12-2-2019

Head Labeling Preference and Language Change

Marcin R. Dadan
University of Connecticut - Storrs, marcin.dadan@uconn.edu

Head Labeling Preference and Language Change

Marcin Dadan, Ph.D.

University of Connecticut, 2019

This dissertation explores a cross-linguistic trend of a diachronic loss of obligatory syntactic movement, which includes the loss of phrasal movement, as in the new observation regarding the unidirectionality of *wh*-dependency changes—always from fronting to *in-situ* (e.g. Old to Modern Japanese, Archaic to Modern Mandarin, Old to Modern Indo-Iranian, Latin to Modern Romance (also Basque)), as well as the loss of head-movement, e.g. V-to-T in English or Swedish.

I propose a unified explanation for these changes based on the preference for head-phrase {H,YP} configurations from the perspective of labeling (Chomsky 2013). I argue that the pressures imposed by Labeling Algorithm to maximize head-phrase configuration and minimize the {XP,YP} as well as {X,Y} merger (which are dispreferred from the standpoint of labeling) make the latter ones fragile and prone to loss.

I extend this analysis to traditional grammaticalization, but also additional phenomena, e.g., change from OV to VO word order and the loss of traditional rightward adjunction.

I also investigate specifiers which are more resistant to diachronic change, in particular cases involving multiple movement as with multiple *wh*-fronting, and show that in such cases the loss of movement goes through a single *wh*-movement stage. I also explore the motivation for the existence of movement in general, discussing its semantic and interface-based triggers.

Additionally, I propose an account of V2 where V2 involves two distinct configurations with distinct syntactic mechanisms and licensing conditions, with only one of them being subject to diachronic loss.

I also explore the connection between historical change and language acquisition by investigating acquisitional errors of omission in the acquisition of reflexive clitics in Polish. I confirm the connection between acquisition and diachronic change by the history of SE-reflexives in Russian, as well as a broader pattern of acquisition with both monolingual and bilingual children who are exposed to both wh-movement and *wh*-in-situ in their input, as well as mixed language varieties that are based on conflicting (movement vs *in-situ*) languages and show that all these phenomena provide support for the labeling-based structural preferences argued for in the thesis.

Head Labeling Preference and Language Change

Marcin Dadan

B.A., University of Łódź, 1999

M.A., University of Delhi, 2010

M.Phil., University of Delhi, 2013

M.A., University of Connecticut, 2016

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2019

APROVAL PAGE

Doctor of Philosophy Dissertation

Head Labeling Preference and Language Change

Presented by

Marcin Dadan, B.A, M.A., M.Phil.

Major Advisor _____

Željko Bošković

Associate Advisor _____

Ian Roberts

Associate Advisor _____

Mamoru Saito

University of Connecticut

2019

ACKNOWLEDGMENTS

First and foremost, I want to thank Željko Bošković for his generosity with time and patience. This dissertation (and so many other things on this journey) would have never happened if it wasn't for meeting Željko during one LISSIM summer school in India where he came to teach NP/DP. I knew I wanted to be close to this larger-than-life character to see what he could see. From then on everything was like this first drive in his BMW during the open house when he had no glasses, asking me to pay attention to the street names so we can reach Andrea's place because he couldn't see well without them; complaining that his BMW is not fast enough. And I stayed. I stayed for the Tuesday seminars of comparative syntax and appointments till late nights, where Željko could tell me exactly what was going on in the data, or why 'that wouldn't work!', telling me what to check and where to look just to find if he was right. He was right. Always with the same excitement and patience.' Write it down for yourself!' And I wrote. Finally, also this thesis.

It would be impossible without other committee members. First of all, Ian Roberts really inspired me not to be afraid of historical syntax and bold ideas. His knowledge, class, and kindness can only be matched by his imagination and work ethos. I was so extremely fortunate that he could be in the right place at the right time to trigger this diachronic spark in me. That changed all my approach to studying language and seeing myself as a linguist. I am extremely grateful for that.

Mamoru Saito will always be my role model when it comes to courage and creativity. I especially admire how open-minded, genuinely curious, and brave when it comes to exploring new ideas and theoretical tools he is. All this paired with the sweetest personality you could ever imagine makes him an amazing advisor, mentor, and a teacher. Not to mention his handwriting, which is really accurate reflection of the way he is (have you noticed how I succeeded not mentioning Željko's handwriting? Well, almost…). I learned so much from his lectures, not only in terms of syntax, but also how to present the material, approach the students and their questions; finally, how to listen to their comments. All this with such an enormous amount of respect. I hope I can be like him one day.

Jonathan Bobaljik and Susi Wurmbrand had a huge influence on my linguistic growing up and thinking about science, in general. I will always admire their intelligence and rigor in pursuit of truth. Their work is extremely important and of such high standards then I could never ask to have better committee members on my Generals papers, as well as teachers. I learned so much from working with them, both by their classes (Jonathan's 'Readings and Research' is a great example of how to teach one of the most important courses for budding syntacticians; Susi's 'Problems in Syntax' seminar was always a chance to learn how to diagnose the structure and not to believe in easy but shallow solutions), and by meeting them to discuss my work. I had to admit that appointments were extremely demanding but I never left without the inspiration or some extremely important lesson. Susi is my unattainable inspiration in one more domain: how to be well organized. The organization of her classes, handouts, slides, and talks, is indeed 'the next-level' and I catch myself thinking 'how she would do it?' while preparing my materials. Having

all the tasks precisely laid down and thought-over with doable timeline made my work for her as an RA extremely smooth and anxiety-free. It really helps not only her but everybody around her. This is another skill I really want to be good at one day. Jonathan and Susi offered me a real linguistic training and I am extremely thankful and feel privileged I could be at UConn when they were there.

UConn has so many amazing professors that I would need a size of another thesis to mention all the things I learned from them and all the things I am grateful for. I was extremely fortunate to be working with Jon Sprouse as his teaching assistant. It gave me a chance to see what a smart, precise, and hard-working professor can do with an introductory linguistic course like 'Ling1010.' Students really love him and his way of designing and presenting the course made my work incredibly easy and rewarding. He is one of the most fair and approachable teachers any of these students (this includes also us, graduate students) will ever have. And they know it. But most importantly, Jon is an amazing scientist whose standards are incredibly high. His comments are also always very practical, showing his insight and common sense. I was really happy to have him as a reviewer of my prospectus and what he said helped me not to lose the big picture and remember about things that matter (even though they don't bear on any particular theoretical account). I also need to thank William Snyder, with whom I spent a lot of time working on my Generals paper on acquisition of SE-reflexives, part of which study found its way to this dissertation. William was always helpful and willing to talk, and his expertise in statistics and inside-out knowledge of, not only the field of language acquisition, but also current syntactic theory, as well as general knowledge of languages, contributed enormously to my work and general knowledge. Same goes for another UConn acquisitionist, Diane Lillo-Martin, who always asks the right questions that go straight into the heart of the matter; hence they always help to see where I am and what is still missing. I want to also thank to Jon Gajewski, whose semantics class was really accessible, and just at the right pace for us at the beginning of our semantic explorations. Meetings with Jon were always fruitful for me as his grasp on semantic theory always helped me to find solutions or evaluate my proposals. Jon's help as a department head also saved my first job, as he was willing to go out of his way to make it happen despite the formal obstacles. I can't omit Andrea Calabrese here who is one of the most knowledgeable and friendly people I ever met, versed in historical linguistics, the classics, morpho-phonology and who is an extremely gifted and passionate cook. The open-house dinners will always be a stuff of legends due to his hospitality and delicious food. Finally, Magdalena and Stefan Kaufmann are great semanticists who are always interested in the work of the students and always ready to help suggesting some literature or offering explanation whenever we were lost (they also saved some of our early semantics-related projects, e.g., our joint work on temporal adverbial clauses). It is also good to have Harry van der Hulst in the department with his broad interests in everything linguistic-related. Additionally, interactions with Jairo Nunes and Eva Bar-Shalom were always interesting and stimulating. Seriously, I feel my time at UConn has had the best faculty anyone could wish for and I am extremely indebted to so many of them. Special regards should be given to Catalina Ritton and Tamara Cohen, both of whom always offered incredible help and support in all formal paperwork.

UConn is a place with great students. Here I met my best friends and collaborators. I will start with my cohort: Renato Lacerda, Yuta Sakamoto, Emma Yang, and Abby Thornton. We really learned the linguistics and the art of surviving in Storrs/Willimantic together, and I felt we made a great team, even though we were so different. I am sure we will remain friends for life and such visits as Yuta's as my dissertation defense will be a common thing regardless of how far we are from each other (Renato is next to defend. I cannot miss it). Next I want to mention my office mates from William's lab (374), btw, I am extremely grateful to him for allowing me to work there in peace and quiet on my Generals paper. I liked it so much that I decided to never leave; similarly, many others and soon the office became bustling with life and conversations. Those who were there with me were: Emma Nguyen (sometimes I think she keeps the linguistics department together), Yoshiki Fujiwara, Pietro Cerrone, Roberto Petrosino, Sabine Laszakovits, and Shuyan Wang. I really had fun with you guys and staying with you was one of the best decisions. I also visited 'Japanese' office quite often, where I could spend hours talking to Koji Shimamura (especially enjoyed drinking and talking music and syntax to Koji), Hiroaki Saito, Ryosuke Hattori, Yuta Sakamoto was always there and obviously Jayeon Park and Yongsuk Yoo. Those last were also my roommates and my best friends and we shared most of the time together: eating, drinking, travelling, watching movies, and yes, also working on joint papers. Other residents of 560 included Adrian Stegovec, Yuta Tatsumi, Hiro (whom I mentioned earlier, but who spent a lot of time in our apartment), Sabine, Brandon Sugrue, and Yuya Noguchi. I need to also mention Koichi Yoshikawa, who frequented linguistics events so much that everybody was surprised that he actually was from a different department;) I also have to include here Zheng Shen. Zhengy always impressed me with his good heart (anyone could really depend on him) and very practical approach (after all, he is the one who gave us lingalert). Other UConn students I want to mention are: Gisli Harðarson, Aida Talić, Neda Todorović, Ting Xu, Irina Monich, Alex Vaxman, Jungmin Kang, Julio Villa-Garcia, Paul Smith, Beata Moskal, Vanessa Petroj (these guys were there when I arrived and made me feel like home here; Gisli literally gave me home for the first days), and next generations of students: Karina Bertolino,Paula Fenger, Ksenia Bogomolets, Christos Christopulos, Pasha Koval, Shengyun Gu, Gabriel Martinez-Vera, Hiromune Oda, Araki Arano, Ivana Jovović, Sarah Asinari, Nick Schrum, Muyi Yang, Si Kai Lee, Teruyuki Mizuno. Also special shout outs go to Kadir Gökgöz, who became a collaborator in our tense/aspect project. My apologies if I am missing somebody here.

My linguistic life really started at the University of Delhi and at JNU, where I was a frequent visitor, and I will always have enormous intellectual debt to all those who introduced me to this discipline. Hence, I will always be grateful to Tanmoy Bhattacharya, Ramakant Agnihotri, Ayesha Kidwai, Tista Bagchi, Vineeta Dayal, Gail Coelho, Shobha Satyanath, Pradeep Kumar Das, and many others for their assistance in my first linguistic steps. Especially thank you to Tanmoy and Ayesha for wonderful and truly exceptional LISSIM summer schools where we had a chance to learn and interact with the biggest names in the field. I owe you both much more than I can express here.

Thank you to all the teachers and students both in India, Europe, and USA that I had a chance to meet and interact with at the conferences and summer school. We are all part of an amazing community and hear about each other's work is truly inspirational and it keeps us interested in language. Special hello to Krzysztof Midgalski, whose work I truly admire and with whom we share Poland-Storrs connections.

Finally, I would like to thank to my family for their patience for my neverending studies and who were bearing with my absence for so long.

September 2019, Gainesville, FL.

CONTENTS

Chapter 1

Introduction

1.1. The scope and the main goals

This dissertation examines a number of cross-linguistic instances of diachronic syntactic change which despite the superficial appearances of being unrelated to each other will be shown to reval an important common thread. The starting point will be a new observation established based on a wide range of languages regarding wh-dependencies, which will be shown to exhibit unidirectionality in language change: the change is always from wh-fronting to wh-in-situ. Furthermore, with languages that display multiple wh-fronting, the loss of wh-movement first goes through a single wh-fronting change. Taking these observations as the starting point, the dissertation examines a number of additional diachronic changes, all of which will be shown to involve the loss of obligatory syntactic movement, both phrasal movement and head-movement.

I will then offer a unified explanation for all these historical changes based on the preference for a particular structural configuration, namely a head-phrase merger {X,YP}, from the standpoint of labeling (Chomsky 2013, 2015). Importantly, while the observation of the loss of syntactic movement was already reported for some phenomena, and in some cases only for particular languages (the details and references below), this dissertation goes further: by unifying a number of such cases and deducing them from one theoretical standpoint, it shows that the fact of the loss of obligatory syntactic movement is not just a peculiar property of particular languages or tied to specific constructions, but rather part of a broader process with its own constraints, which allows us to explain many attested facts and to some extent also to make diachronic predictions. The proposed analysis will, however, raise a number of theoretical and empirical questions, which will

be addressed in the discussion to follow. I will start the discussion by briefly examining some of the phenomena that will be discussed later in the dissertation.

1.1.1. Wh-movement

One of the phenomena that will be discussed in the thesis is *wh*-movement involved in question formation. Synchronically, languages differ with respect to the behavior of *wh*-phrases in interrogative clauses. Some languages, like English, obligatorily displace a *wh*-phrase in questions (1a). English is a *Single-Wh-Movement* language: in multiple *wh*-questions, i.e., questions with more than one *wh*-phrase, only one wh-phrase (namely the highest wh-phrase prior to the movement) moves, with the remaining wh-phrases staying in their base-generated positions (1b vs. 1c)).

(1) a. What did John give to Mary

b. What did John give to whom?

c. *What to whom did John give?

Cheng (1991) sees *wh*-phrases in English as involving quantificational material, which moves not only to bind the variable as a quantificational operator (and restricting the set of possible answers to the question that would make the proposition true), but also to type the clause as a question (i.e. to indicate the Force of the clause) in the first place, as a language like English lacks question particles that could otherwise perform this function.

The counterpart of English (1d) is grammatical in *Multiple-Wh-Fronting* (MWF) languages. Such languages, like Polish (2) or Bulgarian, must move all their *wh*-phrases.

(2) a. Co Jan dał Marii? [Polish]

what John gave Mary

'What did John give to Mary?'

b. *Co Jan dał komu?

what John gave whom

c. Co komu Jan dał?

what whom John gave

'What did John give to whom?'

Thus, MWF languages are characterized by obligatory displacement of all *wh*-phrases.

Cheng (1991) argues that *wh*-phrases in MWF languages are indefinites devoid of any quantificational force. For instance, in Polish non-interrogative contexts, the indefinite core (the same as in *wh*-phrases) combines with the affix *-ś* that provides it with existential closure (e.g., in *kto-ś* 'somebody', where *kto* is 'who'). Polish *wh*-phrases do not combine with a quantificational affix at the word-level (in the lexical structure) the way general indefinites do in Polish (so no affix is added to the wh-phrase), hence in order to serve as a quantificational operator that binds a variable (which is necessary to obtain the question interpretation) they need to move. The movement in question also serves the clause typing function as (at least) one *wh*-phrase must target the left-periphery where it finds itself in the specifier position of the relevant phrase that is responsible for typing the clause as interrogative, i.e. for encoding the Force of the utterance.

Finally, the third (broad) type of languages in terms of the behavior of *wh*-phrases in questions are *wh-in-situ* languages, where *wh*-words do not undergo obligatory displacement with the question words remaining in their base-generated positions. Languages like Japanese and

Mandarin Chinese are a good illustration of this kind of strategy for question formation, as illustrated below.

(3) a. John-ga **hon-o** katta. [Japanese]
John-NOM book-ACC bought
'John bought a book'
b. John-ga **nani-o** kaimasita ka?
John-NOM what-ACC bought.POL Q
'What did John buy?

(4) a. ta xihuan **ni** [Mandarin Chinese]
he likes you
'He likes you'
b. ta xihuan **shei** (ne)?
he likes who Q
'Who does he like?'

Wh-words in these *wh-in-situ* languages are also taken to be indefinites rather than quantificational elements (Cheng 1991, Hagstrom 1998, Aldridge 2009). The source of the quantification in interrogatives is provided by question particles like *ka* in Japanese or *ne* in Mandarin Chinese which act as unselective binders (or quantify over choice functions), thus providing the *wh*-phrases with the required wide scope in a similar fashion to what happens with wide scope taking indefinites in declarative clauses (Reinhart 1998, Hagstrom 1998). The presence of such particles has also been argued by Cheng (1991) to play a role in the lack of obligatory syntactic *wh*-

movement in these languages: the clause can be typed without the necessity of syntactic displacement of *wh*-phrases.

What is especially interesting in the context of different types of languages with regards to the syntactic behavior of *wh*-phrases in question formation is that Japanese and Mandarin Chinese used to have obligatory *wh*-movement. Old Japanese in 8th century CED and Archaic Chinese around 5th-3rd century BCE had productive and obligatory *wh*-fronting and lost it at some point in their history (see also Ogawa 1976, Whitman 2001, Watanabe 2002, Kuroda 2007, Aldridge 2009, 2010, 2011, 2018). Consider examples below from Old Japanese and Archaic Chinese illustrating the fact that both of these languages were *wh*-movement languages (Japanese even being a multiple-*wh*-movement language of the Polish type).

(5) **Nani-wo-ka-mo** mikari-no hito-no ori-te kazasa-mu ? [Old Japanese]
what-ACC-KA-FOC hike-GEN person-GEN pick-CONJ wear.on.the.hair-will
'What should hikers pick and wear on their hair?' Aldridge 2018: 6; MYS 1974)

(6) **He cheng** bu ke? [Archaic Chinese]
what city not conquer
'What city would you not conquer' (Zouzhang, Xi 4; Aldridge 2010:6)

I will show in this dissertation that this is not a peculiarity of Japanese and Chinese but part of a broader pattern. Importantly, while many other languages underwent the same diachronic process as Japanese and Chinese we will see that there is no change in the opposite direction, from wh-in-situ to wh-fronting.

Consider now multiple wh-fronting. Latin actually used to be a multiple *wh-fronting* language (see also Danckaert 2012). Modern Romance languages are not multiple wh-fronting languages[1] and in fact they also allow *wh-in-situ.* A diachronic change took place from Latin to Modern Romance languages with regards to *wh*-fronting in questions that can be described as a gradual loss of movement. Looking at Latin and the way the obligatory *wh*-fronting was lost, we can see that the loss of movement proceeds in stages; *wh*-fronting does not disappear in one go. Compare Latin multiple *wh*-questions in (7) with the same construction in modern Romance (8):

(7) **Quis** **apud quos** **quibus** praesentibus sit acturus [Latin]
who.NOM with who.ACC who.DAT is.present be do.FUT
'Who is going to plead before whom in whose presence?' (Quintilian 11.3.150; Devine and Stephens 2006: 89)

(8) a. *¿ **A quién qué** le pidió Ivan? [Spanish]
to who what CL asked Ivan
b. ¿ **A quién** le pidió Ivan **qué** ?
to who CL asked Ivan what
c. ¿ Ivan le pidió **qué** **a quién**?
Ivan CL asked what to who?
'What did Ivan ask **to whom**? (Reglero &Ticio 2013: 537)

[1] One exception here is Romanian, which I will ignore below.

Sentence (7) shows that Latin was a multiple *wh*-fronting language and all its *wh*-words had to be fronted (Spevak 2010, Danckaert 2012, Ledgeway 2012). This is certainly not true for Spanish, where multiple *wh*-fronting leads to ungrammaticality and the language has to resort either to English-style single *wh*-fronting, or modern Japanese-style *wh-in situ*. This shows that the loss of movement proceeds in smaller steps, i.e., from multiple *wh*-fronting, through single *wh*-fronting, towards a total lack of fronting.

Additionally, the process of losing movement is again unidirectional. It proceeds from losing multiple fronting towards the loss of single fronting and not the other way around.

This striking unidirectionality will be a characteristic of all the diachronic changes where movement is involved discussed in the present work. The analysis developed in this work will provide an account for this unidirectionality, unifying the previously unnoticed fact of cross-linguistic loss of *wh*-movement with a number of other diachronic phenomena. I will show that the loss of *wh*-movement is closely related to the loss of specifiers which are created due to this movement. However, this kind of unidirectionality with the loss of movement is not found only with phrasal movement to specifier positions (as with wh-movement), but also with head-movement. Head-movement has also been observed to show a tendency to be lost in the history of many languages. I will focus on V-to-T movement here.

1.1.2. The Loss of Head-Movement

Historically, Old English and Old Swedish were among languages that had a productive and obligatory V-to-T raising. This is no longer the case and both English and Scandinavian do not raise the lexical verbs to T anymore.

(9) hu se deofol **beswac** syððan eft þa men [Old English]
how the devil deceived afterwards again the men
'how the devil then deceived mankind again' (colwgeat, + ALet_6_[Wulfgeat]:59.22; H&I 2016: 504)

(10) æn min guÞ **brytar** eigh niÞar Þin guÞ [Old Swedish]
if my god breaks not down your god
'If my god doesn't break down your god' (*Codez Bureanus*, c.1350; Heycock & Wallenberg 2013:129)

I will show that in spite of its reliance on the loss of specifiers, the analysis that will be proposed for the loss of wh-movement can actually be extended to this loss of head-movement.

The V-to-T movement was lost in Old Swedish around the same time as this language lost its rich agreement, i.e., in the 17th century. Similarly, English lost its verb movement to T around 1600 CE, i.e., almost in parallel with its verbal morphology (Roberts 1993, Rohrbacher 1994/1999, Haeberli & Ihsane 2016). This indicates that agreement plays a significant role in the diachronic change in question. We will see that the proposed analysis can capture this state of affairs, providing a new perspective on it.

1.1.3. Diachrony, Synchrony, and Types of the Verb-Second Constraint

Another phenomenon that will be discuss in the thesis is Verb-Second (V2). Broadly speaking, the V2 property is a restriction on the placement of the verb in the clause. Namely, in some languages the finite verb has to appear in the second position, where in most cases it is preceded by only one constituent. Therefore, satisfaction of the V2 constraint often requires not one but two movements:

raising of the verb to some higher position in the left-periphery and movement of some XP which precedes the raised verb (the constituent preceding the finite verb after the movement is also referred to as 'prefield', or 'first constituent' for this reason). This is illustrated in (11) from Swedish, where the moved object serves as the prefield constituent for the raised verb:

(11) *Tidningar* **läser** barnen inte. [Swedish]
newspaper read the-children not
'Newspapers, the children don't read.' (Holmberg 2010: 7)

V2, as many other types of movement, has been a subject to diachronic loss and some languages lost this configuration either entirely, or in most syntactic contexts. This has been attested in the history of the Romance languages (Roberts 1993, Vance 1995, 1997, Ledgeway 2008, 2017, Poletto 2013, 2014, Wolfe 2016, 2019), Celtic languages (Willis 1998, 2007), as well as in English (van Kemenade 1987, Kiparsky 1995, Kroch & Taylor 1997, Pintzuk 1991, Fischer et al. 2004, Haeberli and Ihsane 2016, Roberts 2017).

Sentences below illustrate historical V2 constructions in Old French (12) and Old English (13).

(12) Sor ceste poere **edefierai** je m'eglise
on this rock will-build I my-church (Old French; *Queste del Saint Graal (Q)* 101,31; Vance 1995:174)

(13) *þa* **wæs** þæt folc þæs micclan welan ungemetlice brucende…
then was the people the great prosperity excessively partaking
'Then the people were partaking excessively of the great prosperity.' (Or 1.23.3: Fischer et al. 2004:106)

The loss of Verb-Second (V2) in cases like above can be treated as involving both the loss of movement to a specifier position and movement to a head position. The analysis that will be proposed in this work for the loss of wh-movement and V-to-T movement will in fact be extendable to the loss of V2. There is, however, much more to be said here.

Cases of the loss of V2 are attested and we have just seen some examples of this. However, despite being very rare typologically, V2 actually also tends to be diachronically quite persistent (regardless of the fact that some language indeed lost it). This apparent paradox will be explained in the present work along the lines sketched briefly below.

Closer scrutiny of V2 languages and V2 contexts reveals that V2 is definitely a multifaceted phenomenon. Analyzing the semantics associated with V2 contexts will allow us to postulate at least two distinct V2 configurations. One of them is discourse related (I will refer to it as V2D), where the Verb Second configuration is used to express topic or focus. This kind of V2 can be seen in the examples above from Old French and Old English (12-13). The other type of V2 is connected to the Illocutionary force of the utterance (referred to as V2I), where V2 appears to express the semantics of assertion or question (Hooper & Thompson 1976, Wechsler 1990, Brandner 2006, Truckenbrodt 2006, Wilkund et al. 2009, Julien 2015). In this latter type of V2 (V2I), the placement of the verb in the second position can be related to the clause-typing requirement of Cheng (1991). This type of V2 is illustrated by the following example from modern English:

(14) What has Mary bought?

I will argue that it is a particular type of V2 that relates to its potential to be diachronically stable or not. English is in fact a great illustration of this correlation. English has lost its V2 in contexts of V2D; e.g. today, English does not have Verb Second with topicalization (15a-b). However, English still has, albeit in a reduced form, a 'residual' V2, as illustrated in (14) with an interrogative. Hence, the V2 that is still to some extent operative in English is the illocutionary-force related V2 that will be argued in the present work to be diachronically more stable.

(15) a. *That book, has Mary bought.

b. That book, Mary has bought.

I will show how the licensing conditions of each of these types of V2 vary and how the system proposed in the present work accounts for the diachronic situation of each type of V2, hence resolving the apparently paradoxical status of the V2 constraint cross-linguistically that was noted above. I will also discuss the role of prosody in V2 (discussed e.g., in Boeckx 1998, Rice & Svenonius 1998, Bošković 2001, to appear(a), Jouitteau 2018), which I will argue also strongly plays a role in the potential demise or thriving of this constraint in a particular language.

1.1.4. Loss of Movement and Grammaticalization Processes

Some historical changes discussed in the thesis involve what is traditionally referred to as grammaticalization (Lehmann 1995, Harris & Campbell, Lyons 1999, 1995, Heine & Kuteva 2002, Roberts and Roussou 2003). Such changes are often characterized in the literature by a reanalysis of a formerly lexical element as a functional element, very often base-generated in the position that involved movement in the previous stage of the language. Some changes of this sort include reanalyses of *wh*-phrases, relative pronouns, and demonstrative pronouns as complementizers,

which is quite common crosslingusitically. I will show here examples from Germanic (English (16)), Slavic (Old Church Slavonic (17a), Polish (17b)), and Kartvelian (Georgian (18)). These cases are often described as involving a reanalysis of a syntactic specifier as a head, e.g., a change from [Spec, CP] to C (van Gelderen 2004, 2009, 2011, 2015, Willis 2007).

English (16) illustrates this with the *wh*-phrase *how* being reanalyzed as a declarative complementizer (see also van Gelderen 2009, 2015).

(16) Bob Cratchit told them **how** he had a situation in his eye for Master Peter. (English; Willis 2007:434)

Willis (2007) points out that *how* has split into two elements, a *wh*-phrase and an innovative complementizer, both existing side by side. Huddleston & Pullum (2002) show that the origin of the English complementizer *how* is still quite transparent as it is found only as a complement of predicates that allow both interrogative and declarative clauses (e.g., a sentence like **I believed how his cattle were being rustled* does not allow *how* in the function of a complementizer.

Complementizers may also have their origins in relative pronouns (Harris & Cambell 1995, Roberts and Roussou 2003, van Gelderen 2004, 2009). This change happened in Old English, but it is also common in other languages. Thus, Slavic relative pronoun *jen-ż/ž(e)* changed into a relative particle *jiżto/jeżto* (in Polish, around 15th century) later to function as a complementizer *że* (Meyer 2017). This latest change is illustrated with examples from Old Church Slavonic (17a) and Polish (17b-c) (Polish can still use both modern *że* and more archaic *iż* as complementizers 'that').

(17) a. se že reče o dsě **eže** xotěaxǫ prijeti
this PTCL say.AOR.3SG about spirit.M.PRP that.N.ACC shall.3PL receive
i verujǫšti
him.M.ACC believing.PL
'This he said about the spirit which those who believed in him should receive.' (Old Church Slavonic; Meyer, 2017:100)

b. W ZUS nie ukrywają **że** lekarzom trudno udowodnić,
in ZUS not hide that doctors.DAT hard to.convict
iż nadużywają swych kompetencji.
that abuse.3.PL their powers
'In the Social Security Agency, they don't hide it that it's hard to convict the doctors of abusing their powers.' (Polish; *NKJP)*

c. Gdzie by szatan chciał mieszkać w domu$_i$, **co** go$_i$ mamusia […] wybudowała?
Where SUBJ satan wanted to.live in house what it mother built
'Why would satan want to live in the house that was built by the mother?'(Polish; NKJP)

Polish (17c), involving 'what' (Polish *co*), shows grammatizalization of *wh*-phrases as complementizers that is particularly common.

Another case of that sort where the wh-phrase 'what' got grammaticalized into a complementizer is illustrated by example (18) from Georgian, where complementizer *raytamca* has its source in *wh*-expression *ray* 'what'.

(18) da ara unda, **raytamca** icna vin
and NEG want.PRES.3SG that know.PRES.3SG someone
'and he did not want that anyone know' (Georgian; Harris & Campbell, 1995: 298)

All the changes shown above involve a loss of movement. In particular, they involve a reanalysis of a phrasal element from the Spec position of a projection which it originally targeted with movement to the head of the projection in question. As a result, these cases are often described as a 'Spec-to-head' reanalysis and explained as an economy-driven preference for heads rather than phrases. This family of accounts can be found in Roberts & Roussou (2003) or van Gelderen (2004). Neither of these accounts, however, can be applied to unify these traditional grammaticalization cases with the loss of *wh*-movement noted in section 1.1.1. In the latter case, movement to Spec,CP got lost but without a reanalysis of the specifier: *wh*-phrases that used to move simply stay in their base-generated position. Hence the loss of *wh*-movement as in Japanese or Chinese discussed earlier does not involve a reduction of *wh*-phrases to heads; instead, *wh*-phrases preserve their phrasal status in spite of the loss of movement. While the Spec-to-head analyses of the loss of movement in the traditional grammaticalization cases cannot be extended to the loss of wh-movement of the kind discussed in section 1.1.1, we will see that the account of the latter proposed in chapter 2 of the present thesis can be easily extended to the traditional grammaticalization cases, thus unifying all these cases in a systematic way.

The analysis tackling the loss of obligatory syntactic movement allows us to account for additional historical facts, such as certain changes in the word order and the loss of rightward adjunction.

1.1.5. Word Order Change

Diachronically, many head-final languages become head-initial; hence languages undergo change from OV (head-final) to VO (head-initial) word order. A change of this sort has been attested in several branches of the Indo-European family (e.g., in Romance and Germanic), as well as in Finno-Ugric.

English in fact provides an illustration of this change, as Old English word order involved objects predominantly preceding the verb (19) (van Kemenade 1987, Pintzuk 1991, Roberts 1997, 2017)). Modern English, on the other hand, is a strict VO language.

(19) Ac he sceal Þa acfullan **gesibbian**
but he must the quarrelsome reconcile
'But he must reconcile the quarrelsome.' (Old English; Pintzuk & Taylor 2006:249)

Kiparsky (1996) discusses several types of analysis for this historical OV>VO change offered in literature, which involve: (1) hypothesis of language contact being responsible for the change, (2) switch to head-initial VP as a consequence of the loss of inflectional morphology (VO being the word order which is best suited for encoding subject vs. object relation in morphologically poor systems, as proposed e.g., by Sapir 1921), and (3) reanalysis of the derived VO sequences as basic/underlying word order when the evidence for the OV order has become opaque due to the verb fronting in the embedded clause. From the current perspective, one proposal is especially interesting and relevant to the discussion here, namely Kayne's (1994) proposal as to the universality of the VO word order crosslinguistically. In Kayne's (1994) system, all languages have the underlying VO word order, with the OV word order being derived via movement.

Adopting Kayne's system makes it possible to recast the OV to VO word order change in terms of movement, more precisely the loss of movement. Importantly, we will see that the diachronic change from the OV to the VO word order seems to be unidirectional (see e.g. Kiparsky 1996, Roberts 2017b), which bears resemblance to the unidirectionality of the loss of movement. This will then be another case that the analysis argued for in the thesis will be extended to.

1.1.6. The Loss of Rightward Adjunction

Another phenomenon that will be discussed and accounted for in the thesis concerns an interesting change found in Chinese involving adjunction structures, in particular what is traditionally assumed to be rightward adjunction. Pre-Archaic Chinese allowed such adjunction. Importantly, such adjunction is no longer allowed in modern Chinese. Compare Pre-Archaic Chinese with modern Chinese below:

(20) a. wǒ fēi ài qí cái èr yí zhī **[yí yáng]** yě
1SG NEG cherish 3SG value CONJ replace 3SG with sheep PART
'It is not that I attach a great importance to its value [i.e. the value of the ox] and therefore replaced it with sheep.' (Pre-Archaic Chinese; *Mengzi*, Liang hui wang, 4th-3rd c. BC; Djamouri et al. 2013: 586)

b. a. tā yě /měi-tiān / chángcháng lái **(* yě/ *měi-tiān / *chángcháng)**
3SG also/ every-day / often come (also/ every-day / often)
'He also comes every day/often.' (Modern Chinese; Djamouri et al. 2013: 590)

Sentence (20a) from Pre-Archaic Chinese involves post verbal adjunct PP ‘with sheep’. (20b) on the other hand illustrates that modern Chinese does not allow such adjunction anymore. Adjuncts in modern Chinese can only appear pre-verbally, in contrast to the 4th-century BC Chinese, which allowed adjuncts both before and after the verb.

We will see that the analysis developed to account for the diachronic changes discussed in the preceding sections can also be extended to account for the change regarding the direction of adjunction that occurred from pre-archaic to modern Chinese, noted above.

1.1.7. Diachronic Change and Language Acquisition

There is a rich tradition connecting historical changes with language acquisition. In fact, a number of authors have argued that language acquisition is one of the most important forces driving language change (Paul 1880, Lightfoot 1979, 1991, 1995, Faarlund 1990, Clark and Roberts 1993, Yang 2002, van Kemenade 2007, Roberts 2007, Walkden 2014). This work will also explore the connections between language change and language acquisition.

In particular, to this end, I will offer a case study of acquisition of reflexive clitics of the type shown in (21) below in Polish.

(21) Śniło mi **się** że latam.
dreamed me.DAT SE that flying.1SG.
‘I had a dream that I was flying.’

Polish is one of many languages that uses this kind of *se*-clitics in a variety of contexts, including the non-reflexive usage shown in (21). Contexts of use of such constructions involve ‘self-grooming verbs such as ‘wash’ *myć się* or ‘shave’ *golić się*, and reciprocal contexts, but also

middle voice, passive voice, impersonals, anticausatives, or resultatives. The function of such clitics is the subject of vast literature, with the proposals emphasizing their roles in (among other functions) voice and argument structure alternations (Babby & Brecht 1975, Kayne 1975, 1988, Bouchard 1983, Marantz 1984, Manzini 1986, Pesetsky 1995, Reuland 2001, Embick 2004, Schäfer 2008, Medova 2009, Sportiche 2014, Wood 2015, among others). Languages have also a number of constructions with 'inherent' reflexives (or *reflexiva tantum)* where the predicate has to be marked by SE although there is no obvious relation to voice or argument structure alternation (this is our (23) above) .

