thus remain unvalued. Notice now that this cannot capture the anti-freezing effect of D-linked/relative indeterminate pronouns suggested by Bošković. In order for a D-linked/relative indeterminate phrase to move to the matrix Spec,CP, the unvalued operator feature of the invisible head should somehow be "active"; otherwise, the entire phrase would be frozen when the other operator feature is valued in the first embedded Spec,CP to which it moves. Thus, there are two seemingly conflicting requirements on one of the two heads in (68): it should be invisible for Minimal Search in order for the entire phrase not to be frozen in the course of successive-cyclic movement, but it should be visible for Minimal Search in order for the entire phrase to reach the matrix Spec,CP.

In order to resolve this problem, I modify Chomsky's assumption; specifically, I propose that one of the heads in a ⟨H, H'⟩ amalgam is actually invisible for Minimal Search during the derivation but becomes visible at the end of the derivation (see below for a more precise characterization). This proposal is based on Chomsky's (2004) argument that in the case of adjunction of α to β, which creates an ordered pair $\langle \alpha, \beta \rangle$, β is more prominent in the course of the derivation with respect to syntactic relations such as c-command, but at the point of Spell-Out of the ordered pair α becomes equally prominent. Thus, when a D-linked/relative indeterminate phrase moves to the edge of an embedded CP, one of its operator features is valued, but the other one can remain

unvalued, which avoids freezing of the indeterminate phrase in that position. Since the edge of an embedded CP is not sent to Spell-Out when this CP is completed under Chomsky's (2000, 2001) conception of phases, the remaining unvalued operator feature of the relevant head in the amalgam stays invisible. This operator feature becomes visible in the matrix Spec,CP, since it is the end of the derivation and hence the entire structure is sent to Spell-Out.[36] Thus, I conclude that the generalization (65) can be captured by the revised version of head invisibility in an ⟨H, H'⟩ amalgam. This resolves the issues of technical implementation of the deduction that Bošković (2008a) originally suggests.

Note that this configuration is impossible in non-affixal article languages. As discussed in chapter 3, DP always projects above indeterminate NP in non-affixal article languages, and the operator feature of N is valued by the valued operator feature of this DP (see chapter 3 for details). We can then assume that D-linked/relative D in non-affixal article languages also projects above them rather than adjoins to them, and has a valued operator feature ([Q] or [Rel]). Thus, the delay of feature valuation due to invisibility of one of the ⟨D, N⟩ amalgam I proposed for affixal article languages is not possible in non-affixal article languages, hence the lack of selective island sensitivity in non-affixal article languages.

Before concluding this section, I would like to briefly discuss variation in a particular type of selective island sensitivity among affixal article languages. Specifically, Bošković (2008a) notes that extraction of D-linked/relative argument PP out of multiple wh-islands is allowed in Albanian, Bulgarian, Hebrew, and Romanian as illustrated by Bulgarian (69), while it is disallowed in Icelandic, Norwegian, and Swedish as illustrated by Icelandic (70). Bošković leaves open how the difference can be accounted for.

---

36. In the cases of embedded/indirect questions and relative clauses, an indeterminate pronoun moves to Spec,CP of the relevant embedded/relative CP and has its operator feature valued in that position. The indeterminate pronoun could in principle move to the matrix Spec,CP without its operator feature valued in the embedded/relative Spec,CP, but the operator feature could not be valued in the matrix Spec,CP, so that the derivation would crash. Thus, only the derivation in which the indeterminate pronoun moves to the relevant embedded Spec,CP is legitimate in the cases of embedded questions and relative clauses.

(69) a. Na koja masa se čudiš koj složi knigata?

      on which table REFL wonder.2SG.PRES who put the.book

      'On which table did John ask who put the book?'

    b. Masata, na kojato ti se čudiš koj složi knigata

      the.table on which you REFL wonder who put the.book

      'the table on which John asked who put the book'   (Bulgarian)

(70) a.??Á hvaða borð spurði Jón hver hefði sett bókina?

      on which table asked John who had.SUBJ put the.book

    b.?*Borðið sem Jón spurði hver hefði sett bókina á

      the.table that John asked who had.SUBJ put the.book on   (Icelandic)

Interestingly, the former languages do not allow P-stranding, whereas the latter do, as represented by Bulgarian (71) and Icelandic (72), respectively.

(71) *Koj e govorila Anna s?

    who AUX spoken Anna with   (Bulgarian: Merchant 2001:97)

(72) Hvern hefur Pétur talað við?

    who has Peter talked with   (Icelandic: Merchant 2001:93)

This leads us to conjecture that there is a correlation between the presence/absence of selective island sensitivity of D-linked PPs and PP relativization on the one hand and the (un)availability of P-stranding on the other in these affixal-article languages. An intuitive way to capture this correlation could be that there is a structural difference between Ps in P-stranding languages and those in non-P-stranding languages. Bošković (2014) proposes that Ps in P-stranding languages have a richer structure than those in non-P-stranding languages. Given this, the difference between Scandinavian languages and Bulgarian-type languages regarding selective island sensitivity of D-linked PPs and PP relativization may be attributable to the structural difference regarding PPs, that is, F could be Externally Pair-Merged only with structurally simpler Ps in the latter type of

languages. I leave technical implementation of this idea for future research.

4.6 Conclusion of the chapter

In this chapter, I have discussed large-scale pied-piping, which moves an entire clause or an island that contains a compositional indeterminate pronoun rather than extracting only the indeterminate pronoun. Building on Watanabe's (2004b) generalization that languages that allow large-scale pied-piping have indeterminate pronouns, I have argued that the current proposal regarding indeterminate pronouns can capture it in combination with the assumption that a head F that has an unvalued operator feature is merged with the head of the projection to be pied-piped. In addition, I have established the generalization that large-scale pied-piping is only possible with head-final clauses, and proposed that the availability of merger of F to a pied-pipee is constrained by the morphological status of complementizers in a given language observed by Inaba (2011). In the course of the discussion, I have refined the notion of weak heads in Chomsky's (2015) labeling framework and suggested a correlation between the syntactic and morphological properties of heads. I have then shown that Epstein et al.'s (2016) definition of Minimal Search can be extended to capture variation regarding the superiority effect in multiple wh-fronting, and proposed that the operation of Agree can be eliminated from the computational system of language.

Finally, I have discussed an extension of the proposal on weak heads to selective wh-island insensitivity observed by Bošković (2008a). I have claimed that his generalization that selective wh-island insensitivity is a property of affixal article languages can be captured by the current proposal that weak heads Externally Pair-Merged with another head can undergo feature sharing and the assumption a lá Chomsky (2004) that one of the Pair-Merged heads can be invisible until it is sent to Spell-Out.

Appendix: Weak head and phasehood

In footnote 18 in section 4.3.2, I noted that English D, which is assumed to bear an unvalued Case feature, may count as a weak head as per the proposed morphological criterion of weak heads (41). Under Chomsky's (2015) labeling theory I have adopted here, in which weak heads cannot project, this means that the conventional "DP" node is unlabeled until D undergoes feature sharing with T (in the case of subjects) in English (as well as languages where "DP" projects above NPs) and its Case feature is valued. What is of particular relevance here is that, as mentioned in footnote 18, Bošković (2016c, 2018a) argues that an unlabeled syntactic object cannot be a phase because we need to know the label of the syntactic object in order to determine whether it is a phase or not. Under the current proposal, then, the "DP" would not be a phase until its head undergoes feature sharing with another head (e.g., T, R).

This raises a question regarding extraction out of a nominal domain discussed in chapter 1 and 2. Bošković (2005, 2008b, 2012, 2013a) shows that extraction of an adjunct and Left-Branch Extraction of an adjective out of a nominal domain are impossible in languages with definite articles, whereas they may be allowed in languages without definite articles, as exemplified by English (73a) and Serbo-Croation (73b).

(73) a. *From which city$_i$ did Peter meet [t$_i$ girls]?

   b. Iz   kojeg grada je Ivan sreo [djevojke t$_i$]?

      From which city   is Ivan met   girls   (Bošković 2008b)

Bošković's proposal is that the extraction in question is banned by the interaction of the Phase Impenetrability Condition (PIC) and the anti-locality condition. The PIC states that the complement of a phase head becomes inaccessible for further syntactic operations after Spell-Out, as a result of which movement out of a complement of a phase head is blocked after Spell-Out (Chomsky 2000, 2001). In Bošković's (2013a) formulation of the anti-locality condition, movement has to cross at least one full phrase, not a segment. Crucially, Bošković proposes that DP projects above NP in languages with definite articles and this DP is a phase, whereas DP is absent in article-less

languages and NP is a phase. Thus, in languages with definite articles, when an adjunct, which is adjoined to NP, is extracted, this adjunct either has to violate the PIC to satisfy the anti-locality condition if it moves directly out of DP, as in (74a), or has to violate the anti-locality condition to obey the PIC if it moves to Spec,DP before Spell-Out since it crosses just a segment, not a full phrase, as in (74b). In contrast, in article-less languages, where NP is a phase, an adjunct can be extracted without violating the PIC or the anti-locality condition, as illustrated in (75). (LBE of an adjective out of a nominal phrase can be explained in the same way, on the assumption that an adjective is adjoined to NP.)

(74) a.

```
              DP
             /  \
                 D'
                / \
               D   NP  Spell-Out
                       domain
  ✓PIC
  ✗anti-locality   Adjunct   NP
```

b.

```
              DP
             /  \
                 D'
                / \
               D   NP  Spell-Out
                       domain
  ✗PIC
  ✓anti-locality   Adjunct   NP
```

(75)

```
                   NP
                  /  \
              Adjunct   NP
  ✓PIC                 /  \
  ✓anti-locality      N    YP   Spell-Out domain
```
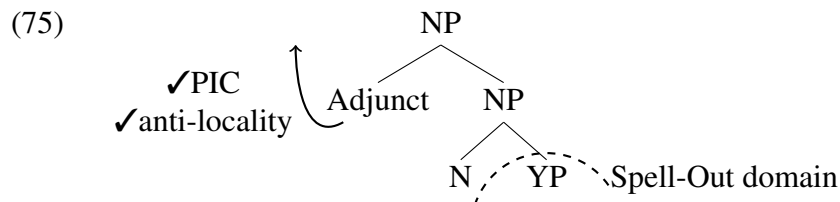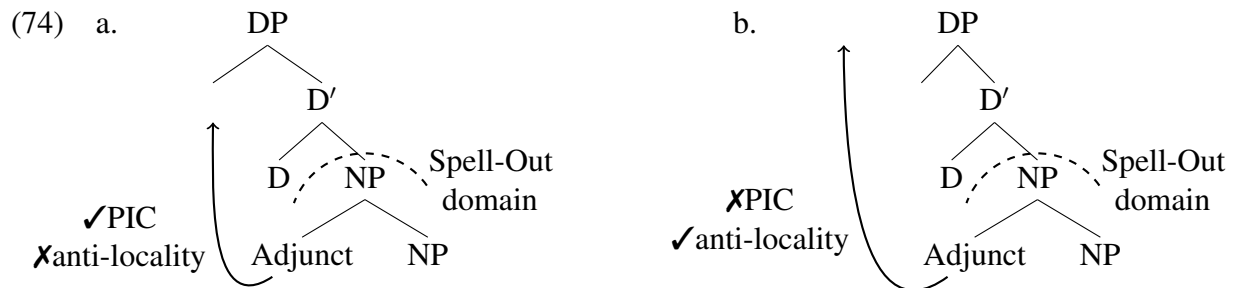
Notice now that the "DP" would actually not project under the current proposal, since D has an unvalued Case feature and hence counts as a weak head, which cannot provide a label on its own. If labeling is a prerequisite for phasehood of a projection as Bošković (2016c, 2018a) argues, "DP" would not count as a phase until valuation of the Case feature. Adjunct extraction (as well as LBE) out of a nominal phrase that takes place before labeling of "DP" may then be incorrectly predicted to be possible, because the phasal complement is sent to Spell-Out upon completion of the phase (Chomsky 2000) and the "DP" phase would not be completed without the "DP" label.

I propose that this problem can be solved once we consider what is essential for the PIC, one of the crucial ingredients for the explanation under discussion. Note first that whether a projection

counts as a phase depends on the nature of its head; namely, a phase head projects a projection that counts as a phase. This follows from the general property of non-minimal projections, i.e., the nature of a projection is determined by the nature of its head. It is worth mentioning here that Bošković (2013a, 2014) argues that the highest projection in the extended projection of a lexical category counts as a phase, the position I have also taken in this dissertation. Regarding the timing of XP becoming a phase, Bošković proposes that XP becomes a phase when a head H that does not belong to the same extended projection as XP is merged with XP. Since a head is crucial for phasehood as mentioned above, Bošković's proposal can be reformulated such that a head X counts as a phase head when a head H that does not belong to the same extended projection as XP is merged above X. This formulation in turn allows us to propose that the complement of a phase head X is sent to Spell-Out when a head H that does not belong to the same extended projection as X merges above X, whether X provides a label or not. Under this formulation, adjunct extraction and LBE out of NP are still banned in languages where D merges above NP, since D is the phase head as the highest projection in the nominal domain, and its complement is sent to Spell-Out as soon as a next head H that does not belong to the same extended projections as D (i.e., the nominal domain) is merged. Thus, the extractions in question are ruled out by the interaction of the PIC and the anti-locality condition as schematized in (74) even if D is a weak head that cannot provide a label on its own.

Chapter 5

A Fine-grained Scale of the NP/DP-language Distinction and the Emergentist View of Parameters

5.1 Introduction

The NP/DP-language distinction discussed in this dissertation was originally proposed as a two-way cut by Bošković's (2008b, 2012); whether a language has a definite article (i.e., DP-language) or not (i.e., NP-language). Talić (2015, 2017) later argued that the two-way distinction is not sufficient and a three-way distinction is needed; non-affixal article languages, where DP always projects, affixal article languages, where DP may be absent when the definite article is absent, and article-less languages, where DP is always absent. I have shown in the previous chapters that the three-way distinction is also observed with the Coordinate Structure Constraint and compositional indeterminate pronouns.

Note, however, that Talić's three-way distinction actually involves a categorical cut in structure building in the nominal domain which is actually similar to Bošković's two-way cut. Under Bošković's two-way cut, the presence/absence of a definite article in a given *language* correlates

with the presence/absence of DP. Under Talić's three-way cut, the presence/absence of a definite article in a given *construction* correlates with the presence/absence of DP. Thus, in both classifications of the NP/DP-language distinction, the presence of a definite article is crucial for projection of DP: when the definite article is present, DP is projected; when it is not present, DP is not projected.

In this chapter, I argue that this distinction, including Talić's three-way distinction, is not sufficient, and that an even more fine-grained distinction needs to be made. In particular, the distinction to be made is not a two-way or three-way "cut", but a "scale" from a canonical DP-language to a canonical NP-language. As illustrations, I discuss Italian, Hungarian, and Greek as languages that show some properties that cannot be captured by the two-way or three-way cut of the NP/DP-language distinction. Importantly, based on these languages, it is argued that mere presence/absence of a definite article in a given language or construction does not straightforwardly correlate with the presence/absence of DP. In particular, DP can be absent even if the definite article is present. To capture this, I propose options for realization of definite articles in the syntax under the Bare Phrase Structure Theory, which capture the otherwise mysterious behavior of these languages. Specifically, I propose that definite articles in Italian, which have been analyzed as being identical with clitics in the literature, can be adjoined to a head in the nominal domain via base-generation, without projecting DP. I extend this to Hungarian and argue that languages can differ with respect to the adjunction site of the definite articles in the nominal domain, which results in various types of DP-languages in the scale. In addition, I propose that Greek definite articles always project DP but can cliticize onto an adjective, which voids locality violations and allows Left Branch Extraction of adjectives, a property associated with NP-languages. Thus, the scale of NP/DP-language distinction can be captured by different options for realization of definite articles in the structure, which are in fact allowed by the current syntactic theory. I also discuss the proposed fine-grained scale of NP/DP-languages distinction from a perspective of the so-called emergentist view of parameters, which conforms to the three-factor design of language proposed by Chomsky (2005). The scale of NP/DP-languages distinction is shown to be an appropriate parameter given what is considered to be the locus of parameterization in minimalism and economy

176

considerations of language acquisition.

The chapter is organized as follows. In section 5.2, I show that Italian behaves differently from other languages in the previous classifications of NP- and DP-languages with respect to adjunct extraction out of a nominal phrase and the distribution of possessive and pronominal reflexives. In section 5.3, I propose that there are different options for realizations of definite articles in the syntactic structure, which I argue capture the behavior of Italian in the relevant respects. In section 5.4, I discuss reflexive possessives and noun-incorporation in Hungarian, proposing that Hungarian allows absence of DP and is less of a DP-language than e.g., English in the NP/DP-language scale. In section 5.5, I demonstrate that Greek is yet another type of a DP-language. It is proposed that the definite article in Greek always projects DP, while DP does not project above NP in indefinite nominal phrases, which is motivated by a number of properties where the relevant nominal phrase behaves differently depending on whether it is definite or not. In addition, I propose that the definite article in Greek can adjoin to A via movement, voiding locality restrictions, and allowing LBE of adjectives, under Bošković's (2013b) rescue-by-PF-deletion mechanism. I also discuss extraction out of the nominal phrase in the presence of the indefinite article, proposing two possible analyses, in both of which the indefinite article in Greek crucially does not project its own functional projection. In section 5.6, I argue that the scale of the NP/DP-language distinction proposed here receives further support from the emergentist view of parameters, a parameter theory that fits the three-factor design of language in the current minimalism. The distinction between canonical DP- and NP-languages is deduced from feature specifications of relevant lexical items as a locus of parameterization and a third-factor principle that requires a learner to postulate as few formal features as possible. Then, the intermediate status of Italian, Greek, and Hungarian is analyzed as emerging from specific properties of the definite articles and/or indefinite articles in these languages that are acquired from the primary linguistic data, which interacts with another third factor principle that generalizes a parameter value to other domains. In addition, the timing of acquisition of definite articles in English and Italian is discussed from the viewpoint of economy of structure building. In section 5.7, I discuss potential relevance of the presence of a fully grammaticalized

indefinite article for projection of DP in indefinite nominal phrases, from a viewpoint of Egyptian Arabic and Basque. Section 5.8 concludes the chapter.

5.2 Italian as a different type from non-affixal and affixal article languages

In this section, I show that Italian, which has been considerd to be a non-affixal article language by Bošković (2008b, 2012) and Talić (2017), behaves differently from prototypical non-affixal article languages and affixal article languages with respect to adjunct extraction out of a nominal phrase and the distribution of reflexive and pronominal possessives, suggesting that Italian is yet another type of a DP-language.

5.2.1 Adjunct extraction out of a nominal phrase

As mentioned in the previous chapters, Bošković (2008b, 2012) establishes the following generalization regarding extraction of an adjunct out of a nominal phrase;

(1)   Only languages without definite articles may allow adjunct extraction out of a nominal phrase.

This is exemplified by (2); the extraction in question is disallowed in English, which has a definite article (2a), while it is allowed in Serbo-Croatian, which lacks a definite article (2b).

(2)   a. *[From which city]$_i$ did Peter meet [t$_i$ girls]?

b.   [Iz   kojeg grada]$_i$ je Ivan sreo [djevojke t$_i$]?
     from which city   is Ivan met   girls   (Bošković 2008b)

Bošković's deduction of (1) is that the extraction in question is banned by the interaction of the Phase Impenetrability Condition (PIC) and the anti-locality condition. The PIC essentially states that only the edge of a phase is accessible to further syntactic operations, hence movement out of a complement of a phase head to a position outside of the phase is blocked (Chomsky 2000,

178

2001). In Bošković's (2013a) formulation of the anti-locality condition, movement has to cross at least one full phrase, not just a segment. In addition, Bošković (2013a, 2014) proposes that the highest projection in the extended projections of a lexical category counts as a phase. Thus, DP is a phase in languages with definite articles, where DP projects above NP, whereas NP is a phase in article-less languages, where DP is absent. Consequently, in languages with definite articles, when an adjunct, which is adjoined to NP, is extracted, this adjunct either has to violate the PIC to satisfy the anti-locality condition if it moves directly out of DP, as in (3a), or has to violate the anti-locality condition to obey the PIC if it moves to Spec,DP since it crosses just a segment, not a full phrase, as in (3b). In contrast, in article-less languages, where NP is a phase, an adjunct can be extracted without violating the PIC or the anti-locality condition, as illustrated in (4).[1]

(3) a.

b.

(4)
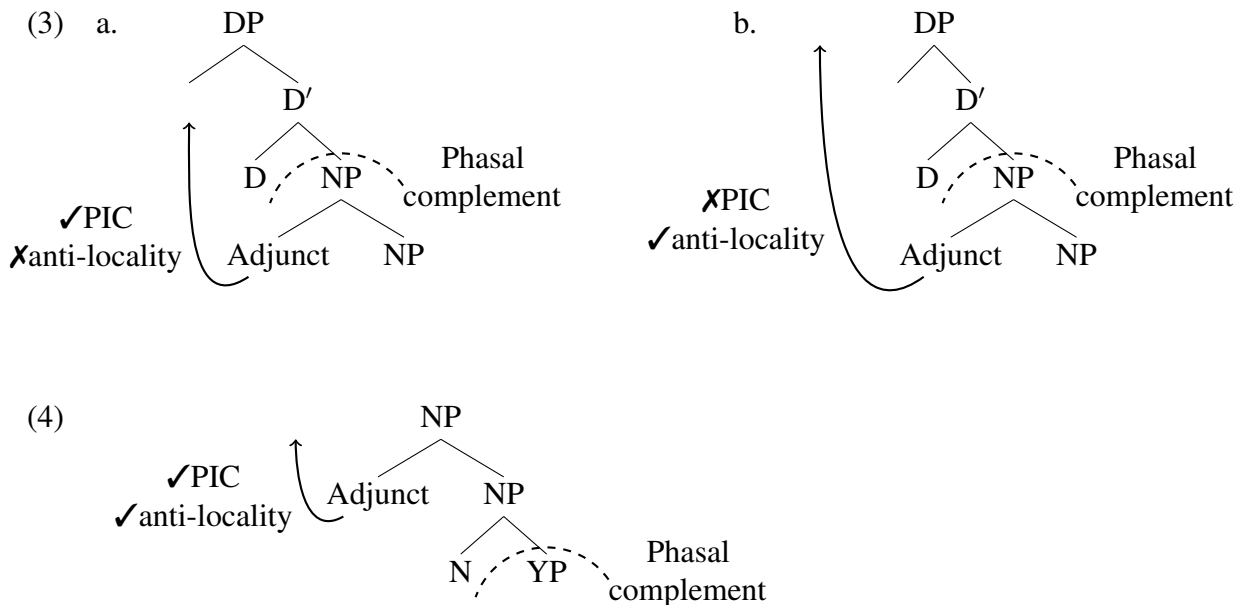
Note that Bošković's generalization has a two-way distinction, i.e., whether a language has a definite article or not. However, Dubinsky and Tasseva-Kurktchieva (2014) show that in Bulgarian, which has affixal definite articles, adjunct extraction out of a nominal phrase is disallowed when the definite article is present with a quantifier or a prenominal possessive, as shown in (5a) and

---

1. It should be noted that Bošković does not argue that there can never be any functional structure above NP in languages without articles; he just argues that there can be no DP in those languages. Thus, projection of some functional projection above NP in article-less languages is not precluded. (Bošković in fact gives such cases.)

(5c), but it is allowed when the article is absent in such environments, as shown in (5b) and (5d).

(5) a. *[Ot koj universitet]$_i$ sreštna-ha nyakolko-**to** studenti t$_i$?
       from which university met-they several-the students
       'From which university did they meet several students?'

  b. [Ot koj universitet]$_i$ sreštna-ha nyakolko studenti t$_i$?
       from which university met-they several students
       'From which university did they meet several students?'

  c. *[Ot koj universitet]$_i$ sreštna-ha nejni-**to** studenti t$_i$?
       from which university met-they her-the students
       'From which university did they meet her students?'

  d. [Ot koj universitet]$_i$ sreštna-ha nejni studenti t$_i$?
       from which university met-they her students
       'From which university did they meet her students?'

<div align="right">(Dubinsky and Tasseva-Kurktchieva 2014)</div>

Appealing to Bošković's deduction of (1) discussed above, Dubinsky and Tasseva-Kurktchieva (2014) and Talić (2017) argue that DP is absent in Bulgarian when the affixal definite article is absent (see Dubinsky and Tasseva-Kurktchieva 2014 and Talić 2017 for more discussion).

In the above cases, the presence of a definite article in a given language or in a given construction correlates with the (un)availability of the extraction in question. Interestingly, however, Bošković (2005:20, fn.27) notes that Italian allows adjunct extraction out of a nominal phrase (only) in the *presence* of the article, as shown in (6).[2,3]

---

2. The judgment is shared by my informants.

3. It should be noted here that the extraction in question is only possible with *di*-phrases (cf. Cinque 1995). Ticio (2003, 2005) observes a similar restriction regarding extraction of PP out of a nominal phrase in Spanish; only a small set of prepositions, including *de* (which is the Spanish counterpart of *di*), can head PPs that can undergo extraction in question. Note that Ticio examines cases where the relevant PPs are arguments of the noun, as seen in (i) (she discusses some cases of adjunct *de*-phrases, but she does not provide examples of the sort we are dealing with here).

(6) a. *[Di che   scaffale]$_i$ Gianni ha letto [libri   t$_i$]?

     of which shelf   Gianni has read books

     'From which shelf did Gianni read books?'

   b. [Di che   scaffale]$_i$ Gianni ha già   letto [i   libri   t$_i$]?

     of which shelf   Gianni has already read   the books

     'From which shelf did Gianni read the books?'

<div align="right">(Bošković 2005, attributed to Giuliana Giusti)</div>

Given that the presence of DP blocks the extraction in question (Bošković 2005), it is implied that DP does not project in (6b) despite the presence of the article. Note that (6) also contrasts with Bulgarian (5), where the relevant extraction is allowed in the *absence* of the article. Thus, the difference regarding adjunct extraction out of a nominal phrase indicates that the definite articles in Italian have a different syntactic status from those in non-affixal article languages such as English and affixal article languages such as Bulgarian. I take this to indicate that Italian belongs to a class different from prototypical non-affixal article languages and affixal article languages; i.e., Italian is another type of a DP-language.

Interestingly, (6b) becomes degraded when an adjective modifies *libri*, as shown in (7).

(7) ?*[Di che   scaffale]$_i$ Gianni ha già   letto [i   grandi libri   t$_i$]?

   of which shelf   Gianni has already read   the large   books

   'From which shelf did Gianni read the large books?'

It then seems that the presence of the adjective somehow forces DP to project in the presence of

---

(i) ¿[De qué   cantante]$_i$ salieron publicadas [las fotos   t$_{iOBJ}$]?
     of which singer   were   published   the photos   (Ticio 2005:238)

However, semantically *di che scaffale* 'of which shelf' in (6) is an adjunct, not an argument, of the head noun *libro* 'book', so it seems plausible to assume that *di che scaffale* is adjoined to NP. Interestingly, Spanish does not allow adjunct extraction out of a nominal phrase with a definite article, unlike Italian, as shown in (ii). (See footnote 31 for more on Spanish.)

(ii)*¿[De qué   estantería]$_i$ leyó María [los libros t$_i$]?
     of which shelf   read Maria the books
     'From which shelf did Maria read the books?'

At any rate, what is important here is that Italian shows a different extraction pattern from English and Bulgarian.

the definite article in a case where otherwise DP is not projected.

To summarize this subsection, I have shown that definite articles in Italian exhibit a different behavior from those in, e.g., English and Bulgarian with respect to adjunct extraction out of a nominal domain. In particular, the extraction in question is allowed in the *presence* of definite articles in Italian, in contrast with English, where it is never allowed, and Bulgarian, where it is allowed in the *absence* of definite articles. Based on Bošković's (2005, 2008b, 2012, 2013a) explanation of the extraction in question, I have taken this as indicating that definite articles in Italian need not project DP unlike those in languages that have non-affixal definite articles and languages that have affixal definite articles. This in turn indicates that Italian is a third type of DP-language, which cannot be captured by the NP/DP-language distinctions proposed by Bošković (2008b, 2012) and Talić (2017).

5.2.2 Pronominal and reflexive possessives

Another domain where Italian is of particular interest concerns reflexive and pronominal possessives. Reuland (2011) and Despić (2011, 2015) observe a correlation between availability of reflexive possessives and definite articles. Specifically, languages with prenominal definite articles do not allow reflexive possessives, as represented by English (8a). In contrast, reflexive possessives are allowed in Serbo-Croatian, which lacks a definite article (8b), and Icelandic, which has postnominal (or suffixal) definite articles (8c).

(8) a. He_i loves *$\textbf{himself's}_i$/$\textbf{his}_i$ neighbors.

    b. [Mnogo izbeglica]$_i$ je napustilo $\textbf{svoje}_i$/*$\textbf{njihove}_i$ kucé.

       many   refugees   is left   self's/their   houses

       'Many refugees$_i$ left their$_i$ homes.'   (Serbo-Croatian, Zlatić 1997:243)

    c. Egil$_i$ vantar bókina $\textbf{sína}_i$/*$\textbf{hans}_i$.

       Egil needs book   self's/his

       'Egil$_i$ needs his$_i$ book.'   (Icelandic, Thraínsson 2007:463)

Among the languages in Despić's survey, those with prenominal definite articles are Afrikaans, Dutch, Frisian, English, German, Italian, Misantla, Totonac, Modern Greek, Portuguese, and Spanish. Those with postnominal definite articles are Bulgarian, Danish, Faroese, Icelandic, Koromfe, Macedonian, Norwegian, Romanian, and Swedish. Those without definite articles are Belorussian, Chinese, Czech, Dolakha Newar, Hindi-Urdu, Japanese, Kannada, Kashmiri, Korean, Latin, Lezgian, Malayalam, Mosetén, Old Church Slavonic, Persian, Polish, Proto-Slavonic, Russian, Serbo-Croatian, Slovak, Slovenian, Sorbian, Tamil, Thai, Turkish, and Ukrainian. Despić formulates the generalization as (9).[4] (Note that (9) is a one-way correlation; there can be languages without definite articles or with postnominal definite articles that lack reflexive possessives. What is important is that there is no language that has prenominal definite articles and reflexive possessives.)

(9) *Generalization on reflexive possessives* (Despić 2011:123)

   If a language has reflexive possessives it either does not mark definiteness at all, or it marks definiteness postnominally.

Importantly, Italian is classified as a language that lacks reflexive possessives by Reuland and Despić as mentioned above. However, Italian actually has the so-called possessive adjective *proprio*, which has the interpretation of a reflexive, as shown in (10). I use the term so-called because *proprio* is used with a definite article and inflects based on the gender and number of the noun it modifies, as seen in (10). It should be noted here, though, that pronominal possessives such as *suo*

---

4. Interestingly, those with postnominal definite articles are all affixal article languages (i.e., definite articles are suffixal), except for Koromfe. As for Koromfe, the alleged postnominal definite articles *hoŋ* and *koŋ* are actually not obligatory with definite nouns, as shown in (i).

(i) də   pa   də   gɪllɛ bi   **a salle** **kebre**
     3SG.HUM give 3SG.HUM self child.SG A plate.SG big.SG
     'He gave **the big plate** to his own son.'   (Rennison 1997:109)

Given Bošković's (2016b) definition of definite articles adopted here, in which definite articles obligatorily occur in a nominal phrase with a definite interpretation, the elements in question are actually not definite articles. We can then revise the generalization (9) as (ii). (Again, (ii) is a one-way correlation; there can be languages without definite articles or with affixal definite articles that lack reflexive possessives.)

(ii) *Generalization on reflexive possessives* (revised version)
     Languages that have reflexive possessives either lack definite articles or have affixal definite articles.

The discussion in text is not affected by the choice of the two versions of the relevant cross-linguistic generalization.

(cf. (11)) 'his' also inflect based on the gender and number of the head noun. In fact, Icelandic reflexive possessives also inflect based on the gender and number (as well as case) of the head nouns just like Italian *proprio*, as exemplified by (12) (see Thraínsson 2007 for the complete paradigm of inflection). It is, then, not implausible to analyze *proprio* as a reflexive possessive of the same sort as the ones mentioned above.

(10) a. Mario$_i$ ha letto il   **proprio$_i$**   libro.
          Mario has read the.MASC self's.MASC book(MASC)

    b. Mario$_i$ ha venduto la   **propria$_i$**   macchina.
          Mario has sold   the.FEM self's.FEM car(FEM)

(11) a. Mario$_i$ ha letto il   **suo$_i$**   libro.
          Mario has read the.MASC his.MASC book(MASC)

    b. Mario$_i$ ha venduto la   **sua$_i$**   macchina.
          Mario has sold   the.FEM his.FEM car(FEM)

(12) a. Egil$_i$ vantar bókina   **sína$_i$/*hans$_i$**.
          Egil.ACC needs book(FEM) self's.FEM/her.FEM
          'Egil needs his book.'

    b. Henni$_i$ þykir bróðir   **sinn$_i$/*hennar$_i$**   leiðinlegur.
          her.DAT thinks brother(MASC) self's.MASC/her.MASC boring
          'She finds her brother boring.'

Note now that Italian allows both *proprio* and the pronominal possessive *suo*, unlike other languages that have reflexive possessives, where only the reflexive possessive is allowed (cf. Icelandic (12)).[5] I take this as another piece of evidence that Italian is a DP-language of a type different from both Icelandic and English.

---

5. As shown in (i), *suo* can refer to a noun that does not c-command it. In contrast, as shown in (ii), *proprio* cannot refer to a noun that does not c-command it. This contrast shows that *proprio* is not a pronominal possessive like *suo* but a reflexive possessive.

The distribution of reflexive and pronominal possessives in Italian is also problematic for Despić's (2011, 2015) account of the generalization in (9), which is based on the notion of phases. Following Canac-Marquis (2005), Heinat (2006), Hicks (2009), Lee-Schoenfeld (2004, 2008), Quicoli (2008) among others, Despić assumes that binding domains are reduced to phases, or more precisely, binding is constrained in terms of Spell-Out domains (Chomsky 2000). Under this view of binding, Binding Condition A and B can be formulated as in (13).[6]

(13)   a.   *Binding Condition A* (phase-based)

        Reflexives need to be bound before they are sent to Spell-Out.

      b.   *Binding Condition B* (phase-based)

        Pronouns must not be bound before they are sent to Spell-Out.

Despić also assumes, following Bošković (2005), that DP projects in the nominal domain in languages with definite articles and it constitutes a phase, whereas DP is absent in languages without definite articles and hence there is no phase in the nominal domain.[7] In addition, he proposes that possessors are base-generated in Spec,PossP, which is located between DP and NP in lan-

---

(i) [L'editore   di Mario$_i$] ha venduto bene il   suo$_i$ libro.
    the.publisher of Mario   has sold   well the his   book
    'Mario's publisher sold his book well.'

(ii)*[L'editore   di Mario$_i$] ha venduto bene il   proprio$_i$ libro.
    the.publisher of Mario   has sold   well the self's   book
    'Mario's publisher sold his book well.'

In addition, the pronominal possessive can be bound by a nominal phrase in a higher clause, whereas *proprio* cannot, as seen in (iii). This is additional evidence that *proprio* is not a pronominal possessive like *suo/sua*.

(iii) a. Mario$_i$ ha detto che il telegiornale ha   parlato   della sua$_i$ casa.
      Mario has said that the news   talked about.the his   house.
      'Mario said that the news talked about his house.'

    b.*Mario$_i$ ha detto che il telegiornale ha   parlato   della propria$_i$ casa.
      Mario has said that the news   talked about.the self's house.
      'Mario said that the news talked about his house.'

6. Here I am simplifying the discussion of Binding Condition A and B. See Zlatić (1997) and Despić (2011) for additional factors relevant for binding of pronouns. See also Charnavel and Sportiche (2016) for a discussion of so-called picture nouns under a phase-theoretic approach to Binding Condition A.
7. Under the contextual approach to phasehood advocated by Bošković (2013a, 2014) adopted in this dissertation, NP constitutes a phase in article-less languages. The choice between these two positions does not matter for the argument here.

guages with definite articles, whereas in languages without definite articles, DP does not project above PossP. Thus, in languages with definite articles like English, the possessor is located in Spec,PossP, which is in the Spell-Out domain of the DP phase, as schematized in (14a).[8] In such a configuration, the reflexive is sent to Spell-Out before the external argument, which is the binder of the possessor, is introduced, and hence Binding Condition A in (13a) is not met. On the other hand, the possessive pronoun in the same position observes Binding Condition B, since it is already Spelled-Out when the binder is introduced.[9] In contrast, in languages that lack definite articles, DP does not project above PossP, so that the possessor is not sent to Spell-Out until the binder (i.e., external argument) is introduced into the structure. Thus, when the binder enters into the derivation, the reflexive *svoje* in Serbo-Croatian in (8b) is bound by the binder before being sent to Spell-Out, meeting Binding Condition A (13a), whereas the pronoun *njihove* in (8b) violates Binding Condition B (13b).

(14)   a.

b.
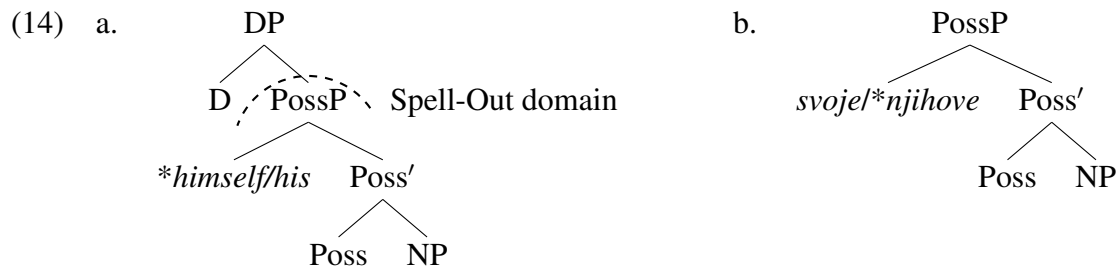
As for languages with postnominal definite articles, Despić (2011, 2015) proposes, building on Fiva (1987) and Delsing (1993), that reflexive and pronominal possessives are base-generated as $Poss^0$ and undergo movement to D^0. This is motivated by the observation that reflexive and pronominal possessives in Scandinavian languages in general are in complementary distribution with non-affixal articles, which are analyzed as D^0, as illustrated in Icelandic (15).[10] In addition,

---

8. Despić (2011, 2015) assumes that there is a phonologically null D head in English possessor constructions. There is, though, another possibility to capture the absence of the definite article in the cases in question. Specifically, I suggest that the genitive marking *'s* is base-generated as D^0, and it undergoes Affix Hopping to $Poss^0$ postsyntactically, just as T undergoes Affix Hopping to v/V in the verbal domain. I will discuss this point in chapter 6.

9. As Despić notes, nothing would change here if pronominal possessors are analyzed as $Poss^0$, since they would still be in the Spell-Out domain of the DP phase.

10. The complemetary distribution of the definite article and (pronominal) possessors is also observed in English. See

Fiva (1987) and Delsing (1993) observe that the reflexive possessive *sin* in Norwegian and the genitive *-s* are in complementary distribution as shown in (16), and propose that *sin* moves from Poss^0 to D^0, which Despić generalizes to other languages with postnominal definite articles. Then, the reflexive and pronominal possessives are outside of the Spell-Out domain of the DP phase as schematized in (17) (note that the head of a phase is not sent to Spell-Out; see Chomsky 2000). Consequently, the reflexive possessives observe Binding Condition A (13a), while the pronominal possessives violate Binding Condition B (13b).

(15) *allar {**hinar þínar/þínar hinar**} þrjár nýju kenningar

      all   the   your/your   the   three new theories   (Thraínsson 2007:117)

(16) a. [mannen med skjegget]**-s**   hus

           man.the with beard.the-GEN house

      b. [mannen med skjegget] **sitt**   hus

           man.the with beard.the self's house

(17)
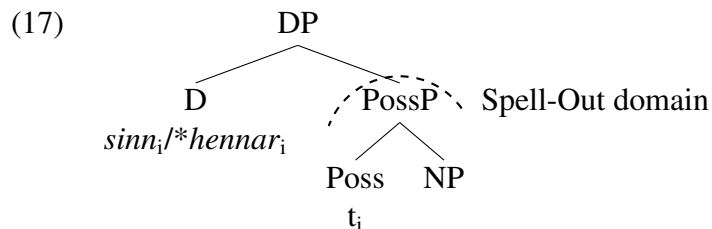
Crucially, the distribution of reflexive and pronominal possessors in Italian cannot be accounted for by Despić's account as it is. Italian is a language with prenominal definite articles, so the possessors should stay in PossP as in (14a). It is, then, predicted that the reflexive possessive *proprio* should not be allowed, contrary to the fact. Note also that we cannot analyze the possessors in question as moving from Poss^0 to D^0 as in (17), since they co-occur with the definite articles unlike those in Scandinavian languages (cf. (10) vs. (15)). Even if they undergo such movement, it would then be mysterious why the pronominal possessor *suo/sua* can be bound by the external

footnote 8 on this.

187

argument in the same clause without violating Binding Condition B (13b). Thus, given the phase-based Binding Condition A and B in (13), which should apply across languages, the distribution of reflexive and pronominal possessives in Italian warrants a different account than those for other languages such as English and Scandinavian languages (as analyzed by Despić).

To summarize so far, I have shown that Italian behaves differently from languages with definite articles discussed by Bošković (2008b, 2012), Talić (2015, 2017), and Despić (2011, 2015) with respect to adjunct extraction out of a nominal phrase and the distribution of reflexive and pronominal possessives. This leads me to conclude that Italian is a different type of DP-language from those in the two-way or three-way cut of the NP/DP-language distinction proposed by Bošković (2008b, 2012) and Talić (2015, 2017), respectively. A question that naturally arises is how we can account for the intermediate behavior of Italian in question. In the next section, I propose a unified account of the distribution of possessors and adjunct extraction out of a nominal domain discussed in the previous subsection. The gist of the proposal is that definite articles in Italian can but need not project DP, the option which I show is not unexpected under the Bare Phrase Structure Theory in minimalism.

5.3 Italian definite articles as clitics

5.3.1 Adjunction of D to N as base-generation

In the traditional DP hypothesis since Abney (1987), definite articles have been considered to project DP in the nominal domain. This position has also been taken by Bošković (2008b, 2012) and Talić (2015, 2017) in order to account for (some of) their generalizations with respect to the NP/DP-language typology. Bošković (2008b, 2012) proposes a two-way cut of NP/DP-languages, which is (mostly) explained by the presence or absence of DP in the nominal domain (which is tied with the existence/lack of definite articles in a given language), as discussed above. For instance, in languages with a definite article, DP projects in the nominal domain, which blocks

adjunct extraction due to the interaction of the PIC and the anti-locality condition.[11] Crucially, for Bošković, DP *always* projects above NP (as the highest projection in the nominal domain) in languages that have definite articles. Talić's three-way cut, which still adopts Bošković's treatment of article-less languages and non-affixal article languages, differs from Bošković's two-way cut in that DP *can* be absent in affixal article languages when a definite article is not required semantically and is absent. Thus, in Bulgarian (5), DP is argued to be absent and hence the relevant extraction is not blocked by the PIC or the anti-locality condition. Crucially, however, in Talić's treatment of definite articles, the presence of a definite article also correlates with the presence of DP in the nominal domain. As discussed in the previous section, this cannot be extended to Italian (6), where a definite article is present but adjunct extraction out of a nominal domain is not blocked. This indicates that the simple presence/absence of DP for affixal article languages under conditions discussed by Talić cannot capture the behavior of Italian, and we thus need a more fine-grained treatment of definite articles in the structure in order to capture the "scale" of NP/DP-language distinction.