The case study presented in the thesis reveals that acquisition of SE-reflexives involves frequent errors of omission of these elements by children (Dadan 2015; see also Rivero and Golędzinowska 2002 for Polish, and Zombolou & Alexiadou 2012 for German). I will show that the errors of omission of these clitics can be related to the diachronic changes discussed earlier, in fact they are all part of the same broader pattern. I will thus present a unified account of all of these changes.

The connection between language acquisition and diachronic change will be confirmed by an actual historical change that took place with Russian SE-reflexives. Polish children omit this reflexive clitic during acquisition; in Russian, on the other hand, this element changed its status diachronically, i.e., it underwent a very interesting structural reanalysis. In particular, while the reflexive clitic SE in Old Russian functioned in basically the same way as in Polish, in modern Rusian it lost the status of an independent element and got reanalyzed as part of the verb (a verbal suffix). I will show how this change in Russian relates to Polish acquisitional errors and will in fact unify these two phenomena in one analysis.

I will also look at other facts from language acquisition and show how the historical facts discussed in the thesis connect to these acquisitional processes. In this context, the loss of *wh*-movement discussed earlier will be reflected in the actual acquisitional facts. Namely, we will see that when facing an ambiguous input, children prefer to avoid *wh*-movement and use *wh-in-situ* in the initial phases of the acquisition. The test case here will be languages that allow both wh-movement and wh-in-situ, like French or Brazilian Portuguese (Hamann 2000, Zuckerman 2001, Plunket & DeCat 2001, Lessa-de-Oliveira 2003). However, acquisition of wh-questions in English will also be discussed. I will also look at bilingual acquisition and adult contact languages. All these cases will be shown to be amenable to a unified account.

As noted in section 1.1., the analysis I will offer will be based on the theory of labeling, which is a mechanism that specifies the nature of more complex syntactic objects. In particular, I will appeal to the version of labeling in Chomsky (2013, 2015). Since this theory is crucial for the account that will be argued for in this thesis, next section will present the relevant background on labeling.

1.2. Labeling

The notion of labeling, i.e. specifying the nature of more complex objects formed during the syntactic derivation, has been present in generative grammar in one way or another since its inception. In Government and Binding (GB), labels were an essential part of structure building, with hierarchical structures built according to a recursive X'-schema. For that purpose, the notions of heads, complements, and specifiers had to be clearly distinguished and built into the universal X'-structure. Overall, the structure building process was quite transparent, with each phrase built according to the same specifications in a top-down procedure with the maximal projection (XP),

intermediate projection (X'), and the head (X), the identity of which was also clearly specified. Hence labels were required to build a hierarchical, recursive structure, and this requirement came with the necessary endocentricity for phrases: each phrase had to have a head with no ambiguity in terms of labeling of the structure.

Early Minimalism (Chomsky 1995) assumed labeling to be a part of the definition of the structure building operation Merge. This binary concatenative procedure selects two syntactic objects, α and β, and forms a set K (another syntactic object) with its label γ.

(22) K= $\{\gamma\ \{\alpha, \beta\}\}$

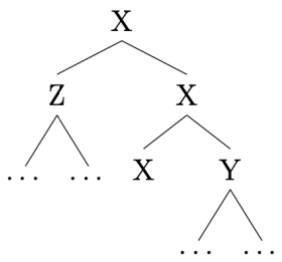
The label of K is determined as soon as the relevant object is built. This label is provided by either α or β, i.e., the identity of the resulting structure is determined by one of the elements in the set. If it is the case that α projects, the resulting structure is likewise labeled by it; its category is determined to be the category of this whole $\{\alpha, \beta\}$ merger. This captures the intuition that when e.g., a verb merges with its complement, the verb projects, creating an object with syntactic properties of the verb phrase; hence it is labeled as such. This is illustrated below: (23a) is an abstract representation of the set involved in a merge, with one of its elements labeling the whole structure; (23b) shows a concrete example of (what would be a traditional) VP: it is labeled by the unaccusative verb 'arrive' which is merged with its single Theme argument. It is the label of the verb 'arrive' that is chosen (and replicated in (23b)) to serve as the identity of this syntactic object.

(23) a. K= $\{\ \alpha\ \{\alpha, \beta\}\}$

b. K= {arrive {arrive, John}}

The element chosen for the label is the head of the projected constituent. As this system dispenses with the X'-based definitions of heads, specifiers and complements, in the proposed Bare Phrase Structure such syntactic relations have to be defined contextually and relationally. A head is taken to be a terminal element drawn from the lexicon, which is minimal (non-projecting) and non-maximal (with the ability to project). The head can combine with a complement, and this head-complement relation is viewed in Bare Phrase Structure as the most local one. This is illustrated in (24), where Y is in this relation with the terminal head X that also provides the label to this structure.

(24)



The spirit of this local head-complement relation as the most basic one is carried over into later Minimalist works, but the place of labeling within this theoretical system has changed.

The desire to simplify the syntactic machinery led to treatment of labeling as independent of Merge. Merge itself remained as a structure-building mechanism, but only in its most basic form, as an operation that combines two syntactic objects. It is the simplest set-forming procedure operating in unlimited and unrestricted fashion. Merge applies freely, with its output being evaluated and sometimes filtered out at the interfaces. In this form, it is one of the few indispensable tools in the derivation.

Labels, on the other hand, have been proposed by some authors to be dispensable. Such proposals for label-free syntax were made by Collins (2002, 2014) and Seely (2006). Both authors

argue for eliminating labels, with the syntactic effects associated with labels being deduced from independently motivated syntactic principles.

The spirit of proposals like Collins (2002, 2014) and Seely (2006) is still present to some extent in the approach to labeling in Chomsky (2013, 2015), where the labels are not required in the syntax itself. Accordingly, labeling is not part of syntactic structure building, hence labeling is kept as independent from Merge. This means that syntactic structures can be unlabeled, with nothing going wrong in that case as far as the syntax proper is concerned. However, for Chomsky labels are still required at the interfaces. The function of labels is to determine the nature of the created objects, which Chomsky argues is needed for the interpretation of the structure at the interfaces. Since narrow syntax itself does not need labels, labeling can then apply after the Transfer to the interfaces.

Chomsky (2013, 2015) thus delegates labeling to a special mechanism, a Labeling Algorithm (LA), which operates at the point of spell-out. LA is based on a minimal-search operation, whose role is to find an appropriate element to provide a label for the derived object. A head is a great candidate to provide a label, as it is a lexical item (LI) that comes from Lexicon with its lexical/syntactic category clearly defined. Since the head H_0, being a lexical item (LI), can immediately provide a label for the whole structure, an object where a head is merged with a phrasal complement {H,YP} constitutes the best-case scenario for labeling, according to Chomsky. In other words, this head gets selected as the label for this object, making labeling rather trivial/straightforward in such cases. This is illustrated in (25).

(25) Head-Labeling

a. {H,YP} head-phrase merger

b. {H{H,YP}} Minimal search of LA: Label → Head H_0

With the merger of two phrasal projections {XP,YP}, it is not obvious which one should serve as (i.e. provide) a label. LA finds the respective heads of {XP,YP}, i.e. X_0 and Y_0, both of which are lexical items (LI) capable of providing the labels, but the search result is too ambiguous to determine which element will actually be taken to label this whole merger. This ambiguity can be resolved in two ways according to Chomsky (2013): either via additional operation of agreement and feature-sharing, or via movement. I will illustrate both options here.

One option to label a symmetrical {XP,YP} configuration is when XP and YP agree, i.e. when they share one prominent feature. Then such a feature can label the resulting syntactic object. *Wh*-questions like the one in (26) is a good example of labeling based on feature sharing.

(26) [$_Q$ What$_i$ [$_{CP[Q]}$ did [you say [$_{CP}$ t_i [$_C$ [$_{TP}$ John bought t_i]]]]]?

The final step of structure building in (26) involves merger of a *wh*-phrase and the matrix $CP_{[+wh]}$. This is a problematic configuration of {XP,YP}, leading to labeling ambiguity. In order to label this {DP_{WH}, CP_{WH}}merger, Labeling Algorithm looks for an available head and it finds not one but two of them, D_0 and C_0. Since it cannot be determined which one of them should be selected as the label, the LA has to look for additional means to resolve this ambiguity. In this particular configuration, the *wh*-phrase and the CP share one prominent feature, i.e., interrogative Q/wh feature. Now this feature can serve as a label for this merger and the resulting structure can be interpreted at the relevant interface.

This additional feature-searching mechanism can also apply in structures like (27) below, where the DP in the subject position finds itself in a problematic configuration {DP,TP}.

(27) [$_\varphi$ John$_i$ [$_{TP[\varphi]}$ [$_{vP}$ t_i [v bought the guitar]]]]

In (27), Labeling Algorithm is again unable to resolve the labeling ambiguity, hence the additional operation of feature sharing is needed. The DP *John* and TP share phi features under traditional Spec-head agreement. The presence of the *φ* features on both of these merged phrases enables labeling of this syntactic object as <*φ*,*φ*>. Note that for a prominent shared feature to serve as a label, the LA must be additionally set to feature-search, not simply LI/head search. (28) below shows a schematic representation of this operation:

(28) Labeling in a Phrase-Phrase configuration: Feature-Sharing

a. {XP,YP} phrase-phrase merger

b.{?{XP $_{[F]}$,YP $_{[F]}$}} Minimal search of LA: ambiguous result (X or Y)

c.{F{XP $_{[F]}$,YP $_{[F]}$}} Feature search→ Feature-sharing; Label=a prominent feature F

Another option for labeling of the {XP,YP} merger is movement. This option is illustrated below. Chomsky (2013), similarly to Bošković (2002, 2007, 2008), assumes that Successive Cyclic movement of the type illustrated in a *wh*-questions like (29) does not involve agreement in intermediate landing positions.

(29) a. What$_i$ did you say [$_?$ t$_i$ [$_{C'}$ that [$_{TP}$ John bought t$_i$]]]?

b. [$_?$ what$_i$ [$_{CP}$ that [$_{TP}$ John bought t$_i$]]]?

The *wh*-phrase *what,* despite landing in the intermediate Spec, CP position, does not enter into a feature-checking relation with the declarative complementizer *that*. This implies that the labeling by feature-sharing cannot apply in such a case. At the stage of the derivation when the symmetrical structure {DP$_{WH}$, CP} is created (as in (29b)), its label cannot be determined, and this merger is unlabeled (as indicated by '?'). Since both of the operations of Labeling Algorithm, i.e., the basic

head/LI-search and additional feature-search/feature-sharing fail in this context, another operation must apply, and this is syntactic movement; in a scenario like this, one of the phrases will have to move. Hence in (29), the *wh*-phrase moves out of this problematic configuration. As one of the ambiguity-causing constituents is removed from the problematic merger (traces, as discontinuous elements are taken to be invisible for LA), that leaves the remaining C_0 (head of the CP) as the only item that can be considered by LA to provide the label for this syntactic object. This is represented in a simplified form in (30):

(30) Labeling in a Phrase-Phrase configuration: Movement

a. {XP,YP} phrase-phrase merger

b. XP…{?{t_{XP},YP}} Minimal search of LA: ambiguous result (X or Y) followed by movement

c. XP… {Y{t_{XP},YP}} Label: Head Y (after XP has moved)

Labeling then also provides motivation for movement away from intermediate positions, i.e. from successive-cyclic movement positions. The wh-phrase has to move away from the intermediate position in (29) for labeling reasons.

Chomsky suggests that not only successive-cyclic movement, but also instances of A-movement, e.g., movement of a subject from the specifier of vP to [Spec,TP], are cases of this kind. (31) below illustrates A-movement of the subject: the subject cannot participate in labeling in its base-position, it undergoes labeling-triggered movement, after which the relevant part of the structure can be labeled as vP.

(31) John$_i$ [$_{vP}$ t_i [$_V$ bought the guitar]]

Hence, both movement of *what* from SpecCP in (29) and movement of *John* from SpecvP in (31) are motivated by labeling. This opens up the possibility that movement in general may be motivated by labeling.[2]

Above I have discussed the ways of labeling complex syntactic objects that are made available by Labeling Algorithm. However, in chapter 2, I will show that these modes of labeling are not all equal. In particular, I will show that there is a preference scale when it comes to the mechanism of Labeling Algorithm, with both labeling through feature-sharing (as illustrated in our (28)) and labeling through movement (30) being less preferred than head-labeling (25), the latter being the most straightforward application of the LA. I will also show that this difference has broad and important consequences for the grammar (including diachronic changes).

1.3. Outline of the dissertation

The dissertation is organized as follows:

Chapter 2 discusses a diachronic trend of the loss of obligatory syntactic movement. The discussion centers around diachronic changes involving wh-questions. In this respect, the chapter discusses the history of (i.e. the diachronic changes in) Japanese, Chinese, Indo-Aryan and Indo-Iranian (i.e. changes from Vedic Sanskrit to Hindi and Persian), Romance (from Latin to modern Romance languages), and Basque. It is shown that all of them display the loss of wh-movement.

[2] Disussing this issue properly would take us beyond the scope of this work. I will only refer the reader to Bošković (To appear(b), 2019a) for discussion on how movement can be motivated by labeling considerations even in cases where it appears to be optional, like e.g., in possessor raising in Serbo-Croatian (i).

(i) ?Markovog$_t$ je on vidio [t_t prijatelja]
Marko's$_{ACC.MASC.SG}$ is he seen friend$_{ACC.MASC.SG}$
'He saw Marko's friend.'

Additionally, the investigation reveals that this change is unidirectional: from wh-fronting to wh-in-situ, not the other way around. A deduction of this general trend is then offered based on Chomsky's (2013, 2015) labeling system. This chapter also discusses other ways of losing specifiers, including cases traditionally referred to as grammaticalization processes, or 'Spec-to-head reanalysis' (as in (16)-(18)) as well as simple loss of specifiers without any reanalysis. The discussion in the latter case involves, among others, Welsh agreeing complementizers and Early Germanic *hwæt*. I also discuss specifiers which are more resistant to a diachronic change, in particular cases involving multiple movement (as in multiple wh-fronting). I show that this class of specifiers is connected to different triggers for movement from those involved in single fronting. This chapter also shows that the labeling-based analysis accounts for another diachronic trend involving loss of movement, namely the loss of head movement (e.g., V-to-T movement). Further, the proposed analysis extends to additional historical changes, which I argue also involve loss of movement: the loss of OV word order and the loss of what is traditionally assumed to be rightward adjunction. Finally, I discuss semantic motivation for the existence of movement as well as other non-syntactic (i.e. phonological and pragmatic) factors that may play a role in this respect.

Chapter 3 is devoted to Verb-Second (V2). I show that this is a multifaceted phenomenon and in particular, I specify two types of V2: discourse-driven and illocutionary force-driven. I show that the former is subject to diachronic change as exemplified by its loss e.g., in English and Romance, whereas the latter is diachronically stable. I provide an account of this difference based on their different labeling conditions. In particular, only the discourse-related V2 is diachronically fragile as its licensing conditions and dependency on movement and feature-sharing make it subject to labeling-related pressures. The role of prosody and semantics in V2 will also be explored in this chapter, especially with reference to the illocutionary-force related V2.

Chapter 4 explores the connection between historical language change and language acquisition. By looking at the process of acquisition of reflexive clitics in Polish, and especially frequent errors of omission attested during early stages of learning of this construction, I argue that the same labeling-based pressures operating in diachronic reanalysis are also present in child language acquisition. I show that the labeling account gives us an explanation of the attested errors. The link between the acquisition and labeling-motivated language change is supported by diachronic facts from Russian, where the reflexive clitic *sja* underwent a reanalysis in the direction predicted by the system advocated here. I further discuss the role of labeling considerations in language acquisition by examining acquisition of wh-questions both in languages with obligatory wh-fronting and languages with optional *wh*-fronting. The conclusions reached based on these cases will be supported by facts from bilingual acquisition and by contact languages.

Chapter 5 concludes the thesis.

Chapter 2

Diachronic Loss of Movement as the Loss of Specifiers

2.1.Introduction

Much work dealing with diachronic syntax discusses the process of change where a lexical element is reanalyzed as a functional element, which is often referred to as grammaticalization. The reanalysis of a lexical element as a functional element can involve the loss of movement, although this is often not the case. Some diachronic changes along these lines involve reanalysis of verbs like 'say' in quotatives, emergence of a definite article from a demonstrative pronoun, or the emergence of complementizers from interrogative *wh*-phrases and relative pronouns (Lehmann 1995, Harris & Campbell 1995, Lyons 1999, Heine & Kuteva 2002, Roberts and Roussou 2003, van Gelderen 2011, 2015, 2018, Axel-Tober, 2017). Example (1) from Georgian illustrates a change of that sort, where a formerly moved *wh*-expression is reanalyzed as a complementizer.

(1) da ara unda, **raytamca** icna vin
and NEG want.PRES.3SG that know.PRES.3SG someone
'and he did not want that anyone know' (Georgian; Harris & Campbell, 1995: 298)

The present work will discuss cases like (1), but as a part of a much broader phenomenon. I will be looking at cases where movement gets completely lost, and there is furthermore no reanalysis of any sort in the formerly moved position. Roberts and Roussou (2003) treat such cases differently.

Their theory of grammaticalization, in fact, cannot be extended to them. [1]However, the approach that will be proposed in this thesis has a natural way of putting together cases like (1), where there is a reanalysis in the formerly moved position, and cases that involve loss of movement without any reanalysis. The focus of this chapter will, in fact, be on cases involving the loss of all sorts of movement, both phrasal and head movement. I will provide a natural explanation that unifies all these cases, as well as cases like (1), in terms of structure building and labeling.

I will start the discussion with cases which are very different from those discussed in Roberts and Roussou (2003), and which are not much discussed in the literature on historical change; namely, the loss of *wh*-movement.

This chapter is organized in the following way: after the discussion of the diachronic loss of obligatory *wh*-movement in many unrelated languages, I will offer a deduction of this diachronic trend based on Chomsky's (2013, 2015) labeling system and argue that the loss of movement actually involves the loss of specifiers, which can be captured by the labeling system. Section 4 discusses other ways to lose specifiers, including cases like (1), and shows that all these changes can be accounted for by the system proposed here. Section 5 discusses specifiers which are more resistant to diachronic change. In particular, I will focus on cases involving multiple movement to the same position in the structure. Section 6 extends the proposed analysis to the loss of head movement. The cases discussed involve the loss of V-to-T movement (T-to-C movement will be discussced in chapter 3). Section 7 shows that the proposed analysis can explain additional historical changes, such as cross-linguistic tendency to reanalyze Object-Verb (head-final) word

[1] Roberts and Roussou (2003) look briefly at some of the cases to be discussed here, namely loss of V2 (chapter 3), loss of V-to-T movement, and word order changes (OV>VO).

order as Verb-Object (head-initial) word order, the opposite being very rare, and the loss of rightward-adjunction. Section 8 concludes this chapter.

2.2.The Loss of Wh-movement

This section presents diachronic data from a variety of languages in which obligatory *wh*-movement has either been lost and replaced by an *in-situ* strategy or has been substantially relaxed and the *wh*-fronting option coexists with optional *wh-in-situ* (with wh-in-situ option becoming more prominent). As apparent optionality may indicate a change in progress (Wallenberg 2016), those languages in which both options in *wh*-questions, i.e., fronting and *in-situ* coexist, may be at the initial stages of losing *wh*-movement as well (as indicated by the wh-in-situ option gaining more prominence). The fact that the data gathered here come from different language families, Sino-Tibetan, Altaic, Indo-European, and an isolate (Basque), argues for the universality of the trend depicted here.

It should be noted that there are some differences in the treatment of the landing site of the traditional wh-movement in the literature, especially in the works adopting split CP. Thus many authors have argued that *wh*-movement is actually movement to [Spec, FocP] (Rizzi 1997). I will not make a distinction between *wh*-movement to Spec, CP versus movement to Spec, FocP, because it does not matter for my purposes, the main point in this section being the loss of obligatory *wh*-movement, regardless of its exact landing site.

We will begin the discussion with two languages: Japanese and Mandarin Chinese, which while today being treated as iconic examples of *wh-in-situ*, were obligatorily *wh*-fronting languages at one point in their history.

2.2.1. Old Japanese

Japanese is often given as a typical example of a *wh-in-situ* language but Old Japanese (around the 8th century) had obligatory *wh*-fronting, together with a system of *wh*-agreement (*kakarimusubi*) between question/focus particle *-ka* and the predicate, which took the nominal ending (Ogawa 1976, Whitman 2001, Watanabe 2002, Kuroda 2007, Aldridge 2009, 2018).[2]

(2) a. Nani-wo-**ka**-mo mikari-no hito-no ori-te kazasa-mu ?
what-ACC-KA-FOC hike-GEN person-GEN pick-CONJ wear.on.the.hair-will
'What should hikers pick and wear on their hair?' Aldridge 2018: 6; MYS 1974)

b. Izuku-yu-**ka** imo-ga iriki-te yume-ni mie-tsuru
where-through-KA wife-NOM[3] enter-CONJ dream-LOC appear-PERF-RT
'From where did my wife come and appear in my dream?' (Watanabe 2002: 182; MYS 3117)

(3) Iduku-ni-**ka** kimi-ga *mi-pune-wo* wa-ga miti-ora-mu?
where- DAT-KA lord-GEN HON-boat-ACC I-GEN wait-be-MOD.RT
'Where shall I be waiting for your boat?' (Aldridge 2018:8; MYS 2082)

[2] This displacement is island sensitive (see Yanagida 1995, Whitman 1997, 2001, Kuroda 2007, Aldridge 2018).

[3] Watanabe (2002) takes *-ga* marked subjects to be nominative, analogically to this case marking in modern Japanese. Aldridge (2009), however, follows Yanagida (2005) and Kuroda (2007), among others, in assuming that subjects bearing *-ga* case particles are genitive. Yanagida (2006) and Yanagida and Whitman (2009) refer to *ga*-marked subjects in kakarimusubi as ACTIVE, while Yanagida (2017) refers to them as AGENTIVE, not Genitive. I am glossing the examples as in the appropriate sources here, leaving the precise character of this marking open since it does not affect the main point made here.

(4) Wa-ga omofu kimi-fa *idukufe-ni* koyofi *tare-to-**ka*** mate-do ki-masa-nu

I-GEN long.for you-TOP where-DAT tonight who-with-KA wait-eve come-HON-neg

'I long for you, but though I wait you do not come. (I wonder) where you are tonight and with who.' (MYS no. 3277; Based on Aldridge 2009:558)

The examples above should leave no doubt that Old Japanese or the Nara period can be considered to be a *wh*-movement language[4]. As far as the landing site of this obligatory *wh*-fronting is concerned, there have been various proposals in the literature.

Watanabe (2002) takes the landing site for this obligatory *wh*-movement to be the FocP, following Rizzi's (1997) split CP system, i.e., he places *ka* marked phrase within the split CP.

(5) [$_{TopP}$ Spec Top [$_{FocP}$ Spec Foc-*XP*-***ka*** [$_{IP}$ Subj VP I]]] (based on Watanabe 2002:183)

The difference in postulated landing sites for these *wh*-elements bears on the interpretation of the position of the subjects. Watanabe (2002) takes *-ga* marked subjects to be nominative, analogically to this case marking in modern Japanese. Taking the spec, TP to be the usual licensing position for nominative subjects, and following the observation of Nomura 1993 that the *wh*-phrase must precede the nominative-marked subject (and it may follow the *ha*-marked topic), he places *ka*-marked XP within the CP, as in many other *wh*-fronting languages.

Aldridge (2009), however, follows Yanagida (2005) and Kuroda (2007), among others, assuming that subjects bearing *-ga* case particles are genitive (besides the *-no* marked elements,

[4] Interestingly, *wh-in-situ* was also available in Old Japanese of Nara period (Ogawa 1976, Watanabe 2002). The existence of these two forms may suggest competition between two ways of licensing *wh*-questions. The fact that the *in-situ* version persisted, while the fronted one was lost, argues in favor of the analysis presented here, where movement is dispreferred option as it leads to a dispreferred configuration from the point of labeling.

the former being assigned to the [Spec, vP], while the latter inside the VP), and they are placed lower in structure, in the specifier of vP. On her analysis, *wh*-fronting in Old Japanese targets a focus position at the edge of vP. If *wh*-fronting targets a position lower than T (still preceding the subject in the Spec of vP), it is possible that a fronted *wh*-phrase can be preceded by constituents other than topics, which Aldridge (2009) argues is possible.

On the other hand, Aldridge (2018) takes Old Japanese *wh*-elements to be targeting left-periphery, [Spec,TP], which she argues to be the focus position due to C-T inheritance.

(6) a. Pototogisu nani-**ka** ki-naka-nu ?
cuckoo.NOM what-KA come-cry-NEG.RT
'Why does the cuckoo not come and sing?'

Finally, Whitman (2001) assumes that Old Japanese *-ka* particle is the head of the CP and the fronted *wh*-elements move into its specifier.

(7) [CP [WH] C-**ka** [IP….]]

Whitman's analysis sees the loss of Japanese overt *wh*-movement, completed by the 15th century (with the decline starting around 12th century), as a shift of the complementizer *-ka* to a clause final position, which is followed by the subsequent loss of the ability to trigger *wh*-movement. For Whitman, the shift to the C-final order in Modern Japanese *wh*-questions involves the loss of movement as well as the loss of a specifier. In fact, the two are closely related, as we will see in this chapter.

At any rate, regardless of the postulated landing site of *wh*-phrases in Old Japanese, what is important for us is the observation that Old Japanese was an obligatory *wh*-movement language,

wh-movement got lost, and is no longer available in modern Japanese. Japanese is in fact not the only language that lost *wh*-movement in its diachrony.

2.2.2. Archaic Chinese

Another language with *wh*-fronting, but only in its historical records, Mandarin Chinese (Feng 1996, Aldridge 2010, 2011). Late Archaic Chinese of the Warring States period (5th to 3rd century BC) had *wh*-movement, contrary to Modern Mandarin Chinese.

(8) a. Wu **shei** [VP qi t_{shei}]?

I who deceive

'Who do I deceive? (A9;Aldridge 2010:8)

b. Gong **shei** yu [yu t_{shei}] ?

you who want give

'Who do you want to give (it) to?' (Zhuangzu 3.2; Aldridge 2010: 7)

c. **He cheng** bu ke?

what city not conquer

'What city would you not conquer' (Zouzhang, Xi 4; Aldridge 2010:6)

Aldridge argues that we are dealing here with phrasal movement, also arguing against the analysis of this movement as an instance of e.g., prosodic cliticization, as in Feng (1996), where the wh-words are argued to cliticize onto the verb (notice that the *wh*-phrase precedes the negation in (8c), i.e., it is not V-adjacent).

Aldridge (2010), in fact, argues that this an instance of A'-movement, which she analyzes as targeting a focus position.[5] This obligatory movement was lost in the Hàn period (by 220 CE).

Another instance of syntactic movement that was also lost around the same time is *suo* relativization (Aldridge 2010).

(9) [ren zhi ***suo*** [vP wei e]]

person GEN REL fear

'what people fear' (*Laozi* 20, Aldridge 2010:28)

Crucially, both these operations involved a syntactic A'-movement, which is no longer available in Modern Mandarin Chinese. Hence, we have another instance of a language with obligatory *wh*-movement, which got lost and replaced by *wh-in-situ* strategy in question formation.

The next sections will show that the diachronic loss of *wh*-movement is also widely attested in Indo-European languages, which has not gained much attention in the literature, despite the history of the Indo-European languages being very well researched, due to many historical records available. I will start the overview with the oldest documented version of Sanskrit as it appears in Rgveda, together with Avestan.

2.2.3. The Oldest Indo-Iranian Languages

Hale (1987) offers an important overview of *wh*-movement in Vedic Sanskrit, Avestan, and Old Persian, which shows that fronting of *wh*-words was obligatory in these languages. We indeed find *wh*-words and relative pronouns in the initial position in overwhelming majority in these early

[5] An element that precedes it functions as a Topic.

Indo-Iranian languages. Fortson (2004) argues that *wh*-movement occurred in all known ancient Indo-European daughter languages, where interrogative expressions move to CP. This situation contrasts with contemporary Indic and Persian, which again suggests a diachronic change of a similar kind as what we have seen in Japanese and Chinese.

For Vedic Sanskrit, the situation seems to be very clear, as 90% of *ká* 'question' expressions (in their variously case-marked forms) appear in the first position in their clause both in Rigveda (RV) (10) and in the Taittiriya Samhita (TS) (11) (Hale 1987). We can assume, therefore, obligatory movement to the left periphery for these *wh*-words in both Vedic poetry (RV) and prose (TS).

(10) a. **kásya** bráhmaṇi jujuṣur yúvānaḥ
whose prayers enjoy youths
Lit: 'whose prayers enjoy the youths?' (RV 1.165.2a; Hale 1987:8)

b. **kúha** sthaḥ **kúha** jagmathuḥ
where 2.DU where gone
'Where are you two? Where have you two gone?'(RV 1.165.2a; Hale 1987:8)

(11) a. **kásya** vāha deva yajnám āgáchanti **kásya** vā ná
'Whose sacrifice do the gods approach and whose not?' (TS 1.6.7.1; Hale 1987:25)

b. **kiṃ** tád yajné yájamanaḥ kurute
'What does the sacrifice do then in the sacrifice?' (TS 2.6.2.2: Hale 1987:25)

Macdonell (1916), discussing relative and interrogative words in his Vedic grammar, shows that they indeed begin the sentence by occupying the first position, 'which in ordinary sentences do not occupy' (Macdonell:286).

(12) **kiṃ** hi sá táir gṛháiḥ kuryāt?

'What indeed should he do with this house?' (ŚB.: Macdonell 1916:286)

The contexts where *wh*-expressions in Vedic Sanskrit are not the first in the clause involve topicalization across the fronted *wh*-word (13), or the placement of the sentence initial particles (14) (Hale 1987, Fortson 2004). As we see in (15), topicalization may involve a part of the pied-piped *wh*-expressions. In all such contexts, only a single constituent can precede the question word.

(13) ráthaṃ **kó** nir avartyat

chariot who prepared

'Who prepared the chariot?' (RV 10.135.5b: Hale 1987: 11)

(14) *atha* **kó** vsda…

'Then who knows…?' (RV 10.129.6d: Hale 1987: 14)

(15) brahma **kó** vaḥ saparyati

priest which you honors (RV 8.7.20c: Hale 1987: 10)

'Which priest honors you?'

Vedic Sanskrit also shows obligatory movement of relative pronouns. Hale observes that out of nearly 600 relative clauses in Rigveda, over 500 have the relative pronoun in the displaced initial

position in the clause. Similarly, in Vedic prose, out of 400 relative clauses in TS, only 20 don't have the relative pronoun in the first position in their clause (Hale 1987).

(16) a. **yám** u dviṣmás tám u prāṇó jahātu

'Whom we hate, let breath leave him.' (RV. 3.33.21d; Hale 1987: 16)

b. **yásmād** yóner udārithā…

'from which womb you arouse…' (RV. 2.9.3c; Hale 1987: 16)

(17) a. **yásmād** evá yóneḥ prajāpatiḥ paśūn ásṛjata

'from which source Prajapati created the cattle' (TS 2.4.11.4; Hale 1987: 27)

b. **yásmiñ** yātā etām iṣṭim nirvápati

'on whose birth he offers this sacrifice' (TS 2.2.5.3; Hale 1987:27)

Similarly to Vedic Sanskrit, Avestan (Av) is shown to have involved obligatory *wh*-word fronting and relative pronoun movement (Hale 1987, Fortson 2004, West 2011). Older Avestan contains 93% of all occurrences of question words in the initial position (Hale 1987).[6]

(18) a. **kat** və̄ xšaʊrəm, **kā** ī̄štiš?

'What (is) your power, what your ability?' (Av; West 2011:5)

b. **kudā** aṣəm vohucā manō xšaʊrəmcā?

'Where (are) Right and Good Thought and Dominion?' (Av; West 2011:5)

c. **kahmāi** mā ʊβarōždūm **kə̄** mā taṣaṭ?

'For whom did you create me? Who fashioned me?' (Av: Hale 1987:32)

[6] The examples in this section show only Older Avestan, however the facts concerning the fronting of wh-phrases and relative pronouns hold for both Older Avestan and Younger Avestan.

Relative pronouns in this language also occur in the initial position in the majority of examples.

(19) a. **yə̄** vå mazdā ahurā / pairī.jasāi vohū manaŋhā

'I who shall serve you, O Wise Lord, with good thinking' (Av: Hale 1987:33)

b. **yāiš** asrūdūm būmiiå haptaiʋē

'through which you are framed in the seventh (part) of the earth' (Av; Hale 1987: 33)

Due to the character of the Old Persian corpus, containing mainly inscriptional decrees of Darius the Great (521-486 BC) and his descendants, we have no instances of real *wh*-questions in this language. However, relative clauses do appear and the obligatory fronting of the relative pronouns is confirmed by the data (see Hale 1987, 1988).

(20) a. avam kāram jadiy **hya** manā naiy gaubatiy

that army smite which mine not calls

'Smite that army that does not call itself mine' (OP; DB 3.15: Hale 1987:57)

b. imaiy martiyā **tyaiy** adakaiy avadā āhatā

these men who then there were

'these are the men who were there then' (OP; DB 4.80; Hale 1987:57)

Wh-fronting to the high left peripheral projection within CP in Vedic Sanskrit and Avestan/Old Persian stands in a contrast with modern-day Indic and Iranian languages such as Hindi or Persian, which are *wh-in-situ* languages. The obligatory *wh*-movement was lost in these languages, similarly to Japan and Chinese. As this process of losing obligatory *wh*-fronting is so common historically, the argument can be made that it is indeed a cross-linguistic trend in need of explanation.

2.2.4. Latin vs. Modern Romance Languages

Staying within the Indo-European family, one of the most interesting examples of syntactic change in the direction of losing obligatory syntactic *wh*-movement has occurred in the Romance languages, where the obligatoriness of interrogative movement in Latin, which was in fact a multiple wh-fronting languages, contrasts with its optionality in modern Romance. More specifically, languages like Spanish, Brazilian Portuguese, and French, having lost the Latin-style multiple *wh*-fronting, are developing *wh*-in-situ, which was not attested in Latin. This section focuses on this contrast between Latin and contemporary Romance languages, changes across time that indicate the loss of the obligatoriness of *wh*-movement in these languages.

Latin was an obligatory *wh*-movement language (see Spevak 2010 citing Panhuis 1982: 61-68 and Babič 1992:61). In fact, Latin had both single (22), and multiple *wh*-fronting (23)[7] (Bennett 1966, Devine and Stevens 2006, Brown, Joseph, and Wallace 2009, Spevak 2010, Danckaert 2012, Ledgeway 2012, a.o.). In other words, Latin was a multiple *wh*-fronting language.

(21) **Quid** habui facere ?
what I.had do.INF
'What did I have to do/should I do?' (Sen.*Contr.* 1.1.19; Ledgeway 2012:137)

(22) a.(…)**Cui** **quando** nupsit?
Who.DAT when she.married
'When and to whom was she married?' (Ter. *Ad.*670; Spevak 2010:198)

[7] Devine and Stephens (2006), and Danckaert (2012) claim that the order of multiple question-words in Latin preserves the hierarchical order of constituents.

b. Ego **quid** **cui** debe-a-m] sci-o.