It is worth mentioning at this point that the correlation between the presence of a definite article and the presence of DP proposed by Bošković and Talić follows the standard analysis since Abney (1987), as mentioned above, which was originally proposed in the framework of the GB theory. In the GB theory, structure building is based on the X′-Theory. Crucially, in the X′-Theory, structure building proceeds in the "top-down" manner; every lexical item obeys the XP-X′-X^0 frame which is an a priori format of structure building. To put it differently, a lexical item generated as X^0 is forced to project X′ and XP, even if it does not take a complement or a specifier.

In minimalism, however, the X′-Theory is deduced from the Bare Phrase Structure (BPS) (Chomsky 1995a). Under BPS, lexical items are simply bundles of features, and the category-based labels are given merely following the convention; in other words, the category-based labels are not presupposed in the BPS Theory (see also Collins 2002, Seely 2006, Chomsky 2013, 2015

---

11. Recall that many of Bošković's generalizations are one-way correlations, and there are often additional factors that are relevant for those generalizations (e.g., ϕ-agreement between an adjective and a head noun in the case of adjective LBE). In other words, the presence/absence of definite articles/DP is just a prerequisite in those generalizations.

on this). Also, the three levels of projection in the X′-Theory, i.e., the XP-level, the X⁰-level, and the X′-level, have no theoretical status, and are determined contextually; the XP-level is defined as an element that does not project further, the X^0-level as an element that is not projecting, and the X′-level as a non-maximal and non-minimal projection. Structure building thus proceeds in a "bottom-up" manner; non-minimal projections only emerge when a lexical item projects by selecting a complement and/or a specifier. Thus, under BPS, a lexical item need not follow the fixed XP-X'-X^0 schema of the X′-Theory and is not forced to project a non-minimal projection. (There can even be elements that are at the same time heads and phrases.)

With this theoretical background in mind, let us now consider Italian definite articles from a morpho-syntactic perspective. As seen in (18), Italian definite articles are morphologically similar with 3rd person accusative pronominal clitics. It is also worth mentioning that historically, both the definite articles and 3rd person accusative clitics developed from the Latin demonstrative *ille/illa/illud*.

(18) a. Clitics

	Masculine	Feminine
Singular	lo	la
Plural	li	le

b. Definite articles

	Masculine	Feminine
Singular	il, **lo**	**la**
Plural	i, gli	**le**

Laenzlinger (1993) in fact proposes that 3rd person accusative clitics and definite articles in Italian (as well as in French and Spanish) are actually the same elements, which he analyzes as D^0. He proposes that the clitics/definite articles take NP as their complement and project DP, but if NP is absent, they undergo head movement and adjoin to a head in the verbal domain as has been standardly assumed in the literature. Uriagereka (1995) also proposes, based on Spanish, Galician, and French, that clitics and definite articles in Romance are identical as D^0 and project DP in the nominal domain but adjoin to a head via head movement in the verbal domain in the case of the pronominal clitic usage. He also shows that definite articles in Galician can also cliticize onto the verb as shown in (19), and proposes that they can adjoin to the verb.

(19) a. Vimos **o** neno.

we.saw the child

‘We saw the child’

b. Vimo-**lo** neno.

we.saw-the child

c. Vimo-**lo**

we.saw-he

‘We saw him.’   (Galician, Uriagereka 1995:84)

It should be noted here that under their proposal clitics/definite articles *always* project DP in the nominal domain, a lá Abney (1987). Notice, however, that projecting DP in the nominal domain is not the only logical possibility to have a definite article in the structure under BPS. As noted above, under BPS, a lexical item is not forced to follow the X′-theoretic schema, and hence can in principle be a maximal and minimal projection that does not take a complement or a specifier. Interestingly, Chomsky (1995b) in fact proposes that clitics are non-branching elements (i.e., minimal-maximal projection) that are both X^0 and XP at the same time (see also Bošković 2002a for evidence to this effect). As for the options of merger of a head (i.e., a non-branching element) into the structure, it is often assumed that adjunction of a head to another head only takes place via head-movement. This, however, is very different from the standard assumption regarding phrasal adjunction. Phrasal adjunction (i.e., adjunction to XP) is standardly assumed to take place either through movement or base-generation. A question that immediately arises is why head adjunction would also not have both options, i.e., why it would also not take place via base-generation; given the relevant options of phrasal adjunction, it is in fact expected that head adjunction can also take place via base-generation. In fact, as discussed in chapter 4, Epstein et al. (2016) and Saito (2020) propose that a head can also be base-generated adjoined to another head (External Pair-Merge in Epstein et al.’s terminology). Building on this, I propose that the definite articles in Italian can be base-generated adjoined to N, without projecting DP. The two options of

merger of the definite article in Italian are schematized in (20).

(20) a. DP projecting above NP   b. D adjoined to N with NP projected
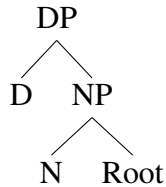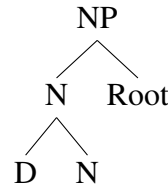
In the next subsections, I demonstrate that this proposal can capture the behavior of Italian definite articles with respect to adjunct extraction out of the nominal domain and reflexive and pronominal possessives.

5.3.2   Adjunct extraction out of a nominal phrase explained

I argue that the above proposal can explain the extraction pattern in Italian discussed in section 5.2.1. Recall that adjunction extraction out of a nominal domain is allowed in Italian only if a definite article is present, as shown in (21), repeated from (6).

(21) a. *[Di che   scaffale]$_i$ Gianni ha   letto [libri   t$_i$]?
          of which shelf   Gianni has read books
          'From which shelf did Gianni read books?'

      b. [Di che   scaffale]$_i$ Gianni ha   già   letto [**i**   libri   t$_i$]?
          of which shelf   Gianni has already read   the books
          'From which shelf did Gianni read the books?'

As discussed in section 5.2.1, under Bošković's (2005, 2008b, 2012, 2013a) proposal, DP blocks the extraction in question due to the interaction of the PIC and the anti-locality condition. As illustrated in (22a), adjunct extraction obeying the PIC violates the anti-locality condition, which requires movement to cross a full phrase. If the extraction obeys the anti-locality condition, it violates the PIC, since NP is not at the edge of the DP phase hence not accessible to a higher

position, as shown in (22b).

(22) a.
```
                DP
               /  \
                   D′
                  / ⌢ \
                 D ˙ NP     Phasal
       ✓PIC        /  \    complement
      ✗anti-locality  Adjunct   NP
```

b.
```
                DP
               /  \
                   D′
                  / ⌢ \
       ✗PIC      D ˙ NP     Phasal
      ✓anti-locality  /  \  complement
                 Adjunct   NP
```
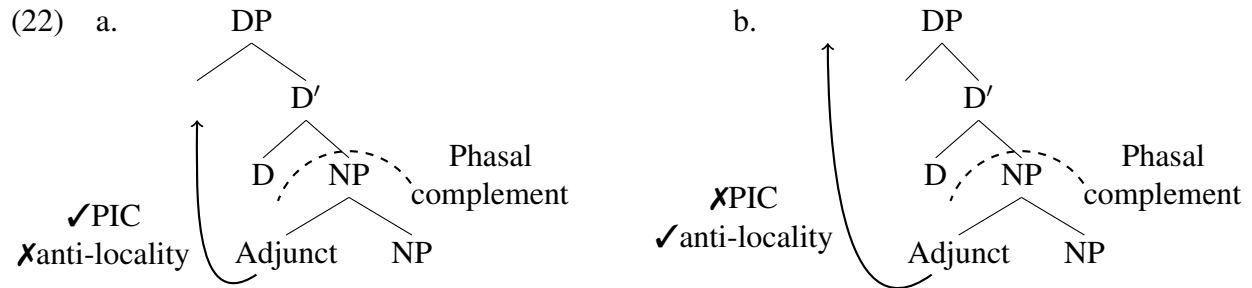
Crucially, under the current proposal, Italian definite articles can be base-generated adjoined to N, with N projecting NP rather than D projecting DP. Then, NP is a phase as the highest projection in the nominal domain, and the adjunct *di che scaffale* 'from which shelf', which is adjoined to NP, can be extracted out of the nominal phrase without violating the PIC or the anti-locality condition, as schematized in (23). Note that the structure in (23) is essentially similar (in the relevant respect) to that in languages without definite articles like Serbo-Croatian, in that DP does not project above NP and the extraction in question is allowed.[12]

(23)
```
                    NP
                  /    \
                NP      PP
               / ⌢ \   di che scaffale     ✓PIC
              N ˙ Root                     ✓anti-locality
             / \      Spell-Out domain
            D   N
            i  (libri)
```
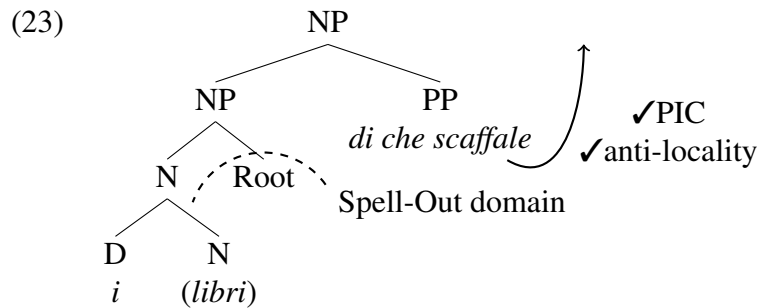
It should be added here that adjunction of D to N in Italian is only possible with the definite article (which is identical to the pronominal clitic). In other cases, DP projects above NP because Italian has definite articles, which is a trigger for DP to project in general, just as in English (see section 5.6.1 for more discussion on this). Thus, when the definite article is absent, as in (21a), DP projects above NP, so that adjunct extraction is blocked, just as in English.

---

12. A question arises regarding the semantic composition of the definite article with the noun and the adjunct PP. See footnote 15 for discussion.

There is a more general point to be made here about phonologically null Ds. It seems that in all languages that have definite articles, if present in the structure (see section 4 and 5 for cases (in Hungarian and Greek) where it is not present), a phonologically null D *always* projects, it is never adjoined to another head in the nominal domain. If this is indeed the case, a possible explanation would be that some PF reflex may be required for the option of head adjunction of D to a nominal head, since without such a reflex we could not tell that D is present in the structure at all given that on the option in question syntactic effects that D might cause would not be there. A phonologically null D would then always project DP, only overt D, which clearly has PF manifestation, could then be head adjoined.

Turning back to Italian, note that the definite article can also project DP in Italian. This option is in fact forced in some cases. Recall now that the extraction in question is degraded when an adjective is present, as shown in (24), repeated from (7).

(24) ?*[Di che   scaffale]ᵢ Gianni ha già   letto [i   grandi libri tᵢ]?
      of which shelf   Gianni has already read   the large   books
      'From which shelf did Gianni read the large books?'

Bošković (2005) explores two possibilities for the structure of a nominal phrase in the presence of an adjective in languages with definite articles. One of the two possibilities is that AP is adjoined to NP, and DP projects above NP and counts as a phase (see also Bošković 2013a). Since Italian is a language with definite articles, it can be assumed that D projects DP above NP in the presence of an adjective. In fact, the definite article *i* precedes the adjective *grandi*, which can be captured given that the adjective is adjoined to NP and the definite article projects DP above NP as Bošković proposes. Under this analysis, the structure of (24) would be something like (25).[13]

---

13. The relative height of the adjective and the PP adjunct in (25) does not matter for the present purpose, since extraction of the PP adjunct would violate the PIC/anti-locality whether it is higher or lower than the adjective as long as they are adjoined to NP.

(25)

```
                    DP
              ┌─────┴─────┐
              D           NP
              i     ┌──────┴──────┐
                    AP            NP
                  grandi    ┌──────┴──────┐
                            NP            PP
                       ┌────┴────┐   di che scaffale
                       N       Root
                    (libri)
```
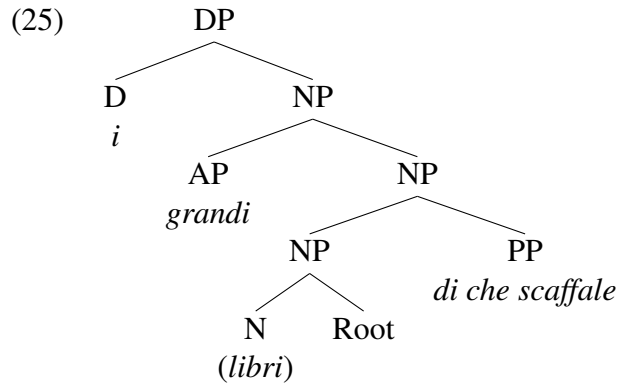
Given this structure, extraction of the adjunct PP either has to violate the anti-locality in order to obey the PIC (i.e., to move to the edge of DP), since the extraction in question would only cross a segment, not a full phrase, or has to violate the PIC in order to obey the anti-locality condition, since the adjunct PP is inside the complement of the DP phase and hence is inaccessible to the structure outside the DP phase. This is schematized in (26). Thus, extraction of an adjunct out of a nominal phrase is blocked in the presence of an adjective.

(26) a.

```
                         DP
                    ┌─────┴─────┐
                              D′
                        ┌──────┴──────┐
                        D         NP ╌╌ Phasal complement
                        i    ┌─────┴─────┐
                            AP           NP
                          grandi   ┌──────┴──────┐
                                   NP            PP
              ✓PIC                           di che scaffale
           ✗anti-locality   ┌────┴────┐
                            N        Root
                         (libri)
```

195

b.

```
                        DP
                       /  \
                     D     NP `` Phasal complement
                     i    / \
                        AP    NP
                      grandi  / \
                            NP    PP
                           / \    di che scaffale
      ✗PIC              N    Root
      ✓anti-locality  (libri)
```
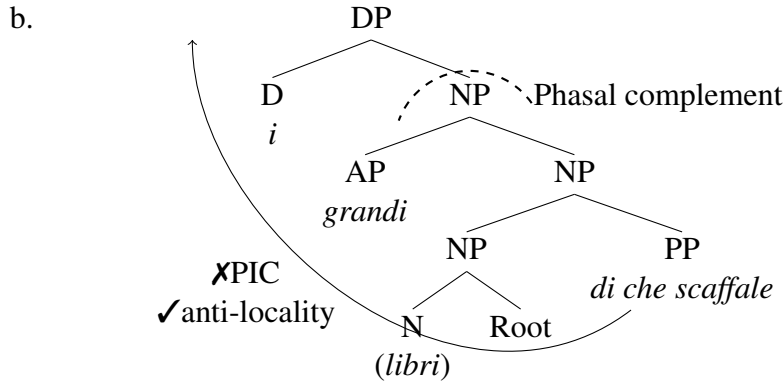
The other possibility Bošković explores, following Abney (1987), is that adjectives project AP between DP and NP and dominate NP in languages that have definite articles, whereas AP is adjoined to NP (or located in Spec,NP) in languages that lack definite articles (see also Despić 2011). Bošković adds that in languages where AP projects above NP, DP must project above AP in order for the entire phrase to function as a nominal argument of a predicate (in other words, AP cannot be an argument of a verb like 'read'). Recall now that under the current proposal, the definite article in Italian *can* be base-generated adjoined to N, but the possibility that it can project DP is not excluded. We can then extend Bošković's proposal to Italian, and assume that AP dominates NP in Italian, and DP projects above AP, as illustrated in (27).

(27)

```
            DP
           /  \
         D     AP
         i    /  \
            A     NP
         grandi  /  \
               N    Root
            (libri)
```
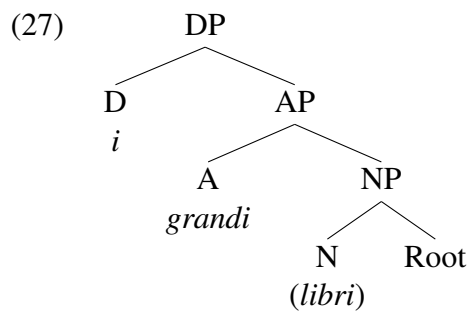
How does this structure disallow adjunct extraction in (24)? My proposal here is that AP constitutes a phase. Under Bošković's (2014) contextual approach to phasehood adopted in this dissertation, the highest projection in the extended projections of a lexical category is a phase. Grimshaw (2000) argues that extended projections are calculated based on the categorial features (see also Biberauer et al. 2014). In Chomsky's (1970) categorial feature specification of lexical

categories, N is [+N, -V] and A is [+N, +V]. Given that D is a functional projection in the nominal domain, it can be considered to be specified as [+N, -V] just like N. Notice now that AP has a different categorial feature specification from N and D. Thus, AP does not count as an extended projection of the nominal domain, and it constitutes a phase under Bošković contextual approach to phasehood.[14] In order to obey the PIC, then, the adjunct, which is adjoined to NP, needs to move to the edge of AP, but this violates the anti-locality condition, since it does not cross a full category. This is illustrated in (28a). On the other hand, if, in order to obey the anti-locality condition, the extraction simply crosses the AP, it will violate the PIC, since the adjunct is not located at the edge of the AP phase, as shown in (28b). Thus, the ban on extraction in question in the presence of an adjective in Italian can also be captured under the assumption that AP projects above NP and DP projects above AP.

(28) a.
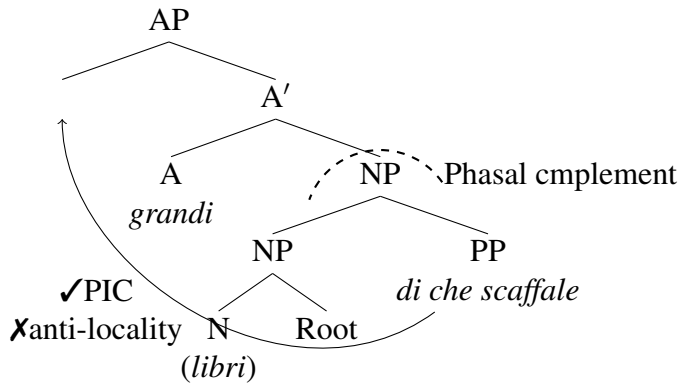
b.

---
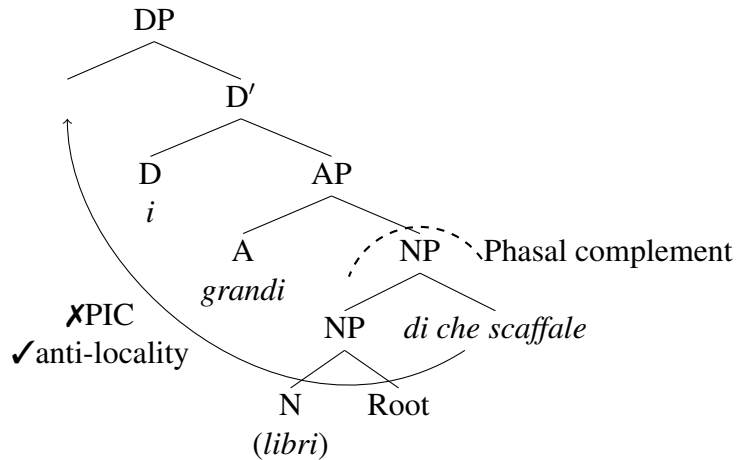
14. DP also constitutes another phase above AP, which is irrelevant here.

The impossibility of adjunct extraction out of a nominal phrase in the presence of an adjective in Italian can thus be captured by the two possibilities in the structure of the nominal phrase in the presence of an adjective in languages with definite articles proposed by Bošković (2005). Although the choice between the two does not matter here, I show in chapter 6 that the structure in which AP projects above NP in (27) is favored, because it can capture other phenomena in which the presence of an adjective plays a crucial role in the presence of definite articles/DP (see also Despić 2011 and Talić 2017 for an argument for this analysis).[15,16]

---

15. A question that arises under the current proposal is why the definite article in Italian cannot be base-generated as adjoined to N in the presence of an adjective. I suggest that this can be explained by the semantic composition of the relevant elements (see Bošković 2012 for relevant discussion). The head noun and the adjective are of type $\langle e,t\rangle$, and the definite article is of type $\langle\langle e,t\rangle,e\rangle$. If the definite article adjoins to the noun in the presence of the adjective, which would be located higher than the complex $\langle D,N\rangle$ head, the definite article would first compose with the noun, yielding type e, and then composition of this complex head with the adjective would yield type t, which cannot serve as an argument for verbs like 'read'. Thus, when an adjective is present, the definite article must merge above AP for semantic reasons, hence cannot adjoin to N.

This, however, raises an issue regarding PP adjuncts in cases like (23). Given that the PP adjunct, which is a restrictive modifier, is of type $\langle e,t\rangle$, just like adjectives, composition of the PP adjunct with the complex $\langle D,N\rangle$ head would result in a type mismatch, so that the adjunction of D to N would be banned. A possible solution to this issue is that the definite article in (23) undergoes covert head movement to avoid the type mismatch. This can be assimilated to Quantifier Raising, which takes place to avoid a type mismatch between a quantifier phrase and its sister in, e.g., *I met every student*, where *every student* undergoes QR.

But then the question is why this covert movement of D cannot take place in the presence of an adjective. My suggestion is that the movement in question is subject to the usual locality restrictions, including the Head Movement Constraint (Travis 1984, Rizzi 1990). Then, when A projects AP above NP, movement of D across A would violate the Head Movement Constraint. Note that this account cannot be extended to the structure in (25), where AP adjoins to NP; in this structure, since A does not project AP above NP, the movement of D would not be blocked. This could be taken to favor the structure in which AP dominates NP over the structure in which NP adjoins to AP, which is consistent with an argument in chapter 6.

16. The definite articles discussed so far have definite interpretation. There are, however, cases where the definite article seems to have no semantic content and exist only for formal reasons. The definite article that occurs with superlatives is one such case (Heim 1999, Sharvit and Stateva 2002). Interestingly, when the adjective in (24) is replaced with the superlative, the sentence improves (24).

(i) ?[Di che scaffale]ᵢ Gianni ha già letto [i grandissimi libri tᵢ]?
　　 of which shelf Gianni has already read the largest books
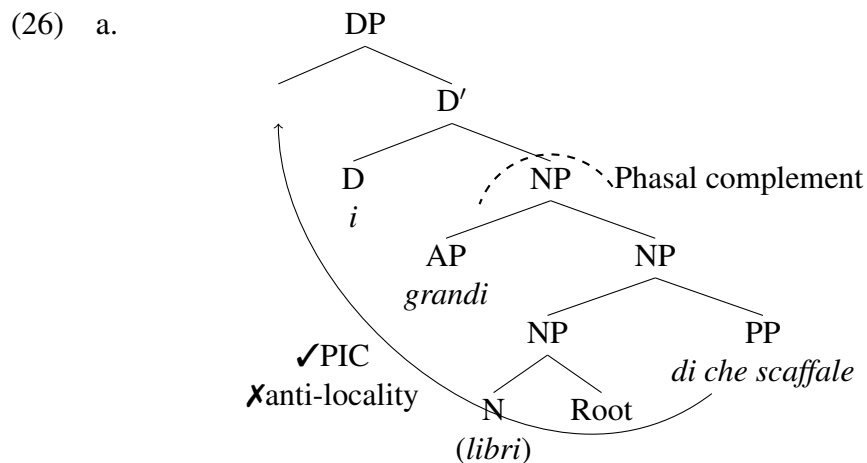　　 'From which shelf did Gianni read the largest books?'

One possibility, which would account for improvement in (i), is that the semantically "expletive" definite article may not project DP (in languages where definite articles in principle do not have to always project DP), unlike definite articles that have definite interpretation (in other words, if the definite article projects, it cannot be expletive). A full investigation of superlatives is left for future research.

Another relevant case of "expletive" definite articles concerns so-called weak definites in cases like *She is listening to the radio*, where *the* does not have definite interpretation. Scholten (2010) shows that in affixal article languages such as Bulgarian, Icelandic, and Romanian, the definite article is absent in weak definite contexts where it is present in English, and Talić (2015, 2017) argues that DP does not project in such cases in those affixal article languages. I leave a full investigation of weak definites for future research.

5.3.3 Reflexive and pronominal possessives explained

In section 5.2.2, I showed that Italian allows both reflexive and pronominal possessives, as seen in (29). This is contrasted with other languages with prenominal definite articles such as English (30a), which allows pronominal possessives but not reflexive possessives, as well as languages that lack definite articles such as Serbo-Croatian (30b) and languages that have postnominal definite articles such as Icelandic (30c), which allow reflexive possessives but not pronominal possessives.

(29)   Mario$_i$ ha letto il   **proprio$_i$/suo$_i$**   libro.
         Mario has read the.MASC self's.MASC/his.MASC book(MASC)

(30)   a.   He$_i$ loves *$\ast$**himself's$_i$/his$_i$** neighbors.

         b.   [Mnogo izbeglica]$_i$ je napustilo **svoje$_i$/*njihove$_i$** kucé.
                many   refugees   is left   self's/their   houses
                'Many refugees$_i$ left their$_i$ homes.'   (Serbo-Croatian, Zlatić 1997:243)

         c.   Egil$_i$ vantar bókina **sína$_i$/*hans$_i$**.
                Egil needs book   self's/his
                'Egil$_i$ needs his$_i$ book.'   (Icelandic, Thraínsson 2007:463)

As discussed in section 5.2.2, Despić (2011, 2015) proposes a phase-based account of the distribution of reflexive possessives. Following Canac-Marquis (2005), Heinat (2006), Hicks (2009), Lee-Schoenfeld (2004, 2008), Quicoli (2008) among others, Despić assumes that the binding domain is defined in terms of phases. I repeat the phase-based formulation of Binding Condition A and B below from section 5.2.2.

(31)   a.   *Binding Condition A* (phase-based)
                Reflexives need to be bound before they are sent to Spell-Out.

         b.   *Binding Condition B* (phase-based)
                Pronouns must not be bound before they are sent to Spell-Out.

In languages with prenominal definite articles such as English, DP projects above PossP, which hosts a possessor, and constitutes a phase. A reflexive cannot be a possessor in such languages, since PossP is the Spell-Out domain of the DP phase and hence the reflexive cannot be bound before it is sent to Spell-Out, as schematized in (32a). In contrast, in languages without definite articles, DP does not project, hence the DP phase is absent. Thus, reflexive possessives are not sent to Spell-Out in the nominal domain and can be bound by an argument in the verbal domain before being sent to Spell-Out, as shown in (32b). Pronominal possessives, on the other hand, violate Binding Condition B in this position, since they are bound by an argument in the verbal domain before they are sent to Spell-Out. As for languages with postnominal definite articles, Despić proposes that reflexive and pronominal possessives are base-generated as $Poss^0$ and undergo movement to D^0, as seen in (33). Thus, they end up being located outside the Spell-Out domain of the DP phase, hence patterning with those in article-less languages.

(32) a.

b.

(33)
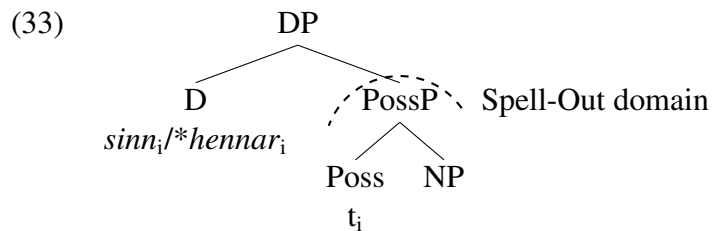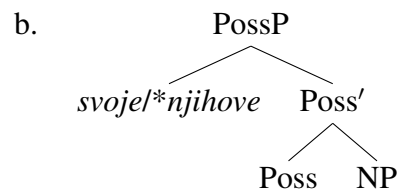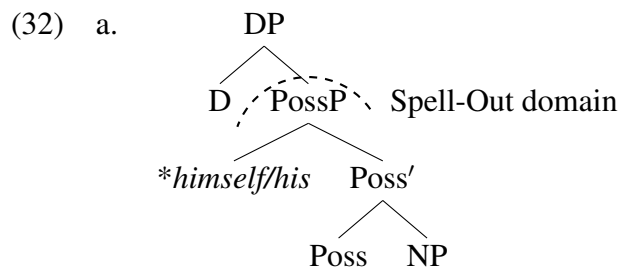
As pointed out in section 5.2.2, Despić's account as it is cannot be extended to Italian, where both reflexive and pronominal possessives are allowed to be bound by a clause-mate argument. Italian is a prenominal definite article language, and hence it would be expected to behave like English, which Despić in fact assumes.

The current proposal can, however, account for the Italian pattern, accommodating Despić's

phase-based account of Binding Conditions. Crucially, under the current proposal, the definite articles in Italian can adjoin to N without projecting DP, as well as merge above NP and project DP. Extending this to possessive constructions, I propose that D can either adjoin to Poss or project DP above PossP. In the former case, PossP becomes a phase in the nominal domain under Bošković's (2014) contextual approach to phasehood, and the reflexive possessive is located outside the Spell-Out domain of the PossP phase, as schematized in (34a). Thus, the reflexive possessive *proprio* obeys Binding Condition A (31a). In the latter case, on the other hand, DP projects above PossP and the pronominal possessive *suo* is inside the Spell-Out domain of the DP phase as shown in (34b), satisfying Binding Condition B (31b).

(34) a.

b.
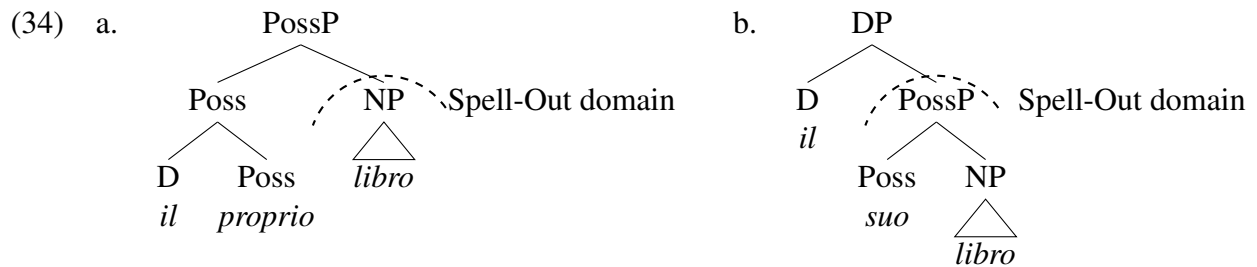
Thus, the availability of both reflexive and pronominal possessives in Italian, which is different from English-type languages where only pronominal possessives are allowed and Icelandic-type languages where only reflexive possessives are allowed (hence indicates the need for a third language type, which is captured by the current system), can be accounted for by the current proposal that definite articles in Italian can adjoin to a head in the nominal domain or project DP under the phase-based account of binding a lá Despić (2011, 2015).

To summarize this section, I have proposed under BPS that Italian definite articles can adjoin to N and Poss in addition to projecting DP. This proposal crucially differs from the treatment of definite articles in the literature, in which projecting DP is taken to be the only option for structure building involving definite articles. I have argued that the current proposal can explain the availability of extraction of an adjunct out of a nominal phrase as well as the possibility of coindexed reflexive and pronominal possessives in the nominal phrases in Italian. In the bigger picture,

empirically we are dealing here with a new pattern that would not fit the previous treatments of the NP/DP-languages distinction, which in turn provides additional evidence for the scale of the NP/DP-language distinction. The proposed account of the behavior of Italian in question captures it by a more fine-grained treatment of structure building involving definite articles, which would otherwise be mysterious under the previous treatment of definite articles in which projection of DP is the only option for structure building for D.

5.4   Hungarian as less of a DP-language

From the viewpoint of the NP/DP-language scale, Hungarian is also worth discussing. Rákosi (2017, 2020) observes that Hungarian allows both personal pronominal possessives and the simplex reflexive possessive *maga* 'oneself' in the presence of the definite article, similarly to Italian, as illustrated in (35a)-(35d). Rákosi also shows that complex reflexives such as *önmaguk* 'themselves' and the reciprocal anaphor *egymás* 'each other' can be used as a possessor reflexive in the absence of the definite article, as shown in (35e)-(35h).

(35) a.   A   fiúk   felfedezt-ék   **az ö**   (kis)   határ-a-i-k-at.
           the boys discovered-3PL the he   little   limit-POSS-PL-3PL-ACC
           'The boys discovered their (little) limits.'

      b. *A   fiúk   felfedezt-ék   **ö**   (kis)   határ-a-i-k-at.
           the boys discovered-3PL he   little   limit-POSS-PL-3PL-ACC
           'The boys discovered their (little) limits.'

      c.   A   fiúk   felfedezt-ék   **a   maguk**   határ-a-i-t.
           the boys discovered-3PL the themselves limit-POSS-PL-ACC
           'The boys discovered their (own) limits.'

d. *A   fiúk felfedezt-ék   **maguk**   határ-a-i-t.

the boys discovered-3PL themselves limit-POSS-PL-ACC

'The boys discovered their (own) limits.'

e. ??A   fiúk felfedezt-ék   **az önmaguk** határ-a-i-t.

the boys discovered-3PL the themselves limit-POSS-PL-ACC

'The boys discovered their (own) limits.'

f.   A   fiúk felfedezt-ék   **önmaguk** határ-a-i-t.

the boys discovered-3PL themselves limit-POSS-PL-ACC

'The boys discovered their (own) limits.'

g. ??A   fiúk felfedezt-ék   **az egymás**   határ-a-i-t.

the boys discovered-3PL the each.other limit-POSS-PL-ACC

'The boys discovered each.other's limits.'

h.   A   fiúk felfedezt-ék   **egymás**   határ-a-i-t.

the boys discovered-3PL each.other limit-POSS-PL-ACC

'The boys discovered each.other's limits.'   (Rákosi 2017:24-25)

What is interesting for the current context is that the reflexive *maguk* is used as a possessor with the prenominal definite article *a*, similarly to Italian *proprio*. Rákosi analyzes this reflexive possessive as an exempt anaphor, essentially treating it as an exceptional case. This is motivated by the observation that the reflexive possessive often has a logophoric interpretation, which is obtained when the relevant reflexive pronoun is embedded in a clause that reports the speech/thought of the matrix subject. However, it is not always logophoric, as seen in (36), where the reflexive *magunk* is not embedded in a clause that reports the speech/thought of the subject.

(36)   Mi csinált-uk **a   magunk** dolg-á-t.

we did-1PL   the ourselves work-POSS-ACC

'We went about our own work.'   (Rákosi 2017:27)

In fact, Rákosi (2017:27, fn. 9) notes that the simple reflexive *maguk* is used as a logophor less frequently than the complex reflexive (e.g., *önmaguk*). In addition, it is not clear why the logophoric usage in possessor constructions is not regularly allowed in other languages with prenominal definite articles like English, where reflexives such as *himself* can otherwise have a logophoric interpretation.

The current proposal that a definite article can adjoin to a nominal head rather than project DP provides another possibility to account for the availability of the reflexive possessive in Hungarian. Notice that both the reflexive and pronominal possessives co-occur with the definite article, just as in Italian, as discussed above. It is thus natural to extend the above proposal for Italian to Hungarian: I propose that the definite article in Hungarian can adjoin to Poss as well as project DP above PossP. When the definite article is adjoined to Poss, PossP becomes a phase and the simple reflexive, which I assume is $Poss^0$, is outside its Spell-Out domain, so Binding Condition A is met. When, on the other hand, the definite article projects DP above PossP, the pronominal possessive is inside the Spell-Out domain of the DP phase, and hence it observes Binding Condition B. Thus, the structure with an anaphor is actually different from the structure with a pronominal possessive in both Hungarian and Italian despite the presence of the definite article, which the proposed system makes possible.

At this point, it is worth mentioning Wang's (2019) work on grammaticalization of indefinite articles in the Beijing dialect of Chinese, Cantonese, Turkish, and Slovenian. Wang proposes that the numeral 'one' in these languages is at an intermediate stage of grammaticalization into an indefinite article. Crucially, at this stage, 'one' is base-generated adjoined to a head in the nominal domain (Cl(assifier) in the Beijing dialect and Cantonese, and N in Turkish and Slovenian) when it is used as an indefinite article, hence does not project its own functional projection. Notice that this is essentially similar to the current proposal on definite articles in Italian and Hungarian; the definite article can be base-generated adjoined to another head, without projecting DP. There is another similarity between the definite articles discussed here and the indefinite articles Wang discusses. In Turkish, *bir* can be interpreted as an indefinite article only when it is adjacent to a

head noun. When it is separated from a head noun by adjectives, it cannot be interpreted as an indefinite article; rather, it is interpreted as a numeral 'one', as shown in (37).

(37) a.   iyi   yeni bir kitap

          good new a   book

          'a good new book'

        * 'one good new book'

    b.   bir iyi   yeni kitap

          one good new book

        * 'a good new book'

          'one good new book'   (Yükseker 2000)

Wang proposes that the indefinite article in Turkish cannot be separated from the head noun because it is adjoined to N; when separated, *bir* can only be used as a numeral. Recall now that the definite article in Italian cannot be adjoined to N (and must project DP) when an adjective modifies the noun, as discussed in section 5.3.2. Thus, abstractly, the presence of an adjective blocks adjunction of the (in)definite article to N in both cases.[17]

   Note here that definite articles in Italian and Hungarian have developed from demonstratives. I suggest that grammaticalization of definite articles in these languages can be analyzed in a similar way as grammaticalization of indefinite articles in the languages discussed by Wang. In particular, the suggestion is that there was an intermediate stage of grammaticalization of definite articles in

---

17. Unlike Turkish *bir*, the Slovenian indefinite article *en* can be used as an indefinite article in the presence of an adjective. (ib) shows that the numeral *dve* 'two' does not yield a specific interpretation. On the other hand, in (ia), *en* can have a specific interpretation, which indicates that it can be used as an indefinite article here, but crucially in the presence of an adjective.

(i) a. Vsi so   kupili eno   novo knjigo.   (specific/non-specific)
         all 3.PL bought a/one.ACC   new   book.ACC
         'They all bought a/one new book.'

     b. Vsi so   kupili dve   novi knjigi.   (*specific/non-specific)
         all 3.PL bought two.ACC   new   book.DU.ACC
         'They all bought two new books.'   (Wang 2019:58)

There are two possibilities to capture the behavior of Slovenian *en*. One is that the grammaticalization of *en* as an indefinite article has proceeded further than that of *bir*, and *en* can project its own functional projection (as well as adjoin to N), similarly to the Italian definite articles. The other is that *en* can adjoin to AP, similarly to the Greek indefinite article as will be argued for in section 5.6.

which the demonstratives in Latin and Old Hungarian were base-generated as head-adjoined to a nominal head. The definite articles in Italian and Hungarian still maintain this option, i.e., head-adjunction as base-generation (the grammaticalization has proceeded further so that the definite articles can also project DP). It is the possible that when the grammaticalization is complete, definite articles do not head-adjoin to a nominal head and only project DP, which is what we find in English (in this respect, it is worth noting that the English definite article has also developed from the demonstrative *se* (or *þe*) in Old English). Wang (2019) in fact argues that grammaticalization of an indefinite article in Mandarin Chinese has proceeded further from the intermediate stage of e.g., Cantonese, and *yi* 'one' as an indefinite article does not head-adjoin to a nominal head but projects its own phrase. Thus, we can have a uniform treatment of grammaticalization of definite and indefinite articles under Wang's proposal, in which head-to-head adjunction as base-generation is crucial.[18]

Let us now return to reflexive possessives in Hungarian. The relevant examples are repeated below. As discussed above, in Hungarian the simple reflexive possessive can be used in the presence of the definite article, as shown in (38a). I have proposed above that the definite articles in Hungarian can adjoin to Poss without projecting DP, so that the simple reflexive possessive can be bound by the subject in a Spell-Out domain in the presence of the definite article in (38a).

(38) a. A   fiúk felfedezt-ék   **a   maguk**   határ-a-i-t.
          the boys discovered-3PL the themselves limit-POSS-PL-ACC
          'The boys discovered their (own) limits.'

      b. ??A   fiúk felfedezt-ék   **az önmaguk** határ-a-i-t.
          the boys discovered-3PL the themselves limit-POSS-PL-ACC
          'The boys discovered their (own) limits.'

---

18. In section 5.6, I will argue that bare NP without a functional projection is the default option of UG, the functional structure above NP being acquired later in language acquisition and grammaticalization. The process of grammaticalization of definite and indefinite articles discussed above can then be taken as additional evidence for the argument that bare NP is the default option of UG.

c. A fiúk felfedezt-ék **önmaguk** határ-a-i-t.

the boys discovered-3PL themselves limit-POSS-PL-ACC

'The boys discovered their (own) limits.'

d. ??A fiúk felfedezt-ék **az egymás** határ-a-i-t.

the boys discovered-3PL the each.other limit-POSS-PL-ACC

'The boys discovered each.other's limits.'

e. A fiúk felfedezt-ék **egymás** határ-a-i-t.

the boys discovered-3PL each.other limit-POSS-PL-ACC

'The boys discovered each.other's limits.' (Rákosi 2017:24-25)

The remaining question is how to capture the complex reflexive in (38c) and the reciprocal anaphor in (38e). Despić (2011, 2015) proposes that English *each other* in *John and Mary saw each other* is (exceptionally) located in Spec,DP, which one might consider extending to (38c) and (38e).[19] It should, however, be noted that the complex reflexive and the reciprocal anaphor do not co-occur with the definite article, unlike the simple reflexive possessive in (38a) and the pronominal possessives in (35a). It is worth recalling in this context that Dubinsky and Tasseva-Kurktchieva (2014) and Talić (2015, 2017) propose that the presence/absence of a definite article correlates with the presence/absence of DP in Bulgarian, an affixal article language (cf. section 5.2.1). As mentioned in chapter 3, MacWhinney (1976) notes that the definite article in Hungarian is treated as a prefix in the traditional Hungarian grammar. It is, then, not implausible to analyze (38c) and (38e) as lacking DP altogether in the absence of the definite article, whereby the complex reflexive and reciprocal anaphor obey Binding Condition A.

This amounts to saying that Hungarian has three options regarding the alleged "DP"; projecting DP by the presence of the definite article, not projecting DP by adjunction of the definite article to Poss, and not projecting DP by omitting the definite article. Note that the first two options are available in Italian, but the third option is not. This indicates that Hungarian is yet another type

---

19. Despić notes that it can be base-generated there or moved from Spec,PossP.

of a DP-language, different from Italian; more generally, Hungarian is yet another type of a DP-language in the NP/DP-language scale. Given the possibility of omission of DP by omission of the definite article, Hungarian is in a sense less of a DP-language than Italian.

The proposal that DP need not project in Hungarian is consistent with the argument regarding compositional indeterminate pronouns in chapter 3. There I established the generalization that languages that have productive compositional indeterminate pronouns either have affixal definite articles or lack definite articles, Hungarian being such a language building on MacWhinney's (1976) treatment. Some of the compositional indeterminate pronouns in Hungarian are given in (39), quoted from Haspelmath (1997).

(39)

	indeterminate	existential	neg-concord	neg-polarity 1	neg-polarity 2
person	ki	vala-ki	sen-ki	akár-ki	bár-ki
thing	mi	vala-mi	sem-mi	akár-mi	bár-mi
place	hol	vala-hol	se-hol	akár-hol	bár-hol
time	mikor	vala-mikor	sem-mikor	akár-mikor	bár-mikor

The gist of the deduction of this generalization I offered in chapter 3 is that DP must project above indeterminate pronouns, which are NPs, in non-affixal article languages whereas it can be absent in affixal article languages and article-less languages, a lá Talić (2015, 2017). Since Hungarian has productive compositional indeterminate pronouns, we were led to analyze Hungarian as lacking DP above compositional indeterminate pronouns. Note also that Italian does not have compositional indeterminate pronouns (see chapter 3 and Haspelmath 1997), which is compatible with the current view that Hungarian is less of a DP-language than Italian.

Yet another property of Hungarian that is interesting in this context is noun-incorporation. Kiefer (1990) and Farkas and de Swart (2003), among others, observe that Hungarian allows bare nominal objects that typically precede the verb, in addition to singular indefinite objects preceded by the indefinite article *egy*, as shown in (40).

(40) a. Éva **újságot**   olvas.

Eve newspaper.ACC read

'Eve is engaged in newspaper-reading.'   (Kiefer 1990:151)

b. Éva olvas egy újságot.

Eve read a   newspaper.ACC

'Eve reads a newspaper.'

Kiefer (1990) observes that the bare singular noun in (40a) is non-referential (and is part of the meaning of the predicate 'newspaper reading'). The non-referential nature of the bare singular noun is shown in (41), where the pronoun *azt* 'it' cannot refer to *újságot* 'newspaper'.

(41) *Jancsi **újságot**$_i$ olvasott és aztán az asztalra tette   (azt$_i$).

Eve   newspaper.ACC read.PAST and then the table.on put.PAST it

'Eve is engaged in newspaper-reading.'   (Kiefer 1990:152)

Kiefer argues that *újságot* in (40a) and (41) is incorporated into V, whereby it is interpreted as part of the complex predicate 'newspaper-reading' (see also Farkas and de Swart 2003 for more discussion of the semantics of noun-incorporation in Hungarian).

Relevant here is that Baker (1988) proposes that noun-incorporation is an instance of head-movement of non-branching N, which would then be impossible if DP projects above N(P), because movement of N to V would be blocked by D due to the Head Movement Constraint (Travis 1984, Rizzi 1990). Relatedly, Baker (1996) observes that polysynthetic languages lack definite articles, and argues that polysynthesis results from noun-incorporation (see also Bošković 2008b for discussion). Given that noun-incorporation is only possible in the absence of DP, the availability of noun-incorporation may well be a cue for a learner of Hungarian not to generalize projection of DP triggered by the presence of the definite article to indefinite nominal phrases.

Thus, Hungarian can be classified as less of a DP-language in the scale of the NP/DP-language distinction than English, in fact even less so than Italian discussed in section 5.3 too. DP can be absent with the definite article dropped or adjoined to a nominal head rather than projecting

above NP, and indefinite nominal phrases are also allowed to be bare NPs. Again, this would be mysterious under the two-way or three-way categorical "cut" of the NP/DP-languages distinction in the previous literature.

5.5   Greek as less of a DP-language

Greek is another DP-language of particular interest from the perspective of the NP/DP-language scale argued for here. In this section, I argue that Greek is yet another type of DP-language, which is different from both Italian and Hungarian in the scale of the NP/DP-language distinction. In particular, I show that Greek exhibits a number of properties of NP-languages, such as adjunct extraction, Left Branch Extraction (LBE), semi-productive compositional indeterminate pronouns, bare singulars, and null objects with sloppy readings. Crucially, these are possible only in the absence of a definite article; when a definite article is present, Greek behaves like a DP-language. I thus propose that DP does not project above NP in indefinite nominal phrases, while it projects above NP in the presence of the definite article. I take this as indicating that Greek is less of a DP-language than English in that bare NP is allowed in Greek, but also it is a different type of a DP-language than Italian or Hungarian in that the definite article always projects DP above NP. It is also pointed out in this section that Greek allows LBE when a definite article is present but the article is left-dislocated together with an adjective. I propose an account of this in connection with a particular type of LBE out of PP observed in Serbo-Croatian. Moreover, I suggest two possible analyses of the indefinite article in Greek. On one analysis, the indefinite article in Greek can be base-generated as adjoined to another head, on a par with the indefinite articles in the languages discussed by Wang (2019) in the context of grammaticalization of indefinite articles. On the other analysis, the traditional indefinite article in Greek is actually an adjective adjoined to NP, just like the numeral 'one' in Serbo-Croatian. I show that both analyses can capture the paradigm of extraction out of a nominal phrase in the presence of the indefinite article.

5.5.1 Adjunct extraction

Greek has definite articles, and has been treated as a DP-language by Bošković (2008b, 2012) in his two-way cut of NP/DP-language distinction. Interestingly, however, Alexopoulou and Folli (2019) observe that Greek allows extraction of an adjunct out of a nominal phrase, like NP-languages such as Serbo-Croatian and unlike canonical DP-languages such as English, as shown in (42).

(42)   [Apo pia   poli]$_i$ ghnorise [koritsia t$_i$] o   Petros?
         from who city   met.3SG girls   the Petros
       'Petros met girls from which city?'   (Alexopoulou and Folli 2019:464)

Note that the indefinite nominal phrase in (42) is plural and hence is not accompanied by an indefinite article. Interestingly, when the indefinite article is present, the extraction in question is also possible, as shown in (43).

(43)   [Apo pia   poli]$_i$ ghnorise [ena koritsi t$_i$] o   Petros?
         from who city   met.3SG a   girl   the Petros
       'Petros met a girl from which city?'

It should be added here that (42) becomes bad when a definite article is present, as seen in (44).

(44) *[Apo pia   poli]$_i$ ghnorise [**ta**   koritsia t$_i$] o   Petros?
         from who city   met.3SG the girls   the Petros
       'Petros met the girls from which city?'

This indicates that there is a contrast between definite and indefinite nominal phrases regarding adjunct extraction. This is further confirmed by the following examples, where the head noun is modified by an adjective:

(45)   a.   [Apo pia   poli]$_i$ ghnorise [psila koritsia t$_i$] o   Petros?
               from who city   met.3SG tall girls   the Petros
             'Petros met tall girls from which city?'

b.   [Apo pia   poli]ᵢ ghnorise [ena psilo koritsi tᵢ] o   Petros?

   from who city   met.3SG a   tall   girl   the Petros

   'Petros met a tall girl from which city?'

c. *[Apo pia   poli]ᵢ ghnorise [ta   psila koritsia tᵢ] o   Petros?

   from who city   met.3SG the tall   girls   the Petros

   'Petros met the tall girls from which city?'

In Greek, demonstratives are accompanied by a definite article. As expected, adjunct extraction out of a nominal phrase with a demonstrative is disallowed, as shown in (46).

(46) *[Apo pia   poli]ᵢ ghnorise [afto to   koritsi tᵢ] o   Petros?

     from who city   met.3SG this the girl   the Petros

     'Petros met this girl from which city?'

Possessive constructions in Greek can be definite or indefinite. Here again, adjunct extraction under discussion is allowed only with indefinite possessums, as shown in (47).

(47) a.   [Apo pia   poli]ᵢ ghnorise [filus   mou tᵢ] o   Petros?

       from who city   met.3SG friends my   the Petros

       'Petros met a friend of mine from which city?'

b.   [Apo pia   poli]ᵢ ghnorise [ena filo   mou tᵢ] o   Petros?

     from who city   met.3SG a   friend my   the Petros

     'Petros met a friend of mine from which city?'

c. *[Apo pia   poli]ᵢ ghnorise [ton filo   mou tᵢ] o   Petros?

     from who city   met.3SG the friend my   the Petros

     'Petros met my friend from which city?'

Thus, one may conclude that whether adjunct extraction out of a nominal phrase is possible correlates with whether the nominal phrase is definite or indefinite.

212

However, it turns out that it is not the semantics of the nominal phrase but the presence/absence of the definite article that matters for the extraction in question. In Greek, a generic interpretation is obtained in the presence of a definite article (Giannakidou 2012). As shown in (48), adjunct extraction out of a nominal phrase that has the generic interpretation is disallowed, although the definite article does not have the usual definite interpretation.

(48) *[Apo pia poli]$_i$ poula [**ta** proionta t$_i$] o Petros?
      from who city sell.3SG the products the Petros
      'Petros sells products from which city?'

Another interesting case concerns superlatives. In footnote 16, I showed that adjunct extraction out of a nominal phrase with a superlative adjective is possible in Italian, based on which I suggested that the definite article that appears with superlatives does not project DP in Italian. The relevant example is repeated here as (49). Importantly, adjunction extraction in such environments is impossible in Greek, as shown in (50).

(49) ?[Di che scaffale]$_i$ Gianni ha già letto [i grandissimi libri t$_i$]?
      of which shelf Gianni has already read the largest books
      'From which shelf did Gianni read the largest books?'

(50) *[Apo pia poli]$_i$ ghnorise [ta pio psila koritsia t$_i$] o Petros?
      from who city met.3SG the SUPL tall girls the Petros
      'Petros met the tallest girls from which city?'

Thus, the generalization that we obtain here is that adjunct extraction is disallowed in the presence of the definite article in Greek, regardless of its interpretation. Given Bošković's analysis of the extraction in question discussed in section 5.3.2, in which DP essentially blocks the extraction in question, this can be taken as indicating that in Greek, DP is projected when the definite article is present, while it is not projected when the definite article is absent.

5.5.2   Left Branch Extraction

Another relevant case concerns Left Branch Extraction (LBE) of adjectives out of a nominal phrase. As noted above, Uriagereka (1988), Corver (1990), and Bošković (2005) establish the generalization that adjective LBE is allowed only in languages that lack definite articles, as exemplified by English (51) and Serbo-Croatian (52).

(51) *Expensive$_i$/Those$_i$ he bought [t$_i$ cars].

(52) Skupa$_i$/Ta$_i$ je video [t$_i$ kola].
      expensive/that is seen   car

Bošković (2005) proposes that in article-less languages, DP does not project above NP, hence AP, which is adjoined to NP, can move out of the nominal phrase without violating the PIC or the anti-locality condition, as schematized in (53).

(53)

✓PIC
✓anti-locality

NP
  AP NP
      N Root Phasal complement
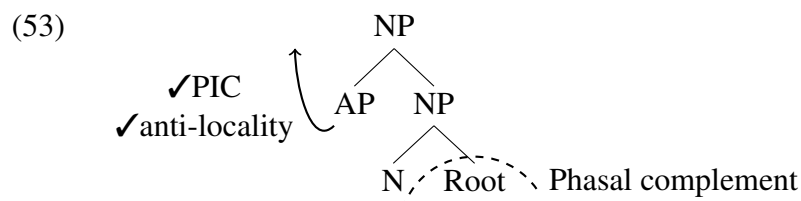
Regarding languages with definite articles, Bošković offers two possible accounts, which were discussed in section 5.3. One is to assume that adjectives are universally adjoined to NP. In languages with definite articles, DP projects above NP and hence extraction of an adjective violates either the PIC or the anti-locality condition, just like adjunct extraction out of a nominal domain as discussed above. This is schematized in (54). The other is to assume that in languages with definite articles AP projects above NP, as discussed in section 5.3.2, whereas they are adjoined to NP in languages without definite articles. Under this account, AP cannot move to the exclusion of NP in languages with definite articles, so that adjective LBE is blocked. This is illustrated in (55).

(54) a.

b.

(55)
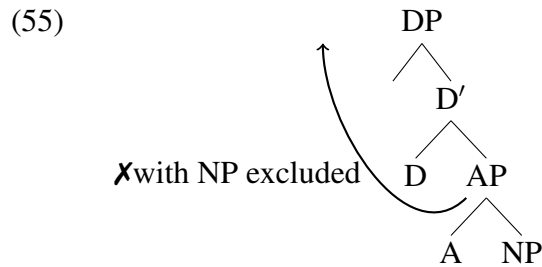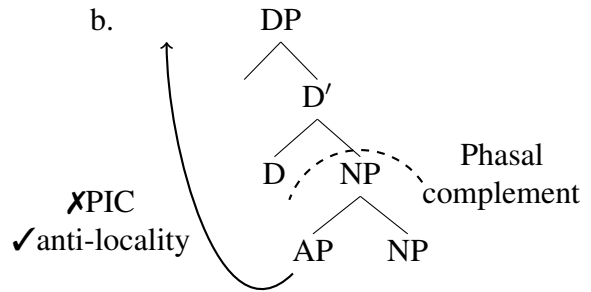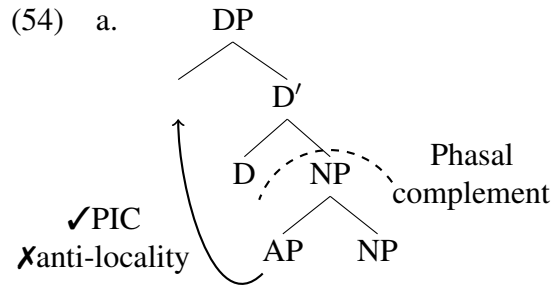
Interestingly, Alexopoulou and Folli (2019) show that Greek allows adjective LBE, although it has definite articles. As shown in (57), the adjective modifying the noun is fronted with contrastive focus (indicated by the small capitals). Note also that the definite article is absent here.

(56) Aghorase [akrivo aftokinito].
bought.3SG expensive car
'She bought an expensive car.' (Alexopoulou and Folli 2019:444)

(57) AKRIVO$_i$ aghorase [t$_i$ aftokinito].
expensive bought.3SG car
'He bought an expensive car.' (Alexopoulou and Folli 2019:463)

In contrast, when the definite article is present, LBE is disallowed, as shown in (59).

(58) Idha [to kokkino forema].
saw.1SG the red dress
'I saw the red dress.'

(59) *[KOKKINO]$_i$ idha [**to** t$_i$ forema].

   red   saw.1SG the dress

   'I saw the red dress.'

It should be added here that the left-dislocated adjective in (57) is not base-generated in the fronted position. As shown in (60a), the left-dislocated modifier and the head noun can be separated by a finite clause boundary. Crucially, as seen in (60b), they cannot be separated by a Complex NP island, which indicates that the modifier is indeed extracted from the nominal phrase inside the island.

(60) a. [AKRIVO]$_i$ mou ipes oti aghorase [t$_i$ aftokinito].

       expensive me(CL) said.2SG that saw.2SG car

       'It is an EXPENSIVE car that you told me that you saw.'

   b. *[AKRIVO]$_i$ ghnorisa [$_{island}$ ti ghineka pou forese [t$_i$ aftokinito]].

       expensive met.1SG   the woman that wore   car

       'It is an EXPENSIVE car that I met the woman who bought.'

Thus, we can conclude that Greek allows LBE in the absence of the definite article but disallows it in the presence of the definite article. Given Bošković's account of LBE mentioned above, this can be taken as indicating that DP does not project above NP in the absence of the definite article, whereas it projects above NP in the presence of the definite article.

Before proceeding, I would like to note here that adjective LBE is possible in the presence of a definite article when the definite article is accompanied by a focused adjective, as shown in (61).[20]

---

20. I acknowledge that there is speaker variation regarding the judgment of cases like (61) (see Bošković 2012). The discussion in the text is based on the grammar of the speakers who accept the relevant constructions. As mentioned in footnote 35, Macedonian and Bulgarian, which have affixal definite articles, also exhibit speaker variation regarding the availability of LBE. The speaker variation is actually not unexpected from the perspective of the argument in this chapter; some speakers of these languages acquire the grammar of more of a DP-language, in which LBE is not possible, whereas other speakers acquire the grammar of less of a DP-language, in which LBE is allowed. This speaker variation can be analyzed as arising from an interaction of the primary linguistic data and third factor principles that instruct the learner to minimize the structure and to generalize a parameter value from one domain to other domains. See section 5.6 for discussion of parameter setting in connection to the NP/DP-language scale advocated here.

(61) [To KOKKINO]ᵢ idha   [tᵢ forema].

the red   saw.1SG   dress

'It is the RED dress that I saw.'   (Androutsopoulou 1998:1)

Note that the extraction in question is also sensitive to islands, as shown in (60).[21]

(62) a. [To KOKKINO]ᵢ mou   ipes   oti idhes   [tᵢ forema].

the red   me(CL) said.2SG that saw.2SG   dress

'It is the RED dress that you told me that you saw.'

b. ??[To KOKKINO]ᵢ ghnorisa [island ti   ghineka pou forese [tᵢ forma]].

the red   met.1SG   the woman that wore   dress

'It is the RED dress that I met the woman who wore.'   (Androutsopoulou 1998:5)

In addition, a nominal phrase with the indefinite article patterns with a nominal phrase with the definite article with respect to LBE. As shown in (63), LBE is impossible when the indefinite article stays in the nominal phrase, but it becomes possible when the indefinite article moves with the left-dislocated adjective.

(63) a. *AKRIVOᵢ aghorase   [ena tᵢ aftokinito].

expensive bought.3SG a   car

'He bought an expensive car.'

b. [Ena AKRIVO]ᵢ aghorase   [tᵢ aftokinito].

a   expensive bought.3SG   car

'He bought an expensive car.'

The distribution of LBE in Greek is thus more complicated than that of adjunct extraction discussed in the previous subsection; LBE is possible (i) when neither the definite nor indefinite article is present, and (ii) when the left-dislocated adjective is dislocated together with the definite

---

21. Androutsopoulou (1998) gives '??' to (62b), but since the island in (62b) is a strong island, it is expected to be much worse than '??'. What is important here is that there is a contrast between (62a) and (62b) (though it is not clear why (62b) is not worse than '??').

or indefinite article. (i) can be explained by the proposal that DP is absent when the definite (or indefinite) article is absent, but (ii) requires a more fine-grained treatment of the definite and indefinite articles. I will discuss LBE in the presence of the definite article in section 5.5.5 and LBE in the presence of the indefinite article in section 5.5.6.

5.5.3 Indeterminate pronouns

Another observation relevant for the NP/DP-language status of Greek is that Greek has semi-productive compositional indeterminate pronouns, as seen in (64).

(64)

	indeterminate	existential	neg-polarity	free choice
person	pjos	ká-pjos	kanénas, kanís	o-pjos-dhípote
thing	ti	ká-ti	tí-pota	o-ti-dhípote
place	pu	ká-pu	pu-thená	o-pu-dhípote
time	póte	ká-pote	poté	o-pote-dhípote

Note that what is derived from compositional indeterminate pronouns are *indefinite* pronouns. Given that productive compositional indeterminate pronouns are allowed only in languages where DP can be absent as argued in chapter 3, the semi-productivity of compositional indeterminate pronouns in Greek can be taken as another case where Greek allows bare NPs that are not dominated by DP when the NPs are indefinite.

It is worth adding here that universal quantifiers are missing in the paradigm of the indefinite pronouns in Greek (this contrasts with NP-languages such as Japanese and Serbo-Croatian, where universal quantifiers are derived from compositional indeterminate pronouns). For instance, the Greek universal quantifier that corresponds to 'everyone' is *kathé-nas*, which is composed of *kathé* 'every' and *énas* 'one', similarly to English *everyone*. Crucially, Giannakidou (2012) notes that *kathé-nas* must be accompanied by the definite article, as shown in (65).

(65) a. **O** kathé-nas éfere apó éna vivlío.

the everyone brought of one book

'Everyone brought one book each.'

b. *Kathé-nas* éfere apó éna vivlío.   (Giannakidou 2012:310)

This can be taken as indicating that DP projects above NP when the definite article is present, which blocks the indeterminate pronoun from being used as an indefinite pronoun (see chapter 3 for relevant discussion). On the other hand, when the definite article is absent, indeterminate pronouns are not dominated by DP and hence can be composed with a quantificational particle and used as indefinite pronouns, as seen in (64). Thus, the paradigm of the indefinite pronouns can be taken as additional supportive evidence that the presence/absence of the definite article correlates with the presence/absence of the DP layer in the nominal domain in Greek.[22]

It should, though, be noted that the free choice items in Greek, which are based on compositional indeterminate pronouns, contain *o*, which is a definite article. A question that naturally arises

---

22. Relatedly, the definite article in Greek can also occur with the distributive quantifier 'every/each', as shown in (ia). The same observation holds for Basque (ib) and Bulgarian (ic), which have affixal definite articles and hence are less of DP-languages (see also section 5.7 for Basque). In Hungarian, the distributive quantifier *valamennyi* 'each' requires the definite article to co-occur (only) when there is an element that precedes it, as seen in (ii).

(i) a. **o** kathe fititis
the every student
'each student'   (Greek, Giannakidou 2004:121)

b. mutil guzti-**a**
boy each-the
'each boy'   (Basque, Etxeberria 2005:42)

c. vsički-**te** momčeta
every-the boys
'each boy'   (Bulgarian, Giannakidou et al. 2009)

(ii) a. (*__**A**) valamennyi [töl-ed kapott] levél rövid volt.
the each from-2.SG received letter short was
'Each letter received from you was short.'

b. *(__**A**) [töl-ed kapott] valamennyi levél rövid volt.
the from-2.SG each received letter short was
'Each letter received from you was short.'   (Hungarian, Szabolcsi 1994:210)

Giannakidou (2004) and Etxeberria and Giannakidou (2010) propose that the definite articles in (ia) and (ib) lack definite interpretation (i.e., do not function as the iota operator) but add a presupposition on the domain of the distributive quantifier (this can be extended to Bulgarian (ic)).

here is why the definite article does not seem to project DP in the case of free choice items, given that indefinite pronouns that are based on compositional indeterminate pronouns are not dominated by DP. Giannakidou and Cheng (2006:151) suggest that "[the free choice item in Greek] is a lexical unit without being semantically or morphologically decomposed in a strict compositional way from all its parts". Greek also has free-choice free relative pronouns, which are identical to the indefinite free choice items and can be used in unconditional clauses (cf. *whatever you bring, I will taste it.*). There are also languages such as Hungarian and Japanese where free choice items and unconditionals use the same morphological ingredients (see Szabolcsi 2019 for Hungarian and Oda 2021b for Japanese). Oda actually suggests that the free choice item in Japanese is a grammaticalized form of the unconditional, which is similar to Giannakidou and Cheng's suggestion mentioned above. Note also that free choice items tend to be indeterminate pronoun-based cross-linguistically regardless of the NP/DP-language status (cf. English *wh-ever*; see also Haspelmath 1997 for cross-linguistic data). It may then not be implausible that free choice items have different internal structure than other indefinite pronouns hence can contain a definite article because they are grammaticalized from unconditionals, which have a clausal structure.

5.5.4   Sloppy reading of null arguments and bare nominals in Greek

The current proposal that Greek allows bare NP in the absence of a definite article can also be supported by the availability of null indefinite objects in this language. Dimitriadis (1994), Giannakidou and Merchant (1997), Panagiotidis (2002), Tsimpli and Papadopoulou (2006) observe that Greek allows null objects, as illustrated in (66a), where the underline indicates the null object. Crucially, the null object in (66a) can refer to a set of dishes that is different from the set of dishes Napoleodas washed, which is the so-called sloppy reading. This is contrasted with clitics, which only have the strict reading (i.e., they refer to the same individual as their antecedent), as seen in (66b).

(66) a. O Napoleodas epline   pjata ke i Nafsika skupise   __.

the Napoleodas washed.3SG dishes and the Nafsika dried.3SG

"Napoleodas washed dishes and Nafsika also dried ones."   (sloppy reading)

b. O Napoleodas epline   pjata_i ke i Nafsika ta_i   skupise.

the Napoleodas washed.3SG dishes and the Nafsika them.CL dried.3SG

"Napoleodas washed dishes and Nafsika dried them."   (strict reading)

(Adapted from Alexopoulou and Folli 2019:477)

As Tomioka (2003) points out, Greek indefinite null objects are similar to null arguments found in languages like Japanese with respect to the availability of the sloppy reading. In Japanese (67), the null argument in the second clause (indicated by the underline) can refer to a car that is different from the car Mary bought.

(67)   Mary-wa   kuruma-o katta. John-mo __ katta.

Mary-TOP car-ACC   bought John-also   bought

'Mary bought a car. John also bought one.'   (sloppy reading)

Interestingly for the current context, following up on Bošković (2012), Cheng (2013) observes that null arguments of the Japanese type are available in Japanese, Korean, Chinese, Turkish, and American Sign Language, but unavailable in English, French, German, Dutch, Spanish, Italian (see also Takahashi 2013 and Sato 2014, 2015 for more languages; cf. footnote 31 for Spanish). Note that the former lack definite articles, whereas the latter have definite articles. Cheng thus establishes the generalization (68) (adapted from his original formulation).[23]

(68)   Null arguments of the Japanese type is available only in languages that lack definite articles.

There are various analyses of null arguments of the Japanese type in the literature; *pro* (e.g., Hoji 1998, 2003, Kurafuji 1999, Tomioka 2003) or ellipsis (e.g., Otaki 2014, Takahashi 2020, Oku 1998,

---

23. Note that this is a one-way correlation; there are languages that lack definite articles and disallow the sloppy reading. See, e.g., Saito (2007), Takahashi (2013), Sato (2014, 2015) for relevance of agreement for the availability of the relevant reading. See also Bošković (2018b) for a perspective which combines the two factors in question (articles and agreement).

Saito 2007, Sakamoto 2017, 2019, Takahashi 2008a,b).[24] It is beyond the scope of this dissertation to fully discuss the issue of null arguments (but see below for some discussion). What is important here is that the licensing mechanism of null arguments with the sloppy reading that is supposed to be available only in article-less languages is available in Greek, which is surprising under Cheng's generalization that is based on Bošković's two-way cut of the NP/DP-language distinction.[25]

The availability of the sloppy reading of null objects in Greek is not unexpected from the perspective of the current proposal on Greek nominal phrases, and more generally, the more fine-grained "scale" of the NP/DP-language distinction argued for here. I have argued that Greek nominal phrases may lack DP; in particular, DP is absent in the absence of the definite article. Thus, Greek is analyzed as allowing bare NPs in such cases similarly to article-less languages like Japanese. It is worth mentioning here that Tomioka (2003) proposes that null arguments in Japanese and Greek under discussion are *pro* NPs of type $\langle e,t \rangle$, which lack the DP layer that would type-shift them to type e. Bošković (2017, 2018b) also proposes under the LF-copying analysis that only elements of type $\langle e,t \rangle$, including bare NPs that lack the DP layer, can be null arguments in question, and he in fact suggests that this can be extended to Greek null objects.[26] In either approach, Greek allows null objects of type $\langle e,t \rangle$ because of the lack of the DP layer that would function as a type-shifter in the case of indefinite nominals. Importantly, unification of Greek null objects and null arguments of the Japanese type makes sense under the current view of the scale of the NP/DP-language distinction; Greek is less of a DP language, which may allow bare NPs including indefinite null objects, unlike canonical DP languages such as English, where DP always projects.

It needs to be added, though, that the relevant null arguments are not as productive in Greek as

---

24. There are two major lines of ellipsis approaches to null arguments of the Japanese type; PF-deletion (Otaki 2014, Takahashi 2020) or LF-copying (Oku 1998, Saito 2007, Sakamoto 2017, 2019, Takahashi 2008a,b). The difference does not matter for the present purposes.

25. There are certain conditions that constrain availability of null objects in Greek. See Giannakidou and Merchant (1997) and Tsimpli and Papadopoulou (2006) for discussion. What is important here is that the relevant null argument is *in principle* available in Greek, which is not expected by the two-way cut of the NP/DP-language distinction (but see Bošković 2018b).

26. It is also worth mentioning that Giannakidou and Merchant (1997) propose an LF-copy analysis of Greek null objects.

in Japanese, which is reasonable since Greek is still a DP-language to some extent. Importantly, Giannakidou and Merchant (1997) observe that when the antecedent is a definite nominal phrase, a null object that refers to it is not allowed; instead, a clitic pronoun must be used, as shown in (69).

(69) a. A: Efere   o   Andreas [ta   vivlia]$_i$?

          brought.3SG the Andreas the books?

          'Did Andreas bring the books?'

      b. B: Ne, ***(ta)**   efere.

          yes   them brought.3SG

          'Yes, he brought them.'   (Giannakidou and Merchant 1997:142)

This contrasts with Japanese, where a null argument can refer to a semantically definite antecedent.

(70) John-wa   kono-hon$_i$-o   yonda. Mary-mo   __$_i$ yonda.

      John-TOP this-book-ACC read   Mary-ALSO   read

      'John read this book. Mary read [it] too.'

Given that, in Greek, only bare NPs that lack the DP layer can be null arguments of the Japanese-type, the unavailability of the null object in (44) can be attributed to projection of DP in the presence of the definite article.[27] Again, this makes sense if Greek is between English and Japanese in the NP/DP-language scale as argued for here; it is less of a DP-language than English in that it allows bare NP in the absence of a definite article, but still a DP-language to some extent, in particular in the presence of a definite article.[28]

---

27. Dimitriadis (1994) observes that a null object with sloppy reading is possible in the presence of the indefinite article.

(i) a. Echis   ena taliro?
         you.have one nickel
         'Do you have a nickel?'

     b. (*To) echo   __
            it   I.have
         'I have (one).'   (Dimitriadis 1994)

This makes sense under the current proposal that DP does not project above NP in the absence of the definite article; the antecedent of the null object is an indefinite nominal phrase, which is a bare NP, so that it can be referred to by the null object (for relevant discussion regarding American Sign Language, see Koulidobrova 2017).

28. Related to this, Greek allows clitic doubling, but only in the presence of the definite article with the doubled

Alexopoulou and Folli (2019) note that Italian is different from Greek when it comes to null objects. Italian does not allow null indefinite objects, as shown in (71).

(71) *Gianni sta cercando   un idraulico ma non trova   __.
     Gianni is looking.for a plumber but not find.3SG
     'Gianni is looking for a plumber, but cannot find one.'

<div align="right">(Adapted from Alexopoulou and Folli 2019:474)</div>

This is expected under the current proposal, under which Italian nominal phrases without a definite article are DPs (see section 5.3). Indefinite nouns in Italian are type-shifted from type $\langle e,t \rangle$ due to the presence of the DP layer, and hence cannot be a null argument in Tomioka's or Bošković's approaches to null arguments of the relevant type. Actually, Alexopoulou and Folli (2019) also propose that Italian nominal phrases are DPs while Greek nominal phrases are smaller than DP (NumP for them), and that Tomioka's (2003) analysis of null arguments can be extended to Greek.[29] Another relevant observation made by Alexopoulou and Folli is that Greek allows bare singular nominals, whereas Italian does not, as shown in (72) (see also Longobardi 1994 on Italian).

---

nominal phrase (Anagnostopoulou 1994). Bošković (2008b, 2012) establishes the generalization that only languages with definite articles may allow clitic doubling, and offers a deduction of it in which the clitic and the doubled nominal phrase Agree with respect to the D-feature, which is present only in languages with definite articles. Building on this, I hypothesize that in Greek only the definite article bears the D-feature and can Agree with the clitic, and in the absence of the definite article, the D-feature is also absent hence clitic doubling is disallowed.

29. For the sloppy reading, Italian uses a pronominal clitic, as shown in (i).

(i) Gianni sta cercando   un idraulico ma non **lo**   trova.
    Gianni is looking.for a plumber but not him.CL find.3SG
    'Gianni is looking for a plumber, but cannot find one.'   (Adapted from Alexopoulou and Folli 2019:474)

Interestingly for the current context, Runić (2014a,b) observes that Bulgarian, Macedonian, Spanish, French, (Brazilian) Portuguese, Romanian, and Greek, all of which have definite articles, disallow the sloppy reading of clitics, whereas Czech, Serbo-Croatian, Slovak, and Slovenian, which lack definite articles, allow the sloppy reading of clitics. She thus establishes the following generalization;

(ii) Only languages without definite articles allow sloppy reading of clitics.

Italian, then, seems to be a counterexample to (ii), since Italian has definite articles (Runić in fact acknowledges this and puts aside Italian in her account). The current proposal may leave room to account for the availability of the sloppy reading in Italian in this case. Under the current proposal, definite articles in Italian need not project DP; they can adjoin to N, so that bare NP (i.e., NP that is not dominated by DP) is in principle allowed. It is, then, not impossible that whatever mechanism licenses the sloppy reading of clitics in article-less languages, where bare NPs are allowed, is available in Italian, which also (in a way) allows bare NPs. This possibility remains to be explored in future research. Another possibility is that the sloppy reading of Italian clitics is attributed to the nature of clitics that is independent of the NP/DP-language distinction. See Ippolito (2017) and Yuan (2018) for discussion.

(72) a. *Greek*

  I   Maria vrike **dada** gia ta   pedhia

  the Maria found nanny for the children

  'Maria found a nanny for the children.'   (Alexopoulou and Folli 2019:442)

  b. *Italian*

  Maria ha   trovato ***(una)** baby-sitter per i   bambini.

  Maria has found   a   baby-sitter for the children

  'Maria has found a babysitter for the children.'   (Alexopoulou and Folli 2019:443)

Alexopoulou and Folli argue that Greek allows bare singular nominals because DP is absent in this language. My proposal can also account for this, since nominal phrases can be bare NPs in the absence of the definite article as discussed above.[30] Thus, one might view their proposal as essentially the same as the current proposal in this respect.

However, there is a crucial difference between their proposal and mine. Under their proposal, *all* nominal phrases in Italian are DPs and *all* nominal phrases in Greek lack the DP layer, which is along the lines of Bošković's (2008b, 2012) two-way cut of the NP/DP-language distinction. Under the current proposal, on the other hand, nominal phrases in Italian are NPs (only) in the presence of a definite article and those in Greek are DPs in the presence of a definite article. My proposal is motivated by the observation that adjunct extraction out of a nominal phrase is possible only in the presence of a definite article in Italian, while it is blocked in the presence of a definite article in Greek (see section 5.3.2 for Italian and section 5.5.1 for Greek). The proposal that DP need not project in Italian in the presence of a definite article is further supported by the availability of reflexive possessives (see section 5.3.3). These observations are mysterious under Alexopoulou and Folli's proposal that treats Italian as a canonical DP-language, where DP always projects. In addition, Alexopoulou and Folli note that Greek shows some properties that are expected for a

---

30. Bare singulars are also possible in the presence of an adjective, as shown in (i), repeated from (56).

(i) Aghorase **akrivo**   **aftokinito**.
    bought.3SG expensive car
    'She bought an expensive car.'   (Alexopoulou and Folli 2019:444)

This is also correctly predicted by the current proposal, since AP is analyzed as being adjoined to NP.

canonical DP-language in Bošković's (2008b, 2012) generalizations, such as the availability of clitic doubling (see footnote 28 on this). I have also shown above that Greek behaves like a DP-language in the presence of the definite article regarding adjunct extraction, LBE, and null objects. Although Alexopoulou and Folli mean to treat Greek as "less of a DP-language", their proposal actually amounts to saying that Greek is an NP language, since the DP layer would be always absent. Thus, I conclude that Alexopoulou and Folli's proposal, or approaches in general that appeal to a total presence vs. absence of the DP layer in a language, cannot be maintained, and that the relevant behavior of Greek can and should be captured by different possible structural options within a single language, namely, the presence/absence of DP that correlates with the presence/absence of the definite article in a nominal phrase. In the bigger picture, all this means that Greek is less of a DP-language than English and is yet another type of a non-canonical DP-language, different from Italian and Hungarian in the scale of the NP/DP-language distinction argued for here.[31]

---

31. Spanish is worth mentioning here. In footnote 3, it was noted that adjunct extraction out of a nominal phrase with a definite article is banned in Spanish, similarly to Greek. The relevant example is repeated in (i). However, when the definite article is absent, such extraction is much better, as shown in (ii). Spanish thus patterns with Greek in this respect, which can be taken as indicating that DP does not project in the absence of a definite article.

(i)*¿[De qué   estantería]ᵢ leyó María [los libros tᵢ]?
      of which shelf   read Maria   the books
   'From which shelf did Maria read the books?'

(ii) ¿[De qué   estantería]ᵢ leyó María [libros tᵢ]?
       of which shelf   read Maria books
    'From which shelf did Maria read the books?'

Interestingly for the current context, Spanish allows bare singulars as seen in (iii) (though they occur only with a limited set of predicates, unlike those in Greek; see, e.g., Espinal 2010 and Riqueros 2013 for Spanish and Alexopoulou and Folli 2019 for Greek).

(iii) Tengo coche.
      I.have car
      'I have a car.'

In addition, Clements (2006) notes that some varieties of Spanish allow null objects with sloppy reading, as shown in (iv). Note that the null objects cannot refer to a definite antecedent, similarly to those in Greek, as shown in (v).

(iv) a. ¿Compraste cafe?
         you.bought coffee
         'Did you buy coffee?'

     b. Sí, compre __.
        yes I.bought
        'Yes, I bought (some).'

5.5.5 Greek definite articles adjoin to A via movement

In section 5.5.2, I noted that LBE is possible even in the presence of the definite article in Greek, if the fronted adjective is accompanied by the definite article. The relevant example is repeated here as (73) from (61). This is contrasted with (74), repeated from (59), where the definite article stays in the nominal phrase.

(73) [To KOKKINO]$_i$ idha   [t$_i$ forema].
      the red   saw.1SG   dress
      'It is the RED dress that I saw.'   (Androutsopoulou 1998:1)

(74) *[KOKKINO]$_i$ idha   [**to** t$_i$ forema].
      red   saw.1SG   the   dress
      'I saw the red dress.'

I have proposed above that DP always projects above NP in the presence of the definite article, which accounts for the ill-formedness of (74) given Bošković's (2005) analysis of LBE discussed in section 5.5.2. A question that immediately arises is why LBE is possible in (73) despite the presence of the definite article.

I propose that the extraction in question in Greek can be accounted for by extending Bošković's (2013b) analysis of a particular type of LBE out of a PP in Serbo-Croatian. In Serbo-Croatian, LBE out of a PP is possible, where an apparent non-constituent undergoes LBE; in (75), the preposition *u* 'in' and the adjective *veliku* 'big' are fronted.

(v) a. ¿Compraste el   libro?
       you.bought the book
       'Did you buy the book?'

   b. Sí, lo=compre / *compre __.
      yes it=I.bought   I.bought
      'Yes, I bought it.'

Spanish thus patterns with Greek with respect to adjunct extraction, bare singulars, and null objects. This can be taken to indicate that Spanish indefinite nominal phrases can be bare NPs that are not dominated by DP, which in turn means that Spanish is less of a DP-language than English.

(75) [U veliku]$_i$ on uđe   [t$_i$ sobu].

      in big   he entered   room

      'He entered the big room.'   (Bošković 2005:30)

Note that (75) becomes bad when the preposition stays in the base position, as shown in (76). This is parallel to the contrast between (73) and (74) in Greek; namely, when the definite article stays in the base position, the extraction in question is disallowed.

(76) *[Veliku]$_i$ on uđe   [u t$_i$ sobu].

       big   he entered in   room

       'He entered the big room.'   (Bošković 2013b:72)

Bošković (2013b) proposes that the AP *veliku* 'big' first moves to the edge of PP and P procliticizes onto the AP. This P+AP then moves to the sentence initial position as a constituent, as schematized in (77).[32]

---

32. Talić (2013) provides evidence for this analysis from a phonological perspective, i.e., accent shift. As shown in (i), the preposition *u* 'in', which is a proclitic, can take over the accent from its host.

(i) u_nòvu → ù_novu
     in_new   in_new

Talić establishes the generalization that the accent shift can take place only if the host is allowed to move independently. As shown in (iia), when there is only one descriptive adjective, the adjective can undergo LBE. In contrast, when there are two descriptive adjectives, LBE is impossible, as shown in (iib) (see Bošković 2005 for discussion of this effect).

(ii) a. Novu$_i$ je on [t$_i$ bratovu   kuću] kupio.
        new   is he   brother's house bought
        'He bought his brother's new house.'

     b.*Novu$_i$ je on [t$_i$ veliku kuću] kupio.
        new   is he   big   house bought
        'He bought a new big house.'

Crucially, the preposition can take over the accent from its host when there is only one adjective, as shown in (iiia), whereas it cannot do so when there are two adjectives, as shown in (iiib). This indicates that the mobility of the adjective is a pre-requisite of the accent shift, which in turn indicates that in such cases, the adjective moves to SpecPP, with the preposition cliticizing to it (this is why the preposition still precedes it).

(iii) a. u_nòvoj bratovoj kući   → ù_novoj bratovoj kući
         in_new brother's house   in_new brother's house

      b. u_nòvoj velikoj kući   → *ù_novoj velikoj kući
         in_new big   house   in_new big   house

Not surprisingly, the same pattern is observed with LBE out of a PP. As seen in (iva), when there is only one adjective, the preposition and the adjective can be extracted together. On the other hand, when there are two adjectives, the adjectives cannot undergo LBE with the preposition, as shown in (ivb).

(77) a.

```
              PP
         ╱         ╲
      AP_i          P′
     veliku       ╱    ╲
      'big'      P      NP
                 u    ╱    ╲
                'in'  t_i   NP
                            △
                           sobu
                          'room'
```

b.

```
              PP
         ╱         ╲
    P_k+AP_i        P′
    u veliku      ╱    ╲
    'in big'     t_k    NP
                      ╱    ╲
                    t_i     NP
                            △
                           sobu
                          'room'
```
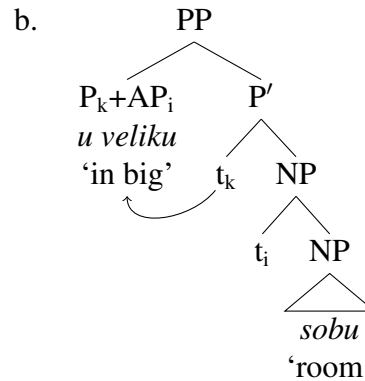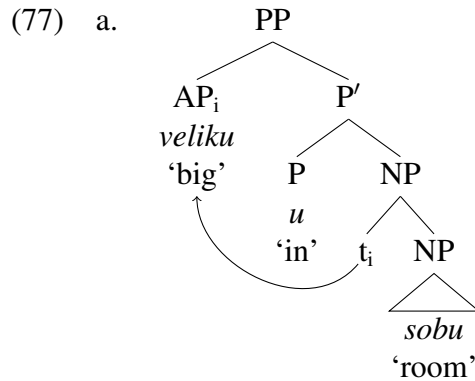
At this point, an issue arises regarding the locality of the movement of AP in (77). Bošković (2013a, 2014) proposes that PP is quite generally a phase (for relevant discussion, see also Abels 2003). This follows from Bošković's (2014) contextual approach to phasehood adopted here, in which the highest projection in the extended projection of every lexical category constitutes a phase. Thus, PP in (77) is a phase and NP is in its Spell-Out domain. The adjective *veliku* 'big' in (77a) must thus move to Spec,PP in order to obey the PIC. Notice, however, that this movement of the adjective from the NP-adjoined position to Spec,PP would violate the anti-locality condition, since it crosses just a segment, not a full phrase (cf. Bošković 2005). One may thus argue that the derivation in (77) should be ruled out.

In order to resolve this issue, Bošković (2013b) proposes that the PIC/anti-locality violation of the movement in question is voided because of the cliticization of the preposition, which is a phase head. Relevant here is Bošković's (2011b) generalization stated in (78).

(78)   Traces do not head islands.

(iv) a. [U novoj]_i je on [t_i bratovoj kući] živio.
        in new   is he   brother's house lived
        'He lived in his brother's new house.'

    b.*[U novoj]_i je on [t_i velikoj kući] živio.
        in new   is he   big   house lived
        'He lived in a new big house.'

Based on all this, Talić concludes that mobility of the adjective is a prerequisite for accent shift, with (75) and (iva) involving LBE of the adjective that carries the preposition; more specifically, the adjective first moves to Spec,PP and then the preposition procliticizes onto the AP, whereby the preposition can take over the accent from the adjective, which then undergoes LBE out of the PP, carrying the P (as a constituent). (See also Bošković 2005, 2013b for a number of parallelisms between LBE out of a PP under discussion and LBE out of a nominal phrase, which further motivate the analysis of PP LBE discussed in text.)