I.NOM what.ACC whom.DAT owe-PR.SUBJ-1.SG know-PR.1.SG

'I know what I owe to whom.' (Sen. Ben. 4.32.4; Danckaert 2012: 245)

c. **Quis quem** fraudasse dicatur ?

who whom defrauded is.said

'Who is said to have defrauded whom? (Pro Rosc Com 21; Devine and Stephens 2006: 89)

d. **Quis** **apud quos** **quibus** praesentibus sit acturus

who.NOM with who.ACC who.DAT is.present be do.FUT

'Who is going to plead before whom in whose presence?' (Quintilian 11.3.150; Devine and Stephens 2006: 89)

As *wh*-phrases in Latin are obligatorily displaced to a high left peripheral position, in most of instances they are clause-initial (see also Ledgeway 2012). In cases where they are preceded by some other elements, those are always topicalized constituents (located in TopP, which is above the projection targeted by multiple *wh*-fronting (Ledgeway 2012, see also Danckaert 2012 on Latin middle field).

Spevak (2010) reports that these 'contextually bound' elements can be found in about 8% of the cases with fronted question-words in her Latin corpus of classical prose (Spevak 2010: 196). Hence, Latin is an obligatory *wh*-fronting language, which did not allow *wh-in-situ*, and required movement of all interrogative elements (Panhuis 1982 refers to the fronted wh-word as 'rheme proper', which is always the initial element in interrogatives).

Modern Romance languages, on the other hand, not only lost the obligatoriness, and even possibility, of multiple *wh*-fronting, but additionally have an option of leaving *wh-in-situ*, which was not possible in Latin, in number of contexts. Modern Romance languages like Spanish

(Jimenez 1997, Uribe-Etxebarria 2002, Reglero 2004, 2007, Reglero and Ticio (R&T) 2013), Brazilian Portuguese (BP) (Pires and Taylor 2007, Zocca DeRoma 2011, Figueiredo Silva and Grolla (FS&G) 2016), or French (Mathieu 2004, Bošković 2000, 2015a, Adli 2006, Cheng and Rooryck 2000, Munaro, Poletto, and Pollock 2001, Oiry 2011, Shlonsky 2012, a.o), are all illustrations of this loss of Latin-style obligatory *wh*-movement and the appearance of wh-in-situ.

Spanish is a good example of a modern Romance language with optional *wh-in-situ* in both short-distance (23a-b) and long distance (24a-b) contexts.

(23) a. ¿**Qué** compró Juan?
what bought John
'What did John buy?'

b. ¿ Juan compró **qué**?
John bought what
'What did John buy?'

(24) a. [Y]tú invitaste a tu fiesta a **quién**?
you invited to your party to who
'Who did you invite to your party?'(based on Reglero 2004:17)

b. [Y] tú te vas a vestir para la fiesta **cómo**?
and you CL are-going to dress for the party how?
'And how are you going to dress up for the party?'

Additionally, in contrast to Latin, Spanish disallows multiple *wh*-fronting. Compare Latin-style (25a), which is ungrammatical in Spanish, with perfectly acceptable (25c), where all *wh*-phrases appear *in-situ.*

(25) a. *¿ **A quién qué** le pidió Ivan?
to who what CL asked Ivan

b. ¿ **A quién** le pidió Ivan **qué** ?
to who CL asked Ivan what

c. ¿ Ivan le pidió **qué a quién**?
Ivan CL asked what to who?
'What did Ivan ask **to whom**? (R&T 2013: 537)

Still, some authors have analyzed *wh-in-situ* in Spanish as involving overt *wh*-movement. One such analysis is Uribe-Etxebarria (2002), where seemingly *in-situ* surface position of *wh*-phrase in Spanish is argued to result from two operations: an overt *wh*-movement to [Spec,CP], followed by remnant IP movement over the displaced *wh*-word to some XP projection above the CP.

Uribe-Etxebarria proposes this overt (two-step) movement derivation for both short- and long-distance *wh*-questions. She observes that wh-in-situ in Spanish has to be sentence final and argues that this follows from the proposed two sequences of movement (i.e., *wh*-movement and subsequent IP remnant movement).

However, there are reasons to reject Uribe-Etxebarria's overt movement analysis and consider Spanish cases like (23b) and (24) to really involve *wh-in-situ.* Reglero (2004, 2007), and Reglero & Ticio (2013) provide evidence that no wh-movement is involved in such configurations.

First of all, overt wh-movement is excluded in cases where the base position of *wh*-phrase is placed within a syntactic island (26a). Grammaticality of *wh-in-situ* in such configurations (unlike when the overt movement is involved (26a)) shows than indeed there cannot be any movement there (23b). This is a problem for Uribe-Etxebarria's analysis, where both (26a) and (26b) involve overt *wh*-movement.

(26) a. *¿ **Con quién** te has enamorado del hombre que vive?
with who you have fallen-in-love of-the man that lives

b. ¿Te has enamorado del hombre que vive **con quién**?
you have fallen-in-love of-the man that lives with who

'Who have you fallen in love with the man that lives with?' (R&T 2013: 509)

Also, *wh-in-situ* survives certain blocking effects which arise with fronted *wh*-phrases. For example, the presence of a possessor blocks the extraction of agents and objects (27a-c), and the presence of agents blocks the extraction of objects (27c). *Wh-in-situ* in such contexts is perfectly fine (28a-c), which indicates that (28a-c) should not be analyzed as involving overt wh-movement on a par with (27).

(27) a. *¿ [**De quién**]$_{ag}$ has leído [varios libros t$_{ag}$ [de Juan]$_{poss}$]?
of whom have-you read several books of Juan

b. *¿ [**De qué**]$_{obj}$ has leído [varios libros t$_{obj}$ [de Juan]$_{poss}$]?
of what have-you read several books of Juan

c. *¿ [**De qué obra]$_{obj}$** conoces [varias traducciones t$_{obj}$ [de Ana]$_{ag}$]?
of what work know-you several translations of Ana

(28) a. ¿Has leído [varios libros t_{ag} [de Juan]$_{pos}$ [**de quién**]$_{ag}$]?
have-you read several books of Juan of whom

b. ¿ Has leído [varios libros t_{obj} [de Juan]$_{poss}$ [**de qué**]$_{obj}$]?
have-you read several books of Juan of what

c. ¿ Conoces [varias traducciones t_{obj} [de Ana]$_{ag}$ [**de qué obra]$_{obj}$**]?
know-you several translations of Ana of what work (R&T 2013: 510-512)

Finally, regarding the Sentence Final Requirement (SFR), Reglero shows that what is responsible for it is a prosodic factor. Spanish is a nuclear stress assignment (NSA) language, which requires focalized elements, including wh-in-situ, to be assigned stress by NSA. This rule assigns stress to the last element within an intonational phrase. As a result, a pause affects the grammaticality of (29a). The *wh*-phrase is the last element in its intonational phrase in (29b), hence can be assigned stress by the NSA rule.

(29) a. *¿Mariá compra **dónde** el pan?
Maria bought where the bread

b. ¿Mariá compra **dónde** # el pan?
Maria bought where the bread
'Where did Maria bought the bread?'(R&T 2013; 519)

Notice also that the requirement for the *wh*-phrases under considerations to be intonational phrase final is absent from other modern Romance languages like Brazilian Portuguese or French, where *wh-in-situ* is not subject to the intonational phrase-final requirement (e.g., there is no need to pause after the *wh*-phrase in (30)).

(30) Tu vois **qui** ce soir?

You see who this evening?

'Who are you seeing this evening? (French; Mathieu 1999)

All these facts indicate that that Spanish indeed has an optional *wh-in-situ*, in this respect it differs from Latin, where the movement of *wh*-phrases was obligatory. This included multiple fronting, which, as we have seen in (25), got lost in Spanish as well as other modern Romance languages.

Brazilian Portuguese (BP), just like Spanish, also allows *wh-in-situ*.

(31) Você viu **o quê?**

you saw what

'What did you see?'(BP; FS&G 2016: 261)

(32) Você disse que ela viu **o quê ?**

you said that she saw what

'What did you say that she saw?' (BP; FS&G 2016: 261)

Again, on a par with Spanish, *wh-in-situ* can occur inside islands, unlike with fronted *wh*-phrases, where extraction out of a relative clause leads to strong ungrammaticality (33b). The contrast between (34a) and (34b) strongly indicates the lack of *wh*-movement (33a).

(33) a. ?Maria admira [o autor que escreveu **que livro**]?

Mary admires the author who wrote which book

b.*__Que livro__ (que) a Maria admira [o autor que escreveu]?

which book (that) Mary admires the author who wrote

Int: 'Which is the book x such that Mary admires the author who wrote x?' (BP; FS&G 2016: 263)

Lopes-Rossi (1996) reports that during the first half of the 19th century the production rate of *wh-in-situ* increased from 0 to 32.4 percent in this language.[8] This shows that Brazilian Portuguese is moving away from the Latin-style obligatory *wh*-movement, hence confirming the cross-linguistic trend discussed here.

Finally, *wh-in-situ* is also quite productive in French. (Adli 2006, Oiry 2011, Shlonsky 2012, among others)

(34) Tu as vu **qui**?

you have seen whom

'Who have you seen?' (French; Bošković 2015a: 252)

(35) Tu penses que Jean va épouser **qui** finalement?

you think that Jean will marry who in the end

'Who do you think Jean will marry in the end?' (French; Adli 2006:175)

8 Interestingly, the rate of head movement to C in questions dropped over the same period as well, as the WH-Verb-Subject word order is reported to have decreased from over 45 percent in the 19th century, to 12.5 percent in the 20th century (Figueiredo Silva and Grolla 2016). This will become relevant later in this section.

Experimental setting has also revealed the possibility of Scope Marking/(Simple)Partial Movement constructions in French, produced both by children, as well as adults, as (Oiry 2011). Sentences in (36) below illustrate the actual adult French examples of the Simple Partial Movement (36a), while (36b) is a Partial Movement structure from child French.[9] Crucially, as Oiry (2011) argues, both of these constructions involve *wh-in-situ*, which is confirmed by the fact that *quoi* kept its in-situ form (not *que*) (see Oiry 2011, 2015).

(36) a. Paul croit que c'est **quoi** que Marie a acheté ?
Paul believes that it-is what that Marie has bought?
'What does Paul believe Mary bought?' (French; Oiry 2015:6)

b. Tu crois **quoi** que Lala elle aime bien **quoi**?
you think what that Lala she like well what
'What do you believe that Lala likes?' (Child French; Oiry 2011: 123)

Fanselow (2006) argues that this kind of Scope Marking constructions are in fact available only in languages with *wh-in-situ*. Hence, it is not surprising that it is attested in French.

The fact that *in-situ* option does not involve movement in this language is confirmed by various diagnostics, one of which is parasitic gaps licensing. Parasitic gaps are licensed by overt wh-movement, as in (37b), but they are not licensed in (37a).

(37) a.*Il a lu **quoi** sans classer?
he has read what without to-file?
Int: 'What did he read without filing?'

[9] Simple Partial Movement does not involve the lexical scope marker in the Matrix clause, unlike in Partial Movement constructions of German type.

b. Qu'a-t-il lu sans classer?

what has-he read without to-file

'What did he read without filing?' Bošković 2015a: 266)

French also has a single-pair interpretation available in questions, but only with wh-phrases in-situ. Bošković (1999, 2002, 2015a) show that cross-linguistically, single-pair answers to questions with multiple wh-phrases are possible only in languages that have a possibility of not moving any wh-phrase to Spec, CP. In other words, overt movement to the specifier of CP blocks the single pair interpretation, and only pair-answer is possible in such cases. Hence, English (38) has only a pair-list answer.

(38) Who bought what?

Languages like Japanese, Chinese, or Hindi, on the other hand allow single-pair answers in similar contexts. Since French (39b) allows a single-pair answer, in contrast to (39b), which indicates that it indeed does not involve overt movement.

(39) a. **Qu**'a-t-il donné **à qui**?

what has he given to whom

b. Il a donné **quoi à qui**?

he has given what to whom

'What did he give to whom? (Bošković 2003: 3)

Finally, we can ask how frequent *wh*-in-situ in a modern Romance language like French is, as compared to the movement option. French actually has more than one overt movement strategy. In spite of this, French *wh-in-situ* is found in nearly 39 percent of interrogative contexts in the

speech of Paris middle-class (Ashby 1977, Coveney 2002, Meyers 2007). Similarly, Behnstedt (1973) finds the Subject-Verb-WH variant in 33 percent of contexts in colloquial speech for middle-class speakers, with 25 percent in the formal contexts for the same speakers. These numbers indicate a real change in comparison to Latin, where occurrence of *wh-in-situ* is close to 0 percent. Coveney (2002) furthermore argues that *wh-in-situ* is on the rise in terms of the frequency of occurrence, which supports the observation explored in the present work about the trend towards the loss of movement (see also Meyers 2007).

This trend is thus observed cross-linguistically. We have already seen its evidence in Altaic, Sino-Tibetan, and Indo-European. A very interesting confirmation of this trend can be found also in Basque, where we can observe a change in progress.

2.2.5. Navarro-Labourdin Basque

Duguine & Irurtzun (2014) show that Navarro-Labourdin dialect of Basque is developing a *wh-in-situ* strategy for matrix questions. Interestingly, this new pattern is attested almost exclusively in the speech of the young speakers of Labourdin Basque. Therefore, apart from a standard *wh*-movement accompanied by T-to-C movement, which yields a V2 clause, this dialect of Basque has presently an additional way of constructing *wh*-interrogatives, i.e., *wh-in-situ*, where neither wh-movement nor V2-like T-to-C movement, is attested. Duguine & Irurtzun (2014) report that this new pattern is considered ungrammatical by the older speakers, which indeed suggests grammatical change in progress.

One example of wh-fronting in Basque is given in (40):

(40) **Zer** *jan* Jonek du ?

what.ABS eat Jon.ERG AUX

'What did Jon eat?'

Basque displays SWH-V-O word order in 'canonical' *wh*-fronted interrogatives, due to inversion that accompanies wh-fronting. The new interrogatives, however, permit other word-order patterns, such as SWH-IO-DO-V or S-IOWH-DO-V, which Duguine & Irurtzun (2014) take as evidence for a genuine *wh*-in-situ strategy, without *wh*-fronting and the inversion. One relevant example is given below:

(41) **Nork** gereziak *jan* ditu?

who.ERG cherries.ABS eat AUX

'Who ate the cherries?' (Duguine & Irurtzun 2014: 4)

The placement of adverbs offers additional confirmation of the analysis without the movement in these constructions, as the order of constituents in questions mirrors the order in statements. In (42), the low adverb of manner (*fite* 'quickly') appears in the immediate preverbal position, following the direct object (42a). The question in (42b) shows the placement of the wh-phrase in exactly the same (i.e., base generated position) with reference to the adverb and other constituents in the language.

(42) a. Jonek **gereziak** *fite* jan ditu

Jon.ERG cherries.ABS quickly eat AUX

'Jon ate the cherries quickly.'

b. Jonek **zer** *fite* jan ditu

Jon.ERG what.ABS quickly eat AUX

'What did Jon eat quickly?'

The possibility of merely the loss of inversion in examples like (41) is excluded by examples like (43) below, where the object would undergo *wh*-movement over the subject but with the verb staying in the lower base position. Such sentences are judged to be ungrammatical by both older speakers (which also reject cases like (42b) and most importantly, also younger speakers of Labourdin Basque, for whom (42b) is perfectly fine, unlike (43).

(43) ***Zer** Jonek *jan* du ?

What.ABS Jon.ERG eat AUX

'What did Jon eat?' (D&I 2014: 5(11))

Therefore, the *wh-in-situ* variant of Basque involves both the lack of *wh*-movement, as well as the lack of C-to-T movement.

The fact that the option for *wh*-questions which does not involve movement is gradually gaining new ground in Basque indicates that we may indeed be dealing here with the loss of *wh*-movement, which manifests itself with the initial optionality, later to be possibly the only option in language.

Concluding this section, we have seen that the loss of *wh*-movement is a diachronic trend cross-linguistically, which is attested in as distant language families as Sino-Tibetan, Altaic, Indo-European, as well as, beyond these. What is particularly important here, and what has not been observed before, is that there are no clear cases where the *wh*-movement strategy is gained. In other words, the direction of the diachronic change here is always from wh-movement to wh-in-

situ, not the other way around. The fact that we never get the other direction indicates that the reasons to lose *wh*-movement, and possibly XP-movement in general, are triggered by factors connected either to Universal Grammar, or by general computational properties of language, referred to as 3rd factor principles in Chomsky (2005). I will indeed pursue this possibility here.

The next section will provide a deduction of the facts regarding the loss of movement we have just seen. The deduction will be based on the labeling system of Chomsky (2013, 2015), and will interpret the loss of movement as actually the loss of specifiers. I will argue that what we have at work here is a labeling-induced preference for the head-complement configuration and dispreference for the Spec-head configuration.

2.3.The Loss of Specifiers as the Labeling-Induced Preference

This section will show that the observed loss of movement can be attributed to the preference for the head-complement configuration during syntactic-structure building.[10] Such a preference is a consequence of a particular mechanism that determines the nature of the syntactic structure, namely the Labeling Algorithm (LA), as in Chomsky (2013, 2015).

Essentially following GB, early Minimalism (Chomsky 1995) saw labeling as part of the definition of Merge, the latter being a binary concatenative procedure that selects two syntactic objects, α and β, and forms a set K (another syntactic object), with its label γ, provided by either α or β. What this means, is that the identity of the resulting structure is determined by one of the elements in the set: a syntactic head that projects. However, later developments led to the treatment

[10] For another perspective on this issue, see Bošković (2019).

of labeling as independent of Merge, reducing the latter to the simplest set–forming procedure capable of operating in free, unlimited, and unrestricted fashion (see Collins 2002). Therefore, labeling is no longer automatic part of Merge. Chomsky (2013) proposes an approach along these lines.

In particular, Chomsky (2013, 2015) delegates labeling to a special mechanism, a Labeling Algorithm (LA), which is based on the Minimal Search (MS) operation that is looking for a minimal item, a head, that can provide a label for the created configuration. Syntactic object (SO), where a head is merged with a phrasal complement {H,YP}, constitutes the best-case scenario in this system, as this head H_0 gets selected as the label for this structure. MS is seeking a head, and in {H,YP} configuration, it finds it easily. Therefore, labeling in a head-complement configuration is trivial.

(44) Head-Labeling

a. {H,YP} head-phrase merger

b. {H{H,YP}} Minimal search of LA: Label → Head H_0

However, when two phrasal projections, {XP, YP}, are merged, Minimal Search cannot find a unique minimal head and it is not obvious what should serve as the label. The search result is too ambiguous to determine the label of this whole merger, as LA finds not one, but two heads, i.e. X_0 and Y_0 that potentially can serve as labels. Therefore, in such cases, something extra has to happen: either Agree and feature-sharing between XP and YP (if they indeed share a prominent feature), or movement of one of the phrases (given that traces/copies are invisible for LA), which means that the moved element is then ignored for labels. This is depicted in (45).

(45) Labeling in Phrase-Phrase Merger

a. {XP,YP} Phrase-Phrase Merger

b.{?{XP , YP }} Minimal search of LA, No Label: Ambiguous result (X_0 or Y_0 = ?)

c.{F{XP $_{[F]}$, YP$_{[F]}$}} Additional operation$_1$: Agree and Feature-sharing

d.{Y{t$_{XP}$, YP}} Additional operation$_2$: Movement

Hence, we see that labeling in a head-complement and phrase-phrase configuration is not equal. In the case of {H,YP}(45), the Minimal Search operation of the labeling algorithm returns the closest syntactic head as the label for the given syntactic object. Minimal Search is a syntactic operation based on locality considerations, which determines the label in the head-complement configuration in a straightforward way.

This is not the case in {XP,YP} (45), where head-identification gives an ambiguous result and something additional has to happen for a label to be determined; either the SO has to be modified by movement or Agree between XP and YP has to be attempted. Hence labeling in these two configurations is not equivalent, only head-complement (44) being automatic.

There is another reason to treat labeling in the head-complement configuration differently. In Chomsky (2013, 2015), unlabeled objects are allowed during the syntactic derivation, but not in the final representation, due to the interface needs. In this regard, Chomsky (2013) proposes that all labeling takes place at the phasal level, at the point of Transfer to the interfaces (since it is the interfaces that need labels). Despite Chomsky seeing labeling in both {H,YP} and {XP,YP} occurring at the point of Transfer (i.e., phasal spell-out), there is evidence that their points of occurrence should be different.

Bošković (2016a) observes that without labels, it is not even possible to determine the points of Transfer. Transfer, or spell-out, is determined by phases. Phases, in turn, are determined by

labels. We cannot know whether an object is a phase without knowing its label, therefore, phase determination requires labeling (phases are objects like CP and DP). At least some labeling then must be done prior to spell-out since without it we may never reach the spell-out point. In particular, Bošković argues that head-labeling (i.e., labeling in the head-complement configuration) is necessary to determine spell-out points. In other words, Bošković notes that the issue in question can be resolved if the result of a head-phrase merger can be labeled immediately.

Head-complement merger should be treated differently also due to reasons of subcategorization, i.e., syntactic selection. Subcategorization is a syntactic requirement to take a complement of a certain kind. But for a head to take a complement, the head must label the resulting object, otherwise there would be no head-complement relation. As a lexically-determined syntactic requirement, subcategorization is checked when the relevant element enters the derivation. This also motivates immediate labeling in the head-complement configuration. This is thus another reason to treat the head-complement merger differently regarding labeling.

We are now ready to return to the diachronic trend to lose *wh*-movement discussed in section 2. We have just seen that head-phrase configuration is preferred from the standpoint of labeling. This is because phrase-phrase merger requires some additional operations and therefore can be considered to be computationally more costly. Since labeling in a head-complement configuration is straightforward and less costly, and indeed can be taken to be a mechanical execution of the LA mechanism, we may expect that there exists a preference in the grammar to maximize the more efficient head-complement merger and minimize the dispreferred phrase-phrase configuration. Notice that *wh*-movement, in fact all Spec-head configurations, involve merger of two phrases. At the point of wh-movement, in e.g., (46), what we are merging is a *wh*-phrase and a CP (the {C, IP} merger is labeled as a CP).

(46) What gift did John buy?

The diachronic change discussed in section 2 is, therefore, operating in the direction of losing the {XP, YP} merger and instead resorting to the preferred head-complement structure. Its effect is thus to maximize the head-complement configuration and to minimize the Spec-head structure.

Diachronic loss of wh-movement is thus the loss of specifiers. It can be explained by the pressure imposed by the labeling algorithm on syntactic structure, and as such it is a structural change that occurs due to this pressure. The labeling preference for head-complement configuration can thus be applied to the observed tendency to lose specifiers and movement that creates these specifiers.

This also explains why we see no counterexamples to this trend, i.e., we don't have diachronic cases of gaining wh-movement, but only losing it. The former would involve adding a phrase-phrase merger, while the latter eliminates it.[11]

2.4.Other Ways to Lose Specifiers

The loss of wh-movement is just one consequence of the loss of specifiers driven by the pressures triggered by labeling requirements. The dispreference for phrase-phrase mergers, {XP,YP}, and at the same time, Spec-head relation, may also bring other changes and structural modifications. This section discusses other ways of losing specifiers, such as reanalyzing the material in the

[11] This is supported by acquisitional data indicating that children actually avoid producing displaced structures and have problems with syntactic movement that creates Specs (Lebeaux 1988, Radford 1990,1996, Labelle 1991, Brattico & Saikkonen 2010, Waldmann 2012). See chapter 4 of the present work for more on language acquisition in this respect.

specifier as a head, or just ceasing supporting it. After all, in the system described there is a pressure to lose a specifier, regardless of whether they involve movement or not.

Therefore, some cases discussed here will indeed still involve movement, whereas others will show both structural reanalysis and the loss of movement. Crucially, all will involve the loss of specifiers. I will show that all of them can be captured in the same way. This will be shown to be an advantage of the current analysis over some others (to be noted below) that cannot account for the full variety of cases discussed here.

2.4.1. No Spec but Movement: Serbo-Croatian *li*

Sometimes a specifier gets lost because a syntactic head just stops supporting it. A case of this kind is attested in Serbo-Croatian (SC) focus/Q clitic *li.*

This second position clitic can be used in yes/no questions that also involve emphatic focus. It is located in C and cliticizes to the element preceding it. Importantly, *li* in Serbo-Croatian cannot appear with unambiguously phrasal elements (compare (47a) and (48a) vs. (47b) and (48b)).

(47) a. *__Čiju ženu *li*__ (Petar) voli?
whose wife Q Petar loves?
'Whose wife does Peter love?'

b. __Koga *li*__ Petar voli?
whom Q Peter loves
'Who does Peter love?' (SC; Bošković 2001:26-27)

(48) a. * Skupe knjige ***li*** (Ana) čita ?
expensive books Q Ana reads
'Does Ana read expensive books?'

b. Skupe ***li*** knjige (Ana) čita ?

expensive Q books Ana reads

'Does Ana read expensive books?' (SC; Bošković 2001:27)

For this reason, this clitic is sometimes described as defective (see Bošković 2001). This defectiveness is seen in its inability to support a specifier. Only a non-phrasal element can precede *li* in SC since such elements are ambiguous head/phrase, hence they can also adjoin to *li* (instead of being in a Spec).

In this respect, Bošković (2001) discusses the different behavior of *li* in Serbo-Croatian and Bulgarian. Interestingly, the one-word restriction of SC *li* does not hold in Bulgarian.

(49) Novata kola ***li*** prodade Petko?

new-the car Q sold Petko

'Did Petko sell the new car.' (Bulgarian)

This indicates that Bulgarian *li* is still located in the specifier of CP, whereas in SC, this option is lost.

This observation is supported by the fact that SC *li* cannot be used in sluicing, which requires Spec-head agreement, i.e., Spec-head configuration (Lobeck 1995, Saito and Murasugi 1990). If *li* can only appear in the head-adjoined contexts and cannot support the specifier any more, the impossibility of sluicing is explained.

(50) Vidi nekoga. * Koga ***li*** vidi ?

sees someone Whom Q sees

'He sees someone. Who?' (SC; Bošković 2001:26-27)

In contrast to SC, Bulgarian *li* does not have such a restriction and allows sluicing. Together with the availability of complex *wh*-phrases with *li* in this language (51), this shows that the structural status of this head is different in the two languages, Bulgarian *li* still supports a specifier, unlike SC *li*, which lost this ability; i.e., it lost the specifier.

(51) Novata kušta ***li***? Kogo ***li ?***
new-the house Q whom Q
'The new house? Whom?' (Bulgarian)

In accounting for the grammatical status of this focus clitic in SC, it is crucial to notice that this construction has a very archaic feeling for the Serbo-Croatian speakers. This suggests that this element is disappearing from the language and the first step on its way out is the loss of the ability to support a specifier. As such, *li* and its loss of the specifier, can be considered an example of gradual disappearing of certain constructions by firstly getting rid of the frail and dispreferred configuration of the phrase-phrase merger; that is, a specifier and the Spec-head configuration.

This is confirmed by Polish, where *li* is in fact no longer productive and used. Today, we have only traces of its existence in some older Polish literature. Below are Old Polish examples:

(52) a. I rzecze Jakob: 'Zdrow-***li*** jest?'
and says Jacob healthy-Q is?
'…and Jacob says:'Is he healthy?' (*Queen Sophia's Bible*; 15th c.; XXIX: 25)
b. 'Jest-***li*** Pan w nas, czyli nics?'
is-Q god in us, or not
'Is god in us or not?' (*Queen Sophia's Bible*; 15th c.; XVII: 21-22)

The case of focus clitic *li* in Serbo-Croatian can then be looked at as another instance of the historical tendency to lose of a specifier, as its existence involves the dispreferred configuration of the phrase-phrase merger.

Serbo-Croatian diachronic loss of a specifier still involves syntactic movement. Movement is in fact necessary since *li* is an enclitic, hence needs some element in front of it to support it. However, there are also cases where we do not even need, and in fact we do not have a movement anymore at all. The next two subsections discuss cases like this.

2.4.2 Loss of Spec and Movement = Spec > Head: Early Germanic *hwæt*

Early Germanic *wh*-phrase *hwæt*, which e.g., appears in Old English where it opens *Beowulf,* can be analyzed as a head, rather than a phrase.

(53) **Hwæt** we Gardena in geardagum Þeodcyninga Þrym
Hw. we Spear-Danes.GEN in year-days.DAT nation-kings.GEN power.ACC
gerfunon hu ða æÞelingas ellen fremedon
heard how then/those.NOM princes.NOM valour performed
'We truly know adbout the might of the nation-kings in the ancient times of Spear-Danes how princes then performed deeds of valour.' (Beowulf 1-3; Bammesberger 2006:3; Walkden 2014:122)

Importantly, *hwæt* can be treated as a part of the process discussed here, with this expression, originally being a *full wh*-phrase that got reanalyzed as a head C, base-generated in this position. Brinton (1996) proposes that *hwæt* has followed the grammaticalization path analogical to *why* or *where,* from interrogative in direct questions to complementizer in indirect question, and finally

becoming a pragmatic marker, a particle, or interjection. In fact, this expression is often translated as *lo, so,* or *you know* (Brinton 1996, Walkden 2014).

Walkden (2014) proposes that *hwæt* in such contexts is an exclamative and argues that *hwæt* can be analyzed on par with other *what*-phrases appearing in 'a non-argumental' or 'exclamative' role (like 'why', 'how', or 'how much). Munaro and Obenauer (1999) discuss many cases of this kind, including Latin *quid* (54a), French *que* (54b) or Pagotto *cossa.* The list of such exclamations may include Dutch *wat,* or Polish *co* (54c). They are also attested in Arabic, Ancient Greek, Norwegian, or Celtic.

(54) a. **quid** venisti? [Latin]
what came.2SG
'Why have you come?' (Plaut.Am.1,1.209; Lewis and Short 1879, Walkden 2014:132)

b. Que ne partez-vous? [French]
what NEG leave-you
'Why don't you leave?' (Munaro and Obenauer 1999:208)

c. Coś tak szybko przyszedł? [Polish]
what.2.SG so fast came?
'Why did you come so fast?'

Therefore, this variant of *what* seems to be a common diachronic development cross-linguistically, also in non-interrogative contexts, but their phrase/head status may vary in individual languages (see Collins 1991, Munaro and Obenauer 1999, Ochi 2004 for different analyses of individual languages). Here, we will focus only on the exclamative variant of *what* appearing in Old English

or Old Saxon (Walkden 2014), where this *what* can be argued to be a base generated syntactic head C.

The source of this exclamation in Early Germanic is seen by Walkden (2014) as involving a reanalysis of discontinuous phrasal elements, such as Old Saxon *huat… guoudes* 'what good things' in (55) below, attested in e.g., in *Heliand* (Old Saxon, 1st half of the 9th century).

(55) **Huat** uualdand god habit **guodes** gigereuuid
hw. ruling G. has good.GEN prepared
'What **good things** Lord God has prepared (for us).' (Heliand 2533-4; Walkden 2014: 143)

The proposed change, i.e., a reanalysis of such discontinuous phrases as two independent elements, connects to the change of the phrasal status of this *what* into a C head. Support for this comes from the fact that *huat* cannot be coordinated with other *wh*-expressions in the language, has to be unstressed, and also crucially prevents verb movement to its usual V2 position. Instead, in these constructions, the verb appears in a later position, analogically to clauses with overt complementizers, as noted by Walkden.

As such, this element can be viewed analogically to modern English *how come* in an analysis like Collins (1991), where *how come* is argued to be an interrogative C head, due to the fact that this expression obeys locality conditions quite strictly (it must have a clause-mate relation with the

element it modifies), cannot appear in multiple *wh*-questions, and does not permit subject-auxiliary inversion on par with *hwæt.*[12]

We can assume, therefore, that *hwæt* underwent a structural reanalysis that resulted in the reduction of the {XP,YP} configuration of the previous stage, producing its non-argumental counterpart. This also reduces *wh*-movement, as *hwæt* is base-generated in its surface position. Brinton (1996) also argues that *hwæt* became syntactically fixed as a complementizer head in this sentence-initial position.[13]

Hence, the development of *hwæt* can be treated as a part of the process discussed here, as this expression, originally being a full *wh*-phrase, got reanalyzed as a C head, base-generated in this position, hence showing a reduction of the {XP,YP} configuration, which as we discussed, is also dispreferred from the standpoint of labeling, hence diachronically unstable. Early Germanic can then be taken to provide additional evidence for the existence of this preference for head-complement configuration, and dispreference for Spec-head merger. The analysis given in section

[12] Additionally, 'what' in the exclamative function discussed here cannot be used in sluicing (also in fragment answers) in many languages (see also Munaro and Obenauer 1999), which can be taken to confirm its status as a head without a specifier, analogically to SC *li* in (47).

i) Wiem że on szybko przyszedł ale nie wiem ***co** / dlaczego [Polish]
know that he quickly came but neg know what/ why?
Int:' I know that he came fast, but I don't know why

[13] Interestingly, a similar process of the loss of a specifier may be attested in contemporary American English in a class of 'aggressively non-d-linked' exclamatives of *what-the-hell* type, where the specifier containing a *wh*-phrase is often omitted. The lack of inversion in such cases can be taken to confirm this expression as the head. This would then be another instance of Spec>Head reanalysis of this sort we have discusses in this section. (I will leave more detailed discussion of this construction for future work.)

i) The fuck you do that for ? (*Triquarterly* (Online): 2016)
ii) The fuck you are talking about? (BK Standoff: 1995)
iii) The fuck you doing,… (Mov:Heist: 2001) (*Corpus of Contemporary English COCA*)

3 can be applied to such cases as well: the previously moved element is still located in the projection where it used to move to, but it is now base-generated in the head position.

2.4.3. Loss of Spec and Movement = Spec > Head: Welsh Agreeing Complementizers and Other Complementizers

Diachronic literature, in fact, provides many more examples of a change where phrasal elements get reanalyzed as part of the syntactic head. Some of the cases discussed involve, e.g., changes from subjects and objects to agreement markers (van Gelderen 2011, 2018), demonstratives in the specifiers of DP being reanalyzed as heads (Lyons 1999, Anderson 2005, van Gelderen 2011), or the emergence of declarative complementizers from interrogative *wh*-phrases, as in Georgian (56a) (Harris & Campbell, 1995; this change is in fact very common cross-linguistically), or from relative pronouns (via relative particles), as in Germanic (56b) (Axel-Tober, 2017), Greek (Roberts & Roussou 2003), or Slavic (56c) (Meyer, 2017).

(56) a. da ara unda, **raytamca** icna vin
and NEG want.PRES.3SG that know.PRES.3SG someone
'and he did not want that anyone know' (Georgian; Harris & Campbell, 1995: 298)

b. Er tháhta odowila tház \ thaz er ther dùriwart wás,
he thought maybe that.ACC that he the gatekeeper was
'Maybe he thought that he was the gatekeeper.' (Old High German; Axel-Tober 2017)

c. se že reče o dsě **eže** xotěaxǫ prijeti
this PTCL say.AOR.3SG about spirit.M.PRP that.N.ACC shall.3PL receive
i verujǫšti
him.M.ACC believing.PL
This he said about the spirit which those who believed in him should receive.' (Old Church Slavonic; Meyer, 2017:100)

Slavic provides many instances of this type, as e.g., Czech and Polish *that*-complementizer (Pol: *że/iż*) emerged as an effect of reanalysis of Old Church Slavonic relative pronoun *i-že/jaže* and subsequently relative particle *iže/eže.* Also, Russian subjunction marker *čto* 'that' emerged via a relative pronoun > relative particle grammaticalization path (Meyer, 2017).

What is important, is that we can see all these changes where phrasal elements change into heads in a broader context, i.e., as the reduction of specifiers.

In this context, Willis (2007) shows that there is reanalysis of preverbal subject pronouns *mi/fe* as main clause complementizers in Welsh, which can be treated quite straightforwardly from the perspective of the loss of specifiers.