One of the striking arguments for (78) comes from the following example from Galician:

(79) a. *De quén_j liches   [_{DP}[_{D′} os [mellores poemas de amigo t_j]]]?

        of whom you.read   the best   poems of friend

        'Who have you read the best poems of friendship by?'

    b.(?)De quén_j liche-**los_i**   [_{DP}[_{D′} t_i [mellores poemas de amigo t_j]]]?

        of whom you.read-the   best   poems of friend

        'Who have you read the best poems of friendship by?'   (Uriagereka 1996:270-271)

(79a) shows that extraction of the wh-phrase out of a definite DP island is banned. Crucially, as shown in (79b), the island effect is circumvented when the definite article cliticizes onto the verb. This falls under (78); the head of the island is a trace in (79b), hence the DP ceases to be an island.

Bošković (2011b) proposes that (78) can be deduced from the rescue-by-PF-deletion mechanism, which is based on Chomsky's (1972) analysis of amelioration of island effects under ellipsis originally observed by Ross (1969). Relevant examples are given in (80).

(80) a. *Ben will be mad if Abby talks to one of the teachers, but she couldn't remember [which (of the teachers)]_i Ben will be mad [_{island} if she talks to t_i].

    b. Ben will be mad if Abby talks to one of the teachers, but she couldn't remember which_i ~~Ben will be mad [_{island} if she talks to t_i]~~.

Chomsky (1972) proposes that a * is assigned to an island when an element is extracted out of it. If the * remains in the final representation, the sentence is ungrammatical. Thus, in (80a), since *which* crosses the adjunct island, a * is assigned to the island and the sentence is bad. However, if the *-marked element, i.e., the island, is deleted in PF, the island effect is voided. Thus, in (80b), the adjunct island is deleted in PF, so that the island violation is circumvented. Bošković extends this proposal to deletion of copies of syntactic objects in minimalism. In addition, he modifies Chomsky's analysis, proposing that a * is not assigned to the whole island but to the head of the island. Thus, when the wh-phrase moves out of the DP island in (79a), the head of the island, i.e., the D head, is assigned a *, which is responsible for the ungrammaticality of (79a). However, in

(79b), D incorporates into V, leaving a trace in the base position, so that the *-marked island head is the trace of D, which means that it is a copy that is deleted in PF. Thus, the island effect is voided in (79b) since the *-marked element is deleted at PF. (Note that this analysis can also capture (80b), where the entire island, including its head, is deleted in PF.)

Returning to LBE out of a PP, Bošković (2013b) argues that the rescue-by-PF-deletion mechanism can also void violations of the PIC and the anti-locality condition, if the relevant phase head is deleted in PF. In (77a), the adjective moves to Spec,PP, which violates the anti-locality condition. The head of the PP phase is then *-marked. This P head undergoes cliticization onto AP, leaving a trace in the base position. Since the copy of P in the base position that is *-marked is deleted in PF, the anti-locality violation is voided. Likewise, if the adjective moves to a position higher than Spec,PP in order to obey the anti-locality condition, the PIC would be violated hence P would be *-marked, but since P undergoes cliticization onto the adjective, the copy of P in the base position that is *-marked is deleted in PF, resulting in amelioration of the PIC violation. Thus, LBE out of a PP in question is made possible by the rescue-by-PF-deletion mechanism. Recall also that if the preposition stays in the base position, movement of the adjective out of the PP is disallowed, as shown in (81), repeated from (76).

(81) *[Veliku]$_i$ on uđe   [u t$_i$ sobu].
        big   he entered in   room
      'He entered the big room.'   (Bošković 2013b:72)

This is straightforwardly captured by Bošković's proposal: since the *-marked phase head, i.e., P, is not a trace here, the * remains in the final representation, resulting in an anti-locality violation (the same holds for the derivation that involves the PIC violation).

Let us now consider how Bošković's analysis discussed above can be extended to Greek LBE in (82), repeated from (73).

(82) [To KOKKINO]ᵢ idha   [tᵢ forema].

      the red   saw.1SG   dress

      'It is the RED dress that I saw.'   (Androutsopoulou 1998:1)

What is important here is that Greek accusative definite articles are morphologically identical to 3rd person accusative clitics, as shown in (83).[33]

(83) a. Clitics

	Masculine	Feminine	Neuter
Singular	ton	tin	to
Plural	tus	tis	ta

      b. Definite articles

	Masculine	Feminine	Neuter
Singular	ton	tin	to
Plural	tus	tis	ta

Building on this, I propose that the definite articles in Greek can cliticize onto the adjective, on a par with the prepositions in Serbo-Croatian discussed above. In the pre-movement structure, the AP 'red' adjoins to the NP 'dress', and the definite article projects DP above the NP. The adjective *kokkino* 'red' moves to Spec,DP, and the definite article *to* procliticizes onto the adjective, as illustrated in (84).[34] The AP that contains the D then moves out of the nominal phrase.

---

33. Genitive clitics and definite articles are also morphologically identical.

34. As noted in footnote 32, Talić (2013) shows that there is a phonological reflex of LBE out of a PP in Serbo-Croatian. An analogue of this in the case of LBE in Greek could be the heavy stress on the extracted element; the element that is extracted receives heavy stress, which may be a phonological reflex of the movement in question.

(84) a.

b.
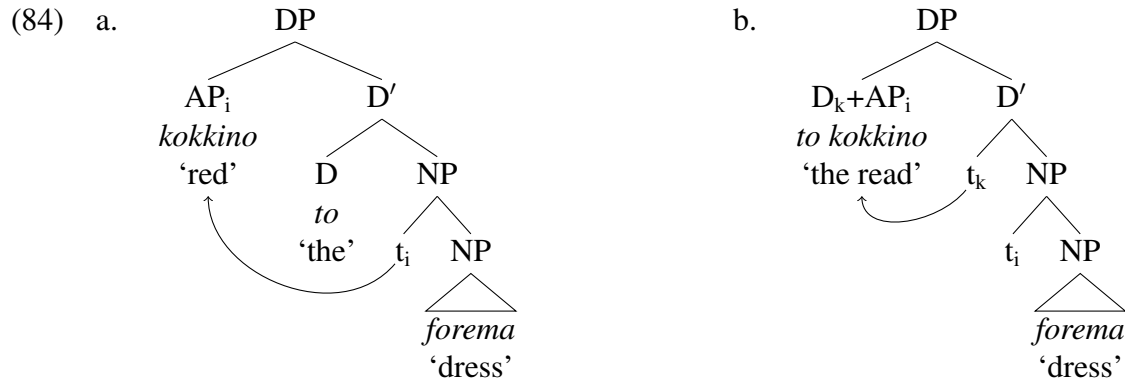
Although the movement in (84a) violates the anti-locality condition, the head of the DP phase, which is *-marked, becomes a trace because of the cliticization onto the adjective. Thus, the anti-locality violation is ameliorated by the rescue-by-PF-deletion mechanism. Recall also that LBE is disallowed if the definite article stays in the base position, as shown in (85), repeated from (74).

(85) *[KOKKINO]ᵢ idha [**to** tᵢ forema].

    red   saw.1SG   the   dress

    'I saw the red dress.'

This can be explained in a similar way as PP LBE in Serbo-Croatian in (81); when the definite article stays in the base position, the *-marked D head is not deleted in PF, so that the anti-locality violation is not voided. Thus, LBE in the presence of a definite article in Greek can be captured by the cliticization analysis of the definite article, on a par with LBE out of a PP in Serbo-Croatian.[35]

It should be noted, though, that there is one difference between LBE out of a PP in Serbo-Croatian and LBE with the definite article in Greek. In Serbo-Croatian, an intensifier adverb is

---

35. This proposal can be extended to some varieties of Bulgarian and Macedonian. Although Bošković (2005) observes that Bulgarian and Macedonian disallow adjective LBE (see also LaTerza 2014 for a survey of Macedonian), Stanković (2019) and Stojković (2019) note that majority of their informants of Macedonian and Bulgarian actually accept adjective LBE (if they are given an appropriate context), as seen in (i) and (ii). It should be added that (i) and (ii) are different from cases like (iiia) observed by Bašić (2005), where the definite article is attached not to the fronted adjective but to the head noun. Bošković (2007a) points out that (iiia) is unacceptable without extraction as shown in (iiib), and concludes that (iiia) does not involve LBE.

(i) *Macedonian*
    Crveni-teᵢ gi kupi [tᵢ čevli]?
    red-the them bought shoes
    cf. Gi kupi [crveni-te čevli]?
    'You bought the red shoes?'   (Stanković 2019:98)

extracted with the adjective that it modifies, as seen in (86). Bošković (2005, 2013b) and Talić (2013) in fact propose that P procliticizes onto the full AP, which can capture the observation that the adverb can move with the adjective it modifies.

(86) a. [U **izuzetno** veliku]$_i$ on uđe [t$_i$ sobu].

     in extremely big he entered room

     'He entered the extremely big room.' (Bošković 2005:30)

   b. *[U veliku]$_i$ on uđe [t$_i$ **izuzetno** sobu].

     in big he entered extremely room

     'He entered the extremely big room.' (Bošković 2005:33)

On the other hand, as shown in (87), an intensifier adverb cannot be accompanied by the definite article and the adjective in Greek, unlike in Serbo-Croatian (the sentence is also unacceptable if *exairetika* 'extremely' receives contrastive stress).

(87) *[To **exairetika** KOKKINO]$_i$ idha [t$_i$ forema].

     the extremely red saw.1SG dress

     'It is the extremely RED dress that I saw.' (Androutsopoulou 1998:1)

If the definite article cliticizes onto AP just like the prepositions in Serbo-Croatian, (87) would be acceptable just like (86a). We can thus conclude that the definite article does not adjoin to the full

---

(ii) *Bulgarian*
  Červeni-te$_i$ kupi [t$_i$ obuvki]?
  red-the bought shoes
  cf. Kupi [červeni-te obuvki]?
  'She bought the red shoes?' (Stojković 2019:348)

(iii) a. Nova ja prodade kolata (toj).
    new it sold car-the he
    'He sold the new car.'

   b.*(Toj) (ja) prodade nova kolata. (Bošković 2007a:fn. 1)

Note that in (i) and (ii), the definite article is attached to the fronted adjective, just like in the Greek cases seen above. This is not unexpected from the viewpoint of the current proposal on Greek; the definite articles in these varieties of Macedonian and Bulgarian can be analyzed as being able to adjoin to A after movement of AP to Spec,DP, just like those in Greek. This suggestion is not so surprising given that Macedonian and Bulgarian are clearly less of DP-languages (e.g., they have productive compositional indeterminate pronouns; most importantly, they have affixal definite articles).

AP, but to the A head.[36] This is a rather natural conclusion given that pronominal clitics, which are identical to the definite articles in Greek, are standardly considered to cliticize onto a head (say, V), not a phrase (VP), in the clausal spine. Note here that the definite article can only cliticize onto A in the nominal domain. This can straightforwardly explain the ill-formedness of (88), where the demonstrative and the definite article are extracted but the adjective stays in the nominal phrase; the definite article cannot cliticize onto the demonstrative.[37]

---

36. It is worth mentioning that the word order in which the intensifier adverb precedes the definite article is not possible (e.g., *exairetika to kokkino forema* 'lit. extremely the red car'). This can be captured by the current proposal that the definite article can adjoin to A but not AP. When there is no adverb, the adjective has no internal structure, i.e., it is ambiguously A and AP (i.e., a head and a phrase) as a non-branching element (cf. Chomsky 1995b), so the definite article can adjon to it. However, when an adverb is present, there is an internal structure in the AP, and adjunction to the adjective would be adjunction within this internal structure, not adjunction to A.

37. Note that the demonstrative can undergo LBE when an LBE-ing adjective is present, as shown in (i). This can be assimilated to multiple LBE in Serbo-Croatian, where the demonstrative and an adjective undergo LBE, as seen in (ii).

(i) [Afto]$_i$ [to KOKKINO]$_j$ eferes [t$_i$ t$_j$ forema].
    this the red bought.2SG dress
    'It is this RED dress that you bought.' (Androutsopoulou 1998:2)

(ii) Onu$_i$ staru$_j$ prodaje [t$_i$ t$_j$ kucú].
    that old sells house
    'S/he is selling that old house.' (Bošković 2016a:21)

Bošković (2016a) proposes that in (ii), the demonstrative and the adjective move one by one. Bošković argues that when there are multiple elements at the edge of a phase, only the highest edge is accessible to a higher domain. However, when the highest edge moves out and leaves a trace, the next highest edge becomes accessible to the higher domain. Thus, in (ii), the demonstrative *onu* 'that' first undergoes LBE to the left periphery, leaving a trace. The adjective then undergoes LBE to a position lower than the demonstrative in the left periphery ("tucking-in" in the sense of Richards 2001). If we assume that the demonstrative *afto* in Greek is base-generated at the edge of the DP phase, Greek (i) can be analyzed in the same way, with any locality-related violations ameliorated through deletion of the original copy of *to*, as discussed above. The assumption that the demonstrative is base-generated at the edge of the DP phase (which is actually not necessary for (i) since any locality-related violation causes by demonstrative movement would be ameliorated due to deletion of the original copy of *to*) is supported by the observation that the demonstrative precedes the definite article in the base position, and it can undergo LBE if the definite article stays in the base position, as seen in (iii).

(iii) a. Ida [afto to forema].
        I.saw this the dress
        'I saw THIS dress (e.g., not that one).'

     b. [Afto]$_i$ ida [t$_i$ to forema].
        this I.saw the dress
        'I saw THIS dress (e.g., not that one).' (Mathieu and Sitaridou 2002)

At any rate, Greek (i) can be analyzed as involving multiple LBE, just like Serbo-Croatian (ii).

(88) * [AFTO to]ᵢ eferes   [tᵢ kokkino forema]

     this   the bought.2SG   red   dress

     'It is this RED dress that you bought.'

It should be stressed here that locality amelioration by cliticization of the definite article is only available when the definite article moves with the adjective. Otherwise, the definite article in Greek always projects DP, which is a phase, as discussed above. Thus, adjunct extraction out of a nominal phrase with a definite article is disallowed, because it would violate the PIC or the anti-locality condition, as noted in section 5.5.1.[38] A relevant example is repeated here as (89) from (44).

(89) * [Apo pia   poli]ᵢ ghnorise [ta   koritsia tᵢ] o   Petros?

     from who city   met.3SG   the girls   the Petros

     'Petros met the girls from which city?'

---

38. As shown in (i), repeated from (45), adjunct extraction out of an indefinite nominal phrase with an adjective is possible, whereas extraction out of a definite nominal phrase with an adjective is impossible (see section 5.5.6 for the case where the indefinite article is present).

(i) a. [Apo pia   poli]ᵢ ghnorise [psila koritsia tᵢ] o   Petros?
     from who city   met.3SG   tall   girls   the Petros
     'Petros met the tall girls from which city?'

  b.*[Apo pia   poli]ᵢ ghnorise [ta   psila koritsia tᵢ] o   Petros?
     from who city   met.3SG   the tall   girls   the Petros
     'Petros met the tall girls from which city?'

The contrast can be captured by the current proposal. In (ia), DP does not project because the definite article is absent, and the adjective (AP) is adjoined to NP. Given Bošković's (2016a) claim that only the highest edge can undergo movement when there is more than one element in the phasal edge (see footnote 37), the PP adjunct should be adjoined to NP above AP (this would not affect the semantic composition, because both the PP adjunct and the AP are of type ⟨e,t⟩ so they are be composed with the head noun via Predicate Modification). The PP adjunct can then move out of the nominal phrase without violating the PIC or the anti-locality condition. In contrast, in (ib), the definite article projects DP above NP, to which the AP and the PP are adjoined, and this DP constitutes a phase. Extraction of the PP adjunct would then violate the PIC or the anti-locality condition.

5.5.6 Extraction out of a nominal phrase in the presence of the indefinite article

As mentioned in section 5.5.1, in Greek, adjunct extraction out of a nominal phrase is possible in the presence of the indefinite article. The relevant example is repeated here as (90). This is contrasted with (91), repeated from (44), where adjunct extraction is disallowed in the presence of the definite article.

(90) [Apo pia poli]$_i$ ghnorise [ena koritsi t$_i$] o   Petros?
       from who city   met.3SG a   girl   the Petros
       'Petros met a girl from which city?'

(91) *[Apo pia poli]$_i$ ghnorise [**ta** koritsia t$_i$] o   Petros?
       from who city   met.3SG the girls   the Petros
       'Petros met the girls from which city?'

Given Bošković's (2005) proposal adopted here, in which adjunct extraction is possibly only if the nominal phrase is bare NP, we are led to analyze the nominal phrase *ena koritsi* 'a girl' in (90) as bare NP.

A question that naturally arises here is the structural position of the indefinite article *enas*. Recall that I have proposed in section 5.3 that the definite article can base-generated as adjoined to N (and Poss), without projecting DP. Extending this to Greek, I propose that the indefinite article in Greek can be base-generated as adjoined to N, without projecting its own functional projection. Then, the nominal phrase *ena koritsi* 'a girl' is a bare NP that is not dominated by a functional projection. Consequently, the adjunct *apo pia poli* 'from which city', which is adjoined to NP, can be extracted out of the indefinite nominal phrase without violating the PIC or the anti-locality condition, as schematized in (92).

(92)

```
                    NP
              _____
            NP            PP
          ___            apo pia poli        ✓PIC
         N   Root                          ✓anti-locality
        ___         Spell-Out domain
       D   N
      ena (koritsi)
```
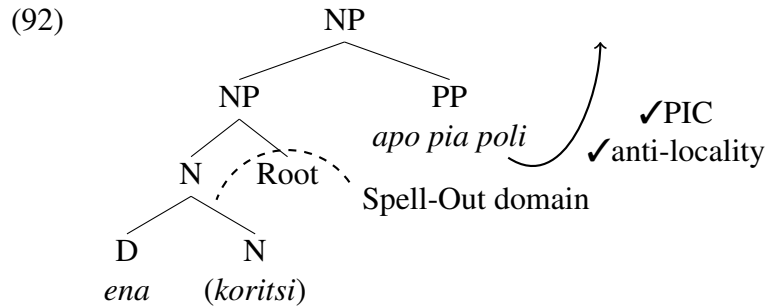
Recall that, as discussed in section 5.4, Wang (2019) proposes that a numeral 'one' that is at an intermediate stage of grammaticalization into an indefinite article is (or can be) adjoined to another head via base-generation. Note here that *ena* can also be used as a numeral 'one' (i.e., the indefinite article has developed from the numeral 'one'). Interestingly, Markopoulou (2000) claims that *ena(s)* is actually not an indefinite article (but an indefinite adjective; see below for discussion of this possibility). In fact, as seen in section 5.5.4, Greek allows bare singulars, i.e., the 'indefinite article' in Greek is not obligatory in contexts where a fully grammaticalized indefinite article such as English *a(n)* is obligatory. Then, it is not implausible to hypothesize that *ena(s)* is not yet fully grammaticalized as an indefinite article, which is compatible with the current proposal that *ena(s)* in Greek can adjoin to N without projecting its own functional projection, given Wang's theory of grammaticalization of indefinite articles. See also section 5.6.1 and section 5.7 for related discussion.

Recall also that LBE of an adjective is disallowed when *ena* stays in the base position but it is allowed when the adjective is accompanied by *ena*, as shown in (93), repeated from (63).

(93) a. *AKRIVO_i aghorase   [ena t_i aftokinito].

        expensive bought.3SG a   car

        'He bought an expensive car.'

      b. [Ena AKRIVO]_i aghorase   [t_i aftokinito].

        a   expensive bought.3SG   car

        'He bought an expensive car.'

There are two possibilities to account for this contrast. One is that *ena* projects its own functional

projection above NP in the presence of an adjective. LBE of the adjective in (93a), which is adjoined to NP, would then violate the PIC or the anti-locality condition. In (93b), on the other hand, *ena* can cliticize onto the adjective via movement, on a par with the definite article discussed in the previous subsection. The locality violation would then be voided because the head of the functional projection projected by *ena* becomes a trace (see section 5.5.5 for relevant discussion). The other possibility is that *ena* can be base-generated as adjoined to AP (see also footnote 17 for the possibility of this option in Slovenian).[39] The adjective that is left-dislocated would then necessarily be accompanied by *ena*.

The second possibility, though, seems to be favored over the first one. As noted in section 5.5.1, adjunct extraction out of a nominal phrase is possible in the presence of an adjective and *ena*, as shown in (94a), repeated from (45b). This is contrasted with (94b), where the extraction in question is disallowed in the presence of the definite article.

(94)   a.   [Apo pia poli]$_i$ ghnorise [ena psilo koritsi t$_i$] o   Petros?

            from who city   met.3SG a   tall   girl   the Petros

            'Petros met a tall girl from which city?'

       b. *[Apo pia poli]$_i$ ghnorise [ta psila koritsia t$_i$] o   Petros?

            from who city   met.3SG the tall   girls   the Petros

            'Petros met the tall girls from which city?'

The ill-formedness of (94b) can be captured by the current proposal that the definite article always projects DP; extraction of the PP adjunct, which is adjoined to NP, would violate the PIC or the anti-locality condition. If *ena* projects its own functional projection above NP in the presence of an adjective just as the definite article projects DP above NP, adjunct extraction in (94a) would be expected to be banned, contrary to the fact. On the other hand, if *ena* is adjoined to AP in (94a) without projecting its own functional projection above NP, the PP adjunct that is adjoined to NP

---

39. Note that the indefinite article *ena* is a minimal and maximal projection at the same time under BPS, because it is a non-branching element that does not project further. Thus, as a maximal projection, it can adjoin to another maximal projection, i.e., AP.

(above *ena* + AP) can be extracted without violating the PIC or the anti-locality condition. It is worth adding here that LBE of an adjective with the indefinite article and an intensifier adverb is possible, as shown in (95). This can be straightforwardly explained by the current proposal, since *ena* can adjoin to the whole AP (note that this is contrasted with the definite article, which cannot be left-dislocated with an intensifier adverb and hence is analyzed as being adjoined to A (via movement), as discussed in section 5.5.5).

(95) [Ena EXAIRETIKA AKRIVO]ᵢ aghorase [tᵢ aftokinito].
      a   extremely   expensive bought.3SG   car
      'He bought an extremely expensive car.'

Thus, it is not implausible that *enas* can be base-generated as adjoined to AP as well as N.

There is, though, yet another possibility that may actually be preferable to the analyses discussed above. Recall that Markopoulou (2000) claims that *ena(s)* is not an indefinite article but an indefinite adjective. In Serbo-Croatian, the numeral 'one' is clearly an adjective, and it has been treated this way in the literature. Interestingly, Serbo-Croatian exhibits the same extraction pattern as Greek in the presence of the numeral 'one'. In Serbo-Croatian, the numeral 'one' precedes the adjective and the head noun, as shown in (96).

(96) Jedna crvena kola
      one   red   car

LBE of an adjective is disallowed when 'one' stays in the base position as shown in (97a). However, it is allowed if *jedna* also undergoes fronting, as shown in (97b).

(97) a. *Crvenaᵢ je kupio [jedna tᵢ kola].
         red   is bought one   car
         'S/he bought one red car.'

240

b. Jedna crvena je kupio [t kola].

    one   red   is bought   car

    'S/he bought one red car.'

In addition, an intensifier adverb can also be extracted with the adjective in such cases, as shown in (98).

(98)   Jedna potpuno crvena je kupio   [t kola].

    one   fully   read   is bought   car

    'S/he bought one fully red car.'

Turning now to adjuncts, adjunct extraction out of a nominal phrase is allowed in the presence of 'one', as seen in (99).

(99)   [Iz   kojeg grada]$_i$ je sreo [jednu djevojku t$_i$]?

    from which city   is seen one   girl

Such extraction is also possible when an adjective is present, as in (100).

(100)   [Iz   kojeg grada]$_i$ je sreo [jednu veliku djevojku t$_i$]?

    from which city   is seen one   big   girl

What is important here is that Serbo-Croatian shows the same extraction pattern as Greek. This can be taken as indicating that *ena(s)* in Greek is not an indefinite article but an indefinite adjective (on a par with the corresponding element in SC), as in fact Markopoulou (2000) proposes.

Let us then consider how the SC paradigm discussed above can be accounted for. Given that 'one' in Serbo-Croatian is an adjective, it is plausible to assume that it is adjoined to NP, just like other adjectives in this language (cf. Bošković 2006a).[40] It should be noted here that the word order of adjectives is not free, and it is semantically constrained (see, e.g., Scott 2002 for a semantic

---

40. Note that Serbo-Croatian 'one' can left-branch extract on its own, and the same appears to hold for one in Greek:

(i) Jena$_i$ je kupio   [t$_i$ kola]. (Serbo-Croatian)
    one   he bought   car
    'He bought one car'

hierarchy of adjectives; see also Bošković 2009a for relevant discussion). The base word order in the nominal phrase in question is (96). As shown in (101), the numeral 'one' in Serbo-Croatian cannot follow an adjective.[41]

(101) *crvena jedna kola

     red   one   car

Bošković (2016a) shows that Serbo-Croatian allows multiple LBE. I suggest that (97b) in fact involves multiple LBE, where more than one NP-adjunct is left-dislocated. An example involving multiple LBE in Serbo-Croatian was given in footnote 37, repeated here in (102).

(102)   Onu$_i$ staru$_j$ prodaje [t$_i$ t$_j$ kucú].

     that old   sells   house

     'S/he is selling that old house.'   (Bošković 2016a:21)

As discussed in footnote 37, Bošković (2016a) proposes that the demonstrative *onu* 'that' and the adjective *staru* 'old' move one by one. Bošković argues that when there are multiple elements at the edge of a phase, only the highest edge is accessible to a higher domain. However, when the highest edge moves out and leaves a trace, the next highest edge becomes accessible to the higher domain. Thus, in (102), the demonstrative *onu* 'that' first undergoes LBE to the left periphery, leaving a trace. The adjective can then undergo LBE, because it is the next highest edge after the movement of the demonstrative (it moves to a position lower than the demonstrative in the left periphery in the "tucking-in" manner in the sense of Richards 2001). Notice now that (97b), which is repeated here as (103), can be analyzed in the same way as (102).

(ii)?Ena$_i$ aghorase [t$_i$ aftokinito]. (Greek)
   one bought   car
   'He bought one car.'

41. Intuitively, cardinality should take scope over a noun and an adjective; when combining 'red' and 'one car' via Predicate Modification, which yields an intersection of two sets, the composed meaning would be something like 'car that is red and one', which is not the interpretation we would normally get from 'one red car'. I leave technical implementation of this for future study.

(103) [Jedna]$_i$ [crvena]$_j$ je kupio [t$_i$ t$_j$ kola].

     one   red   is bought   car

     'S/he bought one red car.'

The numeral 'one', which is an NP-adjunct, first undergoes LBE to the left periphery, leaving a trace. Then the adjective, which is the next highest edge of the NP phase, undergoes LBE, landing in the position lower than 'one', just like the modifiers in (102). The same holds for *potpuno crvena* 'completely read' in (98). In contrast, in (97a), repeated here as (104), the adjective cannot be extracted across 'one', because the adjective is the lower edge and hence it is not accessible to a higher domain.

(104) *Crvena$_i$ je kupio [jedna t$_i$ kola].

     red   is bought one   car

     'S/he bought one red car.'

Thus, the contrast between (97a) and (97b) can be captured by Bošković's (2016a) analysis of multiple LBE.

The treatment of the adjectival numeral 'one' as adjoined to NP can also capture adjunct extraction in the presence of 'one'. A relevant example is repeated here as (105).

(105) [Iz   kojeg grada]$_i$ je sreo [jednu djevojku t$_i$]?

     from which city   is seen one   girl

Here I assume that the PP adjunct has the option of adjoining to NP above 'one'. The PP adjunct can then move out of the nominal phrase, without causing any locality problems.

Recall now that Greek shows the same extraction pattern as Serbo-Croatian in the presence of the numeral 'one'; (i) an adjective can be extracted when 'one' is also extracted, but not with 'one' staying in the base position,[42] (ii) an intensifier adverb can also be extracted with the adjective in the relevant cases, and (iii) an adjunct PP can be extracted in the presence of 'one'. Given

---

42. Note also that the adjective-'one'-noun order is disallowed in Greek, just as in Serbo-Croatian (e.g., *kokkino ena aftokinito* 'lit. red a/one car').

the parallelism, the above analysis of these extractions in Serbo-Croatian can be straightforwardly extended to Greek. This is also consistent with Markopoulou's (2000) proposal that *ena(s)* in Greek is not an indefinite article but an indefinite adjective (recall that the numeral 'one' in Serbo-Croatian is an adjective). It should, however, be noted that the above does not preclude the possibility that that there could still be an option for the numeral 'one' in Greek to be an indefinite article in some cases, but this would not affect the extraction paradigm discussed above.

At any rate, I leave for future research the tease apart the options regarding 'one' in Greek suggested above.

To summarize this section, I have shown that Greek behaves like an NP-language when a definite article is absent, while it behaves like a DP-language when a definite article is present, with respect to adjunct extraction, Left-Branch Extraction, compositional indeterminate pronouns, and null objects with sloppy readings. I have then proposed that Greek allows bare NP that is not dominated by DP in the absence of the definite article. On the other hand, the definite article in Greek always projects DP. In addition, I have proposed that definite articles in Greek are clitics that can adjoin to another head in the nominal domain via movement, which accounts for the observation that LBE out of a definite nominal phrase is possible (only) if the adjective is accompanied by the definite article. I have argued that this is essentially similar to LBE out of a PP in Serbo-Croatian, where the preposition procliticizes onto the adjective and they move together as a constituent. Finally, I have discussed two possible analyses of extraction patterns in the presence of the "indefinite article" *ena(s)*. One is that *ena(s)* can be base-generated as adjoined to N and AP, hence does not project its own functional projection, which captures the observation that Greek behaves like an NP-language in the presence of the indefinite article. Given Markopoulou's (2000) proposal that *ena(s)* is not a (fully grammaticalized) indefinite article, this analysis of *ena(s)* can be taken as additional support for Wang's (2019) proposal that an indefinite article that is at an intermediate stage of grammaticalization is (or can be) base-generated as adjoined to another head. The other possibility I suggested is that *ena(s)* is an indefinite adjective that is adjoined to NP (cf. Markopoulou 2000), on a par with the numeral 'one' in Serbo-Croatian, which is fully an adjective. This anal-

ysis was supported by the observation that Greek and Serbo-Croatian show the same extraction pattern in the presence of 'one'. At any rate, under both analyses, 'one' in Greek does not project a separate functional projection. In the bigger picture, Greek behaves differently from other DP-languages in the NP/DP-language scale such as English, Italian, and Hungarian. This means that Greek is yet another type of a non-canonical DP-language in the scale of the NP/DP-languages distinction argued for here.

Before proceeding, I would like to address an issue regarding the availability of bare NP. Specifically, a question arises as to why English does not allow bare NPs unlike Greek, given that the definite article always projects DP in both languages. Although a full investigation of this issue is left for future research, I would like to note two potential factors that may be relevant for this difference between English and Greek. One is that Greek allows multiple occurrences of a definite article in one nominal phrase (the so-called polydefinite construction), unlike English. Bošković (2012) notes that this casts doubt on the article status of the definite article in Greek, under the definition of definite articles in which a definite article occurs only once in one nominal phrase.[43] In the current context, we can reinterpret this in the way that the presence of a "non-canonical" definite article may in principle allow bare NP in Greek (including outside of definite contexts).

The other possibility concerns the obligatory/non-obligatory presence of the indefinite article. As seen above, Greek allows bare singulars, i.e., the indefinite article need not be used with indefinite singular count nouns. This contrasts with English, where bare singulars are never allowed, i.e., the indefinite article is obligatory for indefinite singular count nouns. As will be discussed in section 5.6, although Bošković (2008b, 2012) does not consider indefinite articles to be rele-

---

43. The polydefinite construction may be assimilated to the long-form endings of adjectives in Serbo-Croatian, which can occur in more than one place in the nominal domain since they can accompany each adjective, and encodes definiteness/specificity (see, e.g., Zlatić 1997, Progovac 1998, Leko 1999, Despić 2011). Then, the "definite article" in Greek may be ambiguous between an adjective marker and a definite article (cf. Mathieu and Sitaridou 2002, who propose that the multiple occurrences of the "definite article" in the polydefinite construction are agreement markers). This would straightforwardly account for the left-branch extraction facts with the definite article since the definite article would actually be part of the AP in the relevant cases (the relevant morpheme in Serbo-Croatian in fact undergoes LBE together with the adjective). It is also worth adding here that the multiple occurrences of the "indefinite article" are not allowed in Greek (Markopoulou 2000). This makes sense given the arguments above that the "indefinite article" in Greek is a head that adjoins to N and AP via base-generation, or an adjective that adjoins to AP; namely, it is not an agreement marker but is a lexical item that adjoins to a nominal element.

vant for his NP/DP-language distinction, it may not be implausible that obligatory/non-obligatory presence of indefinite articles, which are more or less standardly assumed to project their own functional projection (if not identical to D), correlates with obligatory/non-obligatory projection of DP in indefinite nominal phrases. See also section 5.7 for related discussion from the perspective of Egyptian Arabic. At any rate, the definite and indefinite articles in Greek have different morphological status and distribution from those in English, which may be relevant for the difference between these two languages regarding the possibility of bare NP in indefinite nominal phrases.

Regarding the different status of the indefinite article in Greek and English, it may actually not be implausible to hypothesize that if what has been traditionally considered to be an indefinite article never projects a phrase in a language (which is the case in Greek), the language allows indefinite nominal phrases without DP in general. The intuition behind this is that the option of not projecting DP above indefinite nominal phrase is "hinted at" by an indefinite article (in the traditional sense), which is then generalized to other indefinite nominal phrases in the absence of the indefinite article. In this respect, in section 5.6, I will discuss a generalization of a parameter value, i.e., projection of DP above NP in the current context, to all nominal phrases, where the morpho-syntactic nature of definite articles may be relevant. From this perspective, the nature of an indefinite article, more precisely, its not projecting DP above NP, may be a clue for not projecting DP in indefinite nominal phrases in general.

5.6   The emergentist view of parameters, the NP/DP-language scale, and acquisition of definite articles

In this section, I show that the NP/DP-language scale proposed in this work is compatible with the so-called emergentist view of parameters, which conforms to the three-factor design of language proposed by Chomsky (2005). The gist of the argument here is that the "intermediate" behavior of Italian, Greek, and Hungarian in the NP/DP-language scale is expected from a perspective of interactions between the language input and third factor principles that minimize computation

and maximize the effect of language acquisition. I also discuss some observations on acquisition of definite articles in Italian, considering economy of structure building from a viewpoint of the emergentist view of parameters.

5.6.1 The emergentist view of parameters and the NP/DP-language scale

In minimalism, it is hypothesized that the faculty of language is a computational system that is optimized to the interfaces in accordance with language-external general principles, as a part of the biological system. Chomsky (2005) specifically proposes the three-factor design of language, as summarized in (106).

(106) a. Factor 1: The innate endowment, i.e., Universal Grammar

b. Factor 2: Experience, i.e., the primary linguistic data (PLD)

c. Factor 3: Non-domain-specific cognitive optimization principles

Factor 1 is minimal UG, arguably just Merge (Chomsky 2005; see also chapter 4 for deduction of Agree from Minimal Search, a third-factor principle). Given this, there is little or no room for parameters to be encoded in UG, at least in a way that a wide range of linguistic variations can be fully accounted for (see also Boeckx 2011 for a discussion of invariant UG). This is contrasted with the GB theory, in which UG is "rich" in the sense that it consists of a number of principles that are equipped with parameters (i.e., "parameterized UG").

A natural question that arises under this conception of language in minimalism is how to formulate parameters. One view proposed by Chomsky (1995b), who takes the insight of Borer (1984) and Fukui (1986, 1988), is that parameters are reduced to specifications of formal features in the lexicon. This is referred to as the *Borer-Chomsky Conjecture (BCC)* by Baker (2008a,b) and formulated as (107).

(107) *The Borer-Chomsky Conjecture (BCC)* (Baker 2008b:156)

All parameters of variation are attributable to differences in the features of particular items (e.g., the functional heads) in the lexicon.

Note that this falls under Factor 2 of the three-factor design; feature specifications of lexical items are acquired based on the PLD.

Against these backgrounds, Biberauer (2019), Biberauer and Roberts (2017), Roberts and Holmberg (2010), Roberts (2012, 2019) among others propose what they call an *emergentist view of parameters*. Under this view of parameters, UG is underspecified with respect to parametric variation, since UG only consists of the minimal computational system, i.e., Merge. What has been called parameters in the generative literature "emerges" by interactions between the PLD (Factor 2) and Factor 3 in the course of language acquisition. A learner acquires feature specifications from the PLD as a point of parameterization as stated in the BCC, and the acquisition process is facilitated by third factor principles that minimize the effort of acquisition and maximize the effect of acquisition (Biberauer 2019 calls this the "Maximize Minimal Means" model). Thus, the apparent cross-linguistic variation is not prespecified in UG (as in the GB theory), but is a result of acquisition of different formal features, which is conditioned and made efficient by third factor principles.

As for efficiency of language acquisition, the emergentist view adopts two third-factor principles: Feature Economy (Roberts and Roussou 2003) and Input Generalization (Roberts 2007). Feature Economy (FE) is an acquisition bias that minimizes computation by requiring learners to specify features only when there is evidence for them in the PLD. To put it differently, FE instructs a learner to postulate as few formal features as possible based on the PLD (Roberts 2019:93). Roberts's (2019) version of its formulation is given in (108).[44]

(108)   *Feature Economy*

Given a pair of adequate structural representations R, R′ for a substring of input text of the PLD S, choose R iff R has n distinct FFs [= Formal Features] and R′ has $m > n$ distinct FFs.

For instance, Italian shows full specification of ϕ-features on T, whereas Japanese shows no ϕ-

---

44. This is essentially similar to Economy of Representation in early minimalism; see Bošković (1997c), Speas (1994), and references therein.

feature agreement on T. FE requires a learner to only postulate features that are present in the PLD; thus, based on the PLD and FE, a learner acquiring Italian can postulate ϕ-features on T, whereas a learner acquiring Japanese does not postulate ϕ-features in the feature bundle of T. In addition, it is also possible that a functional category, which is merely a bundle of features under BPS, can be absent in a language. If there is no evidence for a particular feature combination in the PLD that would constitute a functional category, a learner simply does not acquire it, because of FE. Thus, the cross-linguistic variation in feature specifications, which is the locus of parameters under the BCC, "emerges" based on the PLD, and functional projections can also be considered to emerge in the course of language acquisition, the lack of them being the default option of UG.[45]

Note now that FE can naturally capture the distinction of canonical DP languages and canonical NP languages, i.e., the two-way cut of the NP/DP-language distinction originally proposed by Bošković (2008b, 2012). Recall that Bošković proposes that languages with definite articles have DP in the nominal phrase, while languages without definite articles lack DP. Definite articles in the PLD can then be considered as a trigger for a learner to postulate the D-feature, i.e., the D head, which is in fact what Bošković (2021a) argues for.[46] Thus, if the PLD contains a definite article, the learner can postulate the D-feature and hence DP, which gives rise to a canonical DP-language (but see below for an additional factor). On the other hand, if the PLD does not contain a definite article, the learner does not postulate the D-feature as per FE, as a result of which the learner acquires a canonical NP language where DP does not project. Thus, we can conclude that the NP-language grammar, where DP does not project, is the default option of UG.

This conclusion is in fact supported by observations in first language acquisition. Specifically, Guasti et al. (2008), Koulidobrova (in press), Matthewson et al. (2001), and Radford (1990) observe that children acquiring a language with definite articles such as English seem to first acquire a grammar in which DP does not project (i.e., NP-language), and then a grammar in which DP

---

45. Progovac (2010) and Progovac and Locke (2009) propose that intransitive predicates antedated transitive predicates in proto-syntax, which can be translated in the current context such that intransitive structure is the default option of UG. This makes sense given the standard assumption in the current syntactic theory that transitive predicates require an additional functional layer (e.g., v in the case of the verbal domain), and hence are structurally "richer" than intransitive predicates.

46. Here I put aside other possible features in the feature bundle of D such as ϕ and Case for expository purposes.

projects (i.e., DP-language). For instance, Koulidobrova (in press) observes that once children learning English acquire the definite article, they stop making mistakes in the use of other D-related items such as demonstratives, although they show different frequency in the input. Interestingly, the definite article appears much more frequently in the input than other D-related items (in fact, the definite article is the most frequent word in English). Bošković (2021a) interprets this as indicating that the definite article is a trigger for acquisition of D(P), which in turn means that the non-D-projecting grammar is the default stage of language acquisition (see also Radford 1990). This is compatible with the emergentist view of the NP/DP-language distinction discussed above, in which bare NP is the default option of UG.[47]

Related to this is the issue of the position of AP relative to NP. As mentioned in section 5.3, Bošković (2005) proposes that the structure in which AP projects above and dominates NP is the default option of UG, and the option of AP being adjoined to NP is only available in languages without definite articles, since AP cannot be an argument of ordinary verbs and hence NP must project in the absence of DP in the relevant languages. From the perspective of the emergentist view, however, the default option should rather be the one in which AP is adjoined to NP, since bare NP that is not dominated by a functional projection is the default option of UG. This consideration leaves room for some languages with definite articles to have a structure in which AP is adjoined to NP, which I in fact argued is borne out regarding LBE in Greek in section 5.5.2.

Another potentially related domain is second language acquisition. Interestingly, Trenkic (2004) observes that L2 learners of English who are native speakers of Serbo-Croatian omit the definite article in English more frequently in the presence of an adjective than in the absence of an adjective in contexts where it is required. Trenkic interprets this as indicating that the definite

---

47. Historical changes in the nominal phrase may be relevant as well. Older Indo-European languages used to lack definite articles and showed properties of NP-languages, and many of the modern Indo-European languages have acquired definite articles and become DP-languages (the extent of DP-language-hood varies across languages, as discussed in this chapter); see, e.g., Bošković (2008b, 2012) and chapter 3 of this dissertation. Colloquial Finnish, which has developed a definite article (Laury 1997), has lost adjective LBE (Franks 2007), which Bošković (2012) takes as an instance of the development of DP. There thus seems to be a general tendency of historical change from NP-language to DP-language, which can be captured by the idea argued for here that the grammar of NP-language is the default option of UG and functional projections are acquired later. (In fact, it seems that this change is uni-directional; but see also section 5.7 for a different perspective of grammaticalization of definite articles.)

article and an adjective might compete for the same position, but there is another possible interpretation of this observation from the current perspective; since their L1 grammar (i.e., the grammar of Serbo-Croatian) has bare NP structure, the default option of merger of adjectives is adjunction to NP, not projection of AP. When an adjective is used in a nominal phrase in their L2 English, this default option is preferred as a default option of UG, hence the definite article, i.e., DP, is more frequently omitted.

The NP-language grammar being the default option of UG may receive further support from code-switching. Petroj (2020) examines code-switching of Romanian, which has affixal definite articles and hence is a DP-language (to some degree), and Serbian, which lacks definite articles and hence is an NP-language, in the context of the NP/DP-languages distinction (here I am simplifying the discussion of Romanian following Petroj 2020). As an illustration, she shows that Serbian verbs can select both Romanian and Serbian nominal phrases as their complement, whereas Romanian verbs can select Romanian nominal phrases but not Serbian nominal phrases. In (109), *examen-ul* 'the exam' is a Romanian nominal phrase, which is a DP, *ispit* 'exam' is a Serbian noun (hence a fully Serbian object), which is an NP, and *ispit-ul* 'the exam' is a combination of the Serbian noun and the Romanian definite article, which is a DP. As seen in (109a), the Romanian verb *trecut* 'passed' can select *examen-ul* and *ispit-ul*, which are both DPs, but cannot select *ispit*, which is an NP. In contrast, the Serbian verb *položila* 'passed' can select all three nominal phrases. Petroj thus establishes the generalization in (110).