(57) a. ...**mi** a 'th rodaf di y
1.SG.IND.SIM PRT 2.SG.ACC put.PRES.1.SG 2SG.AFF.SIM in
'm lle i yn Annwuyn
1.SG.GEN place 1.S.AFF.SIM in Annwfn…
'… I shall put you in my place in Annwfn…'(Middle Welsh; Willis 2007:440)

b. **mi** arhosais (,) **fi**

1SG.IND wait.PAST.1SG 1SG.IND

'I waited, me.' (Early Modern Welsh; Willis 2007: 459)

c**. mi/fe** welais i 'r gêm

PRT/PRT see.PAST.1SG I the game

'I saw the game.' (Modern Welsh; Willis 2007:435)

This reanalysis, ultimately leading to the loss of the specifier, proceeded in several steps, including the loss of preverbal particles, reinterpretation of the expletive element in the subject position as a complementizer, and subsequent extension of this analysis to all other preverbal subjects. The appearance of pronoun-doubling configuration (as in (57b)) and the reanalysis of the post-verbal pronouns as subjects there contributed to the ultimate reduction of the pre-verbal subject that used to be located in a Spec position as the head C in the language (see also discussion of Welsh particles in Roberts 2004).

The labeling account discussed above offers us a way to see all these cases as a part of a broader cross-linguistic pattern, having the same triggering mechanism; namely, the preference for head-complement configuration from the standpoint of labeling. This structure-building preference may have quite far reaching consequences and strong impact on language across time. It enables us to understand why the specifiers are getting lost and why certain syntactic movements got lost with them. But what is particularly important here is that the proposed analysis unifies the two cases of the loss of movement – the cases when the moved element gets reanalyzed as the head of the projection where it used to move, as in (57), and the cases where movement is lost in a way that previously moved element simply stays *in-situ*, as in the case of the loss of movement discussed in section 2.2.

However, obviously not all syntactic movement is lost. In the next section we will discuss why this is the case. We will also see that there is a Spec-involving configuration that is particularly resistant to a loss, which will also be explained from the perspective of the analysis proposed here.

2.5. Different Triggers and Specifiers

This section will offer a discussion as to why not all Spec-creating movement is lost. We have seen that there is a labeling-induced preference for the head-complement configuration over the Spec-head merger. We have seen that in some cases it leads to a loss of movement. The reason why not all movement gets lost is related to a bigger question: why do we have movement in grammar in the first place?

The existence of syntactic displacement is connected to the expressiveness of language (this is in fact, essentially the position taken in Chomsky 1995). Thus, the information-structure encoding often resorts to syntactic movement. Plainly speaking, we need movement to express certain things. Languages widely use movement to express certain semantic and pragmatic notions, such as topic, focus, or scope. We can relate this to interface-driven reasons. In a particular case of interrogatives and *wh*-questions, we can clearly see the interplay of interpretation-related issues with considerations of syntax. I have already discussed some of the observations related to the availability of certain interpretations depending on the syntactic positions of the relevant elements (e.g., the availability of single-pair vs. pair-list answers related to movement to [Spec, CP], as in French), but the list is much longer, and may include the interactions of placement of quantifiers and their scope or plain availability of interrogative interpretation in multiple *wh*-phrases.

Hence, regardless of how strong the labeling-triggered tendency to lose movement is, there is

also a requirement in language to express all these things, and movement serves as a useful tool in this respect.

The displacement may also be closely related to phonological or prosodic notions. Thus movement to *li*, discussed in section 2.4, can also be at least partially triggered by the need to support this enclitic. Regarding this effect, questions also arise here as to interaction of PF with narrow syntax. Works like Richards (2010, 2016) or Mathieu (2016) actually argue that grammar may be much more phonologically driven than we have previously thought. Therefore, movement in language can exist for many non-strictly-syntactic reasons.

Sometimes there could also be a formal reason to move, i.e., movement could also be driven by a requirement to check certain formal features for the derivation to converge. Movement, therefore, can be seen as a way to repair certain inadequacies during the derivation, and as such has a remedial function. Hence it is maintained in many cases.

Feature checking may also help us to capture another fact. As I have shown earlier, Latin was a multiple *wh*-fronting (MWF) language; MWF was lost in modern Romance languages. Spanish, French, or Brazilian Portuguese cannot front multiple question-words anymore and in fact are developing full *wh-in-situ*. The option of *wh*-fronting, albeit single, still exists in these languages side by side with *in-situ.* Based on Romance, (but we will also see evidence from other languages), we can see that multiple *wh*-fronting seems harder to lose, in the sense that we do not go straight from multiple *wh*-fronting to *wh*-in-situ. The path of the loss of *wh*-movement proceeds in stages: from multiple *wh*-fronting, through single *wh*-fronting, and finally to *wh-in-situ.* What is it that makes multiple movement more persistent? I will suggest that the answer to this question lies in the placement of the movement-triggering feature.

I assume that syntactic movement can occur (i.e., be triggered) either due to the feature of the target or due to the feature of the moving element (see also Bošković 1999, 2007, Nunes 2016). Hence the feature that drives syntactic movement, let us call it [uF] here as the precise nature of this feature is not crucial for our purposes, may be placed either on the target of the movement (i.e., its landing position), or on the element that undergoes displacement.

If the movement is driven by the feature of a target, it is enough for one element to move, to satisfy this requirement (e.g., in single *wh*-movement languages, where only one *wh*-phrase moves in questions). As the relevant feature is on the target head, the movement takes place as soon as the movement-driving feature enters the structure; the head is merged, and the movement-hosting specifier is created. This movement is easier to lose, as the existence of the specifier is closely connected to the target of the movement and these two are directly tied. In fact, in Chomsky (2000, 2001) the requirement in question is essentially stated as a requirement for a particular head to have a Spec (this is Chomsky's Edge feature).

On the other hand, when the feature in question (i.e., the movement triggering property) is placed on the moving element, all elements of this kind move, as all of them bear the feature in question. Hence, we are dealing here with the instances of multiple movement. Slavic multiple *wh*-fronting in questions is a great example of this feature-driving movement due to the [uF] placed on the moving element (see Bošković 2007). Diachronically, we have evidence that this movement is very persistent, too; it is harder to lose it. As the displacement-driving feature resides in the moving element(s), there is no direct relationship between this movement-triggering feature and the target of the movement (as discussed in in Bošković 2007, with the target-driven movement we have a requirement for a head to take a specifier, but with the movement-triggering feature on the moving element, there is no actual requirement to be a Spec).

Specifiers with multiple movement are harder to be reanalyzed and lost. This is a consequence of there not being a direct relationship between the feature that resides in the moving element and the head that supports the specifier.

Additionally, the change leading to the reanalysis discussed here would have to occur in the lexical property of the relevant lexical item. With target-driven movement, e.g., single *wh*-fronting, there needs to be a lexical change in the property of one lexical element, namely C_0. But with moving-element driven movement, e.g., with multiple *wh*-fronting, there has to be a lexical change in more than one such item. In fact, the whole class of the elements participating in this construction would have to undergo the change.

Furthermore, as discussed above, this change would have to target a lexical property which is not directly related to the creation of the specifier, unlike in cases of target-driven movement and specifiers, where the change of having or not having [uF] (what Chomsky calls the Edge property) on a certain head is immediately related to having or losing a specifier.

All this is supported by the observation that multiple movement is not lost immediately. This change is more gradual and does not go directly from multiple *wh*-fronting to *in-situ*, but involves an intermediate stage of a single *wh*-fronting, as we have seen in its loss from Latin to modern Romance.[14]

Finally, *wh*-movement may be motivated by semantics. Some languages simply cannot interpret *wh*-expressions *in-situ*, as variables, e.g, by unselective binding or choice function. This

[14] This could involve grammaticalization-like process for the features. Due to the displacement, the feature originally placed on the moving element could be reanalyzed as a feature on the target of the movement, which would mean that fronting of one element would suffice to license it. This can eventually lead to its loss altogether. This would also involve gradual loss of movement for the remaining *wh*-expressions in multiple *wh* constructions, which would lead to the loss of the relevant movement.

would mean that there is more than one way of interpreting *whs* crosslinguistically, i.e., as either moved or non-moved. We actually can see that not all *wh*-expressions have to move, even at LF (e.g. looking at island insensitivity); however, in some languages, displacement is still necessary for interpretation.

Cheng (1991) correlates this with the nature of the *wh*-phrases in a language. One important distinction she makes is connected to whether *wh*-expressions in a particular language have any inherent quantificational force. In *wh*-in-situ languages like modern Japanese or Chinese, *wh*-phrases are indefinites, and all quantification necessary for question semantics and the needed wide scope is provided by quantificational elements like question particles placed in the Force-encoding head in the left-periphery (which also perform the clausal-typing function discussed below). Hence, a Japanese particle *-ka* has a semantic value of an existential quantifier that quantifies over choice functions, applying to sets of individuals expressed by *wh*-expressions and selecting appropriate individuals in the context. No syntactic displacement of *wh*-words is needed in such a language.

On the other hand, in multiple *wh*-movement (MWF) languages all *wh*-phrases have to move. From the current perspective, this can be interpreted as indicating that they cannot be interpreted in situ. We have argued above that movement of multiple *wh*-phrases in these languages is triggered by a feature on the moving element (see Bošković 2007). The presence of this feature may be a formal reflex of them being unable to be interpreted in situ. Cheng (1991) actually argues that *wh*-expressions in these languages are really indefinites without quantificational force (as in *wh-in-situ* languages like Japanese), but they are still forced to move due to the uninterpretable feature each of these *wh*-words carry. At any rate, if *wh*-phrases in such languages cannot be interpreted in situ, the displacement in these cases can be ultimately connected to semantic interpretation.

Turning now to English, Cheng (1991) suggests that *wh*-phrases are inherently quantificational and they themselves have to undergo movement to establish an Operator-variable structure with the trace receiving interpretation of a variable ranging over the individuals. This movement to the left-periphery, to the projection encoding the Force of the utterance, is necessary for the interpretation. Under this view, even *wh*-phrases that stay in situ in English would be moving in LF. Alternatively, *wh*-phrases can be taken to be interpretable in situ in English, but one wh-phrase has to move for what Cheng (1991) refers to as clausal typing, i.e. to type a clause as a question. The clausal typing requirement is ultimately non-syntactic, so this approach would also fit well with the system argued for in this chapter. Another way to look at English in light of multiple *wh*-questions like *Who bought what* is that one *wh*-phrase has to front to the operator position in SpecCP to trigger *absorption* in multiple questions (in the sense of Higginbotham and May 1981). Again, in a system where *wh*-displacement is obligatory from the semantic point of view, it is plausible to assume that the relevant requirements got 'grammaticalized' in the form of an uninterpretable feature in the syntax (see Bošković 2007 regarding how the above suggestions regarding English can be implemented in these terms): movement is obligatory, hence syntax-semantics conspire to derive a correct interpretation by forcing the *wh*-word to move.

At any rate, the intuitive idea behind the discussion above should be clear: in English, fronting of a *wh*-phrase in questions is necessary for interpretive reasons. In languages like Japanese, *wh*-phrases can be interpreted in situ, and in multiple *wh*-fronting languages *wh*-phrases cannot be interpreted in situ, hence they all must move (the cases with movement may involve 'grammaticalization' of relevant requirements in terms of features that drive movement).

Now, under the system postulated here we may expect that since movement is dispreferred from the standpoint of labeling, all *wh*-phrases should acquire the *in-situ* property with time. For this to

happen, however, change in their interpretative possibility must take place, hence some languages may not lose syntactic movement in *wh*-questions because their *wh*-phrases would not be able to be interpreted without the displacement. Part of the reason for why the formal syntactic requirement for *wh*-fronting does not get lost is then essentially because of semantics then (this may not be the only reason, see e.g. Richards 2010, 2016 for potential prosodic reasons).

It does not mean, however, that change is not possible or does not happen. We have already seen many cases in which the loss of wh-movement did occur. In fact, this thesis is devoted to the discovery of this cross-linguistic trend to lose wh-movement (instead of gaining it). From the semantic perspective, this implies that in languages that ceased being *wh*-fronting languages, some change had to occur within the *wh*-phrase itself. Namely, at one point, the semantics of *wh*-words must have changed, and it became possible to interpret it in its base position (possibly with a quantification particle associated with the high scope taking projection in the left-periphery of the clause, possibly in a grammaticalization-like process).

Now, we can ask if it is possible to see this semantics-involving change as an ongoing process, albeit not completed yet. This is the issue of intermediate stages. We have in fact seen that the loss of obligatory syntactic movement proceeds in steps, i.e., from a multiple *wh*-movement to a single *wh*-fronting (before being completely lost and replaced by the *in-situ* option) as from Latin to modern Romance languages. We can ask if there can be other intermediate stages. Interestingly, Rudin (1988) and Bošković (2002) observe that there is a variation within multiple wh-fronting languages, with one group of these languages (e.g., Bulgarian) requiring displacement of all of its *wh*-words to SpecCP, and another group still requiring all wh-phrases to front but to a lower

projection (e.g. Russian).[15] Bošković (2002) gives a number of diagnostics that differentiate the two groups of languages, e.g. Superiority effects and the availability of single-pair readings in multiple questions. Above, I have suggested that the displacement of all *wh*-phrases in MWF languages is connected to their semantic interpretation. This would in fact hold for both types of MWF languages. From the current perspective, the distinction between the two types of MWF languages can be implement as follows: while in both types of MWF languages *wh*-phrases cannot be interpreted *in-situ*, in a language like Bulgarian, which moves all *wh*-phrases to SpecCP, they can only be interpreted in SpecCP, while in a language like Polish, they can be interpreted as long as they are located in an operator, A'-position: this does not have to be interrogative SpecCP. An interesting question then arises whether a situation of this kind can be found in non-MWF languages. We can imagine what this would look like: there would be *wh*-fronting, but not all the way to the Force position, i.e. SpecCP. In other words, there would be *wh*-fronting, but not as high as in English. Do we have any evidence of the existence of such languages? It seems we do. We have reasons to believe that Spanish can be considered to be this kind of language. What is important here is that Spanish actually requires *wh*-fronting in indirect questions (for the reason why this is the case, see Bošković 1998). Consider (58a-b) below:

(58) a. Me preguntaron que mi madre **que cuándo** podría venir
CL. asked that my mother that when could come
'They asked me when my mother could come.' (Villa-Garcia 2012:31)

[15] Note that Bošković (2002) differs from Rudin (1988) in arguing that the second group has obligatory wh-fronting without any fronting to interrogative SpecCP (regarding Russian, see also Stepanov 1998).

b. Hugo me preguntó **que** **cuando** veniamos

Hugo CL asked that when would-come

'Hugo asked me when we would come' (Spanish; Villa-Garcia 2019: 27)

(59) *Hugo asked me that when we would come. (English; Villa-Garcia 2019:27)

(58a-b) above illustrate an interrogative embedded clause with a *wh*-phrase appearing below the complementizer (see also Uriagereka 1988). Villa-Garcia argues that the unavailability of this placement of the *wh*-phrase in English in (59) is not related to the availability of recomplementation, i.e. double complementizer constructions in Spanish, hence it should not be related to the lack of the relevant functional projection in English that Spanish would have. English actually allows two **that**-s and a focal phrase following secondary *that*.

(60) I hope **that** when they are adults **that** AT NO TIME will they forget the work that their parents put into the education. (Villa-Garcia 2019:15 after Haegeman 2012:85)

Hence, the problem with English may lie in the placement of a *wh*-phrase in too low a syntactic position. The fact that Spanish allows its *wh*-phrase to be fronted to such an intermediate projection may indicate that its conditions on licensing and interpreting *wh*-phrases are indeed different from English. We have already seen that Spanish is undergoing a change into a *wh-in-situ* language. *Wh*-in-situ is now possible in direct questions, but not indirect questions. However, even indirect questions are different from those found in English, an obligatory *wh*-fronting language. The configuration in (58) can in fact be another sign of the change occurring in this language. The loss of obligatory *wh*-fronting in Spanish may therefore include the possibility of true *in-situ* in direct question and "intermediate" *wh*-movement in indirect questions. In fact, it is possible that we

dealing with an intermediate stage in Spanish indirect questions: *wh*-fronting is still needed, i.e. the *wh*-phrase cannot be interpreted in situ, but the fronting does not have to go to the highest projection, as it has to in English.

There is also an interesting difference between the interpretation-based movement in *wh*-questions and the *wh*-like-movement in relative clauses. These latter constructions involve consistent movement which I suggest is the case because of the predication relation that needs to be established by the operator movement. The movement in relative clauses is necessary to interpret the relevant relation (the operator needs to be close to the relative head). Hence, the reason why we don't see the loss of movement in this construction in e.g. Spanish and Chinese (see e.g. Huang 1982, Ning 1993, Aoun and Li 2003 regarding Chinese**)** is that this change could not be simply be reduced to the properties of the moving element, like in the case of *wh*-phrases. The operator (overt or covert) in relative clauses must be located in an A'-position so that the syntactic configuration in question can be interpreted.

At any rate, it should be noted that the present system does leave room for the possibility of gaining, or, re-appearing of specifiers. However, under the analysis presented above, such cases would be expected to be interface-driven and attributed to extra-syntactic factors, e.g. prosody, phonology, pragmatics, or information-structure.

One instance of this can be seen in the cyclical change in the expression of negative elements, referred to as Negative or Jespersen's Cycle (Jespersen 1917, van Gelderen 2009, 2011, 2015, Willis et al. 2013). The process involves weakening and strengthening of negative markers, resulting in their structural reanalysis. Jespersen's cycle is usually described as involving 3-6 stages. Below, I give the basic stages illustrated with the examples from the history of English and French:

(61)

	Stage I	> stage II	> stage III	> stage III'	> stage IV
English	ic **ne** secge	> I **ne** seye **not**	> I say **not**	> I do**n't** say	> I do**n't** say **nothing**
	(Old Eng.)	(Middle Eng.)	(Early Modern Eng.)	(Modern Eng.)	(Colloquial Eng.)
French	je **ne** dis	> je **ne** dis **pas**	> je dis **pas**		
	(Old French)	(Middle & Modern written French)	(Colloquial French)		

(Based on Willis, Lucas, and Breithbarth 2013:7)

In stage I, the negative marker expresses predicate negation on its own. Stage II of this cycle involves the weakening of the original negation, as a result of which an additional negative element is introduced. As its function is to strengthen the weakening negation, it is usually merged as a specifier of the negative head. The creation of this specifier is therefore closely related to interface considerations: as a response to phonological erosion of the original negative marker and due to the need to support the semantic and pragmatic expressiveness of the negation. As predicted by the present analysis, many languages go through the next stage, where the newly created specifier is lost, and the negative head becomes the sole negative marker again (stage III).[16]

In fact, van Gelderen (2015), analyzes this stage in Jespersen's Cycle appealing to Chomsky's (2013, 2015) Labeling Algorithm, as well. She proposes that diachronically, specifiers tend to be reanalyzed as heads when the feature-sharing configuration is unavailable. She, therefore, sees this step of losing a specifier as an instance of the Spec > Head reanalysis due to a problem with labeling. Namely, she suggests that in cases like this, the main problem is that the head (here, the

[16] It is worth mentioning that (re-)introduction of specifiers as in stage II does not have to be the only strategy available in Negative Concord languages. Afrikaans uses another head, higher in the structure, as a concord element, contributing to the expression of negation. Biberauer (2009) shows that this higher head is really a C-related Polarity head, whose presence does not have to be related and does not depend on the presence of NegP, hence showing grammaticalization-connected semantic bleaching.

Neg) and its specifier do not share a prominent feature, and as a consequence the label of this merger cannot be determined. Hence, the above-mentioned reanalysis occurs.[17]

We have, however, seen that this stage can be explained by the labeling considerations, with which we are already familiar. The proposal in van Gelderen (2015) can, in fact, be subsumed by the analysis proposed here, i.e., under the head-complement preference in labeling. The account of van Gelderen (2015) really needs two components: the principle of syntactic heads being more economical than phrases, as she proposes (her Head Preference Principle (HPP); van Gelderen 2011), and the loss of feature sharing. But once we assume that feature-sharing is in principle dispreferred, we really do not need anything else. The loss of feature sharing does not need to be additionally stated and motivated. The reason why the change happens is not that we lost feature-sharing, but the actual preference for a {H,YP} merger when it comes to labeling. Therefore, van Gelderen's HPP can be deduced; it does not need to be stated additionally. It should be noted that the relevant elements that are introduced in stage II are very often non-branching, which means that they are ambiguous with regard to their head/phrase status in Bare Phrase Structure; they are then ambiguous enough to be analyzed as a head by the learner.[18] The loss of specifiers in such cases is therefore not surprising.

At any rate, we can observe how interface-related consideration may influence the structure-building process, and even override for some time certain grammatical and computational considerations, such as those connected to labeling. This observation is shared by both van Gelderen (2015) and the present analysis. Both stress that the process of gaining specifiers is driven

[17] Van Gelderen (2018) also argues that head-complement is preferred to feature sharing.
[18] Under the assumption that language change is driven by language acquisition (Andersen 1973, Lightfoot 1991, van Gelderen 2009, 2011, Roberts 2007; see also chapter 4).

by forces outside narrow syntax, such as prosody and pragmatics (see also Kiparsky & Condoravdi 2006, L'Arrivée 2010).

However, van Gelderen's (2015, 2018) observations about the labeling-motivated language change and structural reanalysis cannot account for all the diachronic processes discussed here. More specifically, van Gelderen's Head Preference Principle: 'Be a head, rather than a phrase' (van Gelderen 2015:7), is not obviously relevant to the cases of the loss of *wh*-movement. This is because the relevant element is a phrase in the *in-situ* position, too. The present work, on the other hand, offers a unified analysis of the loss of *wh*-movement and the reanalysis of the material in the specifier as a head.

The following sections will show that the labeling account can offer us a deduction of even more diachronic changes, such as the loss of head movement, the change from the OV to the VO word order, and the loss of what appears to be right-adjunction. These processes had to be treated differently in Roberts & Roussou (2003), but I will show that they can also be deduced from the preference for head-labeling, which will offer us a common thread between the traditional grammaticalization and other diachronic changes, where the same forces may be at play in all of them.[19]

2.6. The Loss of Head-Movement

I will start by showing how the labeling requirement and the mechanics of the labeling algorithm may affect syntactic head movement. We will see that cases that are affected by diachronic change

[19] This is not to say that they have to be the same in all respects (see Roberts & Roussou 2003). I am only concerned here with *why* these changes happen.

are those which are problematic for labeling, which should not surprise us in light of the discussion in the previous sections.

The head-head merger of {X, Y} (where both heads have categorial specifications) raises the same kind of an issue for LA as the symmetrical merger of {XP, YP}. Just like in the case of a merger of two phrases, the result that minimal search returns in this case is also too ambiguous to determine the label for the whole merger. And again, just as with the phrase-phrase merger, something additional has to happen, either Agree or movement.[20] As a result, diachronic changes that would reduce the problematic head-head configuration may be expected, given the approach to labeling discussed above.

One way to reduce this dispreferred configuration is to eliminate head movement that creates such merger of two heads. And indeed, cases of the loss of head movement are well attested in the diachrony of many languages.

One example of such change is the loss of V-to-T movement, as e.g. in English (Roberts 1993) or Mainland Scandinavian (Heycock & Wallenberg 2013), [21] or the loss of the verb-second (V2) (which involves the head movement to C), as e.g. in English (van Kamenade 1987, Kroch & Taylor 1997, Pintzuk 1991, Haeberli & Ihsane (H&S) 2016), French (Adams 1987, Roberts 1993, Vance 1997, Labelle & Hirschbühler 2017, Wolfe 2019), and other Medieval Romance languages (Fontana 1993, 1997, Ledgeway 2008, 2017, Wolfe 2019), or Welsh (Willis 1998). [22]

[20] For the feature-sharing and Agree, as well as various movement options for the head movement case, see e.g. Roberts (1993, 2010), or (Saito 2016).
[21] This includes the change in progress with the loss of this head movement in Faroese (Heycock et al. 2012, Heycock & Wallenberg 2013).
[22] I will discuss V2 in more detail in chapter 3 of the present thesis, putting it aside here.

Indeed, English was able to move the finite verb to T (similarly to French today) until around 1600 CE. Example (62a) below illustrates this with the tensed main verb preceding the negation, and (62b) shows verb raising out of VP (the edge of which is marked by the adverbial) in subordinate clauses (the example (62b) involves a context resisting V2 in Old English (OE)). Similar examples are attested in early ME). Head movement to C was attested as well, as illustrated by (62c). Finally, English displayed object shift (62d), for which, according to Holmberg's generalization (Holmberg 1986: 176, Chomsky 1995), movement of the verb outside of VP is a prerequisite (see e.g., Warner 1997).

(62) a. If I **gave not** this accompt to you.
'If I didn't give this account to you.' (1557: J.Cheke, Letter to Hobby; Roberts 2007:57)

b. hu se deofol **beswac** syððan eft þa men
how the devil deceived afterwards again the men
'how the devil then deceived mankind again' (colwgeat, + ALet_6_[Wulfgeat]:59.22; H&I 2016: 504)

c. What **menythe** this pryste?
'What does this priest mean?' (1466-7:Anon, from J. Gairdner (ed.); Roberts 2007:57)

d. They **tell vs** not the worde of God. (1565: Thomas Stapleton; Roberts 2007: 57)

These cases of the loss of head movement are not surprising from the current perspective. Since the head-merger resulting from head-movement must be labeled by feature sharing or

excorporation, the present analysis, which disfavors these options, can capture the loss of head movement similarly to the to the loss of phrasal movement discussed earlier.

The movement of the main finite verbs is absent in English since at least the 16th century. Interestingly, the movement of the main verb in English correlated with the existence of the relatively rich morphology. Roberts (1993) suggests that what matters in the availability of verb movement is agreement morphology and cites (after Mossé 1968) the paradigms below to show that the present tense in Middle English was indeed rich. Note the agreement ending for grammatical person in the paradigm and a clear distinguishing number marking.

(63) 1sg: sing**e**

2sg: sing**est**

3sg: sing**eth** (south) /sing**es** (north)

1,2,3pl: sing**en** (midland)/ sing**eth** (south)/ sing**es** (north) (Roberts 1993:256)

However, this agreement morphology was lost between the Middle English and Early Modern English. The 1sg marking was lost before 1500 and the plural marking disappeared completely early in the 16th century, as well. With the loss of the overt expression of morphological agreement, the verb raising disappeared as well. Kroch (1989) gives the dates of 1550-1575 as the crucial years of the attrition of the main verb movement together with the loss of agreement paradigms, the rest of the 16th century being the transition period, or change in progress, where both grammars (with the verb movement and without it) coexisted in the speech of the language users.[23]

Similarly to English, Mainland Scandinavian languages used to be robust and consistent in

[23] Haeberli & Ihsane (H&I) (2016) place the beginning of the decline of the verb movement even earlier, in the middle of the 15th century; they also see it as correlating with rich morphology though in somewhat broader terms.

the movement of the finite verb to the TP domain even in subordinate clauses. However, examples like (64) are unavailable in modern Mainland Scandinavian, as it lost obligatory V-to-T movement, leaving the finite verb *in-situ* in such contexts.

(64) æn min guÞ brytar eigh niÞar Þin guÞ
if my god breaks not down your god
'If my god doesn't break down your god' (Old Swedish, *Codez Bureanus*, c.1350; Heycock & Wallenberg 2013:129)

The V-to-T head movement was lost in Old Swedish by the late 17th century, which occurred parallel with the loss of agreement in this language. Similar situation was attested in Danish, where the loss of the rich Old Norse paradigm took place, as well. This resulted in a significant decline in the frequency of the head-movement in question, and ultimately in the loss of this head raising around 1500-1700, occurring simultaneously with other changes in word order in this language.

Hence, we see that the verbal morphology plays a role in the occurrence of the verb movement (Roberts 1993, Rohrbacher 1999). This can be captured in the analysis based on labeling.

The suggestion here is that the presence of the agreement entails the presence of the prominent feature shared by the elements merged together. The reader may remember from the discussion in section 3 that this feature can provide a label for the created configuration and the labeling requirement may be satisfied, albeit in the more costly manner than with head-labeling, i.e., with an additional operation of feature-sharing.

Now, feature-sharing may have morphological realization, but this does not always have to be the case. With the direction of change postulated here, i.e., we may generalize that the loss of

movement is more likely to occur if the morphological reflex of feature sharing is missing. In other words, when feature-sharing is manifested through agreement, the movement is harder to lose. Historically, the relevant morphology usually disappears first, and syntactic movement follows its path (Roberts 2007, Thráinsson 2003). Hence, movement tends to be lost faster when visible morphological indication of feature-sharing involved in the labeling of the object created by the movement is missing (see also Haeberli 2002).

Cross-linguistically we see many instances of this correlation between agreement and movement. One more example comes from Brazilian Portuguese.

Brazilian Portuguese (BP) differs from European Portuguese (EP) in a few important aspects which are related to the agreement system of both languages. Null subjects, for instance, are much more constrained in BP than in EP. Also, only Brazilian Portuguese allows hyper-raising; this option is not attested in EP.

(65) Os meninos$_i$ parecem que t_i fizeram a tarefa
the boys seem.3pl that __ did.3p. the homework (BP; Galves & Kroch 2016: 489)

Both properties of BP, according to Ferreira (2009), are consequences of incompleteness of phi-features of T in this language.

Additionally, Rohrbacher (1999) observes that Brazilian Portuguese differs from European Portuguese (EP) in terms of the inventory of distinctive marking for referential inflectional features. For instance, Brazilian Portuguese has lost distinctive marking on second person plural in the 'direct' addressing (*vós comprais* 'you'all sell'), today having only the 'indirect' form of address

for this person which is syncretic with the third person singular (*vocês compram)*. Hence, EP has distinctive marking in both 1st person and the 2nd person, while in BP 2nd person is never distinctively marked. This morphological impoverishment of BP comparing to EP has reflection in V-to-T raising. Namely, EP has a movement of the verb to the highest inflectional head, whereas BP doesn't. Rohrbacher cites a number of arguments to this effect, one of which concerns the availability of the order Subject-Adverb-Verb in BP (but not in EP or Italian), even if the subject is not dislocated or topicalized. At any rate, the difference between the verb raising possibilities between European Portuguese and Brazilian Portuguese reflects the difference between agreement morphology in both languages.

Moving beyond V-to-T movement in BP and its morphological connotations, Figueiredo Silva & Grolla (2016) report that head-movement to C in questions in BP also got lost. We have already seen in section 2.2.4 that this language is losing its *wh*-movement developing *wh*-in-situ strategies instead. Where *wh*-fronting occurs, however, head-movement to C is being lost, resulting in a change from Wh-V-S word order to Wh-S-V. The authors report a drop in this head-movement from over 45 percent in the 19th century to 12.5 percent in the 20th century.

Another point can be made here regarding agreement. Old Japanese *kakarimusubi*, which we have seen in section 2.2.1. was argued to involve a system of *wh*/focus agreement with the focus particles (*ka, zo, ya*), hence the verb appeared in the adnominal form, which was a manifestation of this agreement (Watanabe 2002). Both the special verb form and the *wh*-movement system were lost by the fifteenth century (however a sharp decline in the frequency of wh-movement is reported to have started earlier, i.e., c. 12 century AD). Traditionally, the loss of movement was analyzed

as a consequence of the change in the morphological agreement (i.e. its loss).[24] This fits well with the suggestion made earlier, that the loss of movement is more likely to occur if the morphological reflex of feature sharing is missing.

In relation to that, we can make a more general claim: in some focus-movement languages, there is actually agreement that is overtly manifested. In the light of what we have seen in this section, such focus movement should also be harder to be lost. The same holds for *wh*-movement, as well as head-movement, as we have just seen: if it is accompanied by overt morphological manifestation, it should be more persistent historically. The loss of *wh*-fronting may be expected to be anteceded by the loss of the overt morphological reflex of the relevant feature-sharing.

Concluding this section, I have argued that the loss of head movement is another instance of the diachronic attempt to minimize the dispreferred and more costly labeling through feature-sharing and movement. Similarly to the loss of the Spec-head configuration, the dispreferred head-head merger can be reduced by the loss of the syntactic head movement, which creates it. I have argued that we are dealing here with the factors analogical to those involve in the loss of specifiers and *wh*-movement; they are related to the workings of the labeling algorithm. We have seen that the labeling-based analysis also enables us to shed light on the role of morphological agreement in diachronic changes that involve loss of movement.

2.7. Additional Consequences of the Loss of Movement

In this section I will examine two cases of diachronic change which can be accounted for in the current system if Kayne's (1994) approach to word order is adopted. The cases in question concern

[24] See, however, Watanabe (2002) (see also Narrog 2019 for an overview of the relevant literature).

the change in word order from the OV word order to the VP word order and the loss of traditional rightward adjunction.

2.7.1 Change from OV to VO Word Order as the Loss of Movement

This section discusses a diachronic change in word order from OV to VO, which can be, and in fact has been, analyzed from the perspective of the loss of movement. The labeling account proposed here helps to add an additional dimension to this discussion, as it allows us to understand what is behind the issue in question from a broader perspective.

It has been observed that there is a diachronic change between the head-final (OV) and head-initial (VO) word orders, and interestingly, the change appears to be unidirectional, i.e., change from OV to VO order is fairly common cross-linguistically, whereas change in the opposite direction is not attested, at least not within the Indo-European family and e.g., Finno-Ugric (see Kiparsky 1996, Biberauer et al. 2009a,b, 2010).We will argue that the labeling account presented here may shed some light on this fact.

What is relevant for us is Kayne's (1994) *antisymmetric* approach to word order, according to which Subject-Object-Verb (SOV) word order is really derived from the 'universal' Subject-Verb-Object (SVO) word order by movement, e.g., by short movement of the direct object to some fairly low position in the extended projection of the verb (see also discussion in Zwart 1993, Roberts 1994, 1997, 2017, Holmberg 2000, Mahajan 2003, Biberauer & Roberts 2006, Jayaseelan 2010, Kayne 2014).[25] A number of authors have given empirical evidence for this movement.

[25] See section 1.1.5 in Chapter 1 for other variants proposed to account for these changes.

Thus Kiparsky (1996) observes that there is very often a strong adjacency requirement within rigid VO languages, in that they require the verb to be strictly adjacent to the nominal complement. This is in contrast to rigid OV languages, where adverbs my freely intervene between the complement and the verb. The movement approach to the OV word order captures this situation, since there can be some projections between the verb and the landing site of the object movement. The authors cited above also give other arguments to this effect.

In this context, the loss of the OV word order is simply the loss of movement that created this configuration. In fact, many authors have adopted Kayne's (1994) idea to this word order change, proposing that this change is a consequence of a loss of movement (see Kiparsky 1996, Roberts 1997, van der Wurff 1997, 1999, Fischer et al. 2001 for English, Hróarsdóttir 2000 for Icelandic, or Ledgeway 2012 and Danckaert 2012 for Latin). [26]

At any rate, what matters for us is that the diachronic change in question is unidirectional (Kiparsky 1996, Biberauer et a. 2009 a,b, 2010). This is not surprising given the current perspective, which captures the OV>VO word order change as a loss of phrase-phrase merger, which is disfavored by the LA. We then have here at work he broader tendency for phrasal movement and disfavored Spec-head configuration to be reduced.

A question also arises here why this change does not always happen. What is relevant in this respect is that we are dealing here with a much broader tendency: change from head-initial to head-final order in general is extremely rare (Kiparsky 1996, Biberauer et al. 2009a,b, 2010, Roberts 2017), which is not surprising from the current perspective.

[26] See Roberts 2007 for overview of different proposasl regarding OV>VO change in English, including typological and other generative accounts.

It should be noted here that the fact that change in the direction of the head-initial order is prevailing cross-linguistically is also not surprising from the perspective of the *Final Over Final Condition* (FOFC) (Biberauer et al. 2014; see also Sheehan et al. 2017). Briefly, FOFC offers an account of some typological gaps in the attested typology of word-orders, e.g., the lack of head-final complementizers in head-initial languages. Under FOFC, this is connected to the ban on having head-initial structures dominated by head-final projections within the same extended projection; therefore head-final phrases can only (immediately) dominate other head-final projections. FOFC can then block an OV-to-VO word order change from occurring. Thus, FOFC would e.g. block such a change in a language with final complementizers. [27,28]

Another factor may be relevant here. Following Bošković (2019b), I suggest that a language will not undergo shift from OV to VO is if the SOV order is really a canonical order in this language, i.e., if it's the most default and unmarked option, used e.g. in a discourse-neutral way to answer a question like 'What happened?' In this respect, it should be noted that 'canonical' word order does not immediately entail base-generation, as languages may have a derived word-order (i.e., involving movement) as their canonical word order (see here Kayne 2010 and Dryer 1992;

[27] It may be worth noting that in in many head-final languages, complementizers are often verbal affixes (this is e.g. the case in Japanese, where inflectional elements are also affixes); PF considerations may require head-finality in such cases given that a verbal affix must be adjacent in PF to the verb (other verbal affixes, which form a prosodic word with the verb, can intervene).

[28] Regarding diachronic change, Biberauer at al. (2009a,b) argue that change from head-final to head-initial structure has to proceed top-down. Hence, in order to avoid FOFC-violating configurations (i.e., head-final head dominating head-initial mergers) the change has to start from the highest phrase within the extended projection. As head-initial phrase can dominate head-final mergers, the change can proceed further towards the bottom in a unified fashion. Change in the other direction (i.e., change of head-initial configurations into head-final configurations), on the other hand, has to proceed bottom-up, i.e., in order to comply with FOFC, the change has to start from the lowest projections and then target all the higher ones, thus avoiding head-final phrasing dominating head-initial phrases (Biberauer et al. 2009b and Roberts 2017 also observe a rare instance of VO to OV change in Ethiopian Semitic languages and argue that the change is actually FOFC-motivated/constrained.

they also discuss cases of languages with canonical word orders involving S-O-Neg-V, which would involve movement even without Kaynean perspective).

Finally, note that in English the OV to VO change took place gradually from around 10th to 12th century (Kiparsky 1996, Roberts 1997, McFadden 2004) with periods of variation in between these dates and certain elements, e.g., negative and quantificational DPs, favoring pre-verbal placement for much longer time than other semantically neutral elements (Pintzuk 1991, Kroch & Taylor 1997, 200). This can be interpreted as indicating that the movement of these elements that resulted in their preverbal placement (yielding the SOV order) was actually semantically motivated (there are many claims for semantically motivated movements in this part of the structure), which would make them more resistant to change in the current system, as discussed above.

At any rate, we can understand the word-order change under consideration in this section under the current labeling-based approach by adopting Kayne's (1994) *antisymmetric* theory, (which also gives us a FOFC-obeying account of the more general diachronic path here).

The current labeling perspective on attested diachronic changes also provides us with a broader perspective on the OV>VO word order change, as well as the overwhelming unidirectionality of this change, in a way that deduces it in a systematic fashion from the dispreference for the phrase-phrase merger. In other words, we are dealing here with a much broader tendency for phrasal movement and disfavored Spec-head configuration to be reduced diachronically.

2.7.2 The Loss of 'Rightward Adjunction' as a Loss of Movement

This section illustrates the application of Kayne's (1994) *antisymmetrical* approach to yet another language phenomenon, i.e., adjunction. In a syntactic theory like Kayne (1994), adjuncts which are traditionally considered to be right-adjoined appear in this position as a result of movement of

other elements to their left. This is so because in Kayne's system there cannot be any rightward adjunction or rightward movement but only leftward-adjunction and leftward movement. Kayne's proposal is appealing theoretically due to its constraining nature. The discussion in this section can be seen as providing evidence for it in that this approach gives us an explanation of certain phenomena that other theories cannot provide. In particular, I will argue here that diachronic data provide us with relevant evidence, where Kayne's analysis of adjunction allows us to account for certain attested instances of language change, hence its advantage over the competitors. Importantly for our purposes, we can see the relevance of Kayne's theory and its explanatory ability once we adopt the labeling-based approach to diachronic reanalysis proposed here. Thus, the two theoretical approaches together will allow us to shed light on certain historical facts.

Under Kayne's proposal, traditional rightward-adjunction movements like Heavy NP shift are analyzed as involving leftward movement of the element that precedes the traditionally rightward-moved element. In other words, traditionally rightward-adjoined elements are derived by leftward movement of elements to their left. Let us consider from this perspective adverb ambiguities like those in (66a) and (66b) below, originally discussed by Andrews (1983).

(66) a. John intentionally twice knocked on the door.

Only possible scope: intentionally > twice = *one intention, to knock twice*

b. John knocked on the door intentionally twice.

Only possible scope: twice > intentionally = *two instances of intentional knocking*

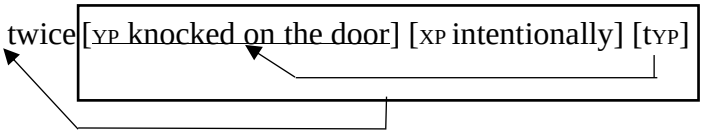
In (66a), *intentionally* must take scope over *twice*. What is interesting here is that in (66b), the adverb to the right (*twice*) takes scope over the adverb to the left (*intentionally*). Cinque (1999) argues that a sentence like (66b) where t*wice* takes scope over *intentionally*, is derived from a

structure like (67b) below, with the movement of YP (*knocked on the door*) across the adverbial *intentionally,* which is followed by the movement of the resulting XP (*knocked on the door intentionally)* across *twice.* The derivation is illustrated in (67c):

(67) a. John knocked on the door intentionally twice. (=twice>intentionally)

b. John (twice) [XP intentionally[YP knocked on the door]]

c. John twice[YP knocked on the door] [XP intentionally] [tYP]



This sequence of two movements captures the scopal relations where *twice* scopes over *intentionally.*

Derivation of seemingly rightward-adjoined elements in a system like Kayne (1994) or Cinque (1999), therefore, involves movement, which as I argue here, is a dispreferred option from the standpoint of labeling. Thus we may expect that rightward adjunction should be diachronically fragile. Thus, just as movement tends to decline, in the same way postverbal adjuncts may diminish diachronically.

This is in fact attested. Djamouri et al. (2013) discuss changes in the distribution of adjunct phrases in Chinese. What is relevant is that such phrases were possible possible in both the post-verbal (68a) and pre-verbal position (68b) in pre-Archaic Chinese (Shang inscriptions 14th -11th c. BC) (68c) illustrates both of these options in one sentence). Today they occur exclusively in the preverbal position in Modern Mandarin (since approximately 3rd century AD).

(68) a. wǒ fēi ài qí cái èr yí zhī **[yí yáng]** yě

1SG NEG cherish 3SG value CONJ replace 3SG with sheep PART

'It is not that I attach a great importance to its value [i.e. the value of the ox] and therefore replaced it with sheep.' (*Mengzi*, Liang hui wang, 4th-3rd c. BC; Djamouri et al. 2013: 586)

b. wáng **[jī dīngsì]** chū

king actual *dingsi* go.out

'The king on this *dingsi* day goes out.' (*Heji 07942:* Djamouri et al 2013:584)

c. qí yī **[yì rì]** qí yán zūn **[yú shì]**

FUT sacrifice next day FUT continue offer at temple

'After having performed a *yi* sacrifice, the next day one will continue to make offerings in the temple hall.' (*Heji 30373;* Djamouri et al. 2013: 587)

To spell out this situation in a little bit more detail, adjuncts in pre-Archaic Chinese, unlike arguments which appear only post-verbally, appear in three positions: preceding the subject, between the subject and the verb, or postverbally (here, however, only one adjunct is permitted) This contrasts with Modern Mandarin, where adjuncts can only appear before the verb; postverbal adjuncts are ungrammatical (69).

(69) a. **tā (yě /měi-tiān / chángcháng)** lái (* **yě/ *měi-tiān / *chángcháng**)

3SG also/ every-day / often come (also/ every-day / often

'He also comes every day/often.' (Djamouri et al 2013: 590)

b. (**zài jiāli / báitiān**) tā (**zài jiāli/ báitiān**) xiūxí (***zài jiāli / *báitiān**)

at home/daytime he at home/daytime rest at home/ daytime

‘(Athome/daytime) he rests (at home/ daytime)

c. wǒ (gěi tā) dāng fānyì (***gěi tā**)

1SG for 3SG act interpreter for 3SG

‘I served as an interpreter for him.’ (Djamouri et al. 2013: 590)

Huang (2006), in fact derives postverbal adjunct phrases in Classical Chinese via VP fronting over the preverbal adjunct phrases. This movement does not occur in all cases and does not cross all adjuncts as we see some adjuncts still in the preverbal position. Certain adjuncts, e.g. *yǔn* ‘indeed’ are always situated above the vP, and this relates to their scopal properties in the clause.

(70) *yǔn* [$_{vP}$ y*ǔ* [ding]]

indeed rain *ding*.day

‘Indeed, it rained on the day Ding.’ (Djamouri et al. 2013: 587)

The moving vP, therefore, would attach below the adjuncts that are required to take scope over the verbal projection, while crossing other adverbs, like *ding* above, which did not have this requirement. At any rate, what is important for us is that Huang (2006) analyzes postverbal adjunct phrases in pre-Archaic Chinese via VP fronting, which got lost in Mandarin Chinese.

If traditional rightward adjunction in cases like these indeed involves movement, we have a systematic account for its diachronic decline, as it involves a dispreferred operation from the labeling perspective.

Djamouri et al. (2013), however, argue against Huang’s analysis based on the lack of multiple adjuncts in postverbal position in pre-Archaic Chinese, multiple adjunction not being limited in

preverbal contexts. They propose to derive this restriction by postulating a VP-shell structure à la Larson (1988), wich accommodes exactly one postverbal adjunct. The observed change in the availability of post-verbal adjuncts in Modern Chinese is then viewed by these authors as a diachronic disappearance of the postulated VP-shell structure over time.

I am adopting here Huang's (2006) analysis in terms of VP fronting for the derivation of post-verbal adjuncts. Assuming VP movement across the adjuncts, the restriction on a single post-verbal adjunct could simply be a consequence of the economy considerations dictating the movement to be as short as possible. We have seen that in cases like (67), where the adverbial expression cannot take a narrow scope with regards to the VP (and possibly the event variable within the VP that the adverb has to bind), verbal projection does not move to a higher position, allowing the adverb to have a wide scope over the VP. Djamouri et al. (2013) report that the situation of occurrence only one adjunct in the post-verbal position continued even after Pre-Archaic Chinese until the total loss of the possibility of post-verbal occurrence of adjuncts in this language. The number of the pre-verbal adjuncts, however, was not limited in any way at any of the stages of this language. If all the adjuncts in pre-Archaic Chinese were base generated in the pre-verbal position, the lack of limitation for the pre-verbal adjuncts is expected, hence the placement of these elements was not constrained by the movement of the verbal projection.

Returning to Djamouri et al.'s (2013) account of the distribution of adjuncts in Pre-Archaic Chinese and the disappearance of post-verbal adjuncts, it should be noted that treating adjuncts essentially as arguments (with the adoption of Larsonian shells), as suggested by Djamouri et al (2013), seems to be problematic given that with the exception of a few postverbal adjuncts like *badly* in (71) below, adjuncts are optional, in contrast to arguments.

(71) Mary treated John *(badly).

A fronting analysis may also allow us to understand the restriction on rightward adjuncts with e.g. sentential adverbs, which cannot appear to the right of the verb (if they do, they have to be accompanied by a pause, which indicates an afterthought).

(72) a. John left yesterday.

b. *John left probably.

c. John left, probably

If we assume XP fronting in cases like these, this can be accounted for since sentential adverbs are located high in the structure. It seems plausible that the remnant XP fronting cannot take place that high.

Another case of diachronic change that is similar to the one from Chinese discussed here concerns extraposition of relative clauses reported in Portuguese (Cardoso 2012), French, Icelandic and English. Looking at the rich corpora of these languages spanning the duration of 500 years, Wallenberg (2016) reports that these languages are in fact undergoing a change towards the loss of extraposed relative clauses. Wallenberg argues that what looks like optionality between the use of extraposed relative clauses versus relative clauses without this operation has all the characteristics of competing grammars, which the author takes as indicative of a change in progress (Clark 1987, 1990, Kroch 1994). This can be treated in the same way as the loss of traditional rightward adjoined adverbs in Chinese.

Given that traditional rightward adjunction in cases like these indeed involves movement, we have a systematic account for their diachronic decline. As it involves a dispreferred operation, it is under pressure to be reanalyzed in an analogical way as other movement operations discussed here. The loss of rightward adjuncts as attested in Modern Chinese, and apparent change in

progress in Portuguese, French, English, and Icelandic can therefore be understood as another instance of the loss of movement, which can be captured by appealing to labeling preferences in the grammar.

2.8. Conclusion

This chapter has investigated the role of the pressures imposed by the labeling algorithm (Chomsky 2013, 2015)[29] on syntactic structure and structural changes that occur as a consequence of such pressures. I have shown that labeling induces a preference in the grammar for head-complement {H, YP} configurations, with the merger of two phrases {XP, YP}, as well as two heads {X, Y}, being dispreferred.

I have shown that a number of cases of diachronic change can be explained and analyzed from this perspective: the loss of *wh*-movement in Japanese, Chinese, Indic, Romance, or Basque, reanalysis of specifiers as heads, as in Early Germanic *wh*-exclamative *hwæt*, Welsh complementizers, or other cases of reanalysis of relative pronouns as base-generated heads, the change from the OV to the VO word order, the loss of the head movement and traditional post-verbal adjunction. Importantly, I have shown that all these cases can be unified, as they involve the loss of configurations that are dispreferred from the labeling perspective. I have also proposed an account of why the loss of multiple *wh*-fronting proceeds through a single *wh*-fronting stage, not going directly to the option without any movement at all.

Overall, the account offered here explains many diachronic facts, which can now be seen as part of a much broader phenomenon and accounted for in terms of a structure-building mechanism

[29] For more proposals on labeling see also e.g. Epstein et al. (2014), Bošković (2016a, b), Rizzi (2016), Saito (2016), Takita et al. (2016).

with labeling playing a central role. Perceiving the trend to lose syntactic movement (both phrasal movement and head-movement) in terms of an attempt to minimize the dispreferred phrase-phrase (i.e., the Spec-head) and the head-head configuration, and to maximize the head-complement merger, which is simpler from the point of view of labeling, leading to straightforward and mechanical head-labeling, thus offers us a unified explanation for many otherwise seemingly unrelated processes.

Chapter 3

Verb Second in Synchrony & Diachrony

3.1.Introduction: V2 or V2s?

While discussing head movement, especially in the context of diachronic changes such as the loss of obligatory syntactic displacement, the Verb Second (V2) phenomenon cannot be avoided. In the broadest sense, V2 languages display a pattern where the finite verb appears obligatorily in the second position in its clause. Additionally, the raised verb has to be preceded by another constituent. The V2 configuration thus involves two movement operations: head movement (a movement of the finite verb to some higher head, often assumed to be C_0) and movement of some XP to the specifier of the projection whose head the verb occupies. Examples in (1) below illustrate this phenomenon in Swedish matrix sentences. Notice the different phrasal elements appearing in the 'prefield', i.e., right before the tensed verb.

(1) a. *Jag* **har** ärligt talat aldrig sett huggormar I den här skogen
I have honestly speaking never seen adders in this here forest
'To be honest I've never seen adders in this forest.'

b. *huggormar* **har** jag ärligt talat aldrig sett i den här skogen
adders have I honestly speaking never seen in this here forest

c. *i den här skogen* **har** jag ärligt talat aldrig sett huggormar
in this here forest have I honestly speaking never seen adders

d. *ärligt talat* **har** jag ärligt sett huggormar i den här skogen
honestly speaking have I never seen adders in this here forest
(Swedish; Holmberg 2015: 342)

As both these operations may involve feature-sharing (we are dealing here with symmetrical {X,Y} and {XP,YP} configurations), hence an option which is dispreferred from the point of labeling, we may expect that V2 configuration would be frail diachronically and prone to reanalysis. Indeed, we do know of cases where V2 was lost diachronically, as e.g., in English (van Kemenade 1987, Kroch & Taylor 1997, Pintzuk 1991, Fischer et al. 2001/2004, Haeberli and Ihsane 2016), French (Adams 1987, Roberts 1993, Vance 1995, 1997, Labelle & Hirschbühler 2017), Portuguese (Ribeiro 1995), and other Medieval Romance languages (Fontana 1993, 1997, Ledgeway 2008, 2017, Wolfe 2019), as well as in Welsh (Willis 1998, Roberts 2004). On the other hand, we know that V2 patterns tends to be quite persistent in many languages, resisting diachronic reanalysis of the type discussed in this chapter. I suggest that this is so because V2 is a multifaceted phenomenon, potentially more than one operation, each with different trigger and licensing conditions. The fact that V2 structures` often display those different operations, both within one language, and across languages, which are unified with the verb placement in the second clausal position, makes it appear as if it was indeed one consistent phenomenon. Hence, the variation in terms of its diachronic status, with some configurations involving the placement of the verb in the second position being more persistent than the others.

I will argue that the phenomenon that we refer to as V2 involves (at least) two distinct configurations with distinct motivations, syntactic mechanisms, and licensing conditions (see also Poletto 2002, Wolfe 2016, 2019); i.e., V2 comes (at least) in two flavors. This distinction is relevant for the argument presented here as it results in two different diachronic patterns. Roughly,

the distinction I will be working with here, appealing to two different flavors of V2s as is proposed in (2):

(2) a. Discourse related V2: **V2D**

b. Illocutionary Force related: **V2I**

The distinction in (2) intends to capture what we will see below, namely that V2 in different languages is connected to different interpretations and contribution (or lack thereof) to the Information-Structure. I will argue that only one of these flavors of V2, namely V2D, is subject to the pressures of the labeling algorithm discussed in the present work as one of the main forces behind diachronic change. Hence this flavor or V2 will be predicted to be diachronically more fragile. The other flavor, V2I, is predicted to be more stable across time because the pressures connected to the labeling requirement are not operative there for the independent reasons.

The distinction in (2) can be illustrated by examples from Modern English, where the discourse related V2D has been lost ((3a) an example of Topicalization, does not involve verb movement to C any more), whereas V2I, illustrated by (3b), still persists, albeit in its 'residual form'. The English facts, in fact, support the proposal made here about the difference in the diachronic sustainability of different V2 patterns, as discussed in more detail below.

(3) a. What **has** John bought?

b. John, Mary likes.

Two important caveats have to be made at his point. First of all, we will see that even though V2D involves the dispreferred feature-sharing operation, its relation to the information-structure of the sentence, hence the semantic/pragmatic component, can make it somehow still less resistant to the

labeling-induced pressures than purely syntactic displacement. Second, many V2 languages involve both of these guises and teasing them apart is extremely difficult as both of these mechanisms may have acquired more grammaticalized status and may be using mixed strategies to satisfy the V2 constraint in the language. Actually, there have been many proposals suggesting that diachronically, languages introduce the V2 constraint into their grammars by means of V2D, which later becomes grammaticalized and become V2I. For proposals of this type see e.g., Kiparsky (1995) or Roberts (1996), which suggest that V2 where the subject-auxiliary inversion occurs only in a subset of structures with strong semantic (i.e., discourse related) profiles (e.g. interrogatives or questions) is actually the oldest Germanic pattern, with the V2 appearing in most other contexts as in other modern Germanic languages being later innovation. If this is the case, the 'residual V2' as it is still attested in English had to transform into a particularly constrained V2I. At any rate, as teasing apart those two flavors of V2 is often a challenging task, I will mark and appeal to the distinction in (2) only where it is crucial for the analysis.

The main claim here is that cases of V2 which are especially resistant to reanalysis must involve V2I, and if the languages that have V2I also have V2D, they may use the mechanisms of V2I to derive it, hence making it diachronically stronger. Now, why is the illocutionary-force related V2 more stable across time? I will propose that while both configurations involve extra-syntactic factors, connected to the interpretive requirements, V2I also involves prosodic mechanisms. Additionally, as discussed below, some languages may be able to employ head-labeling even in cases of V2I, which makes them free from labeling pressures of the kind discussed in chapter 2.

I will appeal here to the analyses of V2 that stress the connection of this verb displacement

(broadly speaking, V-to-C head movement) with particular illocutionary power (Hooper & Thompson 1976, Wechsler 1991, Brandner 2004, Truckenbrodt 2006, Wilkund et a. 2009, Julien 2015), as well as to those accounts that connect second-position effects to prosody in certain phonological constraints that operate at left edges (Boeckx 1998, Rice & Svenonius 1998, Anderson 2000, Bošković 2001, 2018, Jouitteau to appear).

Typologically, V2 pattern is not very common, nevertheless it appears in various language groups and families. The most research has been done on V2 in Germanic (see Vikner 1995, Holmberg 2010/2015 for the overview), but V2 languages are also attested within Celtic: Early Irish, Old Irish (Adger 2006), Middle Welsh (Willis 1998, 2007), Cornish, Breton (Meelen to appear, Jouitteau to appear), Romance: Old French (Adams 1987, Roberts 1993, Platzack 1995, Hulk & van Kemenade 1995, Vance 1995, 1997, Labelle & Hirschbühler 2017, Wolfe 2019), Old Spanish (Fontana 1993), Old Portuguese (Ribeiro 1995), Medieval North Italian dialects (Benincá 2006, Ledgeway 2017, Wolfe 2019), Rhaeto-Romance (Poletto 2002, 2014), and Rumantsch (Anderson 2005), Slavic (Sorbian), Finno-Ugric (Estonian), or Indic: Kashmiri (Bhatt &Yoon 1992, Bhatt 1999, Bhatt & Mushi 2009, Manetta 2011), and in two dialects of Himachali (Kotgarhi, Kochi)(Hendriksen 1986, 1990). Outside Indo-European, V2 has been reported in Karitiana (Tupi-Arikem language) (Jouitteau 2010 and references there) or Papago (Bhatt 1999).

As noted above, V2 is not a uniform phenomenon. A good example of attested differences between V2 languages (and possibly between different operations that come under the V2 term) concerns variation in terms of the availability of V2 in embedded clauses. Despite V2 being considered a main clause phenomenon (even in languages which appear to have V2 in both matrix and embedded clause), we will start our discussion with the embedded clauses, as they are great indicators of some important and exceptional properties of V2 grammars.

All V2 languages display it (in appropriate contexts, again depending on the particular V2 guise) in the matrix clauses. Regarding subordinate clauses, Vikner (1995) argues that V2 languages fall into four main categories: 'well-behaved' V2 languages, 'general embedded' V2 languages, 'limited embedded V2 languages,' and finally 'residual' V2 languages. I will discuss briefly those classes of languages, especially the first three, focusing on the most important characteristics of each class.

Let us start with so called 'well-behaved' V2 languages. German, Dutch, Afrikaans are argued to belong to this group as they display V2 in all contexts lacking complementizers. Since in these languages overt complementizers are in complementary distribution with the raised verb, it has been argued that this indicates that the two occupy the same position, at least in matrix contexts (see e.g., den Besten 1992). This also means that in these languages, the complementizer head is always lexically expressed: either by the directly merged functional head, or the tensed verb. We will return to this observation below.

Certain languages are argued to undergo V2 in both main clauses and embedded clauses. These are referred to as 'general embedded' V2 languages. Icelandic, Yiddish, Kashmiri, and Old French have been claimed to fit this category. Note that these languages (unlike the 'well-behaved' group just mentioned above) display verb-raising in the context of overt complementizers. This has been interpreted as an indication that the verb does not move as high as C, but targets some lower head, potentially I, at least in embedded contexts. Therefore, both can appear simultaneously as they are in different positions. However, even in this class of languages, V2 cannot actually appear in *all* embedded contexts, i.e., even in these languages the embedded V2 appears only with the subset of predicates (e.g., see (7) and especially (8)-(10) below). Hence, they really should be classified as

the group of V2 languages discussed next, i.e., 'limited embedded' V2 languages in Vikner's nomenclature.

The group of 'limited embedded' V2 languages permits verb second only in some classes of embedding verbs, characterized by their potential to express certain semantic content, often defined as *assertion.* This is related to the observation that V2 clauses commit the speaker to the truth of the proposition or they urge the addressee to accept that truth, whereas non-V2 may be interpreted deontically or render speech acts such as orders, requests, wishes, invitations.[1]

Additionally, the presence of V2 in these contexts is often considered to be optional or is argued to give raise to different interpretations depending on whether the verb is raised or not. I will lean to the latter, as we have some evidence that V2 and non-V2 indeed differ in the semantic/pragmatic expressions. This connection between semantic interpretation and the finite verb position is worth exploring from the current perspective, as it is playing an important part in the persistence of the V2 phenomenon. The discussion of the availability of the V2 in some embedded contexts but not others in these languages will serve as an important introduction to the occurrence of parallel semantic effects in main clauses, where V2 is obligatory in the discussed languages, and where we will see semantic effects are even stronger.

1 Truckenbrodt (2006) defines this semantics of V-to-C movement displayed in declaratives and interrogatives in terms of considerations of common ground, in a sense that assertion is just a speaker's addition of some information to an addressee's common ground. On the other hand, Wiklund (2010) proposes to see this unique semantic effect of V2 clauses in terms of evidentiality, i.e., presentation of a speaker's point of view.

3.2. The [Force] of Assertion vs. Discourse

Vikner's 'limited embedded' V2 languages, such as Main Scandinavian languages, allow V2 after verbs of reported assertion (*say, claim)* or after assertive adjectives, such as *true* or *clear*.

(4) a. Folk *sa* at ho **var** ikkje for vakkert kledd. [strongly assertive]
People said that she was n ot too beautifully dressed
'People said that she was not too beautifully dressed.'

b. Det er *klar*t at det **hadde** ikke så mye å si for oss.
it is clear that it had not so much to say for us
'It is clear that it didn't make much difference for us.' (Norwegian; Julien 2015: 164)

This class of predicates is 'strongly assertive' in Hooper & Thompson's (1976) classification of complement taking verbs. Below is the full classification proposed by Hooper & Thompson:

(5) The Hooper & Thompson's (1975) verb classes:
A) 'Strongly Assertive': *say, claim*
B) 'Weakly Assertive': *believe, think, seem, be possible*
C) 'Non-Assertive': *doubt, deny*
D) 'Factive' : *regret, be sad about*
E) 'Semi-Factive': *discover, understand*

Apart from class (A), V2 is attested in complements of other assertion-bearing predicates, i.e., 'weakly assertive' predicates (B) (5a), and 'semi-factive' predicates (E) (5b). All of these contexts involve assertion, with the last class of predicates (E) expressing knowledge or perception on part of the speaker; hence, they also have the illocutionary power of asserting the proposition.

(6) a. Han *trodde* att vi **hade inte** sett den här filmen [weakly-assertive]
he believed that we had not seen the here film-the
b. Jag upptäckte att den här boken **hade** jag inte last [semi-factive]
I discovered that this here book-the had I not read
(Swedish; Bentzen et al. 2007:100-102)

If the predicate belongs to the non-assertive class, however, V2 is ungrammatical. V2 cannot be used in embedded clauses after predicates of classes (C) and (D), as they express overt denial or questioning of the matter that is 'at-issue' in the proposition, or express emotions or subjective attitude (D). Both groups represent non-assertive predicates and they do not allow verb-raising in their complements. Compare sentences (a), without V2, with (b), where the inversion of the subject and the verb indicates verb-raising. Notice also the ungrammaticality of sentences with V2 in languages where the embedded V2 has been argued to be widely available, such as Icelandic.

(7) a. Det er *ikkje* *muleg* at Peter **går** heim nå. [non-assertive]
it is not possible that Peter goes home now
'It is not possible that Peter is going home now.'
b. *Det er *ikkje* *muleg* at nå **går** Peter heim
it is not possible that now goes Peter home
'It is not possible that Peter is going home now.'(Norwegian; Julien 2015:165)

(8) a. Hann *efast* um að hún **hafi** ekki hitt þennan mann
he doubts about that she has not met this man
'He doubts that she hasn't met this man.'

b.* Hann efast um að þennan mann **hafi** hún ekki hitt

he doubts about that this man has she not met

'He doubts that she hasn't met this man.' (Icelandic; Bentzen et al. 2007: 100)

(9) a. Hann *angraði* at hann **hevði** ikki sungið. [factive]

He regretted that he had not sung

'He regretted that he hadn't sung.'

b. *Hann *angraði* at henda sangin **hevði** hann ikki sungið.

He regretted that this song-the had he not sung

'He regretted that he hadn't sung this song.' (Faroese; Bentzen et al. 2007: 100)

(10) a. Hann *sá eftir* að hann **hafði** ekki sungið

he regretted that he had not sung

b. *Hann *sá eftir* að þetta lag **hafði** hann ekki sungið

he regretted that this song had he not sung

'He regretted that he hadn't sung this song.' (Icelandic; Bentzen et al. 2007: 100)

The ungrammaticality of sentences with V2 indicates that such non-assertive contexts ban V2 in more languages than previously thought, showing that V2 is actually more constrained that it has been argued. The same facts hold in both Mainland Scandinavian (as widely argued), as well as in Icelandic and Faroese (arguing against the analyses of these languages as having 'general embedded' V2). This also shows the relevance of the semantics of assertion to the V2 phenomenon. This is confirmed by instances of non-assertive predicates being negated, where V2 becomes grammatical, since a negated non-assertive predicate becomes an assertive predicate.

(11) Det är **ingen** som tvekar på att dom **gör** det alltid
it is nobody that doubts on that they do it always
för att få upp försäljningen
for to get up sale.DEF
'Nobody doubts that they always do it to raise sales.' (Swedish; Julien 2015: 165)

Julien (2015) argues that in a sentence like (11) the embedded assertion made by the speaker becomes available, hence its grammaticality with V2. [2]

The availability of V2 in assertive contexts has been argued to be indicative of the richer structure in the complement of assertion-bearing predicates, which enables raising of the verb to some higher position in the left-periphery of the embedded clause. Analyses of this sort often appeal to Rizzi's (1997) split-CP.[3]

The structure in (12) below illustrates one of the proposals for the cartography of the left-field in general (i.e., some but not all of these projections are argued to be available in some languages in both main clauses and embedded clauses). The inventory below is important for the discussion regarding V2, hence the reader should be familiar with these proposals. The present work does not adopt the idea of cartographic syntax, leaning instead to the mapping approach to syntactic structures (see e.g., van Craenenbroeck 2009 for an overview of the problems connected with cartography); however, since many of the proposals in the V2 literature analyze V2 using

[2] Julien (2015) argues that even factives may allow V2 under an appropriate context, e.g. when the implicit speaker is identical to the actual speaker the assertion becomes available.

[3] Julien (2015) also argues that the assertion-selecting predicates indeed select complements with richer structure, as only these V2 clauses allow indexical shift, which is impossible with structures without V2. Julien's analysis of indexical shift is based on Sigurðsson's (2010) idea of the necessity of so called C/edge linkers appearing only if the CP structure is rich enough.

inventories such as (12), I will use these terms descriptively without committing to this particular approach.

(12) [ForceP [TopP (Contrast) [FocP [TopP (Familiar) [FinP [TP…

What is important for us here is the recognition of the appearance of a projection encoding semantic force of the proposition situated high in the left periphery. In a hierarchy of projections postulated by Rizzi (1997), ForceP and FinP delineate the 'traditional' CP, i.e., they constitute respectively the lowest and the highest projection within complementizer phrase.

ForceP, Rizzi argues, is a projection encoding clausal *Force*, i.e., the *type* of the clause, e.g., it indicates whether we are dealing with interrogative, declarative, or exclamative. This is the projection where overt clause embedding complementizers are merged in some languages, as well where force particles appear in languages which have them. FinP, on the other hand, encodes the finiteness of the clause, evidenced by a difference between finite and non-finite complementizers in some languages. As it has been argued that the most diachronically stable V2 languages (German, Mainland Scandinavian) involve placing of the verb in one of these projections, i.e., in Fin or Force, I will assume that this is indeed the characteristic feature of Illocutionary Force-related V2 languages (V2I). V2D, on the other hand, is associated with Rizzi's TopP and FocP projections, as they, cartographically speaking, host topicalized and focused constituents. This makes their connection to the discourse motivated displacement quite straightforward.

The connection between particular semantics and verb-raising suggests that the predicates allowing assertion in their complement actually select for CP containing Force, where the locus of this semantic feature is expressed. Syntactic manifestation of this, often appealed to in the cartographic accounts of V2 (e.g, Wolfe 2019), is that this richer left-periphery creates more

landing positions for the raised verb, as it has been shown that there is some variation within V2 languages regarding the verb's final target in the left-periphery (both in the main clauses and the embedded clauses).[4] However, the projection of the force of the sentence, which in a non-cartographic approach can be seen as appearance of certain Force-encoding features, hence *Force*, on the appropriate head in the left periphery, is crucial for the V2 phenomenon, as we are about to show. In order not to confuse the particular syntactic projection in Rizzi's split CP with the feature realized on some syntactic head, we will use the term Force when the clause-type is being discussed (versus ForceP or Force$_0$, as would be in Rizzi's style projection)

I will show that V2I languages indeed associate V2 with the Force. The main argument concerns the fact that the presence of Force (in a sense of the presence of a particular feature in the CP structure) can be detected by the placement of the finite verb high in the left-periphery, which often involves placing it in the second position of the clause. Despite variation within V2 languages regarding the phenomenon in question, all of them have one thing in common: V2 in general is attested in the finite contexts only. Appealing to the semantic notion of assertion associated with this phenomenon offers an explanation of this fact, namely that assertion must be anchored in time, which is possible only in finite contexts.

The semantics of assertion associated with V2I is confirmed in embedded clauses with their more independent behavior. Julien (2015) shows that they can function as *Main Point of Utterance* (Simons 2007), i.e., they can constitute an independent answer to a question. Additionally, they do not allow the negation in the matrix clause to scope above the V2 complement. This is illustrated

[4] Poletto (2002) assumes that some V2 languages require verb movement to Force, whereas the others move it to Fin, depending on the strength of the formal feature in each of these projections in a particular V2 language.

with German in (13) below, where the ambiguity connected to the scopal interaction between the negation in the matrix clause and the adjunct clause disappears when the verb is raised to the second-position within the adjunct clause:

(13) Er kommt nicht, weil er (ist) faul (ist)
He comes not because he (is) lazy (is)
'He doesn't come because he is lazy' (German; Holmberg 2015: 361)

i) Non-V2: because is lazy > neg come / neg come > because is lazy

ii) V2: because is lazy > neg come / *neg come > because is lazy

Hence in (13), the V2 version can only have the interpretation that the reason for his absence is his laziness; this is the reading with the verb in the second position scoping over the negation. The other interpretation is blocked.

Having confirmed the semantic effects associated with V2 in embedded clause, which we saw are motivated by the presence of the projection encoding (or expressing) [Force] we will move to the main clauses, where the effects in question are even stronger, as they involve the *requirement* for the [Force] to be overtly expressed, as we are about to see.