(109) a. **Am**   **trecut** {**examen-ul** / *ispit*-**ul**   / **ispit*}

          have.1SG.AUX passed   exam-the   exam-the   exam

          'I passed the exam / the exam / exam.'   (**Bold: Romanian**, *Italic: Serbian*)

    b. **Am**   *položila* {**examen-ul** / *ispit*-**ul**   / *ispit*}

          have.1SG.AUX passed   exam-the   exam-the   exam

          'I passed the exam / the exam / exam.'   (**Bold: Romanian**, *Italic: Serbian*)

                                                 (Petroj 2020:75)

(110)   Romanian verbs must take a DP complement, while Serbian verbs can take either a DP or

an NP complement.   (Petroj 2020:75)

Petroj suggests that the generalization in (110) can be explained if we assume that bare NP is the

default option of the parameter setting in the nominal domain. Just as children modify the bare NP

structure and acquire the DP layer in the course of acquisition of a DP-language, a bilingual speaker

can modify the bare NP structure and add the DP layer in code-switching. Then, a Serbian verb,

which by default selects an NP complement, can also select a Romanian DP by adding the DP layer

in code-switching. In contrast, parameter setting from a non-default to the default by eliminating

structure is difficult (see also footnote 47 on the tendency of historical change from NP-language

to DP-language). Then, a Romanian verb, which by default selects a DP complement, cannot take

an NP complement by removing the DP layer in code-switching. Thus, the Romanian-Serbian

code-switching cases that Petroj examines can be additional supportive evidence for the argument

that the NP-language grammar is the default option (see Petroj 2020 for more data that point to the

same conclusion).

Let us now turn to the other third-factor principle of the emergentist view of parameters, In-

put Generalization (IG) requires that a learner generalizes a value of a parameter to other related

parameters, as formulated in (111).[48]

(111)   *Input Generalization* (Roberts 2019:93)

If a functional head H_i of class C is assigned FF_i, assign FF_i to all functional heads

$\{H_1 \ldots H_n\}$ in C.

IG makes it possible for a learner to determine the value of a number of parameters without receiv-

ing all relevant input, which essentially explains the poverty of the stimulus of language acquisi-

tion. In other words, IG minimizes computation of parameter setting under the BCC by making

---

48. Interestingly, Boeckx (2011) proposes a similar Factor-3 learning principle on independent grounds, which he
calls Superset Bias.

(i) *Superset Bias* (Boeckx 2011:217)
      Strive for parametric-value consistency among similar parameters.

maximal use of input. For instance, rigid head-finality and head-initiality can be captured by IG; once a learner acquires head-finality or initiality in one domain, it is generalized to all other domains.

It should be added here that IG is merely a learning bias, and can be overridden by the PLD. For instance, a learner will first determine that the language s/he is acquiring is rigid head-initial, given the first relevant PLD and IG, but when s/he receives input that contains a head-final order in some domain, s/he will "cancel" the generalization and modify the parameter value for the relevant head, which yields a mixed word order language (see, e.g., Biberauer and Sheehan 2013 and Biberauer et al. 2014 for related discussion). More generally, Biberauer (2019), Biberauer and Roberts (2017) and Roberts (2019) argue that interaction of FE, the PLD, and IG predicts a NO > ALL > SOME path in parameter setting as well as typological classification. The default is absence of a property P (attributed to a formal feature F), i.e., the NO stage, because FE prevents unnecessary features from being postulated. Once a relevant input is available in the PLD, IG generalizes P to all relevant domains, giving rise to the ALL system. Some of the generalized Ps are canceled due to additional PLD that denies existence of P in some domains, which amounts to the SOME stage.[49] Below I argue that the scale of NP/DP-language distinction is essentially an instance of the NO-ALL-SOME distinction, with some qualifications added.

Let us start with canonical NP/DP-languages. Recall that given the BCC and FE, presence/absence of formal features is one instance of parametric variation, which yields canonical DP and NP languages in the NP/DP-language distinction. The trigger for this two-way cut of NP/DP-language distinction is presence/absence of a definite article as a D head, which is a bundle of features under BPS, in the PLD. Note here that in the case of languages with definite articles, a learner acquires a definite article as an instance of D, but presence of D in other nominal phrases (e.g., indefinite nominal phrases) is not necessarily ensured; namely, the learner does not know if DP projects above NP in the absence of definite articles. But in canonical DP-languages such as English, DP

---

49. As Biberauer and Roberts (2017) point out, this way of parameter setting only makes reference to positive evidence, i.e., presence of a cue in the PLD. In other words, given the usual assumption that learners do not make use of negative evidence in language acquisition, a learner does not change a parameter value based on absence of a cue in the PLD.

always projects above NP whether a definite article is present or not. My suggestion here is that IG generalizes presence of D to other nominal phrases. Application of IG in this way is reasonable given that all Ns are in one class C (i.e., being N). Thus, we can consider canonical DP-languages to emerge as a result of the PLD and IG under the BCC.

I would like to add here that IG can also be construed in such a way that the presence/absence of a functional projection in one domain is generalized to other domains, as long as those domains share a feature based on which IG applies. This is actually quite similar to Talić's (2015, 2017) Structural Parallelism discussed in chapter 2, whose formulation is given in (112).

(112)   *Structural Parallelism* (Talić 2015, 2017)

    a.   If a language allows bare lexical structure without a functional layer in the domain of one lexical category, it may allow bare lexical structure in the domain of other lexical categories (e.g., a language can have both bare NP and bare AP).

    b.   If a language never allows bare lexical structure, that is, it always requires a functional layer in the domain of one lexical category, it must have a functional layer in the domain of all lexical categories (e.g., such a language will never have bare NP or bare AP).

As I suggested in chapter 2 and chapter 4, lexical categories can be assumed to have a categorial feature, whose value is specified such as [+N,-V] (i.e., N), as Chomsky (1970) originally proposed. To be more precise, traditional lexical categories are decomposed into Root and a categorizer (Halle and Marantz 1993b, Marantz 1997, Embick 2000, Embick and Marantz 2008 among others); Root has an unvalued categorial feature $\text{Cat}_{[\_]}$ which is valued by a categorizer such as v, n in narrow syntax (in chapter 2, I proposed to extend this to the Conj head, whose categorial feature is valued by the conjuncts). Under this implementation of categorial distinction, the categorial feature is a formal feature, since its value is determined in narrow syntax just like ϕ, Case, and so on. Thus, Structural Parallelism, which generalizes a parameter value, i.e., presence/absence of a functional projection, in one domain to another domain, can be considered to be calculated based on the

categorial feature, which is now understood as a formal feature. It should also be noted that (112a) does not necessarily force another domain to have a functional domain. This is so because the PLD can override this learning bias, just as IG can be canceled by the PLD as mentioned above. Thus, we can conclude that Structural Parallelism can be subsumed under IG.

As mentioned above, generalizing the presence of a functional projection, which is now understood as a parameter value under the BCC and FE, can be canceled upon relevant input in the PLD; a set of "mixed" parameter values can emerge this way. I submit that non-canonical DP languages such as Italian, Greek, and Hungarian discussed in this chapter emerge as a result of the PLD and IG, when the latter is partly overridden by the former. Since those languages have definite articles, there is good reason for a learner to acquire DP given that the definite article is a trigger of projection of DP, as proposed by Koulidobrova (in press). In the case of Italian, the definite articles can adjoin to N and Poss (in the absence of an adjective), and this is "hinted at" by the morpho-syntactic nature of the definite articles; they can cliticize onto (i.e., adjoin to) a verbal head. Thus, the learner of Italian knows that the definite article need not project DP. On the other hand, there is no need for the learner to cancel the presence of DP in indefinite nominal phrases, since there is no cue for this in the PLD (but see below for a potentially relevant factor). Thus, the interaction of the PLD and IG allows a learner of Italian to have the less-DP-language grammar, in which DP is present in the absence of the definite articles and is absent in the presence of them. (But see section 5.6.2 for another possible path of the emergence of the non-canonical DP-language grammar of Italian.)

Greek is even less of a DP-language in the scale of the NP/DP-language distinction. I proposed in section 5.5 that Greek allows bare NP in the absence of the definite article. Greek then seems to resort to the default option of UG in the absence of the definite article, namely, not projecting DP. But here one might wonder why IG does not generalize projection of DP by a definite article to other nominal phrases where the definite article is absent, given that presence of a definite article in the PLD is a trigger for acquisition of DP. One possibility worth considering is the nature of the indefinite article *enas* in Greek. As noted in section 5.5.4, Greek allows bare singular

NPs, which means that the indefinite article is not obligatory, which is contrasted with English and Italian, where the indefinite article is obligatory. Although Bošković (2008b, 2012) does not consider indefinite articles as relevant for his NP/DP-language distinction, it is usually assumed that indefinite articles are a head of some functional projection in the nominal domain, if not identical to D. Recall also that under Bošković's definition, obligatory presence in nominal phrases with definite interpretation is one of the criteria for definite articles, which are a trigger of presence of DP in a language. It may, then, not be implausible that obligatory presence of an indefinite article would also be relevant for projection of DP in indefinite nominal phrases. (It should be added here that the presence of a definite article in a language is still a prerequisite for projection of DP in general. If a language has an indefinite article but lacks a definite article, DP would not project in the language; see Bošković 2009b, who in fact shows that Slovenian, which has an indefinite article but lacks a definite article, exhibits properties of an NP-language). Interestingly, as noted in section 5.5.6, Markopoulou (2000) even claims on synchronic and diachronic grounds that *enas* is actually not an indefinite article (but an indefinite pronominal adjective; see section 5.5.6 for relevant discussion). If presence of an indefinite article is (at least one) requirement for projection of DP in indefinite nominal phrases, the lack of an indefinite article can be considered to be responsible for the absence of DP in indefinite nominal phrases in Greek.[50] I will also discuss this point in section 5.7.

Hungarian, another less-of-a-DP-language that was discussed above, is interesting in this context. In section 5.4, I proposed that the definite article in Hungarian can adjoin to Poss without projecting DP. In addition, I took the presence of productive compositional indeterminate pronouns in Hungarian as indicating that Hungarian can omit the DP layer. This means that IG does

---

50. In footnote 35, I noted that Stojković (2019) and Stanković (2019) report that there are some speakers who accept adjective LBE in Bulgarian and Macedonian. I suggested there that the definite articles in the grammar of those speakers may be able to adjoin to A just like those in Greek, which means that these varieties of Bulgarian and Macedonian can be analyzed as even less of DP-languages than the varieties reported by Bošković (2005) and LaTerza (2014) among others. Interestingly, Geist (2013) claims that Bulgarian *edin*, which has been treated as an indefinite article, is actually not fully grammaticalized as an indefinite article yet, a claim similar to the one about Greek *enas* made by Markopoulou (2000). These similarities can then be taken as supportive evidence that the lack of a fully grammaticalized indefinite article is relevant for absence of DP in indefinite nominal phrases (here "full grammaticalization of an indefinite article" may also be understood as the indefinite article projecting its own projection instead of being head-adjoined to a nominal head; see Wang 2019 and related discussion in section 5.4).

not generalize presence of DP to all instances of nominal phrases in Hungarian despite the presence of the definite article in this language. Here again, a potentially relevant factor can be the indefinite article. Although the traditional Hungarian grammar has treated *egy* as an indefinite article, MacWhinney and Bates (1978:544) note that "Hungarian uses the numeral "one" as an indefinite article when the speaker is drawing attention to non-plurality or non-genericness. In other cases, the Hungarian noun appears without an article." Some examples of bare singular in Hungarian are given in (113). Given the discussion above, this can be interpreted as indicating that *egy* may actually not be an indefinite article, which in turn may indicate the lack of DP in indefinite nominal phrases.

(113) a. Péter **jó**   **tanuló**.

Péter good student

'Peter is a good student.'   (Alberti and Laczkó 2018:1002)

   b. Péter **levelet**   ír.

Péter letter.ACC write.3SG

'Peter writes a letter.'   (Alberti and Laczkó 2018:1002)

   c. Mari **szép**   **levelet**   írt.

Mary beautiful letter.ACC write.PAST

'Mary wrote a beautiful letter.'   (Kiefer 1990:152)

Yet another possibility for a cue for lack of DP in Hungarian is noun-incorporation. As mentioned in section 5.4, Kiefer (1990), Farkas and de Swart (2003), among others, analyze non-modified bare singular such as (113b) as an instance of noun-incorporation. Baker (1988) proposes that noun-incorporation is an instance of head-movement of non-branching N, which would then be impossible if DP projects above NP because movement of N to V would be blocked by D due to the Head Movement Constraint (Travis 1984, Rizzi 1990). Given this, the availability of noun-incorporation in Hungarian may well be a cue for a learner of Hungarian not to generalize projection of DP triggered by the presence of the definite article to indefinite nominal phrases.

To summarize so far, I have argued that the fine-grained scale of NP/DP-language distinction advocated in this chapter nicely fits the emergentist view of parameters, under which parameters are not prespecified in UG but emerge from interactions of the PLD and two third-factor principles, FE and IG. Given FE, which requires a learner to postulate as few formal features as possible based on the PLD, the grammar of a canonical NP-language, in which DP does not project above NP, should be the default option of UG, since D is a bundle of formal features that need to be acquired upon the PLD and hence a point of parameterization under the BCC. The presence of a definite article in the PLD is a trigger for the learner to acquire DP, and it can be generalized by IG to other nominal phrases, which yields canonical DP-languages. IG can, however, be overridden by the PLD, and some lexical items that have peculiar properties need not follow the generalized parameter. Definite articles in Italian and Hungarian have been proposed to be such lexical items. Cancellation of projection of DP in some nominal phrases due to those elements, as well as other potential factors such as optionality of indefinite articles and noun-incorporation, yield less-DP-languages, in which DP may be absent in some cases.

5.6.2 Acquisition of definite articles and economy of structure building

The nature of definite articles also seems to be relevant to language acquisition. Brown (1973) and Warden (1976) observe that the definite article in English is acquired no earlier than age 4, where "acquired" means that it is used in appropriate definite contexts with more than 80% of accuracy. Interestingly, Guasti et al. (2008) observe that the definite articles in Italian (and Catalan) are acquired (i.e., more than 80% accuracy) at age 2;5.4, which is much earlier than acquisition of the definite article in English. The contrast between English and Italian is surprising, given that the syntax of the definite article in English is quite simple, i.e., it always projects DP, whereas those in Italian are syntactically more complicated (i.e., they either project DP or are base-generated adjoined to another head), hence the English definite article seems to be "easier" from the syntactic viewpoint. It is also worth adding here that the definite article is the most frequent word in English, hence it should be the most frequent input for learners of English. A natural question is, then, why

the definite articles in Italian are apparently "easier" than that in English from a viewpoint of language acquisition.

I suggest that this difference can be attributed to the syntactic properties of the definite articles in question and economy of structure building in language acquisition. Under the emergentist view of parameters, bare NP is the default option compared with projection of DP above NP. This is deduced from FE, which postulates as few formal features as possible, and D being a bundle of formal features. I would like to suggest that this economy principle can be generalized beyond formal features. Specifically, if postulating as few elements as possible is a general economy principle, it is not implausible to hypothesize that it is also economical to postulate as few *functional projections* as possible. In other words, bare lexical structures are most economical and hence desirable state of the grammar, and functional projections can emerge only if there is good reason to postulate them given the PLD. Thus, I propose Structure Economy (SE) as (114), based on the formulation of FE in (108).[51]

(114)   *Structural Economy*

   Given a pair of adequate structural representations R, R′ for a substring of input text of the PLD S, choose R iff R has n distinct projections and R′ has $m > n$ distinct projections.

Recall now that I have proposed that the definite articles in Italian need not project DP, in contrast with the definite article in English, which always projects DP. In other words, the definite articles in Italian can maintain the "default" bare NP structure regardless of their presence in the structure. Thus, they can be acquired with SE observed, hence acquired early. On the other hand, the definite article in English always projects DP, so that a functional projection needs to be acquired given the PLD. This is less economical given SE and is dispreferred to be acquired instantly. Still, projection of DP is the only option for the English definite article, so that it is eventually acquired, resulting in the delay of its acquisition compared with the definite articles in Italian.[52]

---

51. Again, this is essentially similar to Economy of Representation in early minimalism; see Bošković (1997c), Speas (1994), and references therein.

52. Anderssen (2007) observes that the affixal definite articles in Norwegian are acquired at the age of 2;0.5, whereas the non-affixal articles in the same language are omitted very frequently in contexts where they are required even at the age of 2;7.8. This can also be captured by SE; affixal definite articles need not project DP so they can be acquired

At this point, one may wonder how SE fits with the NO > ALL > SOME path of parameter setting discussed in section 5.6.1. Under the emergentist view of parameters, a learner starts from the NO-stage as per FE, and proceeds to the ALL-stage. If there is input that can override the ALL-stage setting in the PLD, the learner changes the relevant parameter values, resulting in the SOME-stage. In section 5.6.1, I suggested that the grammar of Italian with respect to the nominal domain corresponds to the SOME-stage grammar, in which DP need not project in the presence of a definite article. This is contrasted with English, whose nominal domain always requires projection of DP (i.e., the ALL-stage grammar). If a learner first acquires the ALL-stage grammar after the NO-stage, why would children learning Italian not first acquire obligatory projection of DP, as a result of which they would postpone acquisition of the definite articles, just like learners of English? To put it differently, why would children learning English not acquire the Italian-type grammar as per SE and hence acquire the definite article earlier?

One possibility is that learners of Italian actually acquire an ALL-stage in which definite articles are always adjoined to a nominal head. This would observe SE by not postulating a functional projection and also observe IG by generalizing the parameter value (i.e., adjunction of D to a nominal head) to all cases in the nominal domain. As mentioned in section 5.6.1, a clue for the adjunction option may be the clitic nature of the definite articles in Italian in the PLD. Later on, the learner would learn that the definite article can also project DP, which would be the SOME-stage. In contrast, the English definite article is not a clitic, so the only option for structure building in the nominal domain is obligatory projection of DP. This is generalized by IG but acquired later because of SE. Alternatively, in an early stage of English in which the definite article is not fully acquired

early, whereas non-affixal ones must project DP so they are acquired late.

Stephany (1997) and Marinis (2003) also observe that definite articles in Greek are acquired by age 2;10, which is much earlier than in English. A question that arises here is why acquisition of the definite articles in Greek is earlier than acquisition of the definite article in English, although the definite articles in both languages always project DP in the nominal domain. A possible speculation is that Greek definite articles are "easier" because they are pronominal clitics at the same time. When used as pronominal clitics, which have been argued to be minimal and maximal projections at the same time (Chomsky 1995b, Bošković 2002a), they adjoin to a head in the clausal spine. Thus, they have an option of not projecting DP by taking an NP complement (they are minimal/maximal projections that do not project further under BPS). Note that this option is not available for the English definite article, which always takes a nominal complement and projects DP. Thus, the difference between Greek and English regarding acquisition of the definite articles may be attributable to the clitic nature of the Greek definite articles.

yet, the definite article may actually be adjoined to a nominal head as per SE. However, since the definite article in English is not a clitic unlike those in Italian, the learner would be required to eventually acquire a grammar in which the definite article always projects DP. This parameter setting to the adult English grammar, which is the ALL-stage, is still postponed by SE, hence acquisition of the definite article in English is delayed compared with acquisition of definite articles in Italian. If this is on the right track, it would mean that the typological classification of NO-ALL-SOME grammars would not necessarily correspond to the order of acquisition of NO-ALL-SOME stages, contra Biberauer and Roberts (2017), who suggest that these two should coincide; namely, although the grammar of Italian, in which DP need not project in the presence of the definite article, is a SOME-stage in the typological classification of NP/DP-languages, the option of the definite articles not projecting DP may be an ALL-stage in the course of acquisition as per IG and SE, as discussed above.

The possibility that definite articles always adjoin to a nominal head in an early stage of language acquisition is also compatible with Wang's (2019) proposal that in the grammaticalization process of an indefinite article from the numeral 'one', there is an intermediate stage in which 'one' is head-adjoined to a head in the nominal domain rather than projecting its own functional projection. As discussed in section 5.4, it is not implausible to hypothesize that definite articles, which have developed from demonstratives, were first head-adjoined to a head in the nominal domain rather than projecting DP. Thus, there may be a parallelism between the historical change in the structure of the nominal domain and acquisition of definite articles; definite articles adjoin to a nominal head in an early stage of acquisition and during grammaticalization, projecting DP at a later stage (see also Dadan 2019, Lightfoot 1979, van Gelderen 2011, Roberts 2007 among others for a more general discussion of the relationship between language acquisition and language change).

5.7 On relevance of grammaticalization of indefinite and definite articles for bare indefinite NPs

In the previous section, I suggested the possibility that the presence/absence of a (fully grammaticalized) indefinite article may be relevant for the availability of bare indefinite NPs which lack a functional projection. In this section, I would like to briefly discuss this from the perspective of Egyptian Arabic, which has prefixal definite articles, building on Soltan (2020), who discusses the NP/DP-language status of Egyptian Arabic. I argue that the behavior of Egyptian Arabic can be taken as supportive evidence for the possibility that the presence/absence of a (fully grammaticalized) indefinite article may correlate with the availability of bare NPs which are not dominated by a functional projection. I also discuss Basque, whose indefinite and definite articles have different distribution from those in languages such as English. I suggest the possibility that Basque may be a language which is "losing" the DP layer in the nominal domain, with the definite article in Basque undergoing grammaticalization.

Soltan (2020) shows that Egyptian Arabic exhibits a number of properties that NP-languages show, although Egyptian Arabic has (prefixal) definite articles. For instance, he observes that adjunct extraction out of a nominal phrase is possible (only) in the absence of the definite article in Egyptian Arabic, which is similar to Bulgarian and Greek, as shown in (115a). If the definite article is present, the interrogative adjunct PP must be embedded in a post-nominal relative clause, as shown in (115b). Extraction out of the definite nominal is banned as seen in (115c).

(115) a. [min ʔanhī balad]$_i$ ʔinta ʔabil-t   [banāt t$_i$]?
            from which country you   met-2.SG.M girls

      b. ʔinta ʔabil-t   [ʔil-banāt [ʔillī [min ʔanhī balad]]]?
          you met-2.SG.M the-girls C   from which country

      c. *[min ʔanhī balad]$_i$ ʔinta ʔabil-t   [ʔil-banāt [ʔillī t$_i$]]?
          from which country you met-2.SG.M the-girls C   (Soltan 2020:239)

This can be taken to indicate that DP is not projected above NP in the absence of the definite article, just as in Bulgarian and Greek.

Another notable observation made by Soltan concerns null objects. Egyptian Arabic allows null objects, which can furthermore have a sloppy reading. Thus, in (116), the null object (indicated by the underline) refers to a different book from the one Mona read.

(116) Mona laʔ-it   kitāb$_i$ wi Huda kamān laʔ-it   __$_i$.
       Monda found.3SG.F book and Huda also   found.3SG.F
       'Mona found a book, and Huda found [a book] too.'   (Soltan 2020:206)

Note that the antecedent of the null object in (116), i.e., *kitāb* 'book', is indefinite. Importantly, if the antecedent is definite, a null object is not allowed to refer to it. Instead, a pronoun clitic must be used to refer to the definite antecedent. This is illustrated in (117).

(117) a. *Mona laʔ-it   ʔil-kitāb$_i$ wi Huda kamān laʔ-it   __$_i$.
           Monda found.3SG.F the-book and Huda also   found.3SG.F
           Intended: 'Mona found the book, and Huda found [it] too.'

       b. Mona laʔ-it   ʔil-kitāb$_i$ wi Huda kamān laʔ-it-**u**$_i$.
           Monda found.3SG.F the-book and Huda also   found.3SG.F-it
           Intended: 'Mona found the book, and Huda found it too.'   (Soltan 2020:206)

Following Cheng (2013) and Bošković (2018b), Soltan proposes that indefinite nominal phrases in Egyptian Arabic are NPs of type ⟨e,t⟩ and hence can be null objects similarly to those in Japanese, whereas definite nominal phrases in the language are DPs of type e and hence cannot be null objects. Based on these contrasts between indefinite and definite nominal phrases, Soltan essentially claims that the NP/DP-language distinction is not a two-way cut as proposed by Bošković (2008b, 2012), and that the relevant distinction can be manifested within a single language.[53] His insight can be straightforwardly integrated into the current proposal; Egyptian Arabic is less of a DP-language than English in the NP/DP-language scale.

---

53. See Soltan (2020) for additional empirical arguments for this claim.

Crucially for the current context, Egyptian Arabic lacks an indefinite article (Abdel-Malek 1972). In the previous section, I suggested based on Greek and Hungarian that obligatory presence of an indefinite article in indefinite singulars, which means that the indefinite article is fully grammaticalized, is necessary for projection of DP in indefinite nominal phrases. In other words, if a language lacks a fully grammaticalized indefinite article, an indefinite nominal phrase can be bare NP that lacks the DP layer. The NP-language properties of indefinite nominal phrases in Egyptian Arabic, i.e., adjunct extraction out of an indefinite nominal phrase and null objects referring to an indefinite antecedent, are then expected; since Egyptian Arabic lacks an indefinite article, DP does not project above indefinite nominal phrases. Thus, the behavior of indefinite nominal phases in Egyptian Arabic, which lacks an indefinite article, can be taken as additional supportive evidence for the correlation between the presence of a fully grammaticalized indefinite article and the availability of bare indefinite NP with no DP layer.

Basque may be interesting in this context. Basque has suffixal definite articles (-*a* for singular and -*ak* for plural). Basque also has compositional indeterminate pronouns, which I argued are possible only in languages where DP can be absent (see chapter 3). In addition, Basque allows null objects which can have a sloppy reading, similarly to Japanese (Takahashi 2007, Duguine 2014). In (118b), the null object can refer to a book different from the one Jon read. These points indicate that Basque is less of a DP-language than English in the NP/DP-language scale.

(118)   a.   Jon-ek   liburu-a   astiro   irakurri du.
             Jon-ERG book-the   slowly   read   AUX
             'Jon read a book slowly.'

        b.   Miren-ek   ere   __   irakurri du.
             Miren-ERG   also   read   AUX
             'Miren read [a book], too.'   (Takahashi 2007)

Basque is also considered to have an indefinite article (*bat*). However, the distribution of *bat* is quite different from that of indefinite articles in languages such as English. Hualde and Ortiz de

Urbina (2003:122) note that "[*bat*] is used much less freely than the indefinite articles of English and other western European languages". In fact, in (118a), the indefinite specific nominal *liburu-a* 'a book' is not accompanied by *bat* but by the definite article *-a*. In Basque, the definite article marks specificity rather than definiteness, hence can be used in indefinite specific contexts instead of *bat* (see below for discussion of the definite article).[54] This can be taken as indicating that *bat* may not be a fully grammaticalized indefinite article, which in turn provides support for the claim that the presence of a fully grammaticalized indefinite article in a language is required for projection of a functional projection above indefinite NPs.

Note also that the distribution of the definite article in Basque is also different from that of definite articles in, e.g., English. As mentioned above, the definite article in Basque encodes specificity rather than definiteness. Interestingly in this context, Greenberg (1978) proposes that grammaticalization of definite articles proceeds in the following way:

(119)   Demonstrative (Stage 0) > Definite article (Stage I) > Specific article (Stage II) > Noun marker (Stage III)

The Basque definite article would then be at Stage II of grammaticalization. Interestingly, Lyons (1999) suggests that the DP structure may be lost in languages at Stage II of grammaticalization of definite articles. Basque may, then, not need to have the DP layer above NPs despite the existence of the "definite articles" (even independently of their affixal status).[55] I would like to investigate this point in future research.

---

54. Hualde and Ortiz de Urbina (2003:122) also note that '[*bat*] often corresponds more directly to 'a certain', rather than merely to 'a(n)', which means that *bat* also has specific interpretation (as well as non-specific interpretation). In the Souletin variety of Basque, where bare nouns can appear in the object position, bare nouns can only be interpreted as non-specific (Etxeberria 2014).

55. Note also that there is a definite article in *liburu-a* 'book-the' in (118a), and yet it can be the antecedent of a null object that has the sloppy reading in (118b), in stark contrast with Greek and Egyptian Arabic, where a nominal phrase with a definite article cannot be an antecedent of a null object. This can then be taken as additional evidence that Basque is losing the DP layer despite the existence of the definite article.

5.8   Conclusion of the chapter

In this chapter, I have argued that there are languages that cannot be captured by a two-way or a three-way cut of the NP/DP-language distinction proposed by Bošković (2008b, 2012) and Talić (2015, 2017), and that we need a more fine-grained distinction, or a scale, of NP/DP-languages. As illustrations, I have shown that Italian, Hungarian, and Greek exhibit some syntactic properties that would be difficult to capture under a two-way or a three-way cut, such as adjunct extraction out of a nominal phrase in the presence of a definite article (Italian), non-complementary distribution of reflexive and pronominal possessives (Italian and Hungarian), noun-incorporation (Hungarian), (semi-)productive compositional indeterminate pronouns (Hungarian and Greek), adjective LBE (Greek), and null objects with the sloppy reading (Greek). The intermediate behavior of Italian and Hungarian is attributed to the syntactic nature of the definite articles. Specifically, I have proposed that the definite articles in Italian and Hungarian can be base-generated as adjoined to a head in the nominal domain without projecting DP, which is allowed under Bare Phrase Structure. Additionally, regarding Greek, where the definite article always projects DP just like English *the*, it can also adjoin to A via movement, in a similar way as cliticization of a preposition onto an adjective in Serbo-Croatian, which ends up making possible adjectival extraction out of the nominal domain, which is not normally found in DP-languages (though this is done in a different way from NP-languages). At any rate, the language-specific features of the definite articles noted above yield the fine-grained scale of the NP/DP-language distinction.

In addition, I have suggested that the presence of a fully grammaticalized indefinite article may be required for projection of DP in indefinite nominal phrases. To put it differently, languages without a fully grammaticalized indefinite article may lack the DP layer in indefinite nominal phrases. This is supported by the behavior of e.g., Greek and Egyptian Arabic, where an indefinite article is not fully grammaticalized or does not exist and indefinite nominal phrases exhibit some properties of NP-languages such as the availability of adjunct extraction out of them and null indefinite arguments. Thus, indefinite articles play a role in projection of a functional projection above NP in indefinite nominal phrases, just as definite articles do in definite nominal phrases.

I have also discussed the scale of the NP/DP-language distinction from the perspective of the emergentist view of parameters, under which UG is invariant, as is assumed in minimalism, and cross-linguistic variation emerges by interactions of acquisition of formal features and third-factor principles. In particular, I have argued that bare NP structure that is not dominated by DP is the default option of UG because of a third-factor principle that postulates as few formal features as possible, and then the DP structure is acquired later by acquisition of relevant features that project DP. In addition, I have proposed that postulating as few projections as possible is a part of a general economy principle, and that this can explain why the definite articles in "intermediate" DP-languages such as Italian are acquired earlier than that in English, which is a full DP-language.

Finally, I have discussed grammaticalization of definite and indefinite articles. In particular, I have suggested that at an intermediate stage of grammaticalization they adjoin to a nominal head via base-generation without projecting its own phrase.

Chapter 6

Article Drop in Non-canonical DP-languages: P-N Affinity, Uniqueness, and Relevance of Modifiers

6.1   Introduction

Under Bošković's (2016b) definition of definite articles in (1) adopted in this dissertation, definite articles are expected to be obligatorily present in a nominal phrase with definite interpretation. As Bošković argues, this follows from Chierchia's (1998) blocking principle, by which presence of a lexicalized semantic operator in a language blocks covert application of the operator at LF.

(1)   DEFINITION: A *definite article* (i) has the meaning of an iota operator, (ii) obligatorily occurs in a nominal phrase with a definite interpretation, (iii) occurs only once in a nominal phrase, and (iv) has a form distinct from demonstratives.

This conception of definite articles leaves a loophole where definite articles would not be present when they do not encode definiteness, which Talić (2015, 2017) in fact observes holds for cases such as superlatives in affixal article languages like Bulgarian.

(2)   Ivan ima naj-dobri**(-te)**   albumi ot U2.

       Ivan has SPRL-good-the albums by U2

       'Ivan has *(the) best albums by U2'   (Pancheva and Tomaszewicz 2012:295-296)

In this chapter, I introduce additional conditions under which definite articles can be omitted. Crucially, however, in those conditions, they are dropped even with definite interpretation, contrary to the expectation under the definition of definite articles in (1). Interestingly, those cases of article drop are observed in non-canonical DP-languages in the NP/DP-language scale argued for in chapter 5. I argue that DP is actually absent in those article drop cases, which is consistent with the current view that DP can be omitted in some cases in non-canonical DP-languages.

There are two major cases that I discuss in this chapter. One of them is article drop in PPs. It is shown that certain PPs in languages such as Romanian and Albanian have a definite interpretation, although the definite article is missing. Interestingly, the definite article is forced to be present when the noun in the relevant PPs is modified by an adjective. Based on Chomsky's (1970) feature specifications of the traditional lexical categories, I propose that there is a P-N affinity, in the sense that P can function as the highest functional element of the extended functional projections in the nominal domain, a lá Grimshaw (2000), Bošković (2016b), Zanon (2020). This P-N affinity is blocked by the presence of AP above NP as proposed in chapter 5 (cf. Bošković 2005), due to the difference in the specification of the categorial feature. In addition, I propose a structural restriction on the feature that is responsible for definiteness, by which the presence of P as the highest functional element in the nominal domain disallows D to be present. I also discuss preposition-article contraction in German as a related case of the P-N affinity, and suggest that DP does not project in this case despite the presence of the (contracted) definite article. This is taken as indicating that German is slightly less of a DP-language than a canonical DP-language like English, though it is more of a DP-language than, e.g., Italian.

The other major case of article drop discussed in this chapter concerns kinship terms. It is shown that the definite article is dropped in a possessive phrase when the possessum is a kinship term in a number of languages, although the entire possessive phrase has a definite interpretation

and other definite possessive phrases with non-kinship nouns have definite articles in the same languages. It is further demonstrated that the definite article is required with kinship possessums in these languages when an adjective modifies the kinship term. I propose that this is due to the semantics of kinship terms and adjectives. The intuition is that those kinship terms are inherently unique, and hence do not require an additional overt iota operator, namely, a definite article. I propose a semantic composition of the kinship possessums that appeals to a difference in semantic types between kinship and ordinary nouns. In addition, this composition becomes unavailable when an adjective is present due to a type mismatch. The type mismatch forces kinship terms to have the same semantic type with ordinary nouns, which in turn requires a definite article to be present as an iota operator for definiteness. I argue that DP is absent in the case of unmodified kinship possessums whereas it is present when the possessum is modified by an adjective, which is consistent with the current proposal that DP can be absent in non-canonical DP-languages.

The chapter is organized as follows. In section 6.2.1, it is shown that the definite article is obligatorily absent in certain PPs in some languages, but it is required to be present when the noun in the PP is modified by an adjective. I propose in section 6.2.2 that P can serve as the highest functional element in the nominal domain and that the presence of P blocks projection of DP. P cannot, however, be the highest functional element in the nominal domain in the presence of an adjective, because AP that projects above NP blocks the affinity between N and P that is calculated based on categorial features. In section 6.3 I discuss preposition-article contraction in German, which I argue has a different structure from non-contracted cases. Here again, the presence of P plays a crucial role for the relevant structural difference. Section 6.4 discusses the syntax-semantics interface of article drop with kinship terms in possessive phrases. In section 6.4.1, it is shown that kinship terms in possessive phrases resist the definite article in non-canonical DP-languages, although ordinary nouns in possessive phrases require the definite article in the same languages. Furthermore, the definite article is required with kinship term possessums when they are modified by an adjective. In section 6.4.2, I propose that kinship terms are by default of type $\langle e,e \rangle$ and do not require a definite article for the definite interpretation, whereas they are of type $\langle e,t \rangle$ in the

270

presence of an adjective in order to avoid type mismatch. This means that DP is absent in the former case, whereas it is present in the latter case. Section 6.5 concludes the chapter.

6.2 Article drop in the presence of P

6.2.1 Data

The first case of article drop with definite interpretation is observed with PPs. Mardale (2006) shows that certain PPs, which are typically locative, resist definite articles in Albanian and Romanian, as seen in (3a) and (3b), respectively. Zwicky (1984) also notes that definite articles are dropped in locative PPs in Yiddish (3c) (see also Verschik 2001 for Estonian Yiddish).

(3) a. Vuri librin   mbi trapezë(**-n**).

put   book.the on table-the

'He puts the book on the table.'   (Albanian, Mardale 2006:4)

b. Mă îndrept către   parc(**-l**).

me head   towards park-the

'I'm heading towards the park.'   (Romanian, Mardale 2006:2)

c. in feld

in field

'in the field'   (Yiddish, Zwicky 1984:120)

This appears to be similar to bare singulars in locatives found in languages like English as shown in (4), which is restricted to a narrow lexical class of nouns (see, e.g., Scholten 2010 and Aguilar-Guevara 2014).

(4) Mary went to school.

However, Mardale (2006) reports that article drop in PPs with definite interpretation is more productive, and also possible with other types prepositions in Romanian, such as direct object marking

pe (5) and the marker of the Goal theme *la* (6).

(5) L-am   văzut pe professor.

     him-have seen   PE professor

     'I saw the professor.'   (Mardale 2006:3)

(6) Dau cărţi   la copii.

     give books to children

     'I give books to the children.'   (Mardale 2006:3)

In addition, Mardale points out that the nouns in the locatives in (3) necessarily receive a definite interpretation; in other words, a non-referential or an indefinite reading is not allowed. This is contrasted with the bare singular in (4), which lacks definite interpretation (see Scholten 2010 and Aguilar-Guevara 2014). In fact, for an indefinite interpretation in the relevant PP in Romanian, an indefinite article must be present, as shown in (7).

(7) Mă îndrept către   *(un) parc.

     me head   towards   a   park

     'I'm heading towards a park.'   (Romanian, Mardale 2006:2)

A question that naturally arises is why the locative PPs in question in Romanian and other languages noted above receive definite interpretation despite the absence of the definite article. An idea I will pursue below is that the prepositions here function as an alternative of definite articles, and that DP is absent in the presence of P in (3), (5), and (6).[1]

Interestingly, these locative PPs require a definite article when the nouns are modified by an adjective.

---

1. It is worth noting here that adpositions can perform other roles in the nominal domain, such as Case. For instance, the preposition *a* in Spanish can be used as a Direct Object Marker, which is essentially a Case marker, in some environments; see, e.g., Jaeggli (1982, 1986). See also Bošković (2006b) for Serbo-Croatian *s(a)*, which is a preposition but can function as a case marker in some environments.

(8) a. Mă îndrept către parc**(-l)** înverzit.

   me head towards park-the green

   'I'm heading towards the green park.' (Romanian, Mardale 2006:2)

 b. [inəm] groys m feld

   in.the big the field

   'in the big field' (Yiddish, Zwicky 1984:120)

Given that omission of the definite article is an indication of the absence of DP in the cases under discussion, the presence of an adjective seems to force projection of DP in (8). It is important to recall at this point that the relevance of presence of a modifier for projection of DP was discussed in chapter 5. There I showed that extraction out of a nominal phrase in Italian is possible when there is a definite article, but this extraction is blocked in the presence of an adjective, as illustrated in (9).

(9) a. [Di che scaffale]$_i$ Gianni ha già letto [i libri t$_i$]?

   of which shelf Gianni has already read the books

   'From which shelf did Gianni read the books?'

   (Bošković 2005, attributed to Giuliana Giusti)

 b. ??[Di che scaffale]$_i$ Gianni ha già letto [i grandi libri t$_i$]?

   of which shelf Gianni has already read the large books

   'From which shelf did Gianni read the large books?'

The gist of the analysis of the contrast between (9a) and (9b) I proposed in chapter 5 is that DP can be absent in the absence of an adjective (by base-generation of D as head-adjunction to N), while DP must be present in the presence of the adjective, which projects AP above NP. The contrast between (3) and (8) can then be assimilated to the similar contrast between (9a) and (9b); abstractly, the presence of a modifier forces projection of DP in both configurations.

6.2.2 Analysis: P-N affinity and extended projections of a lexical category

As mentioned in the previous subsection, P can be considered to function as an alternative of a definite article in the case under discussion. Here I propose that there is a "P-N affinity", in the sense that the prepositions in these cases are part of the extended projection of a nominal domain. In fact, Grimshaw (1990), Bošković (2013a), and Zanon (2020), among others, propose that P can be the highest projection of the extended projections in the nominal domain. The P-N affinity can actually be motivated by the traditional classification of lexical categories proposed by Chomsky (1970), in which N is [+N, -V], A is [+N, +V], V is [-N, +V], and P is [-N, -V]; thus, N and P constitute a natural class as [-V] elements.

It is also worth noting here that the languages mentioned above do not allow P-stranding (see Irimia 2005 for Albanian).

(10) a. *Cine$_i$ ai   vorbit [**despre** t$_i$]?
          what you.have talked   about   (Romanian, Nicolae 2012)

  b. *Vemen$_i$ hot zi [**mit** t$_i$] geredt?
      who   has she with   spoken   (Yiddish, Merchant 2001:96)

Interestingly, Bošković (2016b) proposes that functional heads in general cannot be stranded and that prepositions in non-P-stranding languages are functional elements, whereas those in P-stranding languages are lexical elements (cf. Baker 2003 for the proposal that the functional/lexical distinction is a point of variation with Ps). It is then not implausible to analyze Ps in these languages as functional elements in the extended projection of a nominal domain.

A question that arises here is why omission of D is *forced* in the presence of P. My suggestion here is that D needs to be the highest functional element in the nominal domain, otherwise it could not occur in the structure.[2] To put it differently, the set of features that corresponds to the definite article can be realized as a definite article only if it is the highest element in the nominal domain in

---

2. Under this proposal, *all* in *all the students* should be analyzed as adjoined to DP (rather than projecting its own projection, say, QP), as Sportiche (1988), Benmamoun (1999) and Bošković (2004) in fact propose, since *the* must be the highest element in the nominal domain.

these languages.[3] In the presence of P as the highest projection of the extended projections of the nominal domain, DP would not be the highest projection in the extended projections of the nominal domain in this case.[4] Note also that the languages that allow article drop in PP are affixal article languages, hence less of DP-languages in the scale of NP/DP-language distinction argued for in this dissertation (see chapter 3 for Yiddish). Thus, it is not implausible that D is actually absent in such cases, and the presence of P as the highest functional projection in the nominal domain blocks projection of DP, which needs to be the highest functional projection in the nominal domain.

The next question to be addressed under this proposal is why the bare noun in such cases receives definite interpretation. As mentioned above, Mardale (2006) observes that the bare nouns in the PP in question necessarily receive definite interpretation, despite the absence of the definite article, and hence absence of DP. My proposal here is that P actually contains a feature responsible for definite interpretation, which I dub as Def-feature for ease of exposition. Under the Bare Phrase Structure (BPS) Theory, lexical items that have traditionally been given specific categorial labels are merely bundles of features. As discussed in chapter 5, Chomsky (1995b), building on Borer (1984) and Fukui (1986, 1988), proposes that parameters are reduced to different specifications of formal features in the lexicon (the so-called *Borer-Chomsky Conjecture*). It is then logically possible that Ps in question can in principle have the Def-feature in some languages as a parametric option. I suggest that this option is possible only if P serves as the highest functional element in the extended projections of the nominal domain.[5] Otherwise, the definite article is used as the highest

---

3. In chapter 5, I argued that D can be base-generated adjoined to N, without projecting DP in languages such as Italian. In those cases, D is part of the complex head ($\langle$D, N$\rangle$), which is the highest element in the nominal domain. Thus, the set of features that corresponds to a definite article is realized as a definite article in those cases.

4. Potentially related to this is loss of wh-movement. Ledgeway (2012) observes that in Latin, which was a multiple wh-fronting language, multiple wh-fronting showed superiority effects, which are taken as a diagnostic of wh-movement targeting the highest clausal projection in the literature (Rudin 1998, Bošković 2002b, Richards 2001). In contrast, Modern Romance languages (except for Romanian) have lost multiple wh-fronting, and Bošković (2021b) argues that Spanish wh-fronting does not target the highest clausal projection, based on the observation that an interrogative pronoun can follow a complementizer in an embedded clause (Uriagereka 1988, Rizzi 2001, Villa-García 2015). Interestingly, Spanish has also acquired (restricted) wh-in-situ (see, e.g., Reglero 2007, Reglero and Ticio 2013). Thus, abstractly, the unavailability of the highest position in the C domain for wh-movement has led to (the possibility of) a loss of wh-movement (Bošković 2021b). Notice now that article drop under discussion can be assimilated to this; namely, the unavailability of the highest position in the nominal domain has led to omission of the article.

5. Part of this follows from the current proposal that P can be the highest functional element in the nominal domain only if it takes a [-V] element as its complement; If P is not part of the extended projection of N, it cannot be nominal

functional element in the nominal domain as the locus of the Def-feature. Under this proposal, the P in question and D actually receive a unified treatment from the perspective of the Def-feature; in both cases, the Def-feature needs to be contained in the highest element of the extended projections in the nominal domain.[6]

Let us now turn to the cases where the presence of an adjective forces presence of the definite article in PPs in question, as seen in (8). Given that the presence/absence of a definite article correlates with the presence/absence of DP in the relevant PP, the obligatory presence of the definite article in the PPs in question in the presence of an adjective for definite interpretation indicates that DP is forced to project due to the presence of the adjective. As mentioned in the previous subsection, the situation we see here is essentially similar to extraction of an adjunct out of a nominal phrase in Italian discussed in chapter 5; the presence of an adjective blocks the extraction that would be possible without the adjective. One of the proposals I offered in chapter 5 is that AP projects above NP in the latter case (Bošković 2005), and the extraction in question is blocked due to the interaction of the PIC and the anti-locality condition, where AP is a phase. In the spirit of this analysis, I propose that projection of AP above NP blocks the P-N affinity and forces DP to be the highest functional projection in the nominal domain. Recall that P can be the highest functional projection in the nominal domain because of the P-N affinity, which is calculated based on their categorial features; P is [-N, -V] and N is [+N, -V], so they constitute a natural class as [-V] elements. In other words, the complement of P needs to be [-V] in order for P to be the highest

---

in nature in the first place, so that the Def-feature, which is only assigned to a nominal element, cannot be assigned to P in such cases.

6. Direct Object Marking (DOM) in languages like Spanish may also be relevant here. As mentioned in footnote 1, prepositional elements can function as Direct Object Markers. Interestingly, as is well-known, DOM is cross-linguistically subject to semantic restrictions cross-linguistically, especially definiteness or specificity. Direct Object Markers may, then, receive a uniform treatment with definite articles and the Ps discussed in the text, i.e., DOM may be an element that appears with a feature responsible for definiteness/specificity as the highest element in the nominal domain.

It is worth mentioning here that the accusative case marker in Turkish also encodes specificity, which is similar to DOM in Spanish. The accusative case marker in Turkish may, then, be the highest functional element in the nominal domain. Interestingly in this context, although Turkish lacks definite articles and shows behavior of an NP-language (see, e.g., Bošković and Şener 2014), Turkish lacks productive compositional indeterminate pronouns, which is a characteristic of an NP-language (though as a one-way correlation). In chapter 5, I argued for a scale of the NP/DP-languages distinction, in particular, a scale of DP-language-hood, but there can be a scale on the other side of the distinction, namely, a scale from canonical to non-canonical NP-languages. Turkish may then be less of a canonical NP-language in this scale. I would like to pursue this topic in future research.

functional projection in the nominal domain. Crucially, under this feature-based classification of lexical categories, A is [+N, +V]. Thus, when an adjective is present and projects AP above NP and below PP, AP "intervenes" between PP and NP in terms of the categorial feature; P, which is [-V], is merged above AP, which is [+N], so that PP does not count as the highest projection of the extended projections in the nominal domain, for which the complement of P needs to be [-V]. Since P cannot be the highest functional projection in the nominal domain in this case, D needs to project above AP as the functional projection that carries the definite interpretation. Thus, the definite article, which corresponds to D, must be present in the presence of an adjective.

This account could potentially be extended to capture another case in which the definite article must be present in PP where it is otherwise omitted in the relevant languages. Mardale (2006) observes that in Romanian, when the noun in the relevant PP is marked as plural, it cannot have definite interpretation, unlike its singular counterpart seen above. This is illustrated in (11). Höhn (2014) observes the same point for Basque, an affixal article language, where the definite form of locative is missing in the context of the linker -*ko* in locatives (12a), but it needs to be present when the noun is plural (12b).[7]

(11) Am   pus romane-le pe **rafturi**.
      AUX put novels-the on shelves
      'We/I put the novels on shelves. (NOT: on the shelves)'   (Romanian, Mardale 2006:10)

(12) a. lantegi-∅-ko   tximini-a
          factory-(LOC.DEF.SG)-KO   chimney-the
          'the chimney in the factory'

      b. lantegi-**eta**-ko   tximini-a
          factory-LOC.DEF.PL-KO   chimney-the
          'the chimney in the factories'   (Basque, Höhn 2014:148)

---

7. Note that Basque allows productive compositional indeterminate pronouns as discussed in chapter 3 and 4, which means that it is less of a DP-language in the NP/DP-language scale.

What is crucial here is that the presence of the plural number marking blocks omission of the definiteness article/marking, just as the presence of an adjective blocks omission of the definite article. A possible explanation of this can be that NumP projects above NP in the case of plural, and the categorial feature specification of Num is just [+N], with the specification of [±V] missing. The complement of P would then not be [-V] when NumP projects above NP. This would force the presence of the definite article, i.e., D would then have to be present in the presence of a plural marking in these languages.

To summarize this section, I have introduced the observation by Mardale (2006) and Zwicky (1984) that in Albanian, Romanian, and Yiddish, definite articles are dropped in certain PPs although the bare noun in these PPs receives definite interpretation. In addition, the definite article cannot be dropped when there is an adjective that modifies the noun in the PPs in question. I have proposed that P in such cases serves as the highest functional element in the extended projection in the nominal domain, and it is the locus of the Def-feature that is responsible for the definite interpretation. This is motivated by Chomsky's (1970) feature-based classification of lexical categories, i.e., both P and N are [-V]. P can be the highest functional element in the nominal domain only if it takes a [-V] element as its complement, i.e., NP. When there is an adjective, which projects AP above NP, the complement of P is [+V] since A is [+N, +V] under Chomsky's classification, hence P cannot serve as the highest functional projection in the nominal domain. Thus, in the presence of an adjective in the PP in question, the definite article needs to be present and project DP as the highest functional element in the nominal domain.[8]

---

8. Potentially relevant here is the observation that in some article-less languages, a definiteness/specificity marking appears on an adjective, and it is the only locus of the definiteness/specificity marking in those languages (e.g., Serbo-Croatian, Old English, Lithuanian). The current proposal has a potential to explain this observation. A Def-feature, which is responsible for definiteness/specificity, can in principle be present in those languages, but crucially, cannot project DP, since they lack definite articles hence projection of DP in the nominal domain is impossible in those languages (see chapter 5 for relevant discussion). Since DP cannot project in the nominal domain, the only available option of realization of the Def-feature would be to adjoin the Def-feature to A, without projecting a functional projection (though the Def-feature would be part of the head amalgam that is the highest element in the A domain, cf. footnote 3). (See also Despić 2011, who proposes that the relevant endings in Serbo-Croatian are essentially pronominal.)

6.3 Contraction of definite articles in German and relevance of PP

There is another domain where the presence of P is relevant for the use of the definite article. Specifically, I briefly discuss contracted definite articles in (formal) German, which occur in PPs that contain certain nouns. I propose that they are syntactically different from non-contracted definite articles a lá Schwarz (2009); specifically, contracted definite articles are analyzed as being base-generated adjoined to P, similarly to clitics/definite articles in Romance. This in turn allows us to analyze German as a slightly less of a DP-language in the NP/DP-language scale (but more of a DP-language than Romance).

In German, there are two types of definite articles that occur in PPs. One is what Schwarz (2009, 2019) calls a weak definite article, which is contracted with the preceding preposition. The other is what Schwarz calls a strong definite article, which is not contracted with the preceding preposition.

(13) a. Hans ging **zum** Haus.
Hans went to.the$_{\text{weak}}$ house
'Hans went to the house.'

b. Hans ging zu **dem** Haus.
Hans went to the$_{\text{weak}}$ house
'Hans went to the house.' (Schwarz 2009:7)

(13a) is reminiscent of the contraction of the preposition and the definite article in Romance, as represented by Italian (14).

(14) Gianni è andato **al** mercato.
Gianni is gone to.the market
'Gianni went to the market.'

Unlike the preposition-article contraction in Romance, however, the preposition-article contraction in German is semantically constrained. Let us consider the following examples:

(15) a. In der Kabinettsitzung heute wird ein neuer Vorschlag **vom** {✓Kanzler /
in the cabinet.meeting today is a new proposal by.the$_{weak}$ chancellor /

#Minister} erwartet.
minister expected

'In today's cabinet meeting, a new proposal by the chancellor/minister is expected.'

b. In der Kabinettsitzung heute wird ein neuer Vorschlag von **dem** {#Kanzler
in the cabinet.meeting today is a new proposal by the$_{strong}$ chancellor

/ #Minister} erwartet.
/ minister expected

'In today's cabinet meeting, a new proposal by the chancellor/minister is expected.'

(Schwarz 2019:8)

One of the properties of the weak definite article that is used for the relevant contraction is uniqueness; it is used with a noun that refers to an individual identifiable with knowledge of the world without introduction of the noun in the preceding context.[9] In (15a), it is common knowledge of the world that there is only one chancellor and there is more than one minister in a cabinet meeting. Thus, although there is no preceding context that introduces a chancellor, there is a unique chancellor in this situation, and the weak definite article can be used with the noun *Kanzler* 'chancellor' in this case. On the other hand, since there is generally more than one minister in a cabinet meeting, there is no uniquely identifiable minister without a preceding context, hence the weak definite article cannot be used with the noun *Minister* 'minister' in this case. In contrast with the weak definite article, the strong definite article can only refer to an individual that is introduced in the preceding context; without a preceding context it cannot refer to a uniquely identifiable individual

---

9. I am simplifying the discussion of weak definite articles for expository purposes. See Schwarz (2009) for more extensive discussion. What is important here is that the preposition-article contraction is not freely available, in contrast to Romance.

that the weak definite article would be able to refer to. Thus, in (15b), since there is no preceding context, the strong definite article cannot be used, resulting in infelicity. (15b) becomes felicitous when a preceding context is added, as shown in (16).

(16) a. Hans hat gestern   einen Minister interviewt.

Hans has yesterday a   minister interviewed

'Hans interviewed a minister yesterday."

b. ✓In der Kabinettsitzung heute wird ein neuer Vorschlag von dem   Minister

in the cabinet.meeting today is   a   new   proposal by the$_{strong}$ minister

erwartet.

expected

'In today's cabinet meeting, a new proposal by the minister is expected.'

(Schwarz 2019:9)

Schwarz (2009) proposes that the weak definite article and the strong definite article have different semantics, which is reflected in different syntactic structures. Under the framework of situation semantics (Kratzer 1989), he posits (17a) and (17b) as the denotations of the weak definite article and the strong definite article, respectively.

(17) a. $[\![\text{Def}_{\text{weak}}]\!] = \lambda s \lambda P \iota x[P(x)(s)]$

b. $[\![\text{Def}_{\text{strong}}]\!] = \lambda s \lambda P \lambda y \iota x[P(x)(s) \ \& \ x = y]$

The crucial difference is the presence of an additional argument introduced for the strong definite article (represented as *y* in (17b)), which is responsible for the requirement of the presence of an antecedent in the preceding context. Schwarz suggests that this additional argument (which is covert) is located in the specifier of DP. The structures of nominal phrases with the weak definite article and the strong definite article are given in (18a) and (18b), respectively (the additional argument with the strong definite article is represented as 1).

281

(18) a.   DP
              /   \
         D_weak   NP

b.   DP
        /   \
       1   D′
            /   \
        D_strong NP
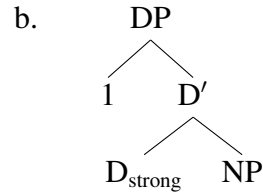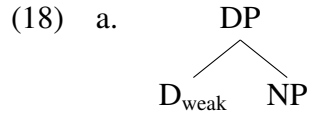
What is interesting for the current purposes is that difference in semantics correlates with a differ-ence in the syntactic structures.

Schwarz does not provide an analysis of the contraction between the preposition and the weak definite article, and merely cites Zwicky's (1982) work as a reference for a possible morphological analysis, in which the weak definite article cliticizes onto the preposition. Zwicky's cliticization is an operation that takes place in PF, which would not affect the syntactic derivation (see also Hinrichs 1986 for discussion). Hanink (2017) and Hanink and Grove (2017) also propose a PF-based analysis of the preposition-article contraction. They modify Schwarz's semantic analysis of the strong definite article and propose that the additional argument represented as 1 in (18b) projects idxP between NP and DP, as schematized in (19).

(19)   DP
        /   \
       D   idxP
            /   \
          idx   NP
           1
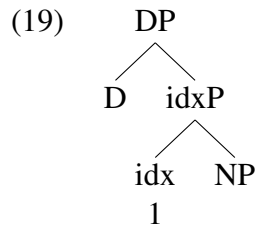
Hanink and Hanink and Grove then propose, under the Distributed Morphology framework (Halle and Marantz 1993b), that P undergoes the post-syntactic operation of Lowering (Embick and Noyer 2001) onto D in the case of weak definite articles, which is realized as the contracted preposi-tion+article, whereas D undergoes Lowering onto idx in the case of strong definite articles, which is realized as the strong definite article without contraction with the preposition. This is illustrated in (20).[10]

---

10. Hanink and Hanink and Grove assume that P vacuously undergoes Lowering to the node in which D was located in (20b), where it is realized as a non-contracted preposition.

(20) a. weak definite articles
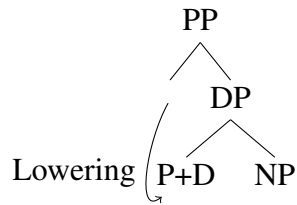
b. strong definite articles
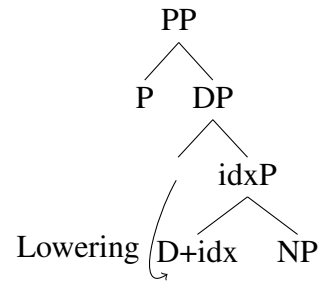
Crucially, under the accounts proposed by Zwicky (1982), Hanink (2017), and Hanink and Grove (2017), the relevant contraction is a PF-operation, and PP projects above DP both in the case of the weak definite articles and in the case of the strong definite articles.

Interestingly, however, Hinrichs (1986) argues against Zwicky's (1982) PF-cliticization analysis, with an argument that can be carried over to Hanink's (2017) and Hanink and Grove's (2017) PF-Lowering analysis. Hinrichs observes that a preposition contracted with a weak definite article cannot be coordinated with a preposition with a strong definite article. The relevant examples in (21) are taken from Puig-Waldmüller (2008).

(21) a. i'm   oder bei'm   Haus
          in.the$_{weak}$ or   at.the$_{weak}$ house

b. *in dem   oder bei'm   Haus
       in   the$_{strong}$ or   at.the$_{weak}$ house

c. *i'm   oder bei dem   Haus
       in.the$_{weak}$ or   at   the$_{strong}$ house   (Puig-Waldmüller 2008:131)

Hinrichs points out that this contrast would be mysterious if the contraction is purely a PF phenomenon as Zwicky (1982) proposes, since the weak definite article is correctly contracted with its preceding preposition. This contrast is also problematic for the PF-Lowering analysis by Hanink (2017) and Hanink and Grove (2017), because under this analysis what is conjoined here would be PPs, and hence the coordinate structure itself should be legitimate given that the conjoined phrases should be of the same category (i.e., the Coordination-of-Likes, see below).

Hinrichs proposes under the GPSG framework that the set of features that corresponds to the definite article is part of the feature composition of the preposition in the case of the preposition-article contraction, whereas the relevant features are not part of the set of the features of the preposition in the non-contraction cases. Notice that this can be easily implemented under the current proposal regarding structure building. In chapter 4 and 5, I proposed, following Epstein et al. (2016) and Saito (2020), that a head can be base-generated as adjoined to another head. The preposition-article contraction can then be analyzed as a result of base-generated adjunction of D to P, which creates a ⟨P, D⟩ amalgam. This amalgam then projects as a complex head. On the other hand, in the non-contracted cases D projects DP above NP due to the richer structure needed for the semantics of the strong definite article, and this DP is dominated by PP. (The additional argument with the strong definite article that Schwarz 2009 argues for can be located in Spec,DP, which forces DP to project, as Schwarz proposes, or project idxP that is selected by D, whereby DP projects between idxP and PP, as Hanink 2017 and Hanink and Grove 2017 propose. The choice between the two positions does not affect the discussion here.) This is schematized in (22).

(22) a. weak definite articles
            ⟨P, D⟩
           /   \
      ⟨P, D⟩   NP
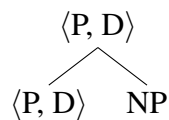
      b. strong definite article
            PP
           /   \
          P   DP
              /   \
             D   NP
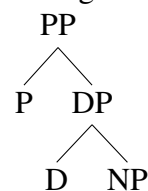
Notice that the traditional "PPs" have different labels in (22a) and (22b): ⟨P, D⟩ and PP. This allows us to explain the ill-formedness of (21b) and (21c) under Bošković's (2020b) implementation of the Coordination-of-Likes in the labeling framework, which I adopted in chapter 2. In Bošković's proposal, all the conjuncts must have the same label. In (21b) and (21c), then, the Coordination-of-Likes is violated, since what is conjoined are ⟨P, D⟩ and PP.

Note also that in (22a), DP is absent in the presence of P. As argued in section 6.2, P can be the highest functional element of the extended projections in the nominal domain. Unlike the cases discussed in section 6.2, the definite article is present in the structure, and P is part of the complex

head, but what is important is that D need not project in this case too.[11] In addition, the Def-feature, which is responsible for definite interpretation, is in the highest element of the nominal domain, i.e., the ⟨P, D⟩ amalgam. Thus, the formal requirement that the Def-feature needs to be contained in the highest element in the nominal domain, which I posited in section 6.2.2, is satisfied in the case under discussion.

It should be added here that the option in (22a) is more generally available in Romance languages, where the preposition-article contraction is not semantically constrained, unlike what is found in German. This is consistent with the proposal in chapter 5 that definite articles in Italian can be base-generated adjoined to N; base-generation of D as adjunction to another head is more generally available in Italian, or Romance languages. In German, in contrast, the contraction in question, i.e., the structure in (22a), is only available if the relevant semantic condition is satisfied in the presence of P. In the context of the NP/DP-language scale advocated in this dissertation, this means that German is slightly less of a DP-language than, e.g., English, but more of a DP-language than Romance languages. This makes sense given that German shows some other properties of a DP-language, such as impossibility of adjunct extraction out of a nominal domain that Italian does not always show (see Bošković 2012).

6.4 Article drop with kinship terms

The second case of article drop with definite interpretation concerns kinship terms. We will see that the definite article must be dropped with a possessum that is a kinship term in languages where other nouns as possessums require the definite article, even when the kinship terms are interpreted

---

11. Interestingly, Schwarz (2009) observes that superlatives require the weak definite article in PPs, as shown in (i).

(i) Auf unserer Reise nach Tibet sind wir natürich auch {**zum** / #**zu dem**} höchsten Berg
    on our trip to Tibet are we of.course also to.the_weak / to the._strong highest mountain
    der Welt gefahren
    the world driven
    'On our trip to Tibet, we of course went to visit the highest mountain in the world.'

In chapter 5, I suggested the possibility that the definite article that occurs in superlatives does not project DP. If this is on the right track, the observation in (i) can be accounted for; in German, which has the weak-strong distinction of definite articles, the weak definite article, which does not project DP, is selected in the presence of the superlative.

as definite. Interestingly, the relevant languages are all non-canonical DP-languages in the NP/DP-language scale advocated in this dissertation. In addition, the article drop with kinship possessums is disallowed when there is an adjective that modifies the kinship term. I propose that there is a syntax-semantics correlation in this domain. Unmodified kinship possessums resist the definite article because of their lexical semantics. This means that projection of DP is blocked, which is consistent with the current view that DP can be absent in non-canonical DP-languages. However, when an adjective is present, kinship terms have a different semantic type, so that the definite article is required for definite interpretation, resulting in projection of DP.

6.4.1 Data

In Bulgarian, the definite article co-occurs with a pronominal possessor, as shown in (23a). Interestingly, Halpern (1995) observes that certain kinship terms resist overt definite articles in such cases, as shown in (23b).

(23) a. statija-**ta** mu
        article-the his
        'his article'

    b. majka-(***ta**) mu
        sister-the his
        'his sister'

This is not limited to Bulgarian. As discussed in chapter 5, Italian definite articles co-occur with pronominal possessors, as seen in (24a). Crucially, they do not appear with certain kinship terms with a pronominal possessive, as shown in (24b).[12]

---

12. As Rákosi (2017) points out, a somewhat weaker version of the restriction in question is observed in Hungarian; the definite article is required with pronominal possessives, as shown in (ia), while omission of the definite article is allowed (but not forced) with kinship terms, as shown in (ib).

(24) a. **la** tua macchina

       the your car

       'your car'

b. (***la**) tua sorella

        the your sister

       'your sister'

Haspelmath (1999) points out that in Somali, Nkore-Kiga, and Vai (in addition to Italian and Bulgarian), kinship terms resist a definite article in the (pronominal) possessive construction, whereas other nouns require one.[13] This is shown in (25)-(27). Note that these three languages are affixal-article languages, hence can be considered to be non-canonical DP-languages in the NP/DP-language scale argued for in this dissertation.

(25) a. saxiib-kay-**gu**

       friend-my-the

       'my friend'

b. hooya-day-(***du**)

       mother-my-the

       'my mother'   (Somali, Haspelmath 1999:236)

(26) a. **e**-kitabo kyangye

       the-book my

       'my book'

(i) a. Szeretem ***(az)** ablak-om-at.
     like.1SG   the window-POSS.1SG-ACC
     'I love my window.'

  b. Szeretem **(az)** anyá-m-at.
     like.1SG   the mother-POSS.1SG-ACC
     'I love my mother.'

13. Note that in Vai (27) the possessor is not pronominal, but the point regarding kinship terms is intact.

b. (***o**)-mukuru wangye

   the-sister   my

‘my sister’   (Nkore-Koga, Haspelmath 1999:236)

(27) a. kaàí-ɛ̌   á   kéŋ-**ɛ̌**

        man-the POSS house-the

        ‘the man’s house’

    b. kaàí-ɛ̌   fǎ-(***ǎ**)

        man-the father-the

        ‘the man’s father’   (Vai, Haspelmath 1999:236)

Haspelmath (1999) further notes that the definite article is optional with pronominal possessors in Brazilian Portuguese and Icelandic, but it cannot co-occur with a kinship term possessum, as shown in (28) and (29). What is important here is that the definite article is disallowed in the possessive construction with a kinship term just as in the languages mentioned above.

(28) a. **(a)** minha casa

        the my   house

        ‘my house’

    b. **(*a)** minha mãe

        the my   mother

        ‘my mother’   (Brazilian Portuguese, Haspelmath 1999:200)

(29) a. hús(**i-ð**)   mitt

        house-the my

        ‘my house’

b.   sonur(**-inn**) minn

      son-the   my

      'my son'   (Icelandic, Haspelmath 1999:200)

Gatt (2004) also observes that in Maltese, an affixal article language, the definite article does not co-occur with a possessum that is a kinship term, although the possessum is interpreted as definite, as seen in (30a). This is contrasted with (30b), where the non-kinship term noun requires a definite article for the definite interpretation. Given that having an affixal article is an indication of being less of a DP-language in the NP/DP-language scale, Maltese can be considered to be less of a DP-language.

(30)   a.   omm   Pietru

          mother   Peter

          'Peter's mother' (definite)

      b.   (**il**)-karozza ta'   Pietru

           the-car   POSS   Peter

          'Peter's car' (definite, when the definite article is present)

          'a car of Peter's' (indefinite, when the definite article is absent)

                                                     (Maltese, Gatt 2004:200)

Soltan (2007) observes a contrast in Egyptian Arabic that makes the same point. In Egyptian Arabic, just as in other Arabic varieties, there are two options to express possessive relations. One is the so-called Free State, in which the possessive relation is expressed by the morpheme *bitaaʕ* (glossed here as ''POSS' following Soltan 2007). The other is the so-called Construct State, in which *bitaaʕ* is absent. Nouns like 'school' can be used in both Free State and Construct State, as shown in (31). Note that the definite article is affixed to the possessum in Free State in (31a). Crucially, a kinship term like 'brother' cannot be used in Contrast State, and hence cannot co-occur with the definite article, as shown in (32).

(31) a. **il**-madrasa bitaaʕ-it ʔaħamad

   the-school POSS-F Ahamad

   'Ahamad's school'

   b. madras-it ʔaħamad

   school-F Ahamad

   'Ahamad's school'

(32) a. *il-ʔax   bitaaʕ ʔaħamad

   the-brother POSS Ahamad

   'Ahamad's brother'

   b. ʔaxuu ʔaħamad

   brother Ahamad

   'Ahamad's brother'

Thus, the general pattern here is that the definite article is disallowed with kinship possessums. As mentioned above, the languages introduced here are all non-canonical DP-languages in the NP/DP-language scale. On a par with article drop in PP discussed in section 6.2, article drop with kinship possessums in these languages can be taken as indicating that the absence of the definite article correlates with the absence of DP.

Interestingly, Halpern (1995) observes that those kinship terms require a definite article when they are modified by an adjective in Bulgarian, as shown in (33). The same effect is observed in Italian, as seen in (34).

(33) a. majka-(***ta**) mu

   sister-the   his

   'his sister'

b.   xubava-**ta** mu majka

pretty-the his mother

'his pretty mother'

(34)   a.   (***la**) tua   sorella

the your sister

'your sister'

b.   **la** tua   sorella intelligente

the your sister   smart

'your smart sister'

This is reminiscent of article drop in PP discussed in section 6.2; abstractly, the presence of an adjective forces presence of a definite article in the environments where otherwise the definite article must be absent.

It should, however, be noted here that the category-based account proposed for article drop in PP in section 6.2.2 cannot be extended to the article drop with kinship term possessums discussed here. The account proposed there is that P can be the highest functional element in the nominal domain when it takes a [-V] element as its complement, and it can function as an alternative of D, allowing (and forcing) omission of the definite article. When there is an adjective that modifies the noun in the relevant PP, AP, which is [+V], projects above NP, so that P does not satisfy the above condition for becoming the highest functional element in the nominal domain and the definite article needs to project DP above AP to be an argument of P (and to ensure the definite interpretation). Notice now that the possessive constructions in question do not have P above the nominal phrase in the first place. Thus, there is no motivation to omit the definite article from the perspective of category-based calculation of extended projections of lexical categories. In addition, it seems plausible to analyze adjectives as projecting AP above NP, but below PossP, where the pronominal possessor is assumed to be located by Despić (2011, 2015), the assumption I also adopted in chapter 5. This is motivated by the modification relation in (33b) and (34b); in

both cases, the properties of the possessums are the intersection of the properties denoted by the adjective and by the head noun. This means that AP first combines with NP, and then the possessor combines above them. Then, the highest projection before the merger of the definite article should always be PossP, whether the adjective is present or not. This means that we cannot appeal to the structure before the merger of the definite article to the possessive phrase.[14]

Turning back to the kinship terms, they receive definite interpretation in the relevant possessive constructions, and hence can be analyzed as requiring the iota operator that picks out a uniquely identifiable individual because of that. A definite article may then be expected to co-occur with them. In fact, a definite article is required when there is an adjective, as seen above. Recall now that the languages introduced above are non-canonical DP-languages in the NP/DP-language scale advocated in this dissertation. This indicates that such article drop is allowed only in less of a DP-language in the NP/DP-language scale, and that there is some correlation between the nature of the definite article and article drop in such cases, an issue I address in the next section.[15]

6.4.2   Analysis: A type-theoretic account

In this subsection I propose a semantic account of article drop in the possessive constructions with kinship terms. The gist of the proposal is that kinship terms yield an individual without an

---

14. There are temporal adjectives such as *former* which can have two interpretations in possessive constructions; for instance, *John's former pants* can be interpreted as (i) an object that John now owns and used to be pants, and (ii) pants that John formerly owned. Larson and Cho (2003) argue that these two readings are derived from two different structures. In particular, for the interpretation (i), *former*, which is of type $\langle\langle e,t\rangle,\langle e,t\rangle\rangle$ (simplified for expository purposes), is composed above the possessor and the head noun (the possessor *John* moves to Spec,DP in syntax, but is interpreted below *former*). In this case, *former* would project AP above PossP and below DP, which in turn may predict that the definite article in Italian should be present with a kinship possessum in the presence of a temporal adjective under this interpretation. It should, however, be noted that in Italian, it is not natural to use *precedente*, which roughly corresponds to 'former', in a possessive construction (Pietro Cerrone p.c.). It remains to be investigated in future research how temporal adjectives (or non-intersective adjectives that have a different semantic type than $\langle e,t\rangle$) would interact with kinship possessums in the languages discussed in text.

15. Unlike the languages mentioned here, English quite generally disallows the definite article to co-occur with a pronominal possessor, regardless of what kind of a noun the possessum is. Another difference is that while the presence of an adjective forces the presence of the definite article with those kinship term/unique possessums in the languages discussed in this section, English exhibits no such effect (i.e., the definite article never co-occurs with a possessor, whether there is an adjective or not). I take these observations as indicating that there is some formal difference regarding the syntax of possessives between the affixal article languages discussed in this section and English. See section 6.4.2 for a possible analysis.

application of the iota operator when combined with a possessor, whereas ordinary nouns yield a property when combined with a possessor, so that the former do not require an overt iota operator, i.e., a definite article, whereas the latter require it.

A more or less standard analysis of the semantics of possessive constructions is that the possessor is of type $\langle e,t \rangle$ and involves a relator variable R that establishes a relation between the possessor and the possessum (Barker 1991, 1995, Jensen and Vikner 1994, Partee 1997, Partee and Borschev 2003, Vikner and Jensen 2002, among others). Under this line of approaches to possessive constructions, ordinary nouns such as *bike* are assumed to be of type $\langle e,t \rangle$, while relational nouns including kinship terms such as *sister* are assumed to be of type $\langle e,\langle e,t \rangle \rangle$, and they are combined with R with or without application of type-shifting depending on the analysis. For instance, Partee (1997) proposes that the Saxon genitive *'s* in English corresponds to R that is of type $\langle e,\langle e,t \rangle \rangle$, and combines with the possessor such as *John* and then an ordinary noun like *bike*.[16] Under her proposal, the value of R, namely, the relation between *John* and *bike*, is determined pragmatically (here "ownership"). On the other hand, Partee proposes that R for relational nouns including kinship terms such as *sister* is of type $\langle e,\langle \langle e,\langle e,t \rangle \rangle,\langle e,t \rangle \rangle \rangle$ and preserves the relation inherently denoted by the relational noun. As an illustration, let us consider (35).

(35) a. bike of John's   (ordinary noun)

      b. sister of John's   (kinship term)

Partee (1997) assumes that R is encoded by the Saxon genitive *'s*, and *of* is semantically vacuous. The composition of (35a) and (35b) proceeds as follows:

(36) a. $[\![(\text{of}) \text{ John's}]\!] = \lambda P \lambda x[P(x) \ \& \ R_i(\text{John})(x)]$

      b. $[\![\text{bike of John's}]\!] = \lambda x[\text{bike}(x) \ \& \ R_i(\text{John})(x)]$

---

16. Partee (1997) assumes that the possessor is type-shifted from $\langle e,t \rangle$ to $\langle \langle e,t \rangle,\langle e,t \rangle \rangle$ in order to combine with the ordinary noun of type $\langle e,t \rangle$, but this can also be achieved via Predicate Modification proposed by Heim and Kratzer (1998). The choice between these two does not matter for the present purposes, and below I adopt Predicate Modification for ease of exposition.

(37) a. ⟦(of) John's⟧ = λRλx[R(John)(x)]

     b. ⟦sister of John's⟧ = λx[sister(John)(x)]

Both (36b) and (37b) are of type ⟨e,t⟩, and either the indefinite article or the definite article will be merged above them. In the case of a prenominal possessor such as the possessor in *John's bike* and *John's sister*, Partee assumes that there is a covert iota operator above the possessive phrase (see below on this). Barker (1995) and Vikner and Jensen (2002) propose that ordinary nouns like *bike* undergo type-shifting and are "coerced" into relational nouns. What is important for the current purposes is that possessive constructions before the application of the iota operator are of type ⟨e,t⟩ under these proposals, whether the possessum is an ordinary noun or a relational noun.

Dobrovie-Sorin (2000, 2002, 2004) proposes a different semantic account of possessive constructions. Specifically, she proposes that nouns that are combined with a possessor are of type ⟨e,e⟩, which is a function from an individual denoted by the possessor to an individual denoted by the entire possessive phrase. She assumes that relational nouns such as *sister* are inherently of type ⟨e,e⟩, and are straightforwardly combined with a possessor. The semantic composition of *John's sister* is given in (38).

(38) ⟦John's sister⟧ = λxιy[sister(x,y)](John)

                = ιy[sister(John,y)]

On the other hand, non-relational nouns such as *bike* are coerced into a relational noun, by adding a relator variable R in Partee's sense via the genitive marker. The semantic composition of *John's bike* is given in (39).

(39) ⟦John's bike⟧ = λxιy[R$_{gen}$(x,y) & bike(y)](John)

                = ιy[R$_{gen}$(John,y) & bike(y)]

What is interesting here is that the composition of the possessor and the possessum yields an element of type e, namely, an individual, without an application of a covert iota operator unlike what we have seen in the proposals by Partee (1997) and others discussed above. This captures the intuition of inherent uniqueness of kinship possessums.

However, under Dobrovie-Sorin's proposal, the composition of the possessor and the posses-sum yields the same type, just as in those by Partee among others. Recall that the question that needs to be addressed in this section is why kinship terms resist a definite article while ordinary nouns require it in possessive constructions in the languages discussed in section 6.4.1. The rele-vant examples from Italian are repeated here as (40).

(40) a. **la** tua macchina

 the your car

 'your car'

 b. (***la**) tua sorella

 the your sister

 'your sister'

If kinship terms and ordinary nouns yield the same semantic composition as Dobrovie-Sorin (2004) and Partee (1997) among others propose, it would be mysterious how the difference between kin-ship terms and ordinary nouns discussed above could be accounted for. One might suggest here that possessive constructions have different syntactic structures in the cases of kinship terms and ordinary nouns, but the question then boils down to where the syntactic difference comes from. In what follows, I argue that there is actually a correlation between the syntax and the semantics of the possessive constructions.

My proposal here is that ordinary nouns and kinship terms have different semantic types, and are composed with the possessor in different ways; what I am proposing is, essentially, a hy-brid account of Partee (1997) and Dobrovie-Sorin's (2004), but crucially coercion is not assumed. Regarding ordinary nouns, I follow Partee (1997) in assuming that they are of type $\langle e,t \rangle$ and com-bined with the relator variable R that takes a possessor as its argument. Here, I slightly modify her proposal in accordance with the structure of possessive constructions proposed by Despić (2011, 2015) that I adopted in chapter 5, where NP is dominated by PossP and PossP is dominated by DP in languages with definite articles. Specifically, R is of type $\langle\langle e,t \rangle,\langle e,\langle e,t \rangle\rangle\rangle$ and takes a property

as its first argument and a possessor as its second argument. The denotation of R is given in (41).

(41) 〚 R 〛 = λPλxλy[R(x,y) & P(y)]

I assume that R corresponds to the head of PossP in Despić's structure. The structure of Italian (40a) is illustrated in (42). The semantic composition of (42) is given in (43), where (43c) can be read as there is a unique car that has a relation with the addressee in the discourse (the relation is pragmatically supplied; in this case, ownership). What is crucial here is that the entire possessive phrase requires the definite article in order to be interpreted as an individual and receive the definite interpretation.

(42)

```
                        DP (3)
                         e
              ┌──────────┴──────────┐
              D                   PossP (2)
              la                   ⟨e,t⟩
            'the'          ┌─────────┴─────────┐
          ⟨⟨e,t⟩,e⟩       tua              Poss′ (1)
                         'your'            ⟨e,⟨e,t⟩⟩
                           e         ┌────────┴────────┐
                                   Poss              NP
                                    R            ┌────┴────┐
                            ⟨⟨e,t⟩,⟨e,⟨e,t⟩⟩⟩  macchina
                                                 'car'
                                                 ⟨e,t⟩
```

(43) a. 〚 (1) 〛 = λPλxλy[R(x,y) & P(y)](car(z))

  = λxλy[R(x,y) & car(y)]

  b. 〚 (2) 〛 = λxλy[R(x,y) & car(y)](you)

  = λy[R(you,y) & car(y)]

  c. 〚 (3) 〛 = λQιz[Q(z)](λy[R(you,y) & car(y)])
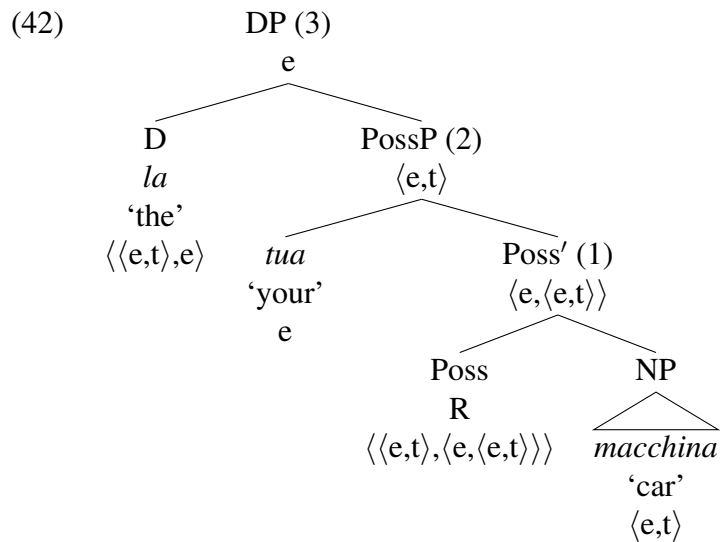
  = ιz[R(you,z) & car(z)]

As for kinship terms, on the other hand, I follow Dobrovie-Sorin (2000, 2002, 2004) in assuming that they are of type ⟨e,e⟩ and inherently encode uniqueness of the individual that has the relevant

property, the relational variable R being absent. Thus, the composition of a kinship term with a possessor yields an individual of type e, with no need for the application of the iota operator. Consequently, the definite article, which is an overt version of the iota operator, is not required, and in fact is disallowed because composition of the definite article ($\langle\langle e,t\rangle,e\rangle$) and the kinship term ($\langle e,e\rangle$) would result in a type mismatch. Given that Italian, and more generally languages where the definite article cannot co-occur with a kinship possessum, are non-canonical DP-languages in the NP/DP-language scale, it is reasonable to conclude that DP is then absent in this case. The structure of (40b) is given in (44), and its semantic composition is given in (45).

(44)   PossP (4)
             e
      ┌──────┴──────┐
    tua   Poss′
   'your'   ┌────┴────┐
     e   Poss   NP
                     sorella
                     'sister'
                     $\langle e,e \rangle$
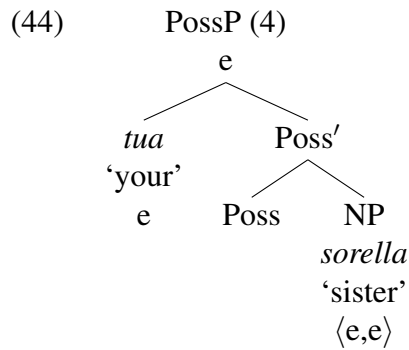
(45)   ⟦ (4) ⟧ = $\lambda x \iota y[\text{sister}(x,y)](\text{you})$

         $= \iota y[\text{sister}(\text{you},y)]$

Thus, the obligatory presence/absence of the definite article, which amounts to the presence/absence of DP, correlates with the semantic difference between ordinary nouns and kinship terms.

A remaining question is why the definite article is required with a kinship possessum in the presence of an adjective. The relevant Italian example is repeated here as (46).

(46)   **la**   tua   sorella   intelligente

       the   your   sister   smart

       'your smart sister'

I propose that the presence of the adjective, which is standardly assumed to be of type $\langle e,t \rangle$, requires the kinship term to also be of type $\langle e,t \rangle$ for a successful semantic composition. In other words,

sorella 'sister' in (46) is of type $\langle e,t\rangle$ rather than of type $\langle e,e\rangle$.[17] This can be implemented either by covert type-shifting or by lexical ambiguity of kinship terms. The choice between these two options has no significance for the current purposes. I assume here that those kinship terms are lexically ambiguous between type $\langle e,t\rangle$ and type $\langle e,e\rangle$ for ease of exposition. The composition of (46) then proceeds essentially in the same way as in the case of the ordinary noun shown in (42) and (43), the only difference being that the adjective *intelligente* 'smart' combines with the noun *sorella* 'sister' via Predicate Modification (Heim and Kratzer 1998). The structure of (46) is given in (47) and its semantic composition in (48). (I ignore the word order of *sorella* and *intelligente* here and simply present the structure where AP dominates NP adopted in chapter 5. What is important here is the semantic composition of those items.)