Proposals appealing to the connection between the obligatory verb movement and its semantic import are well attested in the V2 literature (Hooper & Thompson 1976, Wechsler 1991, Truckenbrodt 2006, Wilkund et a. 2009, Julien 2015, among others). I will adopt some aspects of the proposal in Brandner (2004) that V2I languages are characterized by the requirement to overtly express the force of the sentence. We can translate the intuition behind this proposal as the requirement to overtly fill the syntactic projection responsible for encoding [Force], such as Force$_0$

(or Fin$_0$ head, as languages may vary in this requirement)[5] in Rizzi's cartography, or just the C$_0$ head with the [Force] feature[6], which will also be behind the motivation (for the perseverance) of V2 systems (see also Julien 2015 on V2 involving movement to Force$_0$). Under this approach, it is really not the case that verb raising to C creates the semantics of assertion, but rather the overt presence of the verb in the projection encoding Force is an indication of the presence of this syntactic head, due to its requirement to be overtly manifested (in this respect, see the proposal in Rizzi 1997 and Bošković in press that certain functional projections in split-CP are present only when they are overtly phonologically manifested). In other words, V2 is licensed in this context, instead of creating it by its movement .

Brandner (2004) ties the V2 effect to the proposal in Cheng (1991) regarding the requirement for clauses to be typed (*The Clause Typing Hypothesis*). This gives us a straightforward connection of the V2 requirement for the projection encoding the type of a clause (interrogative, declarative, etc.) to be overtly filled as only then the interpretation of the clause can be determined. Languages may have different means and strategies to overtly express Force: some use overt particles, and the others have to move the tensed verb. Brandner gives an example of Korean as a language of the first type, with the declarative (14a) or interrogative particles (14b) doing the job of overt Force expression:

[5] See also Poletto (2002) who argues that V2 may be attracted to Force or Fin, and Roberts (2004) who argues that whether complementizers are realized in Force or Fin is a matter of crosslinguistic (and sometimes diachronic; Wolfe 2019) variation.

[6] See also Carnie et al. 2000 for a similar constraint, FILL-C in Old Irish, which ensures that C position is always filled by material with phonological content

(14) a. ku-ka seoul-e ka-ass-**ta**

he-NOM Seoul-to go-PAST-DECL

'He went to Seoul.'

b. ku-ka seoul-e ka-ass-**nunya**?

he-NOM Seoul-to go-PAST-INTERR

'Did he go to Seoul?'

Korean particles fulfill the structural requirement that Brandner proposes for the proper [Force] marking: they are in a high position where they can scope over the whole proposition, and they can operate on a proposition with a concrete Tense value. Hence the analogy to a syntactic head like Force$_0$ in the high left-periphery and the finiteness requirement characteristic of V2I contexts.

Languages like German, on the other hand, do not have a specific Force particle in their disposal and they have to displace the finite verb to satisfy the clause-type requirement. This proposal predicts that if a V2I language has some other means to satisfy the overt-Force requirement, we should not see the finite verb moving to this position. This is in fact borne out.

Walkden (2014) observes that Old English displays a verb-final pattern in sentences with speaker-oriented adverbs, such as *soþlice/soðlice* 'truly' and *witodlice* 'certainly'. Walkden notices that in modern English such adverbs (and they include adverbs like *honestly, clearly, or luckily*) are unacceptable in the complement of factive verbs (15a), cannot occur with inversion (15b), or cannot appear in the scope of negation (15c).

(15) a. Bill regrets that Frank (*luckily) discovered the uranium.

b. So fast did Tom (*luckily) run that he got to Texas in ten minutes.

c. Karen has not (*luckily) left. (Walkden 2014:101)

Walkden argues that these adverbs are strongly assertive, hence they are not compatible with any other contexts except those involving assertion (we saw the restriction on assertive contexts with V2 to scope over the negation in (13)). What is more, these adverbs are argued to be responsible for creating a secondary proposition. As such, their role is similar to V2. Indeed, as predicted by the clause-typing/Force or Fin)-filling analysis, OE does not resort to verb movement in the contexts where these adverbs appear, as they are able to satisfy the requirement to express the assertive force of the clause (and arguably overtly lexicalize the appropriate [Force]-encoding projection). The verb tends to appear in the final position in these contexts (see Walkden 2014:102 for precise frequencies of these adverbs with V2 vs verb-late (V4+)).

(16) a. He þa *soþlice* oðre þara flascena þam halgan were brohte
he then truly other.ACC the.GEN bottles. GEN the.DAT holy.DAT man.DAT brought
'He then truly brought one of the bottles to the holy man' (OE:Walkden 2014:103)

b. Zosimus *soþlice* Þa eorðan mid tearum ofergeotende hire to **cwæð**
Z. truly the.ACC earth.ACC with tears.DAT overspilling her to said
'Truly, soaking the earth with his tears, Zosimus said to her…' (OE:Walkden 2014:103)

Similarly, Roberts (2004) argues that sentential particles in Celtic languages (including those in Welsh, the diachrony of which we discussed in the previous chapter) perform the analogous function of expressing the Force of the sentence. Consider (17) with 'root affirmative' particles *fe, mi, y.*

(17) Fe/me welais i John
PRT saw I John
'I saw John' (Welsh; Roberts 2004: 298)

Roberts (2004) offers the following inventory of particles in Welsh within Rizzi's split-CP system, depending on which head they lexicalize.

(18) Force Fin

mai/ai/nad/nid *a/y/fe/mi/bod (Pres/imperf)*

This may also hold in V2D language. This is also supported by the facts from Gothic, where two possibilities to mark Force involve particle insertion (like a Topic particle *uh* (19)) or verb movement to the second position (e.g. in the context of a definite subject (19b) (Eythórsson 1995, Brandner 2004, Walkden 2014, Migdalski 2016).

(19) a. þat-***uh*** samo jah þai waidedjans ….**idweitidedun** imma

this-PRT same also the bandits insulted him

'in the same way the bandits too…insulted him.' (*Matt 27:44*, Eythórsson 1995:66)

b. iþ Jesus **iddj-*uh*** miþ im

but Jesus went-PRT with them

'But Jesus went with them' (*Luke 7:6*; Walkden 2014:108)

Brandner (2004) discusses cross-linguistic variation in the requirement to clause-type the sentence, with languages like Japanese or English, which only need to type non-declarative contexts (*non-explicit marking languages*), and with languages like Korean or Welsh, which are required to type/ express all types of Force (*explicit marking languages*).[7] She sees V2 languages as belonging to

[7] In this context, it may be relevant to ask what is going on in English examples like (i) that lack *that* in the embedded clause.

(i) John think Mary left.

the latter group, referring to this requirement of expression all types of Force as 'overgeneralization' (not all languages of that type are V2 languages, however). Also, Brandner's requirement to express Force overtly is analogous to the proposal in Roberts (2004) that Fin^0 has to have PF realization. In V2 languages this requirement is satisfied by the placement of the finite verb in Rizzi's Fin projection. Finally, Roberts (2004) offers a unidirectional generalization that languages where Fin requires PF-realization in declarative clauses will also have this requirement in non-declarative clauses. This captures Brandner's observation as to the variation between languages regarding the overt realization of Force. This gives us a prediction that languages overtly realizing declarative force, will also necessarily realize interrogative force, but not the other way around. This is confirmed by residual V2 languages like English where interrogative contexts require overt marking (e.g., subject-auxiliary inversion), but where declarative (but again, not interrogative) complementizers can be omitted.

(20) a. What {can} you {*can} do?

b. Jon said (that) he was hungry.

c. Jon asked *(whether) she was hungry.

At any rate, we have shown that V2I languages are characterized by the requirement to have [Force] overtly expressed in a number of contexts, including the declarative, and verb-raising is one way of fulfilling this requirement. In fact, languages having the requirement to express the declarative force are not that rare typologically: languages like English, where the declarative complementizer

One possibility is suggested in Bošković (1997) that in cases like this the lack of overt complementizer suggests the lack of CP (notice the lack of *that-trace effects* in this configuration, as well). Furthermore, given that the embedded clause in (i) is interpreted as a declarative, Bošković (1997) suggests that declarative Force in English is a default option (so this is the interpretation when CP is missing).

can be dropped, are actually more rare (with languages that never mark the interrogative force being extremely rare).

As the occurrence of V2 in declarative or assertive (and in general, in illocutionary) contexts tends to be diachronically quite stable, the next section will try to give some answers to why that could be the case. As the present analysis connects the loss of movement to the labeling-induced preference for head-complement configuration and head-labeling, the goal will be to show why V2I is not subject to such pressures.

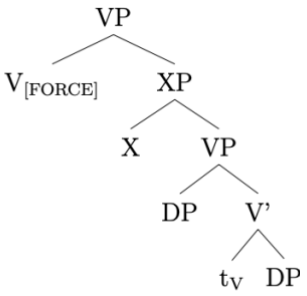
3.3.Verb Second Placement

We have shown in section 2.6. of chapter 2 that head-movement tends to be diachronically lost as it involves the symmetrical and thus ambiguous (from the standpoint of labeling) merger configuration {X,Y}. It is dispreferred because, in contrast to {H,YP} merger, its labeling must involve an additional operation. The question is whether all head movement must result in such an ambiguous structure. If some instances of head-movement may not give rise to such a symmetrical configuration, we would predict that the diachronic forces, attributed to labeling here, would not be operative in such cases. Alternatively, if the structures created by head-movement were free from the labeling requirement, we would also predict that such operations could be diachronically more stable. Since V2I can be argued to be historically quite persistent (at least in the small amount of languages that have this constraint), we can argue that the pressures connected to labeling considerations are indeed not operative in these contexts. This section and the next one will show that V2I is indeed free from the labeling pressures. We will start with the account of the verb-movement operation involved in the explicit [Force] marking requirement and show that this configuration can actually be labeled in the most economical way, i.e., by head-labeling.

Brandner (2004) proposes a system of the Force-marking verb movement in V2, which can serve as a solution to problem of the resistance of V2 to the pressures of labeling. In her system, which is modelled after certain proposals by Koeneman (1995, 2010) and Bury (2002, 2010), the ambiguous {X,Y} configuration is not created here in the first place and the resulting configuration can in fact be head-labeled by the raised V head.

Briefly, Brandner adopts the idea of V2's verb-raising to be an instance of 'self-attachment'. What that means is that the moved verb does not target any existing projection, but instead projects its own VP in the landing site. In a system that dispenses with cartographic pre-packaged and rigid syntactic-structure (argued to constitute a universal functional sequence), there is no designated landing site for V-movement in this case (see also Neelman et al. 2009, Horvath 2010), verb-movement involved in V2 is not a head-adjunction to an existing head. Instead, it targets the root of the syntactic structure in a movement whose landing site guarantees the satisfaction of the overt expression of the Force requirement. This self-attaching head-movement is illustrated in (21) below:

(21)



One of the requirements for [Force] marking is to operate on the already formed propositions, which necessarily have to have a specified [Tense] feature, i.e., they have to be finite. Now, structurally speaking, [Force]-marking in Brandner's system involves scoping over all other verbal

projections, thus it needs to move (or be based generated in case of Force particles) high enough to guarantee its wide scope in the sentence. (I will discuss the [Force] feature given to the verb, below.)

Schematic representation of this head-movement as self-attachment in (22) shows that it essentially results in the extension of the verbal domain. Brandner draws another parallelism here with languages like Korean, where the Force marking particles are attached to the verbal projection in a way indicating the extension of the verbal domain:

(22) [V_0+ T_0+C_0]

ka-ass-nunya

go-PAST-INTERR (Korean: Brandner 2004:110)

Interestingly, Saito (2012) proposes a similar self-attachment head movement to derive Japanese Case valuation (among other issues). Saito (2012) adopts the Merge-based analysis of Japanese Case after proposals in Shimada (2007) and Tonoike (2009), who propose a system which allows head movement to observe the strict extension condition. Briefly, their derivation starts with a complex of heads C-T-v-T, which merge with the object in transitive clause structure (23a). Their system assumes the excorporation of the functional complexes and remerging the heads higher, where they project to create new configurations. Hence, in (23b), the C-T-v excorporates creating a vP, which merges with the subject DP (23d). Next C-T excorporates in (23e) to create a TP. This is followed by the internal merge of the subject with this TP in (23f), and finally in (23g) C excorporates, and remergers, projecting the full CP structure.

(23) a. {C-T-v-V, DP}

b. { C-T-v, {V,DP_1}}

c. { C-T-v, {V, DP_1}}

d. {DP_2 {C-T-v, {V,DP}}}

e. {C-T,{DP_2, {v,{V,DP_1}}}}

f. {DP_2{C-T{DP_2{v{V,DP_1}}}}}

g. {C,{DP_2, {C-T,{DP_2,{v,{V,DP}}}}}}

Saito (2012) adopts this structure building, which allows a nominative object in Japanese attested in sentences like (24) to be valued without the need to movement to [Spec, TP] (which he shows would be problematic).

(24) Hanako-ga rosia-go wakar-u (koto)
Hanako-NOM Russian-NOM understand-Pres fact
'the fact that Hanako understands Russian' (Japanese; Saito 2012: 118)

Saito proposes that *v* in this structure does not value accusative and it not a phase head. His derivation given in (25) starts with the phase head C, then.

(25) a. {T, C} (nominative)

b. {*v*, {T, C}}

c. {V, {*v*, {T, C}}}

d. {DP_1-NOM, {V, {*v*, {T, C}}}}

e. {{DP_1-NOM, V}, {*v*, {T, C}}}

f. {DP_2-NOM, {{DP1-NOM, V}, {*v*, {T, C}}}}

g. {{DP_2-NOM, {{DP1-NOM, V}, *v*}}, {T, C}}

h. {DP_2-NOM, {{DP2-NOM, {{DP_1-NOM, V}, *v*}}, {T, C}}}

i. {{DP$_2$-NOM, {{DP2-NOM, {{DP$_1$-NOM, V}, *v*}}, T}}, C}

In (25b) and (25c) T selects *v* and v selects V. Next, in (25d), this C-T complex merges with the object and values its case as nominative. (25e) is interesting for us here, as the the excorporation of *v*-T-C occurs at this point and the vP which merges with the external argument in (25f) is projected. Its Case is valued as nominative. Next, T-C excorporates in (25g) and the subject raises to TP Spec as in (25h) to satisfy the EPP requirement of T-C. The last step involves another excorporation, this time of C in (25i).

Saito (2012) shows that this Merge-based derivation can account for such puzzling phenomena as Japanese Nominative/Genitive Case alternation, the scope properties of accusative and nominative objects in Japanese, as well as the presence of Case-marking on Japanese PPs.

Importantly, we can see here the application of self-attachment in other constructions and languages outside the V2 derivation.

Returning to V2I and its expression of Force in the self-attached position, if we allow the projection encoding the force of the sentence to be determined relationally, which is the case in a system where information-structure relations are satisfied by mapping the positions to their interpretation (so what matters is simply relative height; the dedicated cartography-style projections are not necessarily involved), such verb-movement can satisfy the requirement to overtly express the [Force] of the sentence. Most importantly, such an operation creates the configuration that can be labeled by a lexical head in a {V, XP} configuration. Since the labeling requirement can be satisfied here in the most economical way, i.e., by a minimal search operation in the narrow syntax, the fact that diachronically V2 structures created that way are resistant to change is not surprising. What is important for our purposes is that despite the presence of syntactic movement, we are dealing here with the preferred option from the point of view of labeling the

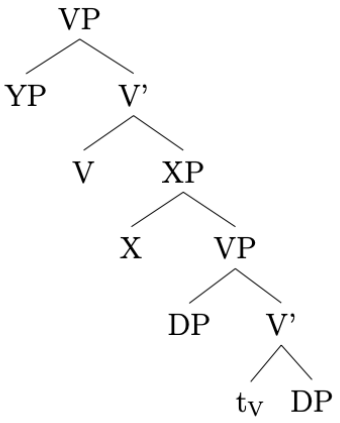
syntactic structure. As expected, then, this configuration is not diachronically fragile. This can in fact be taken as a confirmation that the loss of movement and syntactic specifiers really have basis in labeling considerations; the pressures responsible for diachronic changes described here are not operative outside of the labeling contexts. This also means that syntactic displacement per se does not really have to be considered to be more costly than base-generation, as long as it does not create problems for labeling the resulting structure. In other words, it is not really syntactic movement that is problematic; the problem is that in many cases syntactic movement results in a configuration that is costly from the labeling point of view. However, this is not always the case.

While discussing head movement in chapter 2, we have made the point that head movement that does not have an overt reflex of feature-sharing is more likely to be lost. V2 does not have an overt morphological reflex but exhibits diachronic stability. This can be interpreted as indicating that what we have here is not labeling through feature sharing, which is indeed the case under the current analysis.

Chomsky (1995) assumes that a configuration like (21), which is a reprojection structure, is not a legitimate object due to the ambiguity it creates. The higher projection of V cannot be unambiguously identified to be either base-generated or its copy; it looks like we are dealing here with one maximal projection with two heads. As this violates endocentricity, Chomsky (1995) rejects the possibility of the projection of the target.

Koeneman (1995) (see also Bury 2010) argues that Chomsky's problem is resolved when the reprojected head involves a specifier. Then both the lower copy of the moved V and the fronted copy can be structurally interpreted, each being a head of a separate VP.

(26)



This is how he explains the requirement of V2 to be consisting of two operations: self-attaching head-movement, and subsequent projection of the specifier filled with some moved XP.

Brandner (2004) rejects this idea, as it would amount to the lack of the verb-initial structures, which she assumes to be attested (though there could be a null Op in the specifier position at least in some such cases).

We can add another objection to Koeneman's proposal, namely the availability of Spec-less V2 projections, where an overtly filled higher syntactic head counts as the initial constituent in V2 configurations. As such structurers are attested, e.g., in C-VSO configurations in Welsh or Middle Breton, it cannot be the case that V2 always involves specifiers.

Brander proposes to resolve Chomsky's (1995) ambiguity problem in a different way. She assumes that self-attachment is legitimate only if the raised verb acquires additional features during the process. It is therefore an altered version of its trace so that the two heads are not completely identical. To make this possible, Brandner assumes that the moved verb gains its force value by Spec-head agreement with the moved XP, which in fact determines the force in question. If it is a *wh*-phrase, the force will be interrogative. On the other hand, if the XP in question does not bear any marked value, the [Force] acquires default value of declarative. In her system then, both V2I

and V2D are derived in the same way and the prediction is that they both should be diachronically stable to the same degree. [8]

I will not adopt Brandner's requirement for the Spec-head agreement in determination of the value, instead proposing that the value of the Force is specified on the verb itself, but it requires placing the verb in the high left-peripheral position, as only there, Force can be expressed (semantically, it has to take the widest scope). Therefore, the trigger for the movement is placed in the moving element, the foot of the chain, in a way analogous to the cases of multiple movement discussed in chapter 2. This goes back to the observation that sometimes the requirement for the raised verb to be in the second position is satisfied by elements other than phrases, e.g., syntactic heads such as non-finite verbs in Breton, Yiddish, or Icelandic (this includes covert operators that can be argued to be head-adjoined due to their non-phrasal character). This also applies to instances of [Force] other than declarative, hence it applies to both V2D and V2I in our system. Therefore, instead of adopting any of the analyses postulating obligatory creation of specifiers in the syntax in V2, I will show that the requirement for the forefield to be filled, placing the [Force] expressing verb in the second position, has different basis in V2D and V2I. In the latter, the requirement to fill the prefield has prosodic bases, whereas in the former, the merge of discourse-related XP resorts to the feature-sharing, as the XP merge occurs in syntax.

[8] Following Cechetto & Donati (2015), Chomsky (2008) actually allows the possibility of the moving element projecting, but only in the case of head movement, as in free relatives of the type: *I read what you wrote,* where the wh-element projects after movement forming DP which is the complement of the verb *read.* This situation is restricted to syntactic heads as *I read [what book you wrote]* does not have the interpretation of a free relative.

3.4. The Prefield and Prosody

The requirement for the syntactic constituent to be in the second position in its clause has often been argued to be prosodically motivated. The obvious intuition behind this proposal is that syntax does not operate on the linear ordering, and we cannot find any other requirement appealing to the particular numerical position in the order, other than the second. Simply put, grammars do not count, hence the significance of the second position must be extra-syntactic. Chomsky (2001) suggests that V2 should be treated as a PF phenomenon due to its opaque motivation and the lack of obvious semantic and morphological effects (though see the above discussion). If it is indeed the case that at least the prefield requirement of V2 is not syntactic, but prosodic, and as such not part of the narrow-syntax, the labeling considerations are simply not operative there, and accordingly, the pressures induced by the labeling algorithm will not affect this construction (i.e., satisfaction of the second requirement part of this construction).There is a fair amount of evidence that this is indeed the case and that V2 is at least partially a post-syntactic phenomenon.

The connection of the V2 requirement as being prosodically motivated has been made quite early in the history of linguistic research. Wackernagel (1892) argued that in early Indo-European languages verbs, similarly to auxiliaries and discourse particles, were unstressed elements when located in the second position (see also Anderson 2000). Hendriksen (1990:169) notices that in the accentuated texts of Old Indian, the Vedic, and Brahmana period, the main clause verbs were unaccented in contrast to the predicates of the subordinate sentences, which were accented.[9] Analogically, the matrix clauses began with a strongly stressed word, often a demonstrative

[9] This suggests that the [Force] feature is connected to the lack of stress, where V with the [+Force] feature gets destressed.

pronoun or an adverb, whereas the subordinating conjunction or a pronoun introducing the subordinate clause were unstressed. Hence the main verb, as being unstressed, appeared adjacent to the initial constituent of the matrix clause; hence, such a situation was much more common in the matrix contexts than in the embedded contexts. Thus, Hendriksen (1990) in his discussion of Kotgarhi and Koci (two dialects of Himachali displaying V2 patterns) speculates that the reason for modern V2 languages to display a similar pattern and to obey the verb-second constraint could be that such word-order rules are a common inheritance in India and Western Europe. Alternatively, the same 'rhythmic' causes could be working independently in both language systems across time (and probably in languages outside the Indo-European family). The intuition is that the reason for V2 now and then could be prosodic.

Thus, historically, V2 has been motivated by prosody. Remnant of this is visible even now, as some modern V2 languages display the connection of the verb placement and the prosody in a quite straightforward way. Thus, Rice and Svenonius (1998) (see also Westergaard & Vangsnes 2005, and Westergaard 2009 for a discussion of language acquisition) show that the placement of the finite verb in the Tromsø dialect of Norwegian is clearly phonologically restricted. Specifically, the verb can appear in the second position only if the *wh*-phrase contains at least two-syllables (one foot). Compare (27), where the finite verb appears in the second position as the prefield *wh*-phrase is phonologically heavy enough, with (28a) where the preceding *wh-phrase* is too light, hence V2 is ungrammatical. This sentence is only grammatical with the verb remaining in the base position (28b).

(27) a. Korsen **kom** ho hit?
how came she here
'How did she get here?'

(28) a. *Kor **kom** du fra?
where came you from
'Where did you come from?'
b. Kor du **kom** fra?
where you came from
'Where did you come from?'

Germanic languages provide other clues as to the prosodic motivation for V2. Consider the following German V2 paradigm. As in most cases the verb has to appear in the second position of its clause in (29) (cf. (29b)). This is not the case in (30): such contexts where the finite verb is not second position in its clause provide evidnce for the prosody being involved in V2 effects.

(29) a. Das Buch **hat** die Frau gelesen
the book has the woman read
'The woman has read the book.'
b. *Das Buch die Frau **hat** gelesen
the book the woman has read
'The woman has read the book.' (German; Bošković to appear (a):12)

(30) Wie reich sie auch sie, # ich **heirate** sie nicht
however rich she too may-be I would-marry her not
'However rich she may be, I would not marry her.' (German: Boeckx 1998: 276)

The German sentence in (30) is grammatical despite the verb being third in its clause. Bošković (2001, to appear(a)) argues that what matters for the well formedness of the Verb-Second

configuration is the placement within the *i(ntonational)-phrase* (where the verb indeed must be in the second position). As the syntactic structure is mapped into the prosodic structure, each syntactic clause corresponds to an intonational phrase, the boundary of which corresponds to the CP edge in syntax (see e.g. Selkirk 1986, 2011, Nespor & Vogel 1986, Hayes 1989, Bošković 2001, 2015b). Now, this one-to-one correspondence between syntax and prosody can be altered as some elements (e.g., appositives, parentheticals, prosodically heavy elements) can form their own separate i-phrases. In cases like these we may be dealing with more than one i-phrase within one syntactic clause. Returning to V2, since it is a constraint operating at the level of the intonational phrase rather than the syntactic clause it is obeyed as long as the verb is in the second position there. This is exactly what accounts for the grammaticality of (30): the pause (marked by '#') indicates the i-phrase boundary here and the verb is placed in the second position within its own i-phrase. The pause thus changes the possibilities for the placement of the verb within the syntactic clause because it affects prosodic phrasing.[10]

Bošković (2001, to appear (a)) explicitly refers to the V2 effect as a PF constraint that filters out in PF constructions where the finite verb does not appear in the second position within its intonational phrase[11]. At any rate, since it is affected by the prosody of the clause, the placement

[10] The placement of the verb in V2 constructions, and especially the placement which looks like V3 or V4 (delayed placement) is often referred to as 'glass ceiling': languages (and even particular constructions within the same language) vary as to what can be ignored in calculation of V2. It is sometimes assumed that what can be ignored is elements that are base-generated in their surface position (see Holmberg 2015 and references therein) but Bošković (to appear) shows that this does not have to be the case since there are cases of moved elements that are disregarded for the V2 constraint determination (see in fact the discussion below). What matters here is the prosody-i.e., that the verb is second in its intonational phrase.

[11] See also discussion on V2 being prosodic in nature in Boeckx (1998). Note that Bošković (2001, to appear(a)) does not assume PF movement here. According to Bošković, we are dealing here with a PF constraint which rules out a syntactic output that ends up not conforming to it.

of the verb within syntactic structure may be affected by phonological effects such as pauses, as (30) shows.

Accordingly, as noted in Bošković (to appear (a)), it is also impossible to have a pause *before* the verb in V2, which again supports the placement of V2 being really prosodic in nature (note that there is a stark contrast here with English *however,* which involves a pause after it, as the translation of (31) shows):

(31) a. *Emellertid # **kan** du inte använda en DVD-RAM skiva som startskiva
however can you not use a DVD-RAM disc as a start-up disc
'However, you cannot use a DVD-RAM disc as a start-up disc.'

b. Emellertid **kan** du inte använda en DVD-RAM skiva som startskiva
however can you not use a DVD-RAM disc as a start-up disc
'However, you cannot use a DVD-RAM disc as a start-up disc.' (Swedish; based on Holmberg 2015)

Hence, the delayed placement of the verb (sometimes even quite substantially as in (32)) is fine syntactically as long as it does not delay the V-placement prosodically (i.e., as long as the verb is second in its i-phrase).

(32) I går, vid femtiden, utanför stationen, när jag kom från jobbet, **mötte** jag en
yesterday, at about.five outside the.station, when I came from work met I an
gammal skolkamrat
old schoolmate
'Yesterday at about five o'clock, when I came from work I men an old schoolmate outside the station.' (Swedish; Holmberg 2015)

Regarding (32) with multiple adverbials, Bošković (to appear(a)) shows that an i-phrase boundary is present after the adverbials that are followed by an adverbial. A comma follows each adverbial, but the adverbial and the verb that follows it are situated within one i-phrase, there is no i-phrase boundary between them (see Bošković to appear (a) for details of the prosody here). Hence *mötte* is indeed second in its i-phrase in this sentence.

Bošković (to appear (a)) also discusses, in this context, German (33) and Dutch (34), using them to show that elements that delay the placement of the verb can be either base-generated in the left periphery (as in (33)) or moved there, as in (34) (see Bošković to appear(a) for relevant discussion and the references).

(33) Peter, ich **werde** ihn Morgen sehen
Peter I will him tomorrow see
'I will see Peter tomorrow (German; Holmberg 2015)

(34) a. Die man, die **ken** ik niet
the man him know I not
'I don't know this man.'
b. För två veckor sen, då **köpte** Johan sin första bil
for two weeks ago then bought Johan his first car
'Johan bought his first car two weeks ago.' (Dutch; Holmberg 2015)

Furthermore, Bošković (to appear(a)) observes that when the delayers of V2 are phonologically light, they are parsed as a phonological phrase together with the following element (e.g. a pronoun) and together serve as one initial prosodic constituent. This includes the question particle, the

adverbial, and the adjunct *wh*-phrase followed by the pronoun in (35-37). The verb placed right after this phonological phrase is indeed second within the i-phrase.

(35) Mon han **er** syg?
Q he is ill
'I wonder if he is ill.' (Danish; Holmberg 2015)

(36) Bara han **kommer** smart!
only he comes soon
'If only he'd be here soon.' (Swedish; Holmberg 2015)

(37) Warum du **machst** DINGS?
why you do things
'Why are you doing that?' (German; Walkden 2016)

More generally, Bošković (to appear(a)) observes that the fact that there is a great deal of crosslinguistic variation regarding elements that are ignored in calculating V-2 is easier to capture if this constraint is indeed prosodic in nature, since intonational phrasing of the kind discussed above is actually subject to a great deal of variation (both across languages and particular lexical items). That what matters here is the placement of the verb in the second position in the intonational phrase provides the strongest evidence for the prosody-based nature of Verb-Second.

Other authors have also argued that V2 involves PF movement and/or that it is a PF phenomenon, see e.g. Jouitteau (to appear), Holmberg (2000), and Anderson (2000).

Thus, Jouitteau (to appear) argues that V2 effects are in fact epiphenomenon of other constraints operating in the language. Using Breton, the author claims that V2 in this language

derives from a morpho-phonological obligatory exponence effect observed at the sentence level, which she refers to as *Left Edge Filling Trigger* (LEFT).

Jouitteau argues that Breton requires the presence of a pre-Tense element, and this requirement can be satisfied by insertion of any material that would precede the inflected verb or the overtly marked Fin head. This could involve an XP or a syntactic head X, and this material could be either base-generated (e.g., in the form of an expletive) or moved. If movement occurs, these are always instances of PF movement (Jouitteau specifies that they occur post-syntactically, but pre-PF, as in Rivero 1999 or Adger 2006) displaying the characteristics of the last-resort operations triggered by LEFT. Its effects always occur on the left of the phase being linearized. In fact, Jouitteau argues that LEFT in Breton kicks in due to the requirement for the explicit realization of a pre-Tense element in the language being the effect of the existence of a particular constraint formulated as *Avoid Tense-First*. The occurrence of V2 can be therefore thought as the result of conspiracy of these two principles, especially in V2I languages.

Jouitteau argues that these LEFT-motivated movements may violate syntactic constraints and argues that this can be captured if we relegate them to the post-syntactic component. Hence, these syntactically ill-behaved movements may involve, among others, a violation of the Head Movement Constraint (HMC) (Travis 1984) (38ha) in *Stylistic Fronting* (see also Roberts 2004 on *Long Head Movement*; for arguments that Stylistic Fronting (SF) involves head movement see Holmberg & Platzack 1995, Jónsson 1991, Poole 1996, Santorini 1994), excorporation and ‘do’ support (38b), or verb-doubling (with the pronunciation of the lower inflected copy) (38c) (the bolded elements in (38) are subject to the second requirement).

(38) a. Prenet **em** eus __ levr evit ma breur deh.
bought Fin.have.1SG a book for my brother yesterday
'I have bought my brother a book yesterday.'

b. Prenañ **a** ris ul levr d'am breur deh
to.buy Fin did.1SG a book to my brother yesterday
'I have bought my brother a book yesterday.'

c. Goud **a** ouien eur mare zo lar e oa teo ar vamm!
to.know Fin knew a time is C Fin was fat the mother
'I knew since a long time that the mother was fat!' (Breton; Jouitteau 2018: 10)

Holmberg (2000) also argues that *Stylistic Fronting* in Icelandic occurs post-syntactically, viewing it as a last-resort operation to obey the V2 constraint.

The observation of the PF character of V2 is also evidenced by the fact that there is very little syntactic restriction on which elements can satisfy the prefield requirement of V2. As we have seen in (1), quite a variety of elements can be placed in the preverbal field, including expletives, which are great candidates for PF-motivated insertion (due to the lack of semantic features; English 'do' appearing in *do*-support is in fact often treated in this way). The fact that the specifiers (if involved) of the verbs in V2 are extremely non-picky suggests that we may not be dealing with a feature-sharing operation there. (This is in fact what Bošković to appear (a)argues; see also Blümel 2017) Jouitteau argues that this is the hallmark of LEFT; in order to satisfy this requirement, which is basically phonological or prosodic in nature, a variety of operations can occur: expletives can be merged, or short movement may occur (i.e., whatever material is close to the position required to be filled will be moved post-syntactically), or material can excorporate from the tensed head itself to get inserted in this pre-Tense position (for relevant PF-based discussions of V2 the reader

is also referred to Anderson 2000 and Adger 2006, where even the very placement of the relevant elements, including the verb, may be a matter of a post-syntactic operation).

We should also understand why the initial position in the clause is important from the perspective of the prosody of the clause. We have seen the connection of the V2 with the expression of Force. Adger (2006) argues that [Force] (or a head that encodes it) is typically a category that marks the left edge of a phonological constituent which is larger than the prosodic word, namely intonational phrase. He proposes a constraint according to which some languages require such aligning of the left-edge of an intonational phrase with the right edge of the Force.

There is also an observation that illocutionary clauses correspond to intonational phrases (i-domains) in phonological representation (Selkirk 2005, 2011). This also holds for embedded clauses, as embedded illocutionary clauses appear to have a stronger tendency than embedded standard clauses to be prosodically parsed as i-domains (Selkirk 2011:453, for more on the correspondences between illocutionary force and i-domains see e.g., Ladd 1996, Selkirk 2005, 2009, Dehé 2009).

The initial position within a given i-domain has a special status (Selkirk's 2011 'Strong start effects') and it is certainly visible at the level of i-phrases. This is also confirmed by facts from other languages. Bošković (2015b, to appear(a)) shows thati-phrase boundaries are places where many exceptional phonological processes occur. These effects can be even stronger at the i-phrase boundaries of main clauses. Bošković discusses Japanese case-drop, which occurs only there, i.e., it's a main clause phenomenon. Namely, Japanese case-markers allow stranding under NP-ellipsis in these contexts. When the case-particle is adjacent to the i-phrase boundary it can receive an extra-stress in this position, and this allows the nominal to destressed and elided.

(39) Naomi-mo moo tsuki-masi-ta ka? Naomi-GA mada tsuki-mase-n
Naomi-also already arrive-POL-PAST-Q Naomi-NOM yet arrive-POL-NEG
'Has Naomi already arrived?' 'She has not arrived yet.' (Japanese; Otaki 2011)

There is abstract similarity here with initial V2 (historically). With the Japanese case-drop, a prosidically deficient element is exceptionally stressed, with initial V2, a prosodically non-deficient element was exceptionally distressed, with both processes occurring at i-phrase edges. Hence the association of V2 with the left edge of the i-phrase is non-accidental. As Bošković (to appear(a)) notes, the fact that in many languages V2 is predominantly a main-clause phenomenon can also be explained from the prosodic point of view: this is the place where an utterance boundary and an i-phrase boundary meet (this is in fact what matters for Japanese case-drop). Thus, the interaction of prosodic constraints with the syntactic constraints may be expected in these positions more than anywhere else.

V2 is also reported to be acquired very early, even as early as around age 2 (Wexler 1998).We know that prosodic cues are one of the most important clues in the process of acquisition, enabling bootstrapping, i.e. faster acquisition of such phenomena as word order or the patterns of *wh*-placement. If the V2 placement is indeed not only guided by prosody, but in general a prosodic process, we can perhaps understand how children can master this construction so early.