(47)

```
                    DP (8)
                     e
          ┌──────────┴──────────┐
          D                  PossP (7)
          la                  ⟨e,t⟩
         'the'          ┌───────┴────────┐
       ⟨⟨e,t⟩,e⟩       tua            Poss′ (6)
                       'your'          ⟨e,⟨e,t⟩⟩
                         e        ┌───────┴───────┐
                                 Poss           AP (5)
                                  R              ⟨e,t⟩
                           ⟨⟨e,t⟩,⟨e,⟨e,t⟩⟩⟩  ┌────┴────┐
                                              A        NP
                                         intelligente sorella
                                           'smart'   'sister'
                                            ⟨e,t⟩    ⟨e,t⟩
```
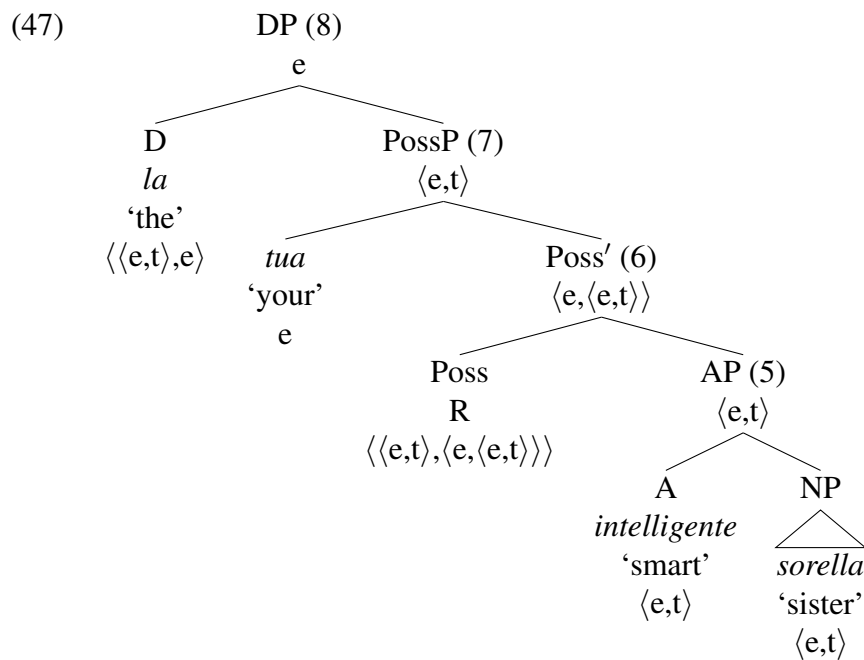
(48) a. ⟦ (5) ⟧ = $\lambda x[\text{sister}(x)\ \&\ \text{smart}(x)]$

    b. ⟦ (6) ⟧ = $\lambda P\lambda x\lambda y[R(x,y)\ \&\ P(y)](\lambda x[\text{sister}(x)\ \&\ \text{smart}(x)])$

---

17. Here I am simplifying the semantics of the relevant adjectives and the kinship terms to a large extent; it would actually be more complex than what is proposed above. What is important here is that the semantic type of the relevant kinship terms is different depending on the presence vs. the absence of an adjective, which is necessary to account for the contrast regarding the presence/absence of the definite article in these two cases. A more detailed technical implementation of the idea pursued here is left for future research.

$$= \lambda x \lambda y[R(x,y) \ \& \ \text{sister}(y) \ \& \ \text{smart}(y)]$$

c. $[\![\,(7)\,]\!] = \lambda x \lambda y[R(x,y) \ \& \ \text{sister}(y) \ \& \ \text{smart}(y)](\text{you})$

$$= \lambda y[R(\text{you},y) \ \& \ \text{sister}(y) \ \& \ \text{smart}(y)]$$

d. $[\![\,(8)\,]\!] = \lambda Q \iota z[Q(z)](\lambda y[R(\text{you},y) \ \& \ \text{sister}(y) \ \& \ \text{smart}(y)])$

$$= \iota z[R(\text{you},z) \ \& \ \text{sister}(z) \ \& \ \text{smart}(z)]$$

Thus, the presence of an adjective, which is of type $\langle e,t \rangle$, forces the possessum to also be of type $\langle e,t \rangle$, even if the possessum is a kinship term, hence the definite article is required in order for the legitimate semantic composition even in languages where non-modified kinship terms resist the definite article in possessive phrases.[18]
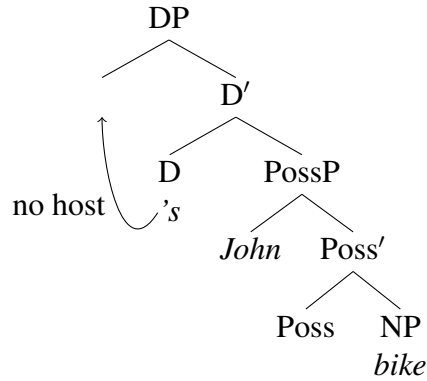
Before concluding this section, it needs to be addressed why English, which is a canonical DP-language where DP is expected to always project, never allows the definite article to occur in possessive phrases, whether the possessum is an ordinary noun or a kinship term, as mentioned in footnote 15. My suggestion, which I briefly mentioned in chapter 5, is that the Saxon genitive *'s* is base-generated as D, and undergoes Affix Hopping onto Poss in PF. In the structure of possessive phrases proposed by Despić (2011, 2015), which I have adopted here, the pronominal possessor is located in Spec,PossP. Despić assumes that non-pronominal possessors like *John* are located in Spec,DP, but it can rather naturally be assumed that they are actually located in Spec,PossP given the semantic composition discussed above.[19] Then, there is no morphological host for the Saxon genitive *'s*, which is base-generated as D, in a phrase like *John's bike*, as illustrated in (49a). The Saxon genitive then undergoes Affix Hopping, getting affixed to *John*.

---

18. In chapter 5, I proposed that Italian definite articles can adjoin to N. One may then wonder why the definite article cannot adjoin to N in the possessive constructions under discussion. I suggest that this can be explained in terms of semantic composition (see also chapter 5); if the definite article, which is of type $\langle\langle e,t \rangle,e \rangle$, is first composed with the head noun and then the adjective is composed with D+N, there would be a type mismatch (i.e., D+N would be of type e, and then composition of D+N with A would be of type t, which cannot serve as an argument of predicates).

19. An exception is *each other*, which needs to be outside of the Spell-Out domain of DP for Binding Condition A, as mentioned in chapter 5; see Despić (2011, 2015).

(49) a.

```
              DP
             /  \
                 D′
                / \
               D   PossP
      no host  's  /  \
                John   Poss′
                       /  \
                     Poss   NP
                            bike
```

  b.

```
              DP
             /  \
                 D′
                / \
               D   PossP
                   /  \
 Affix Hopping  John's  Poss′
                        /  \
                      Poss   NP
                             bike
```
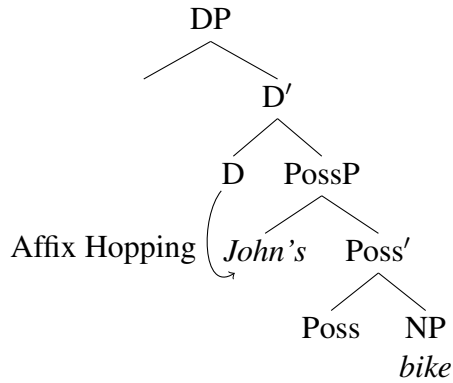
Following Larson and Cho (2003), I assume that *'s* as D is the locus of the iota operator. It then follows that the definite interpretation obtains despite the absence of the overt definite article. Thus, the obligatory absence of the definite article in English possessive phrases can be accounted for by the Affix Hopping analysis of *'s*.

6.5   Conclusion of the chapter

To summarize this chapter, I have demonstrated that there are conditions under which definite articles are dropped even with definite interpretation, which is not expected under Bošković's and Talić's treatment of definite articles. Specifically, the definite article must be dropped in certain PPs in languages such as Romanian, Albanian, and Yiddish, and in possessive phrases with kinship possessums in languages like Bulgarian, Italian, Somali, etc. What is remarkable is that such cases of article drop are observed only in non-canonical DP-languages in the NP/DP-language scale argued for in this dissertation, where DP can be absent in such languages, unlike canonical DP-

languages like English. I have taken the relevant state of affairs as indicating that DP is absent in these cases of article drop. I have also shown that the presence of an adjective forces the definite article to be present in such cases, which indicates that DP must be present in the presence of an adjective, on a par with several cases discussed in chapter 5, where the presence of an adjective also forces the presence of a DP.

Furthermore, I have proposed a syntactic account for article drop in PPs and a semantic account for article drop with kinship possessums. Regarding the former, I have argued that P can be the highest functional element in the nominal domain because P and N share the [-V] categorial feature, and in this case the presence of DP is blocked. However, the presence of an adjective, which projects AP above NP, blocks the relevant relationship between P and N since A is [+V], so that DP is required to project as the highest projection in the nominal domain. I have also discussed preposition-article contraction in German, which only takes place in PPs under certain semantic conditions. The relevant contraction has been analyzed as having the definite article base-generated adjoined to P, the operation discussed in chapters 4 and 5. Based on this, I have suggested that German is slightly less of a DP-language in the NP/DP-language scale than e.g. English. As for the kinship possessums, I have argued that they are of type ⟨e,e⟩ and inherently unique, so that only the possessor, which is of type e, is required and the definite article is disallowed to co-occur with them. However, when an adjective modifies the kinship term, the kinship term has a different semantic type in order to compose with the adjective, and the resulting structure requires the definite article for the definite interpretation to obtain. In both article drop in PP and article drop with kinship possessums, the presence of an adjective, which projects AP, forces projection of DP, the exact motivation depending on the construction (syntactic for the former and semantic for the latter).

Chapter 7

Concluding Remarks

In this dissertation, I have established a number of novel cross-linguistic generalizations and discussed their relevance to the linguistic theory, from the perspective of the NP/DP-language distinction. In particular, I have discussed the concepts of phases, labeling, Agree, Merge, and the theory of parameters in minimalism, building on broad typological investigations of the phenomena discussed in this dissertation. With respect to the NP/DP-language distinction, I have argued that the two-way and three-way distinctions proposed in the previous works like Bošković (2008b, 2012) regarding the former and Talić (2015, 2017) regarding the latter are not sufficient, and that we need a "scale" of NP/DP-language from canonical NP-languages to canonical DP-languages.

In chapter 2, I have established the generalization that only languages that have affixal definite articles or lack definite articles may allow extraction *of* a conjunct out of a coordinate structure. I have then offered a deduction of this generalization based on Bošković's (2014) contextual approach to phasehood and Talić's (2015, 2017) Structural Parallelism. I have also shown that extraction *out of* a conjunct is still disallowed in the languages where extraction *of* a conjunct is allowed. Building on this, I have argued that the CSC, which was originally proposed as a single condition by Ross, should be separated into two independent conditions: the ban on extraction *of* a conjunct (CSC I) and the ban on extraction *out of* a conjunct (CSC II). In addition, I have claimed that the CSC I and the CSC II apply at different places in the grammar: the CSC I is a purely syntactic

condition, and the CSC II is an interface condition.

In chapter 3, I have established the generalization that only languages that have affixal definite articles or lack definite articles may have productive compositional indeterminate pronouns. I haven then offered a deduction of this generalization based on Saito's (2017) analysis of the relevant pronouns in Japanese. I have further shown that this deduction allows us to treat multiple wh-fronting in languages such as Serbo-Croatian and wh-in-situ of the Japanese type in a uniform manner. I have also discussed various types of wh-in-situ, arguing that they can be captured by the proposed system of indeterminate pronouns.

In chapter 4, I established the generalization that large-scale pied-piping is possible in a language only if the language has productive compositional indeterminate pronouns and the projection to be pied-piped is head-final. In order to deduce this generalization, I have proposed under Chomsky's (2015) labeling framework that weak heads (i.e., heads that do not provide a label on their own) are not only syntactically but also morpho-phonologically weak, in that they cannot be realized as free morphemes. In addition, I have proposed a criterion for weak heads, which generalizes the notion of weak heads to all heads that have unvalued features at the point of External Merge. I have shown that this new conception of weak heads captures the above generalization, coupled with Inaba's (2011) observation that head-final complementizers are generally suffixal, whereas head-initial complementizers are generally free morphemes. Moreover, this conception of weak heads allows and requires us to deduce Agree from Minimal Search, which is a third factor principle external to UG, hence minimize UG.

In chapter 5, I have argued that the three-way distinction of NP/DP-languages proposed by Talić (article-less languages, affixal-article languages, non-affixal article languages) is not sufficient either, and that we need a more fine-grained "scale" of NP/DP-language distinction. In particular, I have shown that Italian, which has been treated as a non-affixal article language by Talić (2017), behaves differently from proto-typical non-affixal article languages such as English and proto-typical affixal article languages such as Bulgarian, with respect to adjunct extraction out of a nominal phrase and the availability of reflexive and pronominal possessives. In order to capture

the intermediate behavior of Italian, I have proposed that definite articles in Italian can be base-generated as adjoined to another head in the nominal domain, which is an option theoretically allowed in minimalism. I have also shown that Hungarian also behaves differently from proto-typical non-affixal and affixal article languages, but also from Italian, which means that Hungarian is yet another type of a DP-language in the NP/DP-language scale. I have also discussed Greek, Egyptian Arabic, and Basque from the perspective of the NP/DP-language scale. They exhibit a number of properties of NP-languages in the absence of the definite article, which means that these languages are less of DP-languages than English. Based on these languages I have suggested that the presence of a fully grammaticalized indefinite article may be required for projection of a functional projection in indefinite nominal phrases. I have also argued that the scale of NP/DP language distinction argued for here is an appropriate point of parameterization in minimalism, especially under the emergentist view of parameters advocated by Biberauer (2019), Biberauer and Roberts (2017), Roberts (2019) among others.

In chapter 6, I have discussed cases in which the definite article is omitted in certain PPs and possessive constructions with kinship possessums in less of DP-languages in the NP/DP-language scale argued for here. I have proposed that PP can be the highest functional projection in the nominal domain, which is motivated by the categorial specification originally proposed by Chomsky (1970), and that DP cannot project in the presence of such PPs. As for kinship possessums, following Dobrovie-Sorin (2002, 2004), I have proposed that they are of type $\langle e,e \rangle$; they take a possessor as an argument and yields type e, hence the definite article cannot be composed with the possessor phrase due to a type mismatch. We have also seen that when an adjective is present in these constructions, the definite article must be present, which I have taken as indicating that the presence of an adjective forces projection of DP in such cases. In the case of PPs, there is a syntactic reason for obligatory projection of DP in the presence of AP; projection of AP between NP and PP breaks the relation between NP and PP as extended projections in the nominal domain, due to the feature specifications of those projections. On the other hand, there is a semantic reason for projection of DP in the presence of AP with kinship possessums; in the presence of an adjective, which is of type

⟨e,t⟩, a kinship possessum must be of type ⟨e,t⟩ for a legitimate semantic composition. Since the resulting structure does not yield an individual (of type e), an overt iota operator, i.e., the definite article, must be present.

As a final remark, I would like to note some open issues worth investigating in future research. First, many more languages should be closely examined from the perspective of the NP/DP language scale argued for here. There are many languages that have definite articles, including those discussed by Bošković (2008b, 2012) and Talić (2015, 2017), and it may turn out that they show varying behavior of NP/DP-language-hood, as I have in fact shown regarding, e.g., Italian, which was considered to be a "canonical" DP-language in the previous literature. Relatedly, I have actually oversimplified Bošković's discussion of NP-languages in this dissertation; it should be emphasized that he did not say that there can never be a functional projection above NP in languages without definite articles, he only argued against the presence of DP. In a series of works, he in fact proposed that there is a functional projection above NP in article-less languages with, e.g., non-adjectival numerals, certain quantifiers, as well as case particles in Japanese (see also Takahashi 2011 regarding the last one). We need a closer and more principled investigation of such cases, to determine what kind of elements can project their own phrases and what effects that may have.

Grammaticalization of definite and indefinite articles is another topic to be pursued. In chapter 5, I have discussed grammaticalization of definite and indefinite articles building on Wang (2019); in particular, I have suggested that both definite and indefinite articles (can) adjoin to another head via base-generation at an intermediate stage of grammaticalization. A wider range of languages will need to be examined from this perspective. Relatedly, in chapter 5 I have also suggested the possibility of "losing" a DP layer in the course of grammaticalization of a definite article. Greenberg (1978) proposes that grammaticalization of a definite article proceeds as follows: Demonstrative (Stage 0) > Definite article (Stage I) > Specificity article (Stage II) > Noun marker (Stage 3). The Basque "definite article" actually marks specificity rather than definiteness, so it would be at Stage II of grammaticalization in Greenberg's term. Interestingly, Lyons (1999) suggests that languages that have a Stage II article may be losing a DP layer. In fact, Basque shows some prop-

erties of NP-languages, such as sloppy readings with null objects and compositional indeterminate pronouns. Basque may then be "losing" the DP layer because of grammaticalization of the definite article (even independently of the affixal status of the definite article, but actually the affixal status itself might also be related to the grammaticalization process). It will be worth examining other languages where articles mark specificity such as St'át'imcets to see if the above correlation is found in those languages.

Interestingly, given the two processes of grammaticalization discussed in chapter 5, grammaticalization seems to proceed in the way that the DP layer gets developed going from Stage 0 discussed above to Stage I (through the stage of head-adjunction via base-generation), it can get lost from Stage I to Stage II (and possibly completely lost at Stage III). Hypothetically applying all this to a single language, a language starts from an NP-language (as a default option of UG as discussed in chapter 5) and becomes a DP-language, and then it "goes back" to an NP-language. This may be analogous to Jespersen's cycle, in which a language starts with a single negation and acquires two negations, and then "goes back" to a single negation by losing one of the two negations. We may then be dealing here with a more general process of grammaticalization cycles, a possibility Lyons (1999) in fact suggests. This possibility may be worth pursuing more thoroughly in future research.[1]

---

1. As noted in footnote 47 in chapter 5, the general tendency of historical change seems to be from an NP-language to a DP-language (i.e., development of DP). The Basque case (i.e., loss of DP) is in fact quite rare. This is actually expected under the economy considerations of language acquisition discussed in chapter 5, section 6.

References

Abdel-Malek, Zaki N. 1972. *The closed-list classes of colloquial Egyptian Arabic*. The Hague: Mouton.

Abe, Jun. 2011. Real parasitic gaps in Japanese. *Journal of East Asian Linguistics* 20:195–218.

Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Doctoral Dissertation, University of Connecticut, Storrs.

Abels, Klaus. 2012. *Phases: An essay on cyclicity in syntax*. Berlin: Walter de Gruyter.

Abney, Steven. 1987. The English noun phrase in its sentential aspect. Doctoral Dissertation, Massachusetts Institute of Technology.

Agbayani, Brian, and Chris Golston. 2010. Phonological movement in Classical Greek. *Language* 86:113–167.

Aguilar-Guevara, Ana. 2014. *Weak definiteness: Semantics, lexicon, and pragmatics*. Utrecht: LOT.

Aissen, Judith. 1996. Pied-piping, abstract agreement, and functional projections in Tzotzil. *Natural Language and Linguistic Theory* 14:447–491.

Alberti, Gábor, and Tibor Laczkó. 2018. *Syntax of Hungarian: Nouns and noun phrases, volume II*. Amsterdam: Amsterdam University Press.

Aldridge, Edith. 2009. Short wh-movement in Old Japanese. In *Japanese/Korean linguistics 17*, ed. Shoishi Iwasaki, Hajime Hoji, Patricia M. Clancy, and Sung-Ock S. Sohn, 549–563. Stanford, CA: Center for the Study of Langauge and Information.

Aldridge, Edith. 2018. C-T inheritance and the left periphery in Old japanese. *Glossa: A Journal of General Linguistics* 3:1–22.

Alexopoulou, Theodora, and Raffaella Folli. 2019. Topic strategies and the internal structure of nominal arguments in Greek and Italian. *Linguistic Inquiry* 50:439–486.

Ámason, Kristján. 1985. Icelandic word stress and metrical phonology. *Studia Linguistica* 39:93–129.

Ambrazaitis, Gilbert. 2009. Nuclear intonation in Swedish: Evidence from experimental-phonetic studies and a comparison with German. Doctoral Dissertation, Lund University, Lund.

Anagnostopoulou, Elena. 1994. Clitic dependencies in Modern Greek. Doctoral Dissertation, Salzburg University, Salzburg.

Anderssen, Merete. 2007. The acquisition of compositional definiteness in Norwegian. *Nordlyd* 34:252–275.

Andersson, Lars-Gunnar. 1982. What is Swedish an exception to?: Extractions and island-constraints. In *Readings on unbounded dependencies in Scandinavian languages (Umeå studies in the humanities 43)*, ed. Elisabet Engdahl and Eva Ejerhed, 33–45. Stockholm: Almqvist & Wiksell International.

Andrews, Avery. 2007. Relative clauses. In *Language typology and syntactic description*, ed. Timothy Shopen, volume II, 206–236. Cambridge: Cambridge University Press.

Androutsopoulou, Antonia. 1998. Contrastive focus and remnant DP-movement in Modern Greek. In *Proceedings of the North East Linguistic Society 28*, ed. Pius N. Tamanji and Kiyomi Kusumoto, 1–16. University of Toronto: Graduate Linguistic Student Association.

Aoun, Joseph. 1986. *Generalized binding: The syntax and logical form of wh-interrogatives*. Dordrecht: Foris.

Arano, Akihiko, and Hiromune Oda. 2019. The A-/A′-distinction in scrambling revisited. In *Proceedings of the 36th West Coast Conference on Formal Linguistics*, ed. Richard Stockwell, Maura O'Leary, Zhongshi Xu, and Z.L. Zhou, 48–54. Somerville, MA: Cascadilla Proceedings Project.

Asher, R. E. 1985. *Tamil*. London: Croom Helm.

Baker, Mark. 1988. *Incorporation*. Chicago, Illinois: University of Chicago Press.

Baker, Mark. 1996. *The polysynthesis parameter*. Oxford: Oxford University Press.

Baker, Mark C. 2003. *Lexical categories: Verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.

Baker, Mark C. 2008a. The macroparameters in a microparametric world. In *The limits of syntactic variation*, ed. Theresa Biberauer, 351–374. Amsterdam: John Benjamins.

Baker, Mark C. 2008b. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.

Baker, Mark C., and James McCloskey. 2007. On the relationship of typology to theoretical syntax. *Linguistic Typology* 11:285–296.

Baker, Mark C., and Osamuyimen T. Stewart. 1999. On double-headedness and the anatomy of the clause. Ms., Rutgers University.

Barker, Chris. 1991. Possessive descriptions. Doctoral Dissertation, University of California at Santa Cruz.

Barker, Chris. 1995. *Possessive descriptions*. Stanford, CA: CSLI Publications.

Bašić, Monika. 2005. Split phrases in Slavic. Talk given at Formal Description of Slavic Languages (FDSL) 6, University of Potsdam.

Bayer, Josef. 1996. *Directionality and Logical Form*. Dordrecht: Kluwer.

Bayer, Josef. 2001. Two grammars in one: Sentential complements and complementizers in Bengali and other South-Asian languages. In *The yearbook of South Asian languages and linguistics 2001*, ed. Peri Bhaskararao and K. V. Subbaram, 11–36. New Delhi: Sage Publications.

Beavers, John, and Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, ed. Stefan Müller, 48–69. Stanford, CA: CSLI Publications.

Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14:1–56.

Belletti, Adriana. 2001. 'Inversion' as focalization. In *Inversion in Romance and the theory of universal grammar*, ed. Aafke Hulk and Jean-Yves Pollock, 60–90. Oxford: Oxford University Press.

Belletti, Adriana. 2004. Aspects of the low IP area. In *The structure of CP and IP*, ed. Luigi Rizzi, 16–51. Oxford: Oxford University Press.

Benmamoun, Elabbas. 1999. The syntax of quantifiers and quantifier float. *Linguistic Inquiry* 30:621–642.

Biberauer, Theresa. 2019. Factors 2 and 3: Toward a principled approach. *Catalan Journal of Linguistics Special Issue: Generative Syntax. Questions, Crossroads, and Challenges* 45–88.

Biberauer, Theresa, Anders Holmberg, and Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45:169–225.

Biberauer, Theresa, and Ian Roberts. 2017. Parameter setting. In *The Cambridge handbook of historical syntax*, ed. Adam Ledgeway and Ian Roberts, 134–162. Cambridge: Cambridge University Press.

Biberauer, Theresa, and Michelle Sheehan. 2013. Theoretical approaches to disharmonic word order. In *Theoretical approaches to disharmonic word order*, ed. Theresa Biberauer and Michelle Sheehan, 1–44. Oxford: Oxford University Press.

Bobaljik, Jonathan. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge.

Bobaljik, Jonathan. 2002. A-chains at the PF-interface: copies and 'covert' movement. *Natural Language and Linguistic Theory* 20:197–267.

Boeckx, Cedric. 2011. Approaching parameters from below. In *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, ed. Anna Maria Di Sciullo and Cedric Boeckx, 205–221. Oxford: Oxford University Press.

Bonan, Caterina. 2019. On clause-internally moved wh-phrases: wh-to-foc, nominative clitics, and the theory of Northern Italian wh-in situ. Doctoral Dissertation, Universit'e de Genève.

Borer, Hagit. 1984. *Parametric syntax*. Dordrecht, Holland: Foris Publications.

Bošković, Željko. 1994. D-structure, θ-criterion, and movement into θ-positions. *Linguistic Analysis* 24:247–286.

Bošković, Željko. 1997a. On certain violations of the Superiority Condition, AgrO, and economy of derivation. *Journal of Linguistics* 33:227–254.

Bošković, Željko. 1997b. Superiority effects with multiple *wh*-fronting in Serbo-Croatian. *Lingua* 102:1–20.

Bošković, Željko. 1997c. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Massachusetts: MIT Press.

Bošković, Željko. 1999. On multiple feature checking: multiple wh-fronting and multiple head movement. In *Working minimalism*, ed. Samuel Epstein and Norbert Hornstein, 159–187. Cambridge, Massachusetts: MIT Press.

Bošković, Željko. 2000. Sometimes in [Spec, CP], sometimes in situ. In *Step by step*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 53–156. Cambridge, Massachusetts: MIT Press.

Bošković, Željko. 2001a. On the interpretation of multiple questions. *Linguistic Variation Yearbook* 1:1–15.

Bošković, Željko. 2001b. *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Amsterdam: Elsevier Science.

Bošković, Željko. 2002a. Clitics and nonbranching elements and the linear correspondence axiom. *Linguistic Inquiry* 33:329–340.

Bošković, Željko. 2002b. On multiple wh-fronting. *Linguistic Inquiry* 33:351–383.

Bošković, Željko. 2004. Be careful where you float your quantifiers. *Natural Language and Linguistic Theory* 22:681–742.

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59:1–45.

Bošković, Željko. 2006a. Case and agreement with genitive of quantification in Russian. In *Agreement systems*, ed. Cedric Boeckx, 99–120. Amsterdam/Philadelphia: John Benjamins.

Bošković, Željko. 2006b. Case checking versus case assignment and the case of adverbial NPs. *Linguistic Inquiry* 37:522–533.

Bošković, Željko. 2007a. On the clausal and NP structure of Serbo-Croatian. In *Proceedings of Formal Approaches to Slavic Linguistics 15: The Toronto Meeting 2006*, ed. Richard Compton, Magdalena Goledzinowska, and Ulyana Savchenko, 42–75. Ann Arbor, MI: Michigan Slavic Publications.

Bošković, Željko. 2007b. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38:589–644.

Bošković, Željko. 2008a. On the operator freezing effect. *Natural Language and Linguistic Theory* 26:249–287.

Bošković, Željko. 2008b. What will you have, DP or NP? In *Proceedings of NELS*, ed. Emily Elfner and Martin Walkow, volume 37, 101–114. University of Massachusetts, Amherst: GLSA.

Bošković, Željko. 2009a. More on the no-DP analysis of article-less languages. *Studia Linguistica* 63:187–203.

Bošković, Željko. 2009b. The NP/DP analysis and Slovenian. In *Proceedings of the Novi Sad Generative Syntax Workshop 1*, 53–73. Novi Sad: Filozofski fakultet u Novom Sadu.

Bošković, Željko. 2009c. Unifying first and last conjunct agreement. *Natural Language and Linguistic Theory* 27:455–496.

Bošković, Željko. 2011a. On unvalued uninterpretable features. In *Proceedings of North East Linguistic Society 39*, ed. Suzi Lima, Kevin Mullin, and Brian W. Smith, 109–120. Amherst, MA: Graduate Linguistics Student Association.

Bošković, Željko. 2011b. Rescue by PF deletion, Traces as (non)interveners, and the *That*-Trace effect. *Linguistic Inquiry* 42:1–44.

Bošković, Željko. 2012. On NPs and clauses. In *Discourse and grammar: From sentence types to lexical categories*, ed. Günther Grewendorf and Thomas Ede Zimmermann, 179–242. Berlin: Mouton de Gruyter.

Bošković, Željko. 2013a. Phases beyond clauses. In *The nominal structure in Slavic and beyond*, ed. Lilia Schürcks, Anastasia Giannakidou, and Urtzi Etxeberria, 75–128. Berlin: Mouton de Gruyter.

Bošković, Željko. 2013b. Traces do not head islands: What can PF deletion rescue? In *Deep insights, broad perspectives: Essays in honor of Mamoru Saito*, ed. Yoichi Miyamoto, Daiko Takahashi, Hideki Maki, Masao Ochi, Koji Sugisaki, and Asako Uchibori, 56–93. Tokyo: Kaitakusha.

Bošković, Željko. 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45:27–89.

Bošković, Željko. 2016a. Getting really edgy: On the edge of the edge. *Linguistic Inquiry* 47:1–33.

Bošković, Željko. 2016b. On second position clitics crosslinguistically. In *Formal studies in Slovenian syntax: In honor of Janez orešnik*, ed. Franc Marušič and Rok Žaucer, 23–54. Amsterdam: John Benjamins.

Bošković, Željko. 2016c. On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling. *The Linguistic Review* 33:17–66.

Bošković, Željko. 2017. On clitic doubling and argument ellipsis. In *Proceedings of the 34th West Coast Conference on Formal Linguistics*, ed. Aaron Kaplan, Abby Kaplan, Miranda K. McCarvel, and Edward J. Rubin, 97–106. Somerville, MA: Cascadilla Proceedings Project.

Bošković, Željko. 2018a. On movement out of moved elements, labels, and phases. *Linguistic Inquiry* 49:247–282.

Bošković, Željko. 2018b. On pronouns, clitics doubling, and argument ellipsis: Argument ellipsis as predicate ellipsis. *English Linguistics* 35:1–37.

Bošković, Željko. 2020a. Formalism <u>and</u> not <u>vs</u> Functionalism. Class Handout in Comparative Syntax Fall 2020, University of Connecticut, Storrs.

Bošković, Željko. 2020b. On the Coordinate Structure Constraint, across-the-board-movement,

phases, and labeling. In *Recent developments in phase theory*, ed. Jeroen van Craenenbroeck, Cora Pots, and Tanja Temmerman, 133–182. Berlin: De Gruyter Mouton.

Bošković, Željko. 2021a. Formalism <u>and</u> not <u>vs</u> Functionalism. Ms., University of Connecticut, Storrs.

Bošković, Željko. 2021b. Generalized asymmetry. In *Keynotes from the international conference on explanation and prediction in linguistics (CEP): Formalist and functionalist approaches*, ed. Peter Kosta and Katrin Schlund. Frankfurt am Main: Peter Lang.

Bošković, Željko. 2021c. Merge, Move, and contextuality of syntax: The role of labeling, successive-cyclicity, and EPP effects. Ms., Univesity of Connecticut, Storrs.

Bošković, Željko. in press. On the Coordinate Structure Constraint, islandhood, phases, and rescue by PF-deletion. In *A festschrift for Steven Franks*, ed. George Fowler, James Lavine, and Ronald F. Feldstein. Bloomington, IN: Slavica Publishers.

Bošković, Željko, and Serkan Şener. 2014. The Turkish NP. In *Crosslinguistic studies on nominal reference: With or without articles*, ed. Patricia Cabredo Hofherr and Anne Zribi-Hertz, 102–140. Leiden: Brill.

Bošković, Željko, and Jon Gajewski. 2011. Semantic correlates of the NP/DP parameter. In *Proceedings of North East Linguistic Society 39*, ed. Suzi Lima, Kevin Mullin, and Brian Smith, 121–134. Amherst, MA: GLSA.

Bošković, Željko, and I-Ta Chris Hsieh. 2013. On word order, binding relations, and plurality in chinese noun phrases. *Studies in Polish Linguistics* 8:173–204.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24:591–656.

Brody, Michael. 1995. *Lexico-logical form: A radically minimalist theory*. Cambridge, Massachusetts: Massachusetts Institute of Technology.

Brown, Roger. 1973. *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Bruening, Benjamin. 2007. Wh-in-situ does not correlate with *Wh*-indefinites or question particles. *Linguistic Inquiry* 38.

Bruening, Benjamin, and Thuan Tran. 2006. Wh-questions in Vietnamese. *Journal of East Asian Linguistics* 15:319–341.

Cable, Seth. 2007. The grammar of Q: Q-particles and the nature of wh-fronting, as revealed by the wh-questions of Tlingit. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Cable, Seth. 2010. *The grammar of Q: Q-particles, wh-movement and pied-piping*. Oxford: Oxford University Press.

Canac-Marquis, Réjean. 2005. Phases and binding of reflexives and pronouns in english. In *Proceedings of the 12th International Conference on Hhead-Driven Phrase Structure Grammar*, ed. Stefan Müller, 482–502. Stanford, CA: CSLI.

Charnavel, Isabelle, and Dominique Sportiche. 2016. Anaphor binding: What French inanimate anaphors show. *Linguistic Inquiry* 47:35–87.

Cheng, Hsu-Te Johnny. 2013. Argument ellipsis, classifier phrases, and the DP parameter. Doctoral Dissertation, University of Connecticut, Storrs.

Cheng, Lisa Lai-Shen. 1991. On the typology of wh-questions. Doctoral Dissertation, Massachusetts Institute of Technology.

Cheng, Lisa Lai-Shen, and Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30:509–542.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6:339–405.

Chino, Yukiko, and Ken Hiraiwa. 2014. Coordination and the Head Parameter. Poster presented at the 88th Annual Meeting of the Linguistic Society of America, Minneapolis, Minnesota, January 4.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English transformational grammar*, ed. J. Jacobs and P. Rosenbaum, 184–221. Waltham, Massachusetts: Ginn.

Chomsky, Noam. 1972. Some empirical issues in the theory of transformational grammar. In *The goals of linguistic theory*, ed. Stanley Peters, 63–130. Englewood Cliffs, New Jersey: Prentice-Hall.

Chomsky, Noam. 1995a. Bare phrase structure. In *Government binding theory and the minimalist program*, ed. Gert Webelhuth, 383–439. Oxford: Oxford University Press.

Chomsky, Noam. 1995b. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries. In *Step by step: Essays on minimalist syntax in honor of howard lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and beyond: The cartography of syntactic structures*, ed. Adriana Belletti, volume 3, 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36:1–22.

Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 291–321. Cambridge, MA: The MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33–49.

Chomsky, Noam. 2015. Problems of projection: Extensions. In *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, ed. Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 3–16. Amsterdam: John Benjamins.

Chung, Sandra, and William A. Ladusaw. 2004. *Restriction and saturation*. Cambridge, MA: The MIT Press.

Cinque, Guglielmo. 1995. *Italian syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Cinque, Guglielmo. 2013. *Typological studies: Word order and relative clauses*. London/New York: Routledge.

Citko, Barbara. 2011. *Symmetry in syntax: Merge, Move, and Labels*. Cambridge: Cambridge University Press.

Citko, Barbara, and Kleanthes K. Grohmann. 2001. The (non)-uniqueness of multiple *wh*-fronting. In *Proceedings of Formal Approaches to Slavic Linguistics: The Bloomington Meeting, 2000*, ed. Steven Franks, Tracy Holloway King, and Michael Yadroff, 117–136. Ann Arbor, MI: University of Michigan Press.

Clements, J. Clancy. 2006. Null direct objects in Spanish. In *Functional approaches to Spanish syntax*, ed. J. Clancy Clements and Jiyoung Yoon, 134–150. New York: Springer.

Cole, Peter. 1982. *Imbabura Quechua*. Amsterdam: North-Holland.

Cole, Peter, and Gabriella Hermon. 1994. Is there LF wh-movement? *Linguistic Inquiry* 25:239–262.

Cole, Peter, and Gabriella Hermon. 1998. The typology of wh-movement. Wh-questions in Malay. *Syntax* 1:221–258.

Collins, Chris. 1994. Economy of derivation and the generalized proper binding condition. *Linguistic Inquiry* 25:45–62.

Collins, Chris. 2002. Eliminating labels. In *Derivation and explanation in the minimalist program*, ed. Samuel D. Epstein and Daniel Seely, 43–64. Oxford: Blackwell.

Corver, Norbert. 1990. The syntax of left branch extractions. Doctoral Dissertation, Katholieke Universiteit Brabant.

Cowper, Elizabeth. 1987. Pied-piping, feature percolation and the structure of the noun phrase. *Canadian Journal of Linguistics* 32:321–338.

Dadan, Marcin. 2019. Head labeling preference and language change. Doctoral Dissertation, University of Connecticut, Storrs.

Danckaert, Lieven. 2012. *Latin embedded clauses: The left periphery*. Amsterdam: John Benjamins.

Davis, Henry, and Jason Brown. 2011. On A′-dependencies in Gitksan. In *Proceedings of the 46th International Conference on Salish and Neighbouring Languages (UBC Working Papers in Linguistics 30)*, ed. John Lyon and Joel Dunham, 43–80. Vancouver: UBC Working Papers in Linguistics.

Dayley, Jon P. 1989. *Tümpisa (Panamint) Shoshone grammar*. Berkeley and Los Angeles, California: University of California Press.

Delsing, Lars-Olof. 1993. The internal structure of Noun Phrases in the Scandinavian languages. Doctoral Dissertation, University of Lund.

Despić, Miloje. 2011. Syntax in the absence of determiner phrase. Doctoral Dissertation, University of Connecticut, Storrs.

Despić, Miloje. 2015. Phases, reflexives, and definiteness. *Syntax* 18:201–234.

Despić, Miloje. 2019. On kinds and anaphoricity in languages without definite articles. In *Definiteness across languages*, ed. Ana Aguilar-Guevara, Julia Pozas Loyo, and Violeta V'azquez-Rojas Maldonado, 259–291. Berlin: Language Science Press.

Devine, A. M., and Lawrence Stephens. 2000. *Discontinuous syntax: Hyperbaton in Greek*. Oxford: Oxford University Press.

Dimitriadis, Alexis. 1994. Clitics and object drop in Greek. In *Proceedings of the Sixth Student Conference in Linguistics (SCIL-6)*, ed. Chris Giordano and Daniel Ardron, 1–20. Cambridge, MA: MIT Working Papers in Linguistics.

Dixon, R. M. W. 2003. Demonstratives: A cross-linguistic typology. *Studies in Language* 27:61–112.

Dobrovie-Sorin, Carmen. 2000. (in)definiteness spread: From Romanian genitives to Hebrew Construct State nominals. In *Comparative studies in Romanian syntax*, ed. Virginia Motapanyane, 177–226. Oxford: Elsevier.

Dobrovie-Sorin, Carmen. 2002. From DPs to NPs: A Bare Phrase Structure account of genitives. In *From NP to DP. vol. 2: The expression of possession in noun phrases*, ed. Martine Coene and Yves D'hulst, 75–120. Amsterdam: John Benjamins.

Dobrovie-Sorin, Carmen. 2004. Genitives and determiners. In *Possessives and beyond: Semantics and syntax, university of Massachusetts Occasional Papers in Linguistics 29*, ed. Ji-yung Kim, Yury A. Lander, and Barbara H. Partee, 115–132. University of Massachusetts, Amherst: GLSA.

Dryer, Matthew S. 2004. Typological database. Available at http://www.acsu.buffalo.edu/ dryer/.

Dubinsky, Stanley, and Mila Tasseva-Kurktchieva. 2014. On the NP/DP frontier: Bulgarian as a transitional case. Poster presented at the 89th Annual Meeting of the Linguistic Society of America, Portland, OR.

Duguine, Maia. 2014. Argument ellipsis: A unitary approach to pro-drop. *The Linguistic Review* 31:515–549.

É Kiss, Katalin. 1995. *Discourse configurational languages*. Oxford: Oxford University Press.

Embick, David. 2000. Features, syntax and categories in the Latin perfect. *Linguistic Inquiry* 31:185–230.

Embick, David, and Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39:1–53.

Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.

Engdahl, Elisabet. 1983. Parasitic gaps. *Linguistics and Philosophy* 6:5–34.

Epstein, Samuel D., Hisatsugu Kitahara, and Daniel Seely. 2016. Phase cancellation by external pair-merge of heads. *The Linguistic Review* 33:87–102.

Epstein, Samuel D., Hisatsugu Kitahara, and Daniel Seely. 2020. Unifying labeling under minimal search in "single-" and "multiple-specifier" configurations. *Coyote Papers: Working Papers in Linguistics, Linguistic Theory at the University of Arizona* 22:1–11.

Espinal, M. Teresa. 2010. Bare nominals in Catalan and Spanish: Their structure and meaning. *Lingua* 120:984–1009.

Etxeberria, Urtzi. 2005. Quantification and domain restriction in Basque. Doctoral Dissertation, University of the Basque Country (UPV-EHU), Bilbo.

Etxeberria, Urtzi. 2014. Basque nominals: From a system with bare nouns to a system without. In *Weak referentiality*, ed. Ana Aguilar-Guevara, Bert Le Bruyn, and Joost Zwarts, 335–364. Amsterdam: John Benjamins.

Etxeberria, Urtzi, and Anastasia Giannakidou. 2010. Contextual domain restriction and the definite determiner. In *Context-dependence, perspective and relativity*, ed. François Recanati, Danijela Stojanovic, and Neftalí Villanueva, 93–126. Berlin: Mouton de Gruyter.

Faarlund, Jan Terje. 2009. On the history of definiteness marking in Scandinavian. *Journal of Linguistics* 45:617–639.

Farkas, Donka F., and Henriëtte de Swart. 2003. *The semantics of incorporation: From argument structure to discourse transparency*. Stanford, CA: CSLI Publications.

Fiva, Toril. 1987. *Possessor chains in Norwegian*. Oslo: Novus.

Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, Massachusetts: MIT Press.

Franks, Steven. 2001. The internal structure of Slavic Noun Phrases, with special reference to Bulgarian. In *Generative linguistics in Poland: Syntax and morphosyntax*, ed. Adam Przepiórkowski and Piotr Bański, 53–69. Warszawa: Instytut Podstaw Informatyki PAN. Ms., Indiana University, Bloomington.

Franks, Steven. 2007. Deriving discontinuity. In *Proceedings of FDSL 6.5*, ed. Franc Marušič and Rok Žaucer, 102–120. Frankfurt am Main: Peter Lang.

Franks, Steven, and Tracy Holloway King. 2000. *A handbook of Slavic clitics*. Oxford: Oxford University Press.

Friedmann, Naama, Adriana Belletti, and Luigi Rizzi. 2009. Relativized Relatives: Types of intervention in the acquisition of A′ dependencies. *Lingua* 119:67–88.

Fukui, Naoki. 1986. A theory of category projection and its application. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Fukui, Naoki. 1988. Deriving the differences between English and Japanese: A case study in parametric syntax. *English Linguistics* 5:249–270.

Gajewski, Jon. 2005. Neg-raising: Polarity and presupposition. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Gatt, Albert. 2004. Regular and generic possessives in Maltese. In *Possessives and beyond: Semantics and syntax, university of Massachusetts Occasional Papers in Linguistics 29*, ed. Ji-yung Kim, Yury A. Lander, and Barbara H. Partee, 199–215. University of Massachusetts, Amherst: GLSA.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12:155–184.

Geist, Ljudmila. 2013. Bulgarian *edin*: The rise of an indefinite article. In *Proceedings of FDSL 9, Göttingen 2011*, ed. Uwe Junghanns, Dorothee Fehrmann, Denisa Lenertová, and Hagen Pitsch, 125–148. Frankfurt am Main: Peter Lang.

van Gelderen, Elly. 2011. *The linguistic cycle*. Oxford: Oxford University Press.

Giannakidou, Anastasia. 2004. Domain restriction and the arguments of quantificational determiners. In *Proceedings of SALT XIV*, ed. Robert B. Young, 110–126. Cornell University, Ithaca, NY: CLC Publications.

Giannakidou, Anastasia. 2012. The landscape of Greek quantifiers. In *Handbook of quantifiers in natural language*, ed. Edward L. Keenan and Denis Paperno, 285–346. Dordrecht: Springer.

Giannakidou, Anastasia, and Lisa Lai-Shen Cheng. 2006. (In)definiteness, polarity and the role of wh-morphology in free choice. *Journal of Semantics* 23:135–183.