Prosodic clues are also very influential in language contact situation. We know that syntactic borrowing (i.e., adopting of a particular construction) is quite rare and constrained (Harris & Campbell 1995), however, intonational and prosodic patterns are more prone to be adopted in interlinguistic contexts. Hence the fact that almost all V2 languages developed in areal contact with each other may point towards prosodic considerations enabling successful and quick adoption

of this construction by languages involved in contact. This is confirmed by the development of V2 in Sorbian or Estonian through contact with German (Ehala 2006).

Finally, V2 is really a spoken language phenomenon, not attested in sign languages. In this respect, the prosodic account of this phenomenon makes sense as it is straightforwardly connecting this constraint to the particular modality of spoken languages rather than a deep syntactic property.

We have seen that we have strong indication that V2, especially V2I, may be *not* subjected to the pressures connected to the labeling requirement because it may in fact involve head-labeling (with the XP in the prefield undergoing post-syntactic/PF movement). Its relative diachronic stability indicates that this construction may not be as problematic from the point of view of labeling, as other instances of XP or head-movement discussed in the present work.

As far as the V2D languages are concerned, as the movement in V2 is discourse-driven, it may have to occur in the syntax, and it cannot be postponed. The appropriate [uF] trigerring displacement has to be licensed by movement in such cases (some languages may have an option of base generation, e.g., Classical Latin (see Ledgeway 2008)). Hence, we assume here that V2D involves a V-to-Topic$_0$ movement, which importantly is not a self-attachment as with V2I, together with XP movement of the topicalized element. Both occur in the syntax, give raise to symmetrical merger configurations. If this is indeed the case, we can predict that V2 labeled in this way may show signs of diachronic reanalysis; the (already familiar) pressures of the labeling algorithm are operative in this context. And this is what we observe. V2D tends to be lost diachronically (as we will see in the next section), and if the discourse-driven displacement still occurs in a language, very often only one of the two V2D related movements (XP movement, which is motivated by extra-syntactic need to express certain discourse-related notions (see chapter 2) is maintained. This is what we already saw in English (3).

Bošković (2018, to appear(a)) offers a syntactic account of V2 which also presupposes the lack of labeling-related issues (in the sense of labeling-connected pressures), hence this analysis can be applied to our V2I. Bošković suggests that the observed non-pickiness of the V2-related specifiers hosting the variety of material in the V2's prefield indicates that V2 does not involve feature-sharing (see also Haegeman 1996, Roberts & Roussou 2003, Roberts 2004, Jouitteau 2008) and thus may not be labeled in narrow syntax. In fact, Bošković appeals to the prosodic roots of the second position placement discussed in this section, suggesting that contemporary V2 constraint is an effect of grammaticalization of the purely prosodic requirement of early Indo-European, adopted into syntax as *I-need-a-Spec* requirement. What drives it is an EPP feature, devoid of agreement so that practically any element can fill the Spec of the V2 verb; it is impossible to find one prominent feature (holding for all XPs in the prefield and verbs) that could participate in the agreement relation between the Spec and its sister to provide a label for this configuration. Hence, no feature sharing and no labeling occurs here.

Bošković supports the claim about the lack of a label in the syntax for these constructions with the observation that V2 clauses are syntactically immobile (also Webelhuth 1992, Reis 1997, Wurmbrand 2014, Holmberg 2010/2015).

(40) *weil [CP den Peter mag niemand] allgemein bekannt ist.
since the.ACC Peter likes nobody.NOM commonly known is
'since nobody likes Peter is commonly known' (Wurmbrand 2014:155)

Sentence (40) serves as an illustration of the immobility of V2 clauses. Bošković argues that it is the lack of a label that renders them immobile.

Bošković connects this to the observation that movement out of moved elements is disallowed (Wexler & Culicover 1980, Takahashi 1994, Stepanov 2001, Corver 2017, Bošković 2018, and many more). This is observed in the ungrammaticality of extraction out of subjects (41) and moved objects (42).

(41) ?* I wonder [$_{CP}$ who$_i$ [$_{DP}$ friends of t$_i$]$_j$ [$_{vP}$ t$_j$ hired Mary]]

(42) ?* Who$_j$ did Mary call [friends of t$_j$]$_i$ up t$_i$?

Both (41) and (42) involve movement out of a moved element. Given the VP Internal Subject Hypothesis, subject in (41), has moved from the [Spec,vP] to the [Spec,IP]; Object in (42) has undergone object shift to precede the particle (Lasnik 1999, 2001). On the surface, it looks like that movement out of a moved element poses a problem here. Bošković, however, argues that the problem occurs much earlier.

Bošković deduces the generalization about the ban of movement out of moved elements from the interaction of the phase theory and labeling. One of the differentiating characteristics of phases that distinguishes them from non-phases is that only phases can undergo movement (Chomsky 2000, 2001, Rackowski & Richards 2005, Legate 2014, Bošković 2015b, among others). The moved elements in (41-42) have to be phases then. Now, the Phase Impenetrability Condition (PIC) (Chomsky 2000, 2001) states that extraction out of phases must proceed through their edges, as only those positions are visible from outside of the phase and allow further movement. Therefore, given the PIC, the elements moving out the phases must move to their edge. The mechanisms of the syntactic cycle dictates that the movement to the edge of the phase must occur before the movement of the whole phase. This, however, has consequences for the determination of the label of this phasal projection.

Importantly, movement to the edge of the DP phase in (41-42) is a pure successive-cycle movement. As discussed in chapter 1, such movement does not have any feature-based motivation and no feature is checked in intermediate successive-cyclic positions (see Bošković 1997, 2002, 2007, 2008, Chomsky 2013), i.e., we are dealing here with the same situation as with successive-cyclic movement in long distance *wh*-questions (pure successive-cyclic movement in general does not involve feature-sharing in Chomsky 2013).

(43) a. What$_i$ did you say [$_{CP}$ t_i [$_C$ that [$_{TP}$ John bought t_i]]]?

b. [$_?$ what$_i$ [$_C$ that [$_{TP}$ John bought t_i]]]

What is important here is that since no feature-sharing occurs, no label can be determined. In both cases, the movement to the edge of the phase in effect de-labels this structure (notice the '?' in (43b)). We have already seen in chapter 2 that an element has to be labeled to be able to function as a phase (Phases are objects like vP and CPs--it is the label that tells us which projection is a phase; we need to know which phrase is e.g. a vP or a CP. It follows that unlabeled objects cannot be phases). De-labeling these structures then has consequences in removing their phasal status and rendering them immobile, since only phases can move. We can see then that the movement that is problematic in structures like (41)-(42) is really not the subsequent move out of the moved element but rather the initial movement of the element that will later be extracted out of (due to the way that the syntactic cycle works).

Under Bošković's analysis it is expected that elements which are base-generated at the edge of a moved element or get there through obligatory non-successive-cyclic movement can be extracted out of moved elements (the ban on the movement out of moved elements can then be violated). Bošković gives a number of cases of that sort, one of which is given in (44):

(44) Jovanovu$_i$ je on [$_{NP}$ t$_i$ sliku] vidio t$_j$

John's.ACC is he picture.ACC seen

'He saw John's picture' (Serbo-Croatian; Bošković 2018)

Bošković shows that the possessor in (44) is base-generated at the edge of the object NP (see also Despić 2013). As *Jovanovu* is extracted out of its base position at the edge of the fronted object 'picture' no ungrammaticality occurs. The possessor and the object do undergo feature sharing here, as indicated by the fact that the possessor may stay in this position without extraction (and it also agrees in phi-features and case with the noun).

Returning to the V2, Bošković unifies the immobility of V2 clauses and the movement out of moved elements. The latter holds only for successive-cyclic movement because such movement does not involve labeling, in effect de-labeling the object it targets. On a par with this, movement to Spec,CP of V2 clauses also does not involve feature-sharing, which leaves the clause without a label, rendering it immobile (see also Blümel 2017 on V2 not involving labels).

Bošković speculates that V2 clauses are interpreted even in the event of the lack of a label by a special rule. Alternatively, Bošković (to appear(a)) suggests that V2 structures can be labeled at the spell-out point, by simply ignoring the element in the [CP, Spec] and enabling the other projection in the {XP,YP} merger to label the structure (similarly to traces being ignored for labeling).

At any rate, what is important for us from Bošković's account is that V2 clauses do not involve feature sharing between the Spec and its sister. This supports the current claim that this makes them not subject to labeling-induced pressures, which is indicated by their historical resistance to change.

3.5.Diachrony of V2

Regardless of its diachronic stability, V2 phenomenon has also been subject to historical change and reanalysis in some languages. In this section it will be argued that there are two types of V2, where only one of them is subject to labeling pressures and shows signs of diachronic change (the focus will be on the head movement part of V-2).

We will show that this is indeed the actual state of affairs despite some difficulties in teasing the V2 systems apart (it is possible that in some cases they are really not two separate mechanisms, but they merged into one V2 system (if stable, then derived through self-attachment and prosodic movement).

In this section we offer a very brief overview of some historical facts connected to the loss of V2 in Early Romance and Old English with some proposals to explain these facts.

Despite there being a rich tradition rejecting the V2 hypothesis for Medieval Romance languages (Kaiser 2002, 2009, Zimmermann 2014, Kaiser & Zimmermann 2015, Martins 2002; Eide 2006; Rinke 2009; Sitaridou 2012), Wolfe (2016, 2019)[12] argues that many of them were V2 languages. This observation confirms some earlier work indicating that V2 was indeed operative in Old French (Adams 1987, Roberts 1993, Vance 1995, 1997, Labelle & Hirschbühler 2017), Old Portuguese (Ribeiro 1995, Galves & Kroch 2016)), and other Medieval Romance languages (Fontana 1993, 1997, Ledgeway 2008, 2017), and that these early Romance languages can be possibly analyzed similarly to Germanic V2 languages. Old Neapolitan below illustrates this fact:

[12] We base the discussion on Wolfe 2019, being the most in-depth work analysis of V2 phenomenon in Medieval Romance languages.

(45) de poy queste parole **ademandao** lo messayo licencia
after these words ask.3SG.PST the messenger permission
'following these words the messenger asked permission (to leave)' (Old Neapolitan; Ledgeway 2008:441)

Old Neapolitan displays the characteristics of V2 configuration, namely the inversion of the subject (when overtly expressed) with the verb, with the subject still preceding all the other elements in the clause (as the direct object *licencia* here), when an element different from the subject (here, underlined PP) is fronted. Other characteristics pointing towards V2 in early Romance include the fronting of the direct object with no clitic resumption in variety of contexts, as well as asymmetries between the main and subordinate clauses with regards to the position of the finite verb (see also Adams 1987, Vance 1997). Wolfe takes the existence of a sufficient amount (i.e., comparable to Germanic languages) of the clauses in his corpus to involve the word order in (46), augmented by the evidence of the movement of the finite Verb to the left periphery to be convincing argument to adopt a V2 analysis for the Medieval Romance class.

(46) $XP_{\text{Non-Subject}}$-V_{Finite}

The primary characteristics of the V2 order, e.g., pragmatically marked fronting of the finite verb together with the fronting of pragmatically salient constituents had already been present in later Latin texts (Bauer 1995, Bauer 2009, Spevak 2004, Clackson & Horrocks 2007, Horrocks 2011, Ledgeway 2012, Danckaert 2017, Wolfe 2019). We can assume then that Latin was V2D in those early stages of V2. Wolfe points out that what was responsible for the development of the full bloom and obligatory V2 in later Romance languages was reanalysis of these initially discourse-marked contexts as a neutral word order. The author hypothesizes that this indeed could be a more

general path in many other regular V2 languages. We can add that at that point, V2 either gets reanalyzed as V2I or lost.

Looking at the phenomenon of the finite Verb raising to left-peripheral position in Old Spanish, Old French, Old Piedmontese, Old Occitan, and Old Sicilian, Wolfe (2016, 2019) observes that most of these Medieval Romance languages had undergone a change from a 'relaxed' V2 language with widespread V3, V4, or V5 word orders to the 'strict' V2 language, with the restriction of only one constituent to be able to appear in the prefield. We should add here that the distinction into relaxed vs. non-relaxed placement of V2 in the system advocated here relates to our distinction into V2D vs. V2I, where the latter are more strict than the former in terms of the permissible number of constituent in the prefield. This relates to the placement of finite verb in these systems. V2D involves placement of the verb in a Focus/Topic related position. For V2I it is a projection encoding Force. The latter position in higher, hence the fact that fewer elements can be accommodated preceding the verb follows automatically. Topic and Focus head positions, on the other hand, are placed lower in the structure and therefore more elements can precede the verb. Additionally, the prosodic basis for the 'second position' requirement, discussed earier, is much clearer in V2I constructions than in V2D, where the verb placement relates to the information structure manipulations. This can also explain why V2I is more rigid.

At any rate, the first phase of the V2 in the languages in question (between the 10th and 12th century) is illustrated in examples from the Oldest Italo-Romance (around 960 AD) (47), Old French (48), and Old Spanish (49).

(47) Sao ko kelle terre per kelle fini, que, ki, contene, trenta anni
know.1SG that those lands for those confines thathere contain.3SGthirty years
le **possette** parte sancti Benedicti
them.CL possess.3SG.PST part saint.OBL Benedict.OBL
'I know that those lands, within those borders which are contained here, have belonged to the part of the monastery of St Benedict for thirty years' (*Placito* 1, Ledgeway 2011:215, 960AD)

(48) E pur ço que cist lignages numéément dout si le service Deu
and since that these lineages named should thus the service God
celebrer, besuinz **fud** ke...
celebrate.inf necessary be.3SG.PST that
'Since these lineages had to therefore celebrate the service of God, it was necessary that…' (Old French, *QLR* I, 1, c.1170; Wolfe 2019:163)

(49) SI NOS D'AQUI NON IMOS en paz nunca **bivremos**
if we from-here NEG go.1PL en peace never live.fut.1PL
'If we don't leave here, we'll never live in peace' (Old Spanish, *Alexandre* 254, 1178-1250; Wolfe: 163)

The finite verb is indeed placed later than in the second position in such examples from this period, and it is preceded by topicalized and focused constituents, often appearing together in a sentence, as their movement was obligatory. In Old French (48) the finite verb 'be' is preceded by the topicalized clause '…since these lineages had to therefore celebrate the service of God' followed by the focused 'necessary.' Similarly, (47) and (49) involved such discourse-related constituents

moved into the preverbal positions, altogether with some base generated adjunction phrases, often related to the context of the utterance (which usually don't count for V2 effects even in regular V2 languages of the Germanic type due to the glass-ceiling effect we noted earlier). This indicates that the early Medieval Romance languages had V2D structures.

L'Arrivée (2019) confirms this in his quantitative study of V2 in Old French. Looking at the homogenous corpus of prose legal texts from the Normandy region, the author observes that the period between 1150-1236 is characterized by the categorical discourse-old value of the preverbal elements. Overall, the discourse-old contexts make up 73,6% of all the occurrences in his corpus. L'Arrivée (2019) argues that the loss of discourse-old value triggers the loss of V2. Indeed, this is one option for V2D configuration predicted by the current analysis: its diachronic loss.

Wolfe (2016, 2019) reports that later Old Romance languages see the reduction in the number of the constituents in the prefield. Most V2 clauses involve at most one constituent, V3 orders being rarer and occurring only with attested earlier scene-setting adverbials. V4 orders are almost not attested at all. This is illustrated in (50-52) below:

(50) Et quant il est apareilliez, il **prent**…
and when he be.3SG appear.PTCP he take.3SG
'When he appeared, he took…'(Old French, *La Queste* 129, 1215–1230; Wolfe 2019: 95)'

(51) Unde Brat **levà** la ma(n)
thus Brat raise.3SG.PST the hand
'Brat then raised his hand' (Old Venetian; *Lio Mazor* 51, 1312–1314: Wolfe 2016: 11)

(52) Et luego que llego a la puerta el diablo **abrioge**

and soon that arrive.3SG.PST at the door the devil open.3SG.PST=it

'And as soon as he arrived at the door, the devil opened it" (Spanish, *Lucanor* 204, 1335: Wolfe 2019:116)

The observation that the V2 system in these languages involved the change towards 'more strict', in Wolfe's words, V2 word order seems to be descriptively accurate. However, we have an indication that the prefield elements are still discourse-related, hence Old Romance continues its V2D status. This is confirmed by examples like (53), involving the Romance expletive particle *si*, which precedes the finite verb and follows the topicalized and focalized elements.

(53) Durement en halt si **recleimet** sa culpe

strongly on high SI confess. 3SF his sin

'He confessed his sins aloud.' (Old French, *Roland 2014*, Wolfe 2016:6)

Confirming this, van Kemenade & Selvesen (2018) argue that the Old French of the 13th century still involves topic movement (referred to as Left Dislocation) to the pre-verbal position, which may suggest that the final-verb is still preceded by fronted elements, which are discourse/ information structure related.

Wolfe proposes that the observed constraint that the V2-verb be preceded by only one element (in most cases), visible in the 14th century Old Romance, is due to the change in the position of V2 in those languages. The author proposes that the verb shifted its position to a higher projection.[13]

[13] Wolfe (2019) argues that the shift in question occurred from the Fin to the Force projection in Rizzi's articulated left periphery. Appeaing to cartographic approach, our analysis would rather point to Top or Foc would be more

Now, we can ask what it would exactly mean for the verb in the later Old French, Old Spanish, or Old Venetian to shift into the higher projection, especially if it undergoes movement to this higher projection. Diachronic changes very often involve re-analysis of lexical material in the higher position, as in the grammaticalization process discussed in this chapter. That process, however, involves base-generation, not movement, which is what we are dealing with here. One solution that presents itself here is to assume that the change in question resulted due to the reduced number of the prefield elements. The answer to the question why the prefield in Later Old Romance appears to be reduced is the loss of movement to the high left periphery.

This is confirmed by one important change that occurred in the early Medieval Romance. The 'information' focus movement, which was obligatory in 11th-12th century (similarly to topic movement) was lost and replaced with the post-verbal focus licensing as in Modern French or Spanish (Belletti 2005, Zubizaretta 1998). This is certainly a type of changes we have already seen, and which can be explained by the labeling system and its pressures advocated here. Hence, we can see that the process of moving towards a stricter V2 was strengthened by some loss of movement.

What is important for us here, is that we still have reasons to assume that V2 in later Old

relevant projections that V2 in Old Romance could target, as it displays the characteristics of V2D, hence connection to the discourse-relevant projections rather than Force or Fin, connected to the illocutionary force.

Romance was discourse-driven V2, hence V2D, which suggests that it still had to be labeled in syntax. Hence, its later loss can be connected to the labeling-triggered dispreference for Spec-head labeling resorting to feature-sharing.[14] [15]

This is supported by a similar change that occurred in the same discourse-related contexts from Old Portuguese (OP)/ Classical Portuguese (CLP) to European Portuguese (EP). Old Portuguese, as a V2 language, displayed a high frequency of sentences with subjects either following the verb or being null, with other constituents appearing in pre-verbal positions (Ribeiro 1995, Galves & Kroch 2016). Sentece (54) with OVS word order and an adverbial following the verb indicates that the verb is high in the clause

(54) E esta vertude de paceença ouve este santo monge Lobertino mui compridamenete
'And this virtue of patience had this holy monk Libertino very fully' (DCG 1.5.5; Galves & Kroch 2016:491)

Interestingly, OP/CLP 's verb movement often involved lack of XP movement to the prefield, hence V2 in this language involved spec-less CP. We can assume, then, that this OP/CLP V2 involved the pattern we are already familiar with from chapter 2, namely the loss of specifiers. The loss of V-to-C movement followed as we observe its substantial decline in the transition from CLP to EP, with the modern language not having such general V-to-C movement in such discourse-related contexts, and in general in declarative contexts**.**

[14] Alternatively, if indeed V2D of the early Romance languages changed into V2I, which resulted in the disappearance of the labeling problem, its loss could be connected to changes in the prosodic component of the language under the current approach (Mathieu 2016:266).

[15] There is also potentially an important connection between the loss of V2 in Old French and the loss of null subjects and null topics (see Adams 1987, Vance 1997, Roberts 1993, 2007,Yang 2002, Wolfe 2016, 2019) on analyses relating these two phenomena.

Finally, another case of the loss of V2 is attested in English (van Kemenade 1987, Kroch & Taylor 1997, Pintzuk 1991, Fischer et al. 2001, Haeberli & Ihsane 2016). Old English and Middle English were considered to be V2 languages and the loss of this constraint is reported to occur by approximately the 15th century (Fischer et al. 2001:132-7).

Importantly, English lost V2 (the head movement part of V2) only in discourse-marked V2D contexts, the 'residual' V2 that Modern English still has today is V2I, as it appears e.g., in the context of questions, but not in Topic or Focus. This supports our argument as to the diachronic fragility of the contexts involving less economical labeling options, such as feature sharing involved in the derivation of V2D.

The Old English V2, apart from the interrogative (55) and negative contexts (56), could also appear with with the temporal (discourse-anaphoric) adverbs *þa*`, *þonne* 'then', and *nu* 'now', whose main function was to sequence the main events in the discourse.

(55) for *hwam* **noldest** Þu ðe sylfe me gecyðan Þæt…
for what not-wanted you yourself me make-known that…
'wherefore would you not want to make known to me yourself that…' (*ÆCHom* I, 1.14.2; Fischer et al. 2004:106)

(56) *Ne* **sceal** he naht unaliefedes don
not shall he nothing unlawful do
'He shall not do anything unlawful.' (*CP* 10.61.14; Fischer et al. 2004:106))

(57) *þa* **wæs** þæt folc þæs micclan welan ungemetlice brucende…
then was the people the great prosperity excessively partaking
'Then the people were partaking excessively of the great prosperity.' (Or 1.23.3: Fischer et al. 2004:106)

When V2 appears in declarative clauses, it happens in approximately 70% of cases (Haeberli 2002:250). What is more, V3 clauses are very commonly attested and fronted constituents are very common too. The immediately preverbal constituent, however, usually has the interpretation of a familiar topic and is rarely prosodically prominent (Walkden 2014, 2017, Speyer 2010).

(58) a. *Þis ylce galdor* **mæg** mon singan wið smeogan wyrme
this same charm may man sing against penetrating worm
'One can sing this same charm against a penetrating worm.' (colacnu,Med_3_[Grattan-Singer]:27.1.132; Walkden 2017:72)

b. Fyr ic sende on eorþan
fire I send to earth
'I send fire to earth.' (cowsgosp,Lk_[WSCp]:12.49.4719; Walkden 2017: 72)

Additionally, OE displayed the abundance of verb-late clauses (Pintzuk & Haeberli, 2008), indicating that V2 is not subject to a strong constraint of the type of modern Germanic languages (excluding English), which are V2I.

All these facts may indicate that we are dealing here (in OE and ME) with a V2D language, where V2 arises only in very particular contexts with semantic effects closely connected to the information-structure of a sentence. This indicates that similarly to Old Romance, V2 in English

was licensed and labelled in narrow-syntax via feature-sharing. As these languages displayed V2 only in operator-related contexts, we assume that the prominent feature involved in Spec-head agreement (and head-to-head adjunction) was available to provide the label for the structure. Since labeling of these V2D constructions occurred through the dispreferred strategy (from the point of labeling algorithm which had to resort to additional operations), the pressures connected with labeling were operative and the diachronic decline and ultimate loss of V2 in English can be accounted for.

The loss of V2D in English proceeded in a gradual manner. The V2 in topic-initial constructions shows a sharp decline in the late 14th century, and in the 15th century inversion of nominal subjects declines further. Fischer et al. (2004) report surveys of inversion in sentences introduced by *then, now, there, here, so, yet* and *therefore*, which show the percentages between eighty-five and ninety percent for the late 14th century, twenty-eight to thirty-nine per cent for the mid-15th century, finally ten per cent and lower for the end of the 15th century.

Hence was saw that English had both V2D (in declaratives) and V2I (e.g., in operator contexts: *wh*, negation, with some adverbials). The former was lost but the latter survived.

Overall, we can see that by postulating the distinction into two types of V2 constructions, discourse-related (V2D) and illocutionary force related (V2I), we can account for the loss of V2D and diachronic stability of V2I by appealing to the labeling system, which disprefers the feature-sharing option, which we have shown is involved in V2D but not V2I.

3.6.Conclusion

Verb-Second is a complex and multifaceted phenomenon, subject to rich cross-linguistic variation, despite its rarity among languages. What is especially interesting is its perseverance across time,

with the instances of diachronic reanalysis restricted to very specific subset of V2, which we have argued here is almost exclusively limited to discourse-related movement, i.e. V2D.

As such, its diachronic decline and loss is predicted by the labeling-based analysis proposed here. Other proposals attempting to explain the loss of V2 in English appeal to many other factors, such as the availability of V-to-T movement, disappearance of null-subjects and requirement for TP to be overtly filled (Fuss 2008, Migdalski 2016). Kiparsky (1996:19) proposes that the availability of V-to-T in embedded contexts was also responsible for the change in the word order in Old English from OV to VO (similarly in Icelandic and Yiddish). I have argued in chapter 2 that this change also involves the loss of movement, which can be attributed to labeling considerations.

This indeed gives us three diachronic processes, one occurring after the other, each of them being really the loss of syntactic movement. This gives us the following time-line of the diachronic changes:

1) The loss of OV (OV >VO)(12th -14th century)
2) The loss of V2 (14th century)
3) The loss of V-to-T (up to 17th century)

As all of these changes involve the loss of movement, the fact that the one triggered diachronic reanalysis of the other is quite interesting and can be perceived as the confirmation of the existence of diachronic *drift*, i.e., language change occurring over an extended period of time, showing a particular (labeling-driven as argued here) direction (see also Sapir 1921, Roberts 2007, 2017). Such language change is often attributed to some economy considerations. The current work has proposed a particular labeling-based implementation of these economy considerations. At any rate, in the next chapter I will discuss the link between language acquisition and diachronic change.

Appendix: V2 in Kashmiri

In the context of different types of V2 discussed in this chapter, it is interesting to look at V2 in Kashmiri, which displays some unique properties.

Kashmiri is a split ergative language with an underlying SOV (head-final) word order. What already makes this language to some extent unique is that it is one of the rare non-Germanic languages that displays the V2 constraint. The verb final word order can be seen in this language in constructions that do not exhibit V2, such as non-finite contexts (1a), but also in adjuncts and relative clauses (1b).

(1) a. me chu [tem-sund batI **ran-u-n**] khar-aan
I.DAT aux he-of food cook-Inf dislike
'I don't like his cooking food.
b. yus laRk tsoT **khyv-aan** **chu**…
which boy bread eat-N.PERF aux
'Whichever boy is eating bread…'

The main verb is also in the clause-final position when the auxiliary undergoes V-2 movement, which is typical for V2 languages.

One of the most detailed and comprehensive analyses of V2 in Kashmiri is Bhatt (1999) (which is based on Bhatt & Yoon (1992)). The author proposes that V2 in Kashmiri targets the MoodP projection, which, he argues, is responsible for clause-typing. Bhatt (1999) proposes that the C_0 category can be decomposed into two functions, namely clause-typing and subordination. In some

languages these two are conflated into one projection (e.g., in English), but in other languages, these two components of the complementizer can be distinguished as they are separate lexemes. Bhatt (1999:152) argues that the latter case is attested in Korean, where subordinators and 'mood markers' can be distinguished (see also the discussion of Korean force particles in section 3.2. of this chapter; in fact, we have identified the very same particles as Force-related). Now, we can directly translate Bhatt's analysis into the system of V2I endorsed here; Bhatt's 'Mood' is really the projection encoding the Force of the utterance in our system. This gives us an account of Kashmiri as a V2I language: this V2 is illocutionary force-based and the finite verb fronts (re-merges, as we have argued) to satisfy the requirement of clause (or force)-typing (as proposed by Cheng 1991). Let us see then see how the properties of Kashmiri V2 can be explained by the association of this V2 with force.

First of all, Bhatt (1999) (see also Manetta 2011) observes that in most cases, the Kashmiri prefield cannot consist of both topic and focus. This is expected in a strict V2 language, which allows only one element to precede the fronted verb. Hence (2) below is ungrammatical:

(2) *gari bI **goo-s** vakth-as peth
home 1SG. went.PST-1SG time-DAT on
'Intended: 'As for home, I went there on time.' (Manetta 2011: 23)

Not all cases of 'delayed' placement of the verb are ungrammatical. Something that looks like V3 is attested in interrogatives, with a *wh*-phrase in the immediate prefield of the verb even if the *wh*-phrase is preceded by another element, which usually has a topic interpretation.

(3) rameshan, *kyaa* **dyutnay** tse?

Ramesh.ERG what.NOM gave you.DAT

'As for Ramesh, what is it that he gave you?' (Bhatt 1999:107)

In fact, Bhatt (1999) argues that this is the most unmarked order in interrogatives in Kashmiri, where the *wh*-phrase immediately follows the topic. What is characteristic of this configuration is that the initial topic is followed by a pause. This indicates that it constitutes its own intonational-phrase, independent from the i-phrase of the main clause. This means that the verb is still second within its i-phrase, hence the V2 constraint is satisfied (see section 3.4).

This is confirmed by another construction where the verb again seems to be in the 3rd position, but only apparently. This time the first element in the immediate pre-verbal position does not have to be interrogative. However, the situation is analogical to (3), with the initial topic being dislocated in the clause-initial position and additionally linked to a demonstrative resumptive pronoun inside the clause.

(4) [Su laRk$_i$] pananyivi$_j$ achav **vuch** rameshan$_j$ temis$_i$

That boy self's eyes saw Ramesh.NOM he.DAT

'As for that boy, it is with his own eyes that Ramesh saw him.' (Bhatt 1999:102)

Again, V2 is satisfied in this configuration, the verb being placed in the projection of the Force, which is the highest head in the clausal spine. We can assume that the topic phrase is base-generated adjoined to this projection (the resumptive pronoun indicates the lack of movement of the topic here).

Returning to *wh*-questions in Kashmiri, we have here additional support for the association of V2 with the Force. In multiple *wh*-questions, at least one *wh*-phrase has to front to the clause-initial

position. The remaining *wh*-phrases can stay in-situ, or optionally can also front into the preverbal position.

(5) a. *kus* **dii** *kemyis kyaa*?
who give.FUT whom what
'Who will give what to whom?

b. *kus kemyis kyaa* **dii**?
who whom what give.FUT
'Who will give what to whom?

The requirement of at least one *wh*-word to move strongly indicates the connection between V2 and clause-typing (notice that this holds even in residual V2 languages like English), with the *wh*-movement targeting the projection of Force. As for multiple *wh*-movement (5b), on a par with Rudin's (1988) treatment of multiple wh-frongint in Bulgarian, I suggest that all these *wh*-phrases form one adjunction structure where they are all parsed as one constituent in the i-phrase, making the verb appear in the second position, as we have seen it before in this chapter.[1]

Bhatt (1999) also gives examples of cases where a dummy pronoun *yi* ('this', 'it') is inserted into the prefield. The author claims that this pronoun's distribution in such contexts is much more restricted than that of expletive elements in Icelandic (*Það*) or German (*es*) V2; however, the fact that it constitutes a legitimate prefield in V2 shows that its function is related to the Verb Second satisfaction.

[1] The alternative analysis of Bulgarian-style MWF, proposed in Koizumi (1994) is that the fronted *wh*-phrases are located in separate specifiers (Manetta 2011 actually adopts this analysis for Kashmiri. The issue of adjunction vs. multiple specifiers, however, is not crucial for our purposes: if the multiple-Spec configuration can be parsed prosodically as one constituent, then it really does not matter which of these two analysis is adopted (for relevant discussion see in fact Bošković 2001, who also notes that that the does not make a relevant prosodic difference).

(6) *yi* **oos** akh baadshah

EXPL was one king

'Once upon a time, there was a king.'

Kashmiri V2 is of a 'symmetrical type' in that it holds in both matrix and subordinate clauses, and the conditions on its use are the same in both of these domains. It appears in both declaratives (7) and interrogatives. (8a-b) below illustrate its use in indirect questions, where the finite verb follows a *wh*-phrase (8a), which also can be preceded by a topic (8b).

(7) me buuz ki rameshan **vuch** raath shiila

I heard that Ramesh saw yesterday Sheila

'I heard that it was Ramesh who saw Sheila yesterday.' (Bhatt 1999:98)

(8) a. Miiraayi ch-a pataa ki *kəm-is* **di-ts** mohan-an kitaab

Mira Aux-3SG know that who-DAT give-PAST.F.SG Mohan-ERG book

b. Miiraayi ch-a pataa ki mohan-an *kəm-is* **di-ts** kitaab

Mira Aux-3SG know that Mohan-ERG who-DAT give-PAST.F.SG book

'Mira knows who Mohan gave a book to.' (Manetta 2011:23)

Sentences like (7) and (8) are used by Bhatt (1999) to motivate his analysis of the split of the complementizers into subordinators and Mood (our Force). Namely, the element *ki*, here translated as 'that', in Bhatt's analysis is an example of a subordinator, independent from the projection expressing the type of the clause (again, our Force), where the finite verb moves in Kashmiri

(similarly to Korean *ko* which is also independent from Force particles). Hence, when *ki* is present, the *wh*-phrase appears between *ki* and the verb in the second position (notice that *ki* is also not considered in V2 calculation, which means that prosodically it is parsed either with the preceding verb or the following wh-phrase). This treatment of *ki* in fact accounts for many of its properties, not only in Kashmiri, but also in other Indic languages like Hindi (not a V2 language).

One additional interesting property of V2 in Kashmiri is the fact that this construction is often associated with a cleft-like interpretation where the pre-verbal position is associated with focus, also bearing focal stress. Sentence (9) below even contains a focus particle *ti*, which is attached to the first constituent preceding the verb.

(9) Huun-**ti** chu behna broNh panin jay goD saaf karan
dog-even is sit before self's place first clean do-NPerf
'Even the dog cleans his place before sitting.' (Bhatt 1999)

The XP in the specifier of the Force projection can thus function as a focus. We already saw that this is the position that is required to be filled by *wh*-phrases, which also associate with focus in many languages (see e.g., Kotek 2014). Does this mean, however, that Kashmiri is a V2D language? Not necessarily, as V2 in this language cannot involve topics (in contrast to V2D languages like Old English or Medieval Romance). Bhatt (1999) specifically argues that Kashmiri does not tolerate topicalization in its prefield. If, e.g., objects appear in the pre-verbal position, they cannot be interpreted as discourse-old elements. The author supports this claim by citing examples involving universal quantifiers in the prefield, which are unlikely to be topics (May 1977:146-147).

(10) sooruyikenH **khyav** rameshan

everything ate Ramesh

'Ramesh ate everything.' (Kashmiri; Bhatt 1999:87)

Additionally, Bhatt (1999:86) notes that elements which often resist focalization ('inherently unfocused' in his words), like indefinites, are often awkward in these constructions.

(11) ?* kaNh **oosuyi** tset shaanD-aan

someone was you looking

'Someone was looking for you.'(Bhatt 1999:86)

However, it is not the case that V2 in Kashmiri is a necessarily used to express focus (or that Kashmiri is a focus-movement language), as some V2 constructions do not involve focus interpretation. (12) below illustrates this with V2 involving a subject in the prefield which is devoid of focal stress/interpretation (recall that Kashmiri is an SOV language, hence the verb immediately following the subject indicates its fronting).

(12) Azkal **cha** rameshas shiila khosh karan

these.days is Ramesh Sheila happy do-N.Perf

'These days Ramesh likes Sheila.'

We can postulate, therefore, that Kashmiri is a V2I language, where the Force position is also utilized in expressing focus in this language.

This kind of situation is found in other languages too, in fact with a mix of focus and interrogativity. One such case is Bulgarian. As discussed in Bošković (2001), the specifier of the question particle *li*, which indicates the interrogative force of the sentence, can be filled either by

a *wh*-phrase or by a non-*wh-phrase***.** Example (13) illustrates the latter case: as discussed by Bošković (2001), the element preceding *li* here gets a cleft-like interpretation (a more appropriate translation would be 'Is it the expensive car that Petko sold?'), which means that it is focalized, with the focus associated with a Spec of the Force indicating projection.