Giannakidou, Anastasia, and Jason Merchant. 1997. On the interpretation of null indefinite objects in Greek. *Studies in Greek Linguistics* 17:141–155.

Giannakidou, Anastasia, Lilia Schürcks, and Urtzi Etxeberria. 2009. Domain restriction and the definite determiner: The view from Slavic. Talk presented at the Workshop on NP/DP: Formal descriptions of Slavic Languages.

Goodall, Grant. 1987. *Parallel structures in syntax*. Cambridge: Cambridge University Press.

Greenberg, Joseph H. 1978. How does a language acquire gender markers? In *Universals of human language vol. 3: Word structures*, ed. Joseph H. Greenberg, Charles A. Ferguson, and Edith A. Moravcsik, 47–82. Stanford, CA: Stanford University Press.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, Massachusetts: MIT Press.

Grimshaw, Jane. 2000. Locality and extended projections. In *Lexical specification and insertion*, ed. Hans Bennis, Martin Everaert, and Eric Reuland, 183–207. Amsterdam: John Benjamins.

Groat, Erich, and John O'Neil. 1996. Spell-Out at the LF interface. In *Minimal ideas*, ed. Werner Abraham, Samuel David Epstein, Hóskuldur Thráinsson, and C. Jan-Wouter Zwart, 113–139. Amsterdam: John Benjamins Publishing Company.

Grohmann, Kleanthes K. 2003. *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins.

Grosu, Alexander. 1973. On the nonunitary nature of the Coordinate Structure Constraint. *Linguistic Inquiry* 4:88–92.

Guasti, Maria Teresa, Anna Gavarró, Joke de Lange, and Claudia Caprin. 2008. Article omission across child languages. *Language Acquisition* 15:89–119.

Haegeman, Liliane. 2010. Referential CPs and DPs: An operator movement account. *Theoretical Linguistics* 36:111–152.

Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and the composition of the left periphery*. Oxford: Oxford University Press.

Hagstrom, Paul. 1998. Decomposing questions. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Halle, Morris, and Alec Marantz. 1993a. Distributed Morphology and the pieces of inflection. In *The view from Building 20: Essays in linguistics in honour of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: The MIT Press.

Halle, Morris, and Alec Marantz. 1993b. Distributed morphology and the pieces of inflection. In *The view from building 20*, ed. Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, Massachusetts: MIT Press.

Halpern, Aaron. 1995. *On the morphology and placement of clitics*. Stanford, CA: CSLI Publications.

Hanink, Emily. 2017. The German definite article and the 'sameness' of indices. In *University of pennsylvania working papers in linguistics 23.1*, ed. Kajsa Djärv and Amy Goodwin Davies, 63–72. University of Pennsylvania.

Hanink, Emily, and Julian Grove. 2017. German relative clauses and the severed index hypothesis. In *Proceedings of west coast conference on formal linguistics 34*, ed. Aaron Kaplan, Abby Kaplan, Miranda K. McCarvel, and Edward J. Rubin, 241–248. Somerville, MA: Cascadilla Proceedings Project.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.

Haspelmath, Martin. 1999. Explaining article-possessor complementarity: Economic motivation in noun phrase syntax. *Language* 75:227–243.

Haspelmath, Martin. 2004. Coordinating constructions: An overview. In *Coordinating constructions*, ed. Martin Haspelmath, 3–39. Amsterdam: John Benjamins.

Haspelmath, Martin. 2007. Coordination. In *Language typology and syntactic description volume II: Complex constructions, 2nd edition*, ed. Timothy Shopen, 1–51. Cambridge: Cambridge University Press.

Heck, Fabian. 2008. *A theory of pied-piping*. Mouton de Gruyter.

Heck, Fabian. 2009. On certain properties of Pied-Piping. *Linguistic Inquiry* 40:75–111.

Heim, Irene. 1999. Superlatives. Ms., Massachusetts Institute of Technology.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA: Blackwell.

Heinat, Fredrik. 2006. Probes, pronouns, and binding in the Minimalist Program. Doctoral Dissertation, University of Lund.

Hermon, Gabriella. 1984. *Syntactic modularity*. Dordrecht: Foris.

Hicks, Glyn. 2009. *The derivation of anaphoric relations*. Amsterdam: John Benjamins.

Hinrichs, Erhard W. 1986. *Verschmelzungsformen* in German: A GPSG analysis. *Linguistics* 24:939–955.

Hiraiwa, Ken. 2009. A note on the typology of head-internal relativization. In *Proceedings of Conference on Language Documentation and Linguistic Theory 2*, ed. Peter Austin, Oliver Bond, Monik Charette, David Nathan, and Peter Sells, 123–131. London: SOAS.

Hirayama, Hitomi. 2018. Revisiting a null pronominal accout for parasitic gaps in Japanese. *Glossa: A Journal of General Linguistics* 3:1–33.

Höhn, Georg F.K. 2014. Contextually conditioned allomorphy and the Basque locative. In *Proceedings of the 21st conference of the student organization of linguistics in europe (ConSOLE XXI)*, ed. Martin Kohlberger, Kate Bellamy, and Eleanor Dutton, 146–170. Leiden: Leiden University Centre for Linguistics.

Hoji, Hajime. 1998. Null object and sloppy identity in Japanese. *Linguistic Inquiry* 29:127–152.

Hoji, Hajime. 2003. Surface and deep anaphora, sloppy identity, and experiments in syntax. In *A reference guide*, ed. Andrew Barss, 172–236. Oxford: Blackwell.

Horvath, Julia. 1996. *Focus in the theory of grammar and the syntax of Hungarian*. Dordrecht: Foris.

Hualde, José Ignacio, and Jon Ortiz de Urbina. 2003. *A grammar of Basque*. Berlin: Mouton de Gruyter.

Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Inaba, Jiro. 2009. Some asymmetries between initian and final subordination markers. *Bulletin of the Research Center for the Technique of Representation, Hiroshima University* 5:33–55.

Inaba, Jiro. 2011. The morphosyntax of constituent ordering patterns. *The Hiroshima University studies, Graduate School of Letters* 71:43–72.

Ippolito, Michela. 2017. Indefinite pronouns. In *Proceedings of North East Linguistic Society 47*, ed. Andrew Lamont and Katerina Tetzloff, 99–108. University of Massachusetts, Amherst: GLSA.

Irimia, Monica-Alexandrina. 2005. Types of secondary predication. *Toronto Working Papers in Linguistics* 25:20–29.

Ishihara, Shinichiro. 2003. Intonation and interface conditions. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Ishii, Toru. 1999. Cyclic spell-out and the that-t effects. In *Proceedings of the west coast conference on formal linguistics 18*, ed. Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 220–231. Somerville, MA: Cascadilla Press.

Izvorski, Roumyana. 1995. Wh-movement and focus-movement in Bulgarian. In *Proceedings of ConSOLE II*, ed. Regine Eckardt and Veerle van Greenhoven, 54–67. The Hague: Holland Academic Graphics.

Izvorski, Roumyana. 1996. The syntax and semantics of correlative proforms. In *Proceedings of North East Linguistic Society*, ed. Kiyomi Kusumoto, 133–148. Harvard University and MIT: Graduate Linguistic Student Association.

Jaeggli, Osvaldo. 1982. *Topics in Romance syntax*. Dordrecht Holland: Foris Publications.

Jaeggli, Osvaldo. 1986. Three issues in the theory of clitics: Case, doubled NPs, and extraction. In *Syntax and semantics 19: The syntax of pronominal clitics*, ed. Hagit Borer, 15–42. Orland: Academic Press.

Jenks, Peter. 2018. Articulated definiteness without articles. *Linguistic Inquiry* 49:501–536.

Jensen, Per Anker, and Carl Vikner. 1994. Lexical knowledge and the semantic analysis of Danish genitive constructions. In *Topics in knowledge-based NLP systems*, ed. Steffen Leo Hansen and Helle Wegener, 37–55. Copenhagen: Samfundslitteratur.

Jentsch, Helmut. 1980. *Die sorbische Mundart von Rodewitz/Spree*. Bautzen: Domowina.

Johannessen, Janne Bondi. 1998. *Coordination*. Oxford: Oxford University Press.

Johnson, Kyle. 2002. Restoring exotic coordinations to normalcy. *Linguistic Inquiry* 33:97–156.

Kang, Jungmin. 2014. On the absence of TP and its consequences. Doctoral Dissertation, University of Connecticut, Storrs.

Kasai, Hironobu, and Shoichi Takahashi. 2002. Coordination in Japanese. In *Proceedings of FAJL 3*, ed. Maria Cristina Cuervo, Daniel Harbour, Ken Hiraiwa, and Shinichiro Ishihara, 19–32. Cambridge, MA: MIT Working Papers in Linguistics.

Kato, Takaomi. 2005. A case against the representational approach to the Coordinate Structure Constraint. In *Proceedings of North East Linguistic Society 35*, ed. Leah Bateman and Cherlon Ussery, 307–321. Amherst, MA: Graduate Linguistic Student Association.

Kato, Takaomi. 2006. Symmetries in coordination. Doctoral Dissertation, Harvard University, Cambridge, MA.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, Massachusetts: MIT Press.

Kayne, Richard S. 2004. Here and there. In *Lexique syntaxe, et lexique-grammaire/syntax, lexis & lexicon grammar: Papers in honour of Maurice Gross*, ed. Christian Leclère, Éric Laporte, Mireille Piot, and Max Silberztein, 253–275. Amsterdam and Philadelphia: John Benjamins.

Kazenin, Konstantin I., and Yakov G. Testelets. 2004. Where coordination meets subordination: Converb constructions in tsakhur (daghestanian). In *Coordinating constructions*, ed. Martin Haspelmath, 227–239. Amsterdam: John Benjamins.

Kehler, Andrew. 1996. Coherence and the Coordinate Structure Constraint. In *Proceedings of the 22nd annual meeting of the berkeley linguistics society*, ed. Jan Johnson, Matthew L. Juge, and Jeri L. Moxley, 220–231. Berkeley, California: Berkeley Linguistics Society, University of California.

Kiefer, Ferenc. 1990. Noun incorporation in Hungarian. *Acta Linguistica Hungarica* 40:149–177.

Kishimoto, Hideki. 1992. LF pied piping: Evidence from Sinhala. *Gengo Kenkyu* 102:46–87.

Kishimoto, Hideki. 2005. Wh-in-situ and movement in Sinhala questions. *Natural Language and Linguistic Theory* 23:1–51.

Ko, Heejeong. 2005. Syntax of *Why-in-situ*: Merge into [Spec,CP] in the overt syntax. *Natural Language and Linguistic Theory* 23:867–916.

Kobuchi-Philip, Mana. 2007. Individual-denoting classifiers. *Natural Language Semantics* 15:95–130.

Koulidobrova, Elena. 2017. Elide me bare. *Natural Language and Linguistic Theory* 35:397–446.

Koulidobrova, Elena. in press. DP or not DP: Acquisitional evidence. *Language Acquisition* .

Kratzer, Angelika. 1989. An investigation in the lumps of thought. *Linguistics and Philosophy* 12:607–653.

Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In *The Proceedings of the Third Tokyo Conference on Psycholinguistics (TCP 2002)*, ed. Yukio Otsu, 1–25. Tokyo: Hituzi Syobo.

Kroeger, Paul. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford, CA: CSLI.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kurafuji, Takeo. 1999. Japanese pronouns in dynamic semantics: The null/overt contrast. Doctoral Dissertation, Rutgers University, New Brunswick, NJ.

Kuroda, S.-Y. 1965. Generative grammatical studies in the Japanese language. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Laenzlinger, Christopher. 1993. A syntactic view of Romance pronominal sequences. *Probus* 5:241–270.

Lambova, Mariana. 2001. On A-bar movements in Bulgarian and their interaction. *The Linguistic Review* 18:327–374.

Larson, Richard, and Sungeun Cho. 2003. Temporal adjectives and the structure of possessive DPs. *Natural Language Semantics* 11:217–247.

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15:235–290.

LaTerza, Ivana. 2014. The DP category and Serbian nominal structure. Doctoral Dissertation, Stony Brook University, New York.

Laury, Ritva. 1997. *Demonstratives in interaction: The emergence of a definite article in Finnish*. Amsterdam: John Benjamins.

Ledgeway, Adam. 2012. *From Latin to Romance: Morphosyntactic typology and change*. Oxford: Oxford University Press.

Lee-Schoenfeld, Vera. 2004. Binding by phase: (non-)complementarity in German. *Journal of Germanic Linguistics* 16:111–171.

Lee-Schoenfeld, Vera. 2008. Binding, phases, and locality. *Syntax* 11:281–298.

Leko, Nedžad. 1999. Functional categories and the structure of DP in Bosnian. In *Topics in South Slavic syntax and semantics*, ed. Mila Dimitrova-Vulchanova and Lars Hellan, 229–252. Amsterdam: John Benjamins.

Li, Audrey Yen-Hui. 1992. Indefinite *Wh* in Mandarin Chinese. *Journal of East Asian Linguistics* 1:12–155.

Lightfoot, David. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.

Lightfoot, David. 1999. *The development of language: Acquisition, change, and evolution*. Malden, MA: Blackwell.

Lin, Jonah. 2005. Does wh-in-situ license parasitic gaps? *Linguistic Inquiry* 36:298–302.

Lin, Vivian. 2001. A way to undo A movement. In *Proceedings of the 20th West Coast Conference in Formal Linguistics*, ed. Karine Megerdoomian and Leora Anne Bar-el, 358–371. Somerville, Massachusetts: Cascadilla Press.

Longobardi, Giuseppe. 1994. Reference and proper names. *Linguistic Inquiry* 25:609–666.

López, Luis. 2012. *Indefinite objects: Scrambling, choice functions, and differential marking*. Cambridge, MA: The MIT Press.

Lord, Carol. 1993. *Historical change in serial verb constructions*. Cambridge: Cambridge University Press.

Lötzsch, Ronald. 1968. Einige Auswirkungen das Purismus auf die grammatische Normierung slawischer Schriftsprachen. In *Sorabistiske přinoški k VI. mjezynarodnemu kongresej slawistow w Praze 1968*, 21–36. Bautzen: Domowina.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.

MacWhinney, Brian. 1976. Hungarian research on the acquisition of morphology and syntax. *Journal of Child Language* 3:397–410.

MacWhinney, Brian, and Elizabeth Bates. 1978. Sentential devices for conveying givenness and newness: A cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behavior* 17:539–558.

Maki, Hideki. 1995. The syntax of particles. Doctoral Dissertation, University of Connecticut, Storrs.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st Annual Penn Linguistic Colloquium*, ed. Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark, and Alexander Williams, 201–225. Philadelphia, PA: University of Pennsylvania.

Mardale, Alexandru. 2006. On the use of the definite article with prepositions in Romanian. halshs-00556191.

Marelj, Marijana. 2011. Bound-variable anaphora and left branch condition. *Syntax* 14:205–229.

Margolis, Rebecca. 2011. *Basic Yiddish: A grammar and workbook*. London: Routledge.

Marinis, Theodore. 2003. *The acquisition of the DP in Modern Greek*. Amsterdam: John Benjamins.

Markopoulou, Despina Chila. 2000. The indefinite article in Greek: A diachronic approach. *Glossologia* 11:111–130.

Martin, John N. 1975. Karttunen on possibility. *Linguistic Inquiry* 6:339–341.

Mathieu, Eric, and Ioanna Sitaridou. 2002. Split wh-constructions in Classical and Modern Greek. *Linguistics in Potsdam* 19:143–182.

Matthewson, Lisa, Tim Bryant, and Thomas Roeper. 2001. A Salish stage in the acquisition of English determiners: Unfamiliar "definites". In *The Proceedings of SULA: The Semantics of Under-Represented Languages in the Americas, University of Massachusetts Occasional Papers in Linguistics 25*, 63–71.

Merchant, Jason. 2001. *The syntax of silence: sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.

Mitrović, Moreno. 2011. The syntax of coordination in Sanskrit. Master's thesis, University of Cambridge, Cambridge.

Morita, Hisashi. 2009. Covert pied-piping in Japanese *wh*-questions. *English Linguistics* 26:374–393.

Moritz, Luc, and Daniel Valois. 1994. Pied-piping and specifier-head agreement. *Linguistic Inquiry* 25:667–708.

Moro, Andrea. 2000. *Dynamic antisymmetry*. Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.

Munn, Alan Boag. 1993. Topics in the syntax and semantics of coordinate structures. Doctoral Dissertation, The University of Maryland.

Myrberg, Sara, and Thomas Riad. 2015. The prosodic hierarchy of Swedish. *Nordic Journal of Linguistics* 38:115–147.

Newell, Wayne. 1974. *Kukec*. Indian Township, Maine: Wabnaki Bilingual Education Program.

Nicolae, Alexandru. 2012. P-stranding in a language without P-stranding? the case of sluicing in Romanian. Paper presented at Societas Linguistica Europaea, Stockholm 2012.

Nishigauchi, Taisuke. 1990. *Quantification in the theory of grammar*. Dordrecht: Kluwer.

Nishiyama, Kunio. 2013. Decomposing demonstratives and wh-words. *JELS* 30:159–165.

Noonan, Michael. 2007. Complementation. In *Language typology and syntactic description vol. II, complex constructions*, ed. Timothy Shopen, 52–150. Cambridge: Cambridge University Press.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Linguistic Inquiry Monographs. Cambridge, Massachusetts: MIT Press.

Nygaard, Marius. 1917. *Bemerkninger, rettelser og supplementer til min Norrøsyntax*. Kristiania: Videnskabsselskapets skrifter.

Oda, Hiromune. 2017. Two types of the Coordinate Structure Constraint and rescue by PF deletion. In *Proceedings of NELS 47, volume 2*, ed. Andrew Lamont and Katerina Tetzloff, 343–356. University of Massachusetts, Amherst: Graduate Linguistics Student Association.

Oda, Hiromune. 2021a. Decomposing and deducing the Coordinate Structure Constraint. *The Linguistic Review* 38:605–644.

Oda, Hiromune. 2021b. Japanese free choice items as unconditionals. *Natural Language Semantics* 29:281–338.

Ohno, Yutaka. 1989. *Mo*. In *Papers on quantification. NSF grant report, Department of Linguistics, University of Massachusetts, Amherst*, ed. Emmon Bach, Angelika Kratzer, and Barbara Partee, 224–250. GLSA.

Oka, Toshifusa. 1993. Shallowness. In *MIT working papers in linguistics 19: Papers on Case and Agreement II*, ed. Colin Phillips, 255–320. Cambridge, MA: MIT Working Papers in Linguistics.

Oku, Satoshi. 1998. A theory of selection and reconstruction in the Minimalist Program. Doctoral Dissertation, University of Connecticut, Storrs.

Otaki, Koichi. 2014. Ellipsis of arguments: Its acquisition and theoretical implications. Doctoral Dissertation, University of Connecticut, Storrs.

Panagiotidis, Phoevos. 2002. *Pronouns, clitics and empty nouns: 'pronominality' and licensing in syntax*. Amsterdam: John Benjamins.

Pancheva, Roumyana, and Barbara Tomaszewicz. 2012. Cross-linguistic differences in superlatives movement out of nominal phrases. In *Proceedings of the 30th West Coast Conference on Formal Linguistics*, ed. Nathan Arnett and Ryan Bennett, 292–302. Cascadilla Press.

Pană Dindelegan, Gabriela. 2016. *The syntax of Old Romanian*. Oxford: Oxford University Press.

Partee, Barbara H. 1997. Uniformity vs. versatility: The genitive, a case study. appendix to Theo M. V. Janssen, compositionality. In *The handbook of logic and language*, ed. Johan van Benthem and Alice G. B. ter Meulen, 464–470. Amsterdam: Elsevier.

Partee, Barbara H., and Vladimir Borschev. 2003. Genitives, relational nouns, and argument-modifier ambiguity. In *Modifying adjuncts*, ed. Ewald Lang, Claudia Maienborn, and Cathrine Fabricius-Hansen, 67–112. Berlin: Mouton de Gruyter.

Pereltsvaig, Asya. 2008. *Copular sentences in Russian*. New York: Springer.

Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and clausal architecture: Syntactic derivation and interpretation*, ed. Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 262–294. Amsterdam: John Benjamins Publishing Company.

Peterson, David A., and Kenneth VanBik. 2004. Coordination in Hakha Lai (Tibeto-Burman. In *Coordinating constructions*, ed. Martin Haspelmath, 333–356. Amsterdam: John Benjamins.

Petroj, Vanessa. 2020. Ignoring language barriers: Romanian-Serbian code-switching. Doctoral Dissertation, University of Connecticut, Storrs.

Polinsky, Maria, and Robert Kluender. 2007. Linguistic typology and theory construction: Common challenges ahead. *Linguistic Typology* 11:273–283.

Postal, Paul M. 1998. *Three investigations of extraction*. Cambridge: The MIT Press.

Postma, Gertjan. 1994. The indefinite reading of WH. In *Linguistics in the netherlands 1994*, ed. Reineke Bok-Benneman and Crit Cremers, 187–198. Amsterdam: John Benjamins.

Progovac, Ljiljana. 1998. Determiner phrase in a language without determiners. *Journal of Linguistics* 34:165–179.

Progovac, Ljiljana. 2010. Syntax: Its evolution and its representation in the brain. *Biolinguistics* 4:234–254.

Progovac, Ljiljana, and John L. Locke. 2009. The urge to merge: Ritual insult and the evolution of syntax. *Biolinguistics* 3:337–364.

Puig-Waldmüller, Estela Sophie. 2008. Contracted preposition-determiner forms in German. Doctoral Dissertation, Universitat Pompeu Fabra, Barcelona.

Quicoli, A. Carlos. 2008. Anaphora by phase. *Syntax* 11:299–329.

Radford, Andrew. 1990. *Syntactic theory and the acquisition of English syntax*. Cambridge, Massachusetts: Basil Blackwell Ltd.

Rákosi, György. 2017. The definite article and anaphoric possessors in Hungarian. *Linguistica Brunensia* 65:21–33.

Rákosi, György. 2020. With or without the definite article: On the syntax of anaphoric possessor strategies in Hungarian. In *Approaches to Hungarian volume 16: Papers from the 2017 Budapest Conference*, ed. Veronika Hegedüs and Irene Vogel, 116–136. Amsterdam: John Benjamins.

Reglero, Lara. 2007. Wh-in-situ interrogatives in Spanish. *Probus* 19:267–297.

Reglero, Lara, and María Emma Ticio. 2013. A unified analysis of wh-in-situ in Spanish. *The Linguistic Review* 30:1–47.

Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20:335–397.

Rennison, John. 1997. *Koromfe*. London/New York: Routledge.

Reuland, Eric. 2011. *Anaphora and language design*. Cambridge, MA: The MIT Press.

Richards, Norvin. 1999. Complementizer cliticization in Tagalog and English. In *Toronto working papers in linguistics 16, vol. 2: Proccedings of the sixth meeting of the Austronesian Formal Linguistics Association*, ed. Carolyn Smallwood and Catherine Kitto, 297–312. Toronto: University of Toronto.

Richards, Norvin. 2000. An island effect in Japanese. *Journal of East Asian Linguistics* 9:187–205.

Richards, Norvin. 2001. *Movement in language: interactions and architectures*. Oxford: Oxford University Press.

Richards, Norvin W., III. 1997. What moves where when in which language? Doctoral Dissertation, Massachusetts Institute of Technology.

Riqueros, Jos'e. 2013. Spanish nominal(ization) patterns. Doctoral Dissertation, University of Connecticut, Storrs.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer Academic Publishers.

Rizzi, Luigi. 2001. On the position "int(errogative)" in the left periphery of the clause. In *Current studies in Italian syntax: Essays offered to Lorenzo Renzi*, ed. Guglielmo Cinque and Giampaolo Salvi, 267–296. Amsterdam: Elsevier.

Rizzi, Luigi. 2004. Locality and left periphery. In *Structures and beyond: The cartography of syntactic structures*, ed. Adriana Belletti, 223–251. Oxford: Oxford University Press.

Rizzi, Luigi. 2010. On some properties of criterial freezing. In *The complementizer phase: Subjects and operators*, ed. E. Phoevos Panagiotidis, 17–32. Oxford: Oxford University Press.

Roberts, Ian. 2007. *Diachronic syntax*. Oxford: Oxford University Press.

Roberts, Ian. 2012. Macroparameters and minimalism: A programme for comparative research. In *Parameter theory and linguistic change*, ed. Charlotte Galves, Sonia Cyrino, Ruth Lopes, Filomena Sandalo, and Juanito Avelar, 320–335. Oxford: Oxford University Press.

Roberts, Ian. 2019. *Parameter hierarchy and Universal Grammar*. Oxford: Oxford University Press.

Roberts, Ian, and Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In *Parametric variation: Null subjects in minimalist theory*, ed. Theresa Biberauer, Anders Holmberg, Ian Roberts, and Michelle Sheehan, 1–57. Cambridge: Cambridge University Press.

Roberts, Ian, and Anna Roussou. 2003. *Syntactic change: A minimalist approach to grammaticalisation*. Cambridge: Cambridge University Press.

Rochemont, Michael. 1986. *Focus in generative grammar*. Amsterdam, The Netherlands: John Benjamins Publishing Company.

Ross, John. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Ross, John Robert. 1969. Guess who? In *Chicago Linguistics Society*, ed. Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, 252–286. Chicago, Illinois.

Rudin, Catherine. 1998. On multiple questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6:445–501.

Runić, Jelena. 2014a. A new look at argument ellipsis: Evidence from Slavic. In *Proceedings of North East Linguistic Society 43*, ed. Hsin-Lun Huang, Ethan Poole, and Amanda Rysling, volume 2, 91–102. University of Massachusetts, Amherst: GLSA.

Runić, Jelena. 2014b. A new look at clitics, clitic doubling, and argument ellipsis. Doctoral Dissertation, University of Connecticut, Storrs.

Ruys, Eddy G. 1993. The scope of indefinites. Doctoral Dissertation, Utrecht Univeristy, Utrecht.

Sag, Ivan A., Gerald Gazdar, Thomas Wasow, and Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3:117–172.

Saito, Mamoru. 2007. Notes on East Asian argument ellipsis. *Language Research* 43:203–227.

Saito, Mamoru. 2016. (A) Case for Labeling: Labeling in languages without phi-feature agreement. *The Linguistic Review* 33:129–175.

Saito, Mamoru. 2017. Japanese wh-phrases as operators with unspecified quantificational force. *Language and Linguistics* 18:1–25.

Saito, Mamoru. 2018. Kase as a weak head. In *Heading in the right direction: Linguistic treats for Lisa Travis, mcgill working papers in linguistics 25.1*, ed. Laura Kalin, Ileana Paul, and Jozina Vander Klok, 382–390. Montreal: McGill University.

Saito, Mamoru. 2020. On the causative paradoxes: Derivations and transfer domains. *Nanzan Linguistics* 15:25–44.

Saito, Mamoru, and Keiko Murasugi. 1999. Subject predication within IP and DP. In *Beyond principles and parameters: Essays in memory of osvaldo jaeggli*, ed. Kyle Johnson and Ian Roberts, 159–182. Dordrecht: Kluwer Academic Publishers.

Sakamoto, Yuta. 2017. Escape from silent syntax. Doctoral Dissertation, University of Connecticut, Storrs.

Sakamoto, Yuta. 2019. Overtly empty but covertly complex. *Linguistic Inquiry* 50:105–136.

Sato, Yosuke. 2014. Argument ellipsis in Colloquial Singapore English and the anti-agreement hypothesis. *Journal of Linguistics* 50:365–407.

Sato, Yosuke. 2015. Argument ellipsis in Javanese and voice agreement. *Studia Linguistica* 69:58–85.

Schaarschmidt, Gunter. 1984. Theme-rheme structure and the article in Sorbian. *Working Papers of the Linguistics Circle of the University of Victoria* 4:75–90.

Schachter, Paul. 1977. Constraints on coordination. *Language* 53:86–103.

Schachter, Paul, and Fe T. Otanes. 1972. *Tagalog reference grammar*. Berkeley, CA: University of California Press.

Scholten, Jolien. 2010. Weak definites and bare singulars: Their syntactic peculiarities. Master's thesis, Utrecht Univeristy, Utrecht.

Schwarz, Florian. 2009. Two types of definites in natural language. Doctoral Dissertation, University of Massachusetts, Amherst, MA.

Schwarz, Florian. 2019. Weak vs. strong definite articles: Meaning and form across languages. In *Definiteness across languages*, ed. Ana Aguilar-Guevara, Julia Pozas Loyo, and Violeta V'azquez-Rojas Maldonado, 1–37. Berlin: Language Science Press.

Scott, Gary-John. 2002. Stacked adjectival modivication and the structure of nominal phrases. In *Functional structure in DP and IP*, ed. Guglielmo Cinque, volume 1 of *The Cartography of Syntactic Structures*, 91–122. Oxford University Press.

Seely, Daniel. 2006. Merge, derivation, c-command, and subcategorization in a label-free syntax. In *Minimalist essays*, ed. Cedric Boeckx, 182–217. Amsterdam: John Benjamins.

Sharvit, Yael, and Penka Stateva. 2002. Superlative expressions, context, and focus. *Linguistics and Philosophy* 25:453–505.

Shen, Zheng. 2014. Locality and the superlative conspiracy. Ms., University of Connecticut, Storrs.

Shimoyama, Junko. 2006. Indeterminate phrase quantification in Japanese. *Natural Language Semantics* 14:139–173.

Shiobara, Kayono. 2017. A phonological approach to left branch condition: Evidence from exceptions in Japanese. In *Proceedings of FAJL 8*, ed. Ayaka Sugawara, Shintaro Hayashi, and Satoshi Ito, 143–152. Cambridge, MA: MIT Working Papers in Linguistics.

Simpson, Andrew, and Tanmoy Bhattacharya. 2003. Obligatory overt wh-movement in a wh-in-situ language. *Linguistic Inquiry* 34:127–142.

Soltan, Usama. 2007. On the individual-property contrast in Egyptian Arabic possessive nominals. In *Perspectives on Arabic Linguistics 20*, ed. Mustafa Mughazi, 71–86. Amsterdam: John Benjamins.

Soltan, Usama. 2020. On null objects in Egyptian Aarabic. *Brill's Journal of Afroasiatic Languages and Linguistics* 12:204–259.

Speas, Margaret. 1994. Null arguments in a theory of Economy of Projections. In *University of Massachusetts occasional papers in linguistics 20: Functional projections*, 179–208. Amherst, MA: Graduate Linguistics Student Association.

Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19:425–449.

Stanković, Branimir. 2019. DP or NP?: The case of Serbian Southeastern dialects and Macedonian Language. In *SinFonIJA 6 Proceedings (the 6th International Conference Syntax, Phonology*

and Language Analysis (SinFonIJA 6), held at the Faculty of Philosophy of University in Niš, Serbia, September 26–28, 2013), ed. Branimir Stanković and Aleksandra Janić, 83–118. Faculty of Philosophy, University of Niš.

Starke, Michal. 2001. Move dissolves into merge: A theory of locality. Doctoral Dissertation, University of Geneva.

Stepanov, Arthur. 1998. On Wh-fronting in Russian. In *Proceedings of the North East Linguistic Society*, ed. Pius N. Tamanji and Kiyomi Kusumoto, 453–468. University of Toronto: Graduate Linguistic Student Association.

Stepanov, Arthur, and Wei-Tien Dylan Tsai. 2008. Cartography and licensing of *wh*-adjuncts: a cross-linguistic perspective. *Natural Language and Linguistic Theory* 26:589–638.

Stephany, Ursula. 1997. The acquisition of Greek. In *The crosslinguistic study of language acquisition, volume 4*, ed. Dan Isaac Sorbin, 183–333. Mahwah, NJ: Lawrence Erlbaum Associates.

Stjepanović, Sandra. 1999. Scrambling: Overt movement or base generation and LF movement? *Journal of Slavic Linguistics* 7:305–324.

Stjepanović, Sandra. 1999. What do second position cliticization, scrambling, and multiple *wh*-fronting have in common? Doctoral Dissertation, University of Connecticut, Storrs.

Stjepanović, Sandra. 2003. A word order paradox resolved by copy deletion at PF. *Linguistic Variation Yearbook* 3:139–177.

Stjepanović, Sandra. 2014. Left Branch Extraction and the Coordinate Structure Constraint. In *NELS 44: Proceedings of the 44th annual meeting of the North East Linguistic Society, volume 2*, ed. Jyoti Iyer and Leland Kusmer, 157–170. University of Massachusetts, Amherst: Graduate Linguistic Student Association.

Stojković, Jelena. 2019. The fleeting DP in Bulgarian and Macedonian: The view from left-branch

extraction. In *Structure removal (Linguistische Arbeits Berichte 94)*, ed. Andrew Murphy, 343–384. Leipzig: Universität Leipzig.

Sumangala, Lelwala. 1992. Long distance dependencies in Sinhala: The syntax of focus and Wh questions. Doctoral Dissertation, Cornell University, Ithaca, NY.

Szabolcsi, Anna. 1994. The Noun Phrase. In *The syntactic structure of Hungarian, Syntax and Semantics 27*, ed. Ferenc Kiefer and Katalin É Kiss, 179–275. New York: Academic Press.

Szabolcsi, Anna. 2019. Unconditionals and free choice unified. In *Proceedings of Semantics and Linguistic Theory 29*, ed. Katherine Blake, Forrest Davis, Kaelyn Lamp, and Joseph Rhyne, 320–340. Washington, D.C.: LSA.

Takahashi, Daiko. 1993. Movement of wh-phrases in Japanese. *Natural Language and Linguistic Theory* 11:655–678.

Takahashi, Daiko. 2002. Determiner raising and scope shift. *Linguistic Inquiry* 33:575–615.

Takahashi, Daiko. 2006. Apparent parasitic gaps and null arguments in Japanese. *Journal of East Asian Linguistics* 15:1–35.

Takahashi, Daiko. 2007. Argument ellipsis from a cross-linguistic perspective: An interim report. Paper presented at GLOW in Asia VI, The Chinese University of Hong Kong.

Takahashi, Daiko. 2008a. Noun phrase ellipsis. In *Oxford handbook of japanese linguistics*, ed. Shigeru Miyagawa and Mamoru Saito, 394–422. Oxford University Press.

Takahashi, Daiko. 2008b. Quantificational null objects and argument ellipsis. *Linguistic Inquiry* 39:307–326.

Takahashi, Daiko. 2013. Argument ellipsis in Japanese and Malayalam. *Nanzan Linguistics* 9:173–192.

Takahashi, Daiko. 2020. Derivational argument ellipsis. *The Linguistic Review* 37:47–74.

Takahashi, Masahiko. 2011. Some theoretical consequences of Case-marking in Japanese. Doctoral Dissertation, University of Connecticut, Storrs.

Takahashi, Masahiko, and Kenshi Funakoshi. 2013. On PP left-branch extraction in Japanese. In *Proceedings of the 36th Annual Penn Linguistics Colloquium (University of Pennsylvania Working Paper in Linguistics 19.1)*, ed. Kobey Shwayder, 237–246. Philadelphia, PA: University of Pennsylvania.

Talić, Aida. 2013. Extraordinary complement extraction: PP-complements and inherently case-marked nominal complements. *Studies in Polish Linguistics* 8:127–150.

Talić, Aida. 2015. Adverb extraction, specificity, and structural parallelism. *The Canadian Journal of Linguistics* 60:417–454.

Talić, Aida. 2017. From A to N and back: Functional and bare projections in the domain of N and A. Doctoral Dissertation, University of Connecticut, Storrs.

Tarpent, Marie-Lucie. 1987. A grammar of the Nisgha language. Doctoral Dissertation, University of Victoria, Victoria, BC.

Taylor, Ann. 1990. Clitics and configurationality oin Ancient Greek. Doctoral Dissertation, University of Pennsylvania.

Thráinsson, Höskuldur. 2007. *The syntax of Icelandic*. Cambridge: Cambridge University Press.

Ticio, María Emma. 2003. On the structure of DPs. Doctoral Dissertation, University of Connecticut, Storrs.

Ticio, María Emma. 2005. Locality and anti-locality in Spanish DPs. *Syntax* 8:229–286.

Todorović, Neda. 2016. On the presence/absence of TP: Syntactic properties and temporal interpretation. Doctoral Dissertation, University of Connecticut, Storrs.

Tomioka, Satoshi. 2003. The semantics of Japanese null pronouns and its cross-linguistic implications. In *The interfaces: Deriving and interpreting omitted structures*, ed. Kerstin Schwabe and Susanne Winkler, 321–339. Amsterdam: John Benjamins.

Tran, Thuan. 2009. Wh-quantification in Vietnamese. Doctoral Dissertation, University of Delaware.

Travis, Lisa. 1984. Parameters and effects of word order variation. Doctoral Dissertation, Massachusetts Institute of Technology.

Trenkic, Danijela. 2004. Definiteness in Serbian/Croatian/Bosnian and some implications for the general structure of the nominal phrase. *Lingua* 114:1401–1427.

Tsai, Wei-Tien Dylan. 1994. On economizing the theory of A-bar dependencies. Doctoral Dissertation, MIT.

Tsai, Wei-Tien Dylan. 1999. On lexical coutresy. *Journal of East Asian Linguistics* 8:39–73.

Tsimpli, Ianthi Maria, and Despina Papadopoulou. 2006. Aspect and argument realization: A study on antecedentless null objects in Greek. *Lingua* 116:1595–1615.

Uegaki, Wataru. 2018. A unified semantics for the japanese Q-particle *ka* in indefinites, questions and disjunctions. *Glossa: A Journal of General Linguistics* 3:1–45.

Ultan, Russell. 1978. Some general characteristics of interrogative systems. In *Universals of human language*, ed. Joseph H. Greeberg, volume 4, 211–248. Stanford, CA: Stanford University Press.

Ortiz de Urbina, Jon. 1989. *Parameters in the grammar of Basque*. Dordrecht: Foris.

Ortiz de Urbina, Jon. 1993. Feature percolation and clausal pied-piping. In *Generative studies in Basque linguistics*, ed. José Ignacio Hualde and Jon Ortiz de Urbina, 189–219. Amsterdam: John Benjamins.

Uriagereka, Juan. 1988. On government. Doctoral Dissertation, University of Connecticut, Storrs.

Uriagereka, Juan. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26:79–124.

Uriagereka, Juan. 1996. Determiner clitic placement. In *Current issues in comparative grammar*, ed. Robert Freidin, 257–294. Dordrecht: Kluwer.

Verschik, Anna. 2001. On the dynamics of article use in Estonian Yiddish. *Folia Linguistica* 35:337–369.

Vikner, Carl, and Per Anker Jensen. 2002. A semantic analysis of the English genitive. interaction of lexical and formal semantics. *Studia Linguistica* 56:191–226.

Villa-García, Julio. 2015. *The syntax of multiple-*que *sentences in Spanish*. Amsterdam: John Benjamins.

Wang, Shuyan. 2019. Reconsideration of *yi* 'one' and classifiers in Mandarin Chinese. Ms., University of Connecticut, Storrs.

Warden, David A. 1976. The influence of context on children's use of identifying expression and references. *British Journal of Psychology* 67:101–112.

Watanabe, Akira. 1992a. Subjacency and S-structure movement of *wh*-in-situ. *Journal of East Asian Linguistics* 1:255–291.

Watanabe, Akira. 1992b. Wh-in-situ, subjacency, and chain formation. In *MIT occasional paper in Linguistics 2*. Department of Linguistics and Philosophy, MIT: MITPWL.

Watanabe, Akira. 2002. Loss of overt wh-movement in Old Japanese. In *Syntactic effects of morphological change*, ed. David Lightfoot, 179–195. Oxford University Press.

Watanabe, Akira. 2004a. Indeterminates and determiners. In *Proceedings of the Workshop on Altaic Formal Linguistics 1*, ed. Aniko Csirmaz, Youngjoo Lee, and Mary Ann Walter, 390–405. Cambridge, MA: MIT Working Paper in Linguistics.

Watanabe, Akira. 2004b. Parametrization of quantificational determiners and head-internal relatives. *Language and Linguistics* 5:59–97.

Watanabe, Akira. 2009. A parametric shift in the D-system in Early Middle English: Relativization, articles, adjectival inflection, and indeterminates. In *Historical syntax and linguistic theory*, ed. Paola Crisma and Giuseppe Longobardi, 358–374. Oxford: Oxford University Press.

Webelhuth, Gert. 1992. *Principles and parameters of syntactic saturation*. Oxford: Oxford University Press.

Williams, Edwin. 1978. Across-the-board rule application. *Linguistic Inquiry* 9:31–43.

Williamson, Janis S. 1984. Studies in Lakhota grammar. Doctoral Dissertation, University of California, San Diego, San Diego, CA.

Willim, Ewa. 2000. On the grammar of Polish nominals. In *Step by step*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 319–346. Cambridge, Massachusetts: MIT Press.

Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20:399–467.

Yang, Muyi. 2021. A phase-based analysis of Sinhala Q-particle: New evidence from focus and verbal morphology. Ms., University of Connecticut, Storrs.

Yatabe, Shuichi. 2003. Does scrambling in Japanese obey the Coordinate Structure Constraint? In *Nihon gengogakkai dai-126-kai taikai yokoo-shuu [Proceedings of the 126th meeting of the Linguistic Society of Japan]*, 262–267. Kyoto: Linguistic Society of Japan.

Yatsushiro, Kazuko. 2009. The distribution of quantificational suffixes in Japanese. *Natural Language Semantics* 17:141–173.

Yoshida, Masaya, Chizuru Nakao, and Iván Ortega-Santos. 2015. The syntax of *Why*-Stripping. *Natural Language and Linguistic Theory* 33:323–370.

Yuan, Michelle. 2018. Dimensions of ergativity in Inuit: Theory and microvariation. Doctoral Dissertation, MIT, Cambridge, MA.

Yükseker, Hitay. 2000. Bir 'one'. In *Studies in turkish linguistics: Proceedings of the tenth international conference in Turkish linguistics*, ed. Sumru Özsoy, Didar Akar, Mine Nakipoğlu-Demiralp, Eser E. Erguvanlı-Taylan, and Ayhan Aksu-Koç, 165–172. Istanbul: Boğazçi University Press.

Zanon, Ksenia. 2015. On hybrid coordination and quantifier raising in Russian. Doctoral Dissertation, University of Indiana, Bloomington.

Zanon, Ksenia. 2020. Focus association with ONLY in Russian, Chinese, and English. Talk given at the Cambridge Linguistics Society.

Zentz, Jason Arik. 2016. Forming *wh*-questions in Shona: A comparative Bantu perspective. Doctoral Dissertation, Yale University, New Haven, CT.

Zlatić, Larisa. 1997. The structure of the Serbian Noun Phrase. Doctoral Dissertation, University of Texas at Austin.

Zoerner, Cyril Edward, III. 1995. Coordination: The syntax of &P. Doctoral Dissertation, University of California, Irvine.

Zwart, Jan-Wouter. 2005. Some notes on coordination in head-final languages. In *The linguistics in the Netherlands*, ed. Jenny Doetjes and Joost van de Weijer, 232–241. Amsterdam: John Benjamins.

Zwart, Jan-Wouter. 2009. Relevance of typology to minimalist inquiry. *Lingua* 119:1589–1606.

Zwicky, Arnold M. 1982. An expanded view of morphology in the syntax-phonology interface. In *Preprints of the plenary session papers: XIII international congress of linguists*, 168–178. Tokyo.

Zwicky, Arnold M. 1984. "reduced words" in highly modular theories: Yiddish anarthrous locatives reexamined. *Ohio State Uuniversity Working Papers in Linguistics* 29:117–126.