(13) Novata kola li prodade Petko?
new-the car Q sold Petko
'Did Petko sell the expensive car?' (Bošković 2001:35)

Returning to Kashmiri, alternatively, we can postulate that Kashmiri has both V2I and a subset of V2D, the latter used for expression of focus only. That only a subset of discourse-related readings can be utilized in V2 in this language can then be taken to indicate that we are dealing here with a gradual loss of this V2 variant, which is what happened in English or Medieval Romance. This would provide a confirmation of the relative diachronic stability of V2I in comparison to V2D, which is historically a subject to reanalysis.

At any rate, V2 in Kashmiri can be captured within the general account of V2 developed here and its properties don't have to be so exceptional as suggested in some analyses of V2 in this language (see Bhatt 1999, see also Holmberg 2010/2015 on an overview of some of the properties of Kashmiri V2).

Chapter 4

Head Labeling Preference and Acquisitional Reanalysis

4.1.Introduction: Language Acquisition and Diachronic Change

This chapter discusses the head-labeling preference as attested in connection with the diachronic change with respect to language acquisition. Looking at the acquisition of reflexive SE pronouns *się* in child Polish, and especially at some errors that children make during the course of acquisition of this construction, I will argue that many of these non-adult utterances can be attributed to the pressures imposed by Labeling Algorithm (LA) (Chomsky 2013, 2015) on syntactic structure. As we have seen in chapter 2, appealing to the labeling system allows us to explain many diachronic changes, which seem unrelated at first sight. Now we will see that certain language acquisition facts can also be explained from the same perspective. This important link between language change and language acquisition allows us to explain at least some diachronic processes as a result of diachronic reanalysis involved in the L1 acquisition process (as in Roberts 2007). To support this claim, I will further discuss some diachronic changes involving SE-reflexives in Slavic. I will show that one particular analysis of these reflexives which postulates their base generated position to be a specifier (Kayne 1974, 1986, Marantz 1984, McGinnis 2004, a.o.) allows us to make sense out of these facts.

This chapter shows that reflexive clitics in Slavic, both in the attested diachronic changes as well as acquisitional facts, give us an opportunity to see the postulated connection between language change and its acquisition, and thus reduce the former to the latter. Additionally, we provide evidence that acquisitional errors offer us a valuable insight into the syntactic settings of

Universal Grammar and structural pressures operating therein. I will show that the discussed acquisitional facts are in fact part of a broader paradigm showing the labeling preference for head-complement configuration as operative during the acquisitional process. This will be supported with the facts from bilingual acquisition and contact languages.

This chapter is organized as follows: in section 4.2., we will introduce some theoretical assumptions about the connection between language acquisition and language change. Section 4.3. brings up a case study where some early errors performed by children acquiring Polish can be explained from the perspective of the head-labeling preference. Section 4.4. shows the connection between acquisitional errors and the actual processes occurring in the diachrony of the SE-reflexives in Russian. Section 4.5. discusses other cases of acquisitional facts that support the analysis proposed here, drawing from both monolingual acquisition (looking at the acquisition of the obligatory *wh*-movement language (4.5.1.), and optional *wh*-movement language (4.5.2.)), as well as bilingual acquisition (4.5.3.). I will also show similar facts from contact and mixed languagaes (4.5.4.). All these cases provide the support to the observation that there is a an early preference for the lack of movement, which can be deduced from the labeling system advocated here. Section 4.6. concludes the chapter.

4.2. Diachronic Change and Language Acquisition

Linguistic change can be defined in generativist terms as two distinct grammars, G1 and G2, in a historical relation with one another. But how is it possible that a distinct grammar G2 may emerge in light of the claims that language acquisition is deterministic, and the speakers converge on the *same* grammar without any instances of imperfect learning or spontaneous innovation? (see e.g., Longobardi 2001).

The idea of diachronic change being closely related to language acquisition (Lightfoot 1979, 1991, 1995, Faarlund 1990, Clark and Roberts 1993, Yang 2002, van Kemenade 2007, Roberts 2007, Walkden 2014) allows us to see the linguistic change over time as a result of the attempt of the new generation of speakers to approximate the grammar of the previous generation, but producing a system that is to some degree dissociated from the underlying linguistic source that served as its input (accessed via Primary Linguistic Data (PLD)). Speakers do not always converge on the exact grammar they are exposed to (i.e., the gramgmars of their caretakers). This is because learners don't have the direct access to the grammar (I-language) of their parents and caretakers, but only to its output: i.e., to the corpus (E-language) (or multiple corpora and E-languages) that the grammar generates. In relation to that, Andersen (1973) argued that language transmission proceeds as an *abductive inference*; it starts from 'an observed result, invokes a law, and infers that something may be the case' (Andersen 1973:775). Andersen's (1973) *abduction* principle presented as his 'Z-model' below:

(1) Grammar 1 → Output 1

Grammar 2 → Output 2 (The Z-model; Andersen 1973: 767)

There is a possibility that Grammar 1 will not be perfectly replicated, but instead some mismatches may arise ('errors of abduction'). These mismatches may occur due to a number of factors, mainly having something to do with the actual data that learners are exposed to (Primary Linguistic Data (PLD)), and their interpretation filtered by some broader principles of Universal Grammar (UG), and to some extent also computational economy. This is also where our labeling preference will come into play, as we will show below. Importantly, the factors responsible for language change are the same as those that are involved in the successful language transmission. In the end,

languages do change because the 'new' grammar results from the population of the speakers converging this slightly altered grammatical system. Hence, the same factors are in play in both processes and they may result in convergence or change. Specifically, it is the interaction of UG with other cognitive factors that is responsible for such variable results and ultimately variation and change.

We have seen that in Andersen's model of abduction as in (1), the Grammar 2 is created based on the Grammar 1, via the Output 1. Some authors find it problematic, however. Here is why. If Grammar 1 is able to produce the Output 1, and if Grammar 2 is created based on this output, then why are the two grammars different? There must be something in Output 1 that triggers Grammar 2 without affecting Grammar 1. This is aptly referred to as the Regress Problem (Roberts 2007).

But the paradox can be only apparent, as there are ways to explain why Grammar 1 can give the input for the innovative reanalysis (that will become Grammar 2) without being affected by this innovation: additional factors must be causing some change in the Output 1, hence the reinterpretation of the grammar becomes possible. The character of the evidence must have changed hence the different interpretation of the grammar becomes possible. The change in question can involve both extra-syntactic changes, e.g., changes in phonology, prosody, or morphology and lexicon as well as other syntactic changes, occurring independently. This could result from language contact or some prescriptive pressures and other sociolinguistic factors.[1]

One attempt to capture how certain changes to the PLD result in a greater change to the grammatical system is Lightfoot's (1979) *Transparency Principle.* Lightfoot shows that changes to the PLD may result in a reanalysis because the way they appear in Output 1 makes them

[1] Kroch (2000) suggests that there could also be some change in the learners themselves (e.g., different age groups); another potential source for the change we are discussing.

unlearnable. Using the example of the development of modals in English, Lightfoot argues that this class of verbs accumulated so many 'exception features' in the late ME grammar that they basically became too opaque for the learners and had to be reanalyzed. Hence the appearance of multiple formal features, exceptions, and extra operations can be argued to be dispreferred by acquirers, who lean towards simplicity and *transparency*.

This supports the observation that learners are driven by principles of simplicity and economy. These economy-based biases are visible in both the maintenance of the grammatical systems, as well as the change. As guiding forces in the acquisition and change, they have been captured in the literature in various proposals.

Some of these include Clark & Roberts' (1993) *Fitness Metric*, Roberts & Roussous' (2003) *Feature Economy*, van Gelderen's (2004) *Head Preference* and *Late Merge* principles. The main idea behind such generalizations is that cases like Lightfoot's modals, as discussed above, really give us the insight into the direction of the resolutions of the encountered opacities and ambiguities in the corpus. The expected direction is to go away from the opaque and non-optimal structures, towards more transparent and economical ones, as expected if the learners biases here really have roots in Chomsky's 3rd factor contributing to the optimal design of language.

Roberts & Roussou (2003) appeal to the notion of *markedness* to capture this particular observation, with a structure R being more marked than R' if involving more formal features. They propose that when dealing with ambiguous or opaque data, acquirers use the postulated principle of *Feature Economy* as defined in (2)

(2) Given two structural representations R and R' for a substring of input text S, R is preferred over R' iff R contains fewer formal features than R' (Roberts 2017:145, based on R&R 2003:201)

Walkden (2014:40) argues that the issue of potential ambiguity and multiple analyses does not even arise in such acquisitional contexts, as only one analysis is then available. It is therefore selected based on the given the Primary Linguisic Data (PLD) that the acquirers face. If any difference between two grammars occurs, it is because two different generations were exposed to two different sets of the PLD. Walkden bases his argument for the exclusive function of the PLD in language change on the principle of determinism that guarantees the observed (predominant, but not exclusive, as we have seen) cross-generational convergence, which itself presupposes that learners favor economy and simplicity. If indeed the structures that learners encounter must have already been disambiguated by them (or never seemed ambiguous to start with), then we can assume that what they use in the process of arriving at a certain (possibly innovated) interpretation will be mainly based on the principles of Universal Grammar, i.e., the innate language faculty (plus the principles of the economical computations).

The idea of certain preferences on part of the learners may also explain changes attested over a longer period of time than just between two generations. Sapir (1921) argued that only by looking at longer history of language, we can see that certain changes are operating over many centuries, and most importantly, they are going in a particular direction. Sapir referred to this process as *drift*, and used it to explain linguistic variation, which he believed cannot be accounted for just by geographical, social, or individual distance between the speakers.

The directional nature of language change has often been invoked in typological studies. Lehmann (1973) proposed that 'mixed' systems with regards to certain grammatical property, i.e., typological inconsistencies attested in languages, are actually indication of the drift-like change

in-progress[2]. Here, change from OV to VO, also described in chapter 2, is taken to be an example of this, with Yiddish still *drifting* to the present day in the direction of the consistent VO language. This change, therefore, seems to be unidirectional. The same actually holds for *wh*-movement which in Modern Romance languages is drifting in the direction of *wh-in-situ* from their Latin *wh*-movement stage.

Sometimes it arises as an effect of multiple smaller changes (*cascading* changes of Biberauer and Roberts 2008), which nevertheless display consistent direction. Another example of a similar unidirectional change is progression of morphological marking that languages go through. Here, the attested tendency is to go from isolating, through agglutinating, to inflectional, but never the other way around. To confirm this, Croft (2003) argues that an apparent bi-directional change often turns out to be two unidirectional changes. Roberts (2007, 2017) shows that some expression of Sapir's idea is necessary to account for the fact that certain combinations of parameters are not attested cross-linguistically, and something must be responsible for grammars to clump together synchronically. The grammatical system is predicted to be drifting in the given direction, then, a process which could be disrupted only by some externally triggered change. This may include the push triggered by the expressiveness of language (the idea behind the motivation of movement expressed in section 2.5. of chapter 2), which counterbalances the UG-driven direction. Hence, the need to express certain discourse-related effects, often at the price of the complication of the syntactic derivation, may also result in syntactic change (see chapter 2 for relevant discussion, prosodic factors may also play a role here). The fact that syntactic movement still exists shows that, as also argued by Roberts (2007), innovation towards computationally dispreferred structures

[2] Other explanations of the existence of two different systems involve postulating grammar competition between two variants (both in the speech community as well as in an individual) as proposed by Kroch (1989) or Yang (2002).

needs to be allowed, otherwise 'we predict that all languages are tending towards a steady state, from which they will not be able to escape (…)' (p.265) [3]

The preference for head-labeling in head-complement configuration and dispreference for symmetrical phrase-phrase and head-head mergers in the labeling system postulated here may be argued to result from economy and simplicity considerations, especially since it can account for some of the diachronic changes displaying a drift-like nature towards the loss of dispreferred configurations. Since diachronic change and language acquisition are closely related and labeling in {XP,YP} and {X,Y} mergers has been shown to be dispreferred, we should be able to see indications of this not only in the language history, but also in the process of acquisition. The next sections show that this is indeed the case.

4.3. A Case Study: SE-Reflexives in Child Polish

This section shows that a similar pattern to diachronic change (i.e., an indication of certain configuration being dispreferred due to the labeling considerations) is attested in language acquisition, where children either avoid the dispreferred mergers, or these less economical configurations are acquired later, both basically showing their 'marked' nature from the labeling point of view acquisitionally.

Reflexive clitics with non-strictly reflexive usages are attested in a variety of languages and constructions. Since in the languages considered here they are often realized as *se*, I will refer to them as SE-reflexives/clitics. Much-investigated cases include SE-clitics of Romance languages

[3] Additionally, Roberts (2007) argues that the innovation towards the marked (in R&R's 2003 system, featurally very rich) values must be assumed to be possible, as otherwise we wouldn't have any explanation for the existence of all the marked (or labeling dispreferred from our perspective) parameter values attested in many languages of the world.

(Manzini 1986, Kayne 1988, Sportiche 2014), Germanic (Reuland 2001, Schäfer 2008, Wood 2015), and Slavic (Rivero & Milojević Sheppard 2003, Medova 2009). Polish has around 13 constructions with SE-reflexive marking, e.g.,: anticausatives, grooming verbs (body reflexives), resultatives, impersonals, middles, or inherent reflexives (*reflexiva tantum).* Examples illustrating these four last classes are given in (3) below:

(3) a. Te samochody prowadzi **się** łatwo. [impersonal]
these cars.ACC drive.3.M.SG. SE easily
'One drives these cars easily.'

b. Te samochody prowadzą **się** łatwo. [middle-voice]
these cars.NOM drive.PL SE easily
'These cars drive easily.'

c. Jan boi **się** smierci. [inherent reflexive]
John fears SE death. GEN
'John is afraid of death.'

One important and common group of SE-marked predicates cross-linguistically is the group of transitivity-alternating structures, where the same predicate (like 'drown' in (4) below) can be used in a transitive structure with two thematic NPs (transitive) or with one full NP with a SE-reflexive. We will argue that in constructions like (4), the transitive variant (4a) alternates with the unaccusative one (4b). Many Polish anticausatives belong to this group.

(4) a. Jan *utopił* wrogów w jeziorze.
John drowned enemies.ACC in lake
'John drowned enemies in the lake.'

b. Jan *utopił* **się** w jeziorze.

John drowned SE in lake

'John drowned in the lake.'

Not all unaccusatives in Polish are created by such 'transitivity alternation'. There is a group of unaccusatives which cannot appear with SE-marking. For this reason, I will refer to them as bare unaccusatives. These bare unaccusatives do not have transitive alternants, either, and their sole argument is an internal argument.

(5) Jan utonął (***się)** w jeziorze.

John drowned SE in lake

'John drowned in the lake'

We can already here notice a connection between the external argument and SE. Namely, SE appears only in those structures involving the projection of a specifier hosting the external argument (more on this below).

We will see below that there is a family of analyses where alternating unaccusatives like (4) consistently project vP (or VoiceP as in Kratzer 1996) where the external argument is generated, i.e., the vP is still projected in (4b), analogically to (4a). The transitive syntax is then not lost in (4b) – structures with SE project both the external and internal argument position. What is important, however, is that the thematic structure of the alternating (4b) is analogous to the non-alternating (5) in that the subject NP, John, in both of these examples is really semantically the

undergoer of the action, and not the agent; hence it starts as an internal argument of the predicate 'drown' in both cases.[4] All this will be discussed in more details below.

I now turn to the acquisition of SE in Polish. Polish children do not reach the adult grammar of SE immediately. There is some variation between the constructions, with impersonal constructions being acquired much later. What is important for us is that analyzing the longitudinal corpora of spontaneous-speech data from Polish children, we can see clearly that the initial stages of acquisition of SE (up to around age 3) are characterized by frequent errors of omission of this element. This happens despite the fact that reflexive clitics are very productive in Polish, i.e., children get the abundance of the SE-marked constructions in the input.

The present study is based on the results from Dadan (2015) which focuses on the relationship between the acquisition of the inherent reflexives (*reflexiva tantum*) as in (3c) and other types of SE-reflexives in Polish. It is based on the longitudinal corpora of spontaneous-speech data from Polish children available in CHILDES (MacWhinney 2000). Database used was the Polish subsection of the Slavic corpus consisting of two corpora: 'Szuman' (data from eight children: four girls and four boys), and 'Weist-Jarosz'(data from four children: one girl and three boys).

Analyzing the data of these twelve children I found that Polish children acquire SE-marked constructions by age 3, with some variation between constructions, e.g. impersonal constructions being acquired somewhat later. However, close investigation of the corpora collected in CHILDES reveals a common error, namely the omission of SE-clitics in contexts where they are obligatory in adult Polish.

[4] See also Haider (1995) for the claim that the reflexive pronoun in German anticausatives bears the external θ-role. We will discuss German later in this section.

Some of the attested errors which involve omission of SE, across the whole acquisitional period of these 12 children, are given in (6-10) below (importantly, children's utterances below do not contain the reflexives placed here in parentheses with the indication that their absence is ungrammatical in adult Polish):

(6) śniło *(się)

dreamed SE

Int: 'I had a dream' (Jaś; 2;3, Szuman, *jas203*)

(7) pogniewała *(się)

got-angry (SE)

Int: 'She got angry'(Wawrzon, 2;6, Weist-Jarosz, *waw09*)

(8) co *(się) kąpie…

which (SE) bathes

Int: 'The one who is bathing' (Kasia, 1;10, Szuman, *kas110)*

(9) już *(się) wy-spały

already (SE) PRF-slept

Int: 'they already had enough sleep' (Basia; 2;1; Szuman, *bas201)*

(10) a. Zepsuła *(się) lampa.

broke SE lamp

Int: 'The lamp broke' (Basia, 2;0, Szuman, *bas200*)

b. Inusia będzie *(się) bawić na Krzemionkach wiaderkiem…

Inka be-FUT (SE) play at Krzemionki bucket.INSTR

Int: ‘Inka will be playing with the bucket at Krzemionki’ (Inka; 2;1, Szuman, *ink201*)

c. ciemno robi *(się) wieczorem

dark getting (SE) evening

Int: ’It is getting dark in the evening.’ (Krzyś; 2;1; Szuman, *krz201*)

d. które *(się) otwierają drzwiczki

which (SE) open door

‘which door open’ (Wawrzon, 3;4, Weist-Jarosz, *waw18*)

e. pali *(się)

it’s.on (SE)

Int: ‘[the light] is on’ (Bartosz;1;8, Weist-Jarosz, *bar04*)

Children omit SE in many contexts and constructions, including psych-verbs (6), inherent reflexives (reflexiva tantum) (7), body-grooming reflexives (8), telic/perfectivizing reflexives (9). However, the most prevalent is the omission in anticausatives/ alternating unaccusatives (10).

The early omission of SE-reflexives in child Polish has been confirmed by another study based on the data from CHILDES. Rivero & Golędzinowska (2002) investigated the acquisition of SE-reflexives and the impersonal modal verb *wolno* ‘allowed’. The authors report that children acquiring Polish SE clitics go through three stages, the first of which (around the age 0;10-1;10) is characterized by the systematic omission of the *się* clitic in all constructions. During the second stage (1;10-2;05), children begin producing the clitics in some, but not all cases. Stage 3 (2;05-3;06 and later) sees adult-like use of SE, with only sporadic omissions.

Table 1: Developmental stages in reflexive clitic constructions (R&G 2002: 62)

Child	Age	Stages
Inka (I)	0;10-3;6	I, II,III
Wawrzon (W)	2;2-3;2	I, II,III
Kubuś (K)	2;1-2;6	II,III
Bartosz (B)	1;7-1;11	II
Marta (M)	1;7-1;10	II

In the light of the analysis presented here, one account of SE-reflexives in the literature is very relevant to the data at hand. Kayne (1975, 1986), Bouchard (1984), Marantz (1984), Pesetsky (1995), McGinnis (2004) and Embick (2004) argue that reflexive clitics are merged as the external argument, in Spec,vP. Some accounts assume base generation in that position, while others assume movement from a VP-internal position. What is important for us is that regardless of the details, under this approach SE is merged in a specifier. From our perspective, the most appealing aspect of this group of analyses is the possibility of unification of the variety of constructions with SE and capturing the acquisitional and historical facts (more on diachronic changes in connection to SE-reflexives in the next section). Crucially, the idea that we adopt here from this line of research is that the derivation of SE-reflexives involves the creation of a specifier (here, specifier of vP).

Under this analysis, the full NP that surfaces in the subject position, i.e., [Spec, TP] in these constructions is base generated as an internal argument with SE occupying [Spec,vP] as in (11) below.

(11) $[_{TP}\ NP_i\ T\ [_{vP}\ SE\ v\ VERB\ [_{VP}\ t_i\ ]]]$

Under this analysis, constructions with SE resemble unaccusatives or passives (see also Reinhart 1997). A similarity along these lines is attested in many languages where the correspondences between reflexives and non-active voice are quite clear. Marantz (1984) reports that in Albanian, Dyirbal, French, Greenlandic, Icelandic (*-st* used with both reflexive and passive) or Russian, the SE-reflexive verb is homophonous with the passive and with the marked intransitive, e.g. anticausative. Reflexives also display many characteristics of unaccusative structures, such as participle agreement in French (Kayne 1986) or auxiliary selection (*essere* as in unaccusatives) in Italian (Burzio 1986). Raising of the internal argument to the higher subject position also provides the antecedent for the reflexive, and thus the Principle A of the Binding Theory (Chomsky 1981) is satisfied. It has been also observed that reflexives do not appear with verbs which do not have the external θ-role. In Romance, for instance, raising is ungrammatical with reflexives (see Bouchard 1984, Sportiche 1998, Pesetsky 1995) (note that raising verbs don't assign the external θ-role), and in Polish, as we have seen in (5), bare unaccusatives (which do not project the agentive vP) can never appear with SE. Additionally, in languages with Genitive of Negation (GON), where only internal argument gets GON, SE does not get it.

To elaborate on that last part, the phenomenon of GON is illustrated in (12):

(12) a. Jan widział **Marię**
John.NOM saw Mary.**ACC**
'John saw Mary.'

b. Jan *nie* widział **Marii**
John.NOM neg saw Mary.**GEN**
'John didn't see Mary.'

As true reflexive pronouns undergo a similar change in the presence of negation (we can see that in languages that actually have separate forms for true reflexives depending on grammatical case, like Old Church Slavonic in (13) (accusative *sę* (13a) changing into genitive *sebe* (13b)), a SE-reflexive does not undergo such a change in (14).

(13) a. sъpasi **sę** samъ i ny
save myself.ACC and us[two]
b. iny sъpase, ali **sebe** *ne* možetъ sъpasti
others saved but himself.GEN neg can save
'he saved others, but he cannot save himelf' (Old Church Slavonic; Lunt:160)

(14) a. ašte ***sę*** bi (ne) rodilъ
'if he had (not) been born' (Old Church Slavonic; Lunt 2001: 160)
b. ne divi ***sę***
'don't be surprised' (Old Church Slavonic; Lunt 2001: 160)

This follows if SE in SE-reflexives occupies the external argument but not an internal argument position.

In fact, we can see it clearly in impersonals with SE where the full NP, which is an internal argument, undergoes the change, but not SE. This is illustrated with Polish below, where the internal argument changes from accusative (15a) to genitive when the sentence is negated (15b).

(15) a. Te samochody prowadzi **się** łatwo.
these cars.ACC drive.3.M.SG. SE easily
'One drives these cars easily.'

b. Tych samochodów nie prowadzi się łatwo.

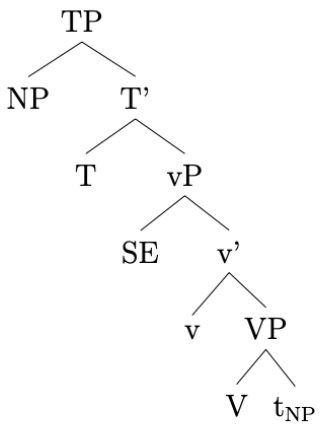
these cars.**GEN** neg drive.3.M.SG. SE easily

'One doesn't drive these cars easily.'

The connection of SE and the external argument position is therefore well grounded.

Now, the observed errors of omission during the acquisition of SE get a straightforward account in the theory of labeling presented here. As the derivation of the SE-reflexives involves generating the specifier, the omission errors in child language are another instance of an attempt to avoid the creation of the Spec (here [Spec, vP] as in (16) below), hence can be considered to be an attempt to minimize the dispreferred configuration.

(16)



There are theories postulating that SE-reflexive, merged in the syntactic position of the external argument, is actually an expletive element (see e.g., Schäfer 2008 who assumes the whole projection where the external argument is merged (VoiceP for him) to be expletive (non-thematic) in the SE-marked anticausatives). Schäfer (2008) argues that a reflexive pronoun like German *sich* is a good candidate for an expletive as it is referentially defective. If indeed SE could be treated as

an expletive without any semantic import, it should not be surprising that acquisitionally such elements can be delayed due to their placement in a dispreferred syntactic configuration.

Interestingly, in many languages anticausatives do not have to involve the reflexive pronoun, hence, the specifier. German, Romance, or Balkan anticausatives come in two flavors, only one of them using the SE-pronoun. The German examples of marked (i.e., involving the reflexive) (17a) and unmarked (i.e., devoid of the reflexive) (17b) anticausatives are given below:

(17) a. Die Tür offnete *(sich)
the door opened REFL
'The door opened.' (German; Schäfer 2008:1)
b. Die Vase zerbrach (*sich)
the vase broke REFL
'The vase broke.'

In fact, Zombolou & Alexiadou (2012) investigate children's acquisition of such anticausative expressions, comparing three types of predicates in German: SE-marked reflexive predicates like 'wash' (grooming reflexives), SE-marked anticausatives (17a), and unmarked (non-involving SE) anticausatives (e.g., *schmelzen* 'melt', (17b) above). Using experimental conditions, the authors found that children 3;00-4;03 still had problems with the SE-marked grooming reflexives and used the SE-marking correctly only in 28% of cases. The problem with SE-marked anticausatives was even more severe, as children were adult-like in only about 11% of cases. In contrast, in unmarked anticausatives, they were 100% correct. In general, the mistake is always an incorrect omission of SE, not producing it when it is not attested in adult speech (or attaching it wrongly).

The observation that children rather omit SE than overuse it makes sense from the current perspective given that SE is located in the specifier. Omitting SE then involves omitting a specifier. This situation is analogous to the loss of *wh*-movement discussed in chapter 2, i.e., moving away from the dispreferred labeling configuration. The acquisitional errors as attested in Polish children are therefore not surprising under the head-labeling preference advocated in the present work. The next section shows that the acquisitional reanalysis of SE-reflexives, as attempted in child Polish, in fact may result in a diachronic reanalysis, as we will see is the case in Russian.

4.4.Diachronic Reanalysis in Russian Reflexives

The link between the labeling-motivated non-adult utterances during the acquisition of SE-reflexives and the actual language change can be seen in the diachrony of the reflexive clitics *sja* in Russian, which is one of the languages where the observed connection between the reflexives, passives and anticausatives (as in Marantz 1984) is clearly visible. (18) below illustrates that.

(18) Dver' otkryl**s'a**
Door open-PAST-REFL
'The door opened.' (Russian; Marantz 1984:184)

In fact, Old Church Slavonic sometimes used passive and reflexive interchangeably and different manuscripts show variants of these two. This indicates their structural correspondence, giving rise to a similar meaning.

(19) a. synъ člověčьskъi predastъ sę
b. predanъ bǭdetъ
'The Son of Man will be betrayed.' (Lunt 2001:Mt 26:2, Zo)

Reflexive suffix -s*ja*, similarly to Polish *się* today, in Old Church Slavonic and Old Russian used to be a free morpheme (Zaliznjak 2008). (20) below shows that *sę /sja* did not have to be adjacent to the verb.

(20) Na gore eže **sja** nyne *zovetь* Ugorъskoje.
on hill which SE now call Ugorskoe
'On the hill, which is now called Ugorskoe.' (Madariaga 2017; *Laurential Chronicle*, 8)

These elements behaved like clitics or weak pronouns; they could precede or follow the verb and other elements (e.g., prepositions). However, as we have seen in (21), in contrast to pronouns, contentful NPs, and true reflexive pronouns, they did not surface with the Genitive of Negation which affects internal arguments in Russian.

(21) a. ašte ***sę*** bi ne rodilъ
'if he had not been born' (Old Church Slavonic; Lunt 2001: 160)
b. ne divi ***sę***
'don't be surprised' (Old Church Slavonic; Lunt 2001: 160)

By Middle Russian, this free and independent form went through a substantial phonological and syntactic reduction and changed into an affix (Zaliznjak 2008). Today it also surfaces as merely a palatalized -*s'* in certain environments.

(22) Imeni moego strašili**sъ**
name my.GEN feared
'They were afraid of my name.' (Madariaga 2017; *Tale of Yeruslav Lazarevich*, 330)

Strictly speaking, the clitic SE no longer has a life of its own, it has become a morphological part of the verb and cannot be targeted by independent syntactic processes anymore (it is referred to as 'postfix'in traditional grammars) (Zaliznjak 2008, Kuznetsova and Nesset 2015, Madariaga 2017).

Hence, today Russian SE-reflexive is very different from Polish *się*, the latter being an independent element generated in the specifier of vP, as we have seen above.

The historical change that the Russian SE went through can be explained in terms of the already familiar labeling pressures. We have shown that Polish SE occupies the Spec of vP. I assume that this analysis applies to the predecessors of Russian *sja*, as well (we have already seen that *sja* did not display GON, assigned to internal arguments only). If SE in Old Church Slavonic and Old Russian was a specifier, then the reanalysis of this element as part of the verb is expected under the labeling considerations, which puts pressure on specifiers to be lost. We have already seen many instances of such a change in chapter 2; it is not surprising then that Russian underwent this very common process as well. Importantly, in the case under consideration, this operation is not just a head-adjunction, but really involves a lexical incorporation of *sja* to the verbal head, where SE-reflexive becomes part of the verb.

Additionally, the fact that the internal argument in Russian *sja* construction may be assigned accusative case (Kuznetsova and Nesset 2015, Madariaga 2017) is very telling. The psych-verb construction involving SE-reflexive illustrates this:

(23) a. On boitsja *ženy*
he fears wife.GEN

b. On boitsja *ženu*
he fears wife.ACC
'He is afraid of his wife.' (Madariaga 2017: 3)

The option shown in (23), with the NP object of the reflexive psych-verbs getting a structural accusative case assigned instead of the genitive case, is available for many speakers. This option is an innovation in modern Russian; thus case configuration is not available historically in Old Church Slavonic or modern Polish. Hence, we can postulate that its innovative character is connected to the reanalysis that we have just discussed, and which occurred in Russian, but not in Polish SE. We are dealing here with a change that is still in progress, as it is indicated by cross-speaker variation and the perception of the accusative case in such contexts as a colloquialism (the association with the informal register). What could be going on here is that *sja* as part of the verb first still 'absorbed' the structural case-assigning ability of the verb, which is no longer the case for some speakers, hence the object in (23b) can now be accusative. Alternatively, this could be a Burzio's generalization effect, where 'he' in (23b) is actually the external argument (as *sja* became part of the predicate); since the external θ-role is assigned, the accusative case can also now be assigned.

In this respect, it is interesting to look at certain changes with psych-verbs that occurred between Middle English and Modern English (Mölid-Falke 2012, van Gelderen 2014), where the object Experiencer was reanalyzed as the subject Experiencer.

In the 14th century, the verb *færan* 'fear' in English meant 'to frighten', with the other meaning of 'fear/respect' emerging in the 14th century:

(24) Þa bodan us *færdon*
The messengers us frightened
'The messengers frightened us.' (van Gelderen 2014: 106)

These psych-verbs with Experiencer Objects could also appear with the reflexive pronouns, which could result in ambiguity as to which element is the Experiencer argument. Van Gelderen (2014) argues that (25) below could mean 'I frighten myself that…', or 'I fear that…', the former being object Experiencer, but the latter having the Experiencer as the subject:[5]

(25) I fere me ye haue made a rode for your self

'I fear/frightened myself you have made a rod for yourself' (van Gelderen 2014:110)

Such a potential ambiguity as to the syntactic position of the Experiencers, whose role could be assigned either to objects (the reflexive pronoun here) or subjects (which would be the pronoun 'I' in (25)), could play a role in the diachronic reanalysis of these predicates in the direction towards the Experiencer role being assigned to the subject (see also Möhlig-Falke 2012).

By the 16th century, the Experiencer objects with *fear* got lost in English. These predicates can only have the Experiencer subjects then, as the translation of *fered* as 'frightened' in (26) below indicates.

(26) Leue son, why hastou fered vs?

'Dear son, why have you frightened us?

(Middle English Dictionary; van Gelderen 2014: 106)

Haspelmath (2001) argues that the subject Experiencer may undergo a further change where this new subject may be reanalyzed as the Agent. The object Experiencer, on the other hand, can be renewed in a language, as it occurred in English, where it appears in loaned structures and

[5] Another way of paraphrasing this in Modern English is 'it is frightening to me that' vs. 'I fear that…'

predicates like *anger*, *worry,* or currently with *love* , *fear,* or *know,* where the change also involves shift in the aspectual properties of these predicates, namely from stative to being able to express, and appear in, non-stative/dynamic contexts.

(27) Wall Street is fearing a bloodbath (COCA Magazine 2007; van Gelderen 2014: 119)

The direction of the diachronic change goes as indicated in (28) below, and never the other way around. Hence, we are dealing here with another unidirectional change here.[6] [7]

(28) Object Experiencer > Subject Experiencer (> Agent) (Haspelmath 2001)

Returning to Russian examples like (23), changing the status of the SE-reflexive so that it no longer occupies the subject position was a prerequisite for the Object Experiencer becomin a Subject Experiencer.

At any rate, what is important for our purposes is that the diachronic behavior of SE in Russian and other Slavic languages like Polish can be captured if SE gets reanalyzed from being a Spec to being part of the verbal head, thus representing another case of the loss of a specifier, which can be captured under the labeling approach argued for here. What is especially interesting from our

[6] The elements in questions are likely specifiers in which case the change would be independent from labeling considerations discussed here.

[7] Van Gelderen (2014) suggests that object Experiencers get reanalyzed as subject Experiencers due to reasons connected to language acquisition: the fact that Themes appear really early in language acquisition (even before Agents and other theta-roles (Ryan 2012)) makes them natural grammatical objects, more natural than Experiencers.

perspective is a language acquisition- diachronic change parallelism. We have seen here a loss of SE-as-a-Spec both in language acquisition, where SE simply gets dropped, and in the diachrony in Russian, where SE gets reanalyzed as part of the verbal head.

The case of reflexives in Russian is a great illustration of the connection between the observed acquisitional facts and the actual historical change, both of which can be explained from the perspective of labeling. Importantly, the analysis presented here does not predict that all children will necessarily have to drop SE, but explains the behavior of those who do, despite the abundance of the evidence for the obligatory presence of this clitic in the language input they receive.

The acquisitional path of SE-reflexives is not an isolated phenomenon, but rather part of a broader paradigm. The next section discusses more cases of this sort, which support the present analysis.

4.5.Head Labeling Preference and Language Acquisition

The labeling-related preference for the head-complement structure (head-labeling) and dispreference for symmetrical mergers that need some extra syntactic machinery is visible in language acquisition more generally, i.e., beyond the cases of SE-reflexives discussed so far in this chapter.

4.5.1. Head labeling Preference and the Acquisition of *Wh*-movement Languages

One of the best-known proposals about children's grammar displaying characteristics that get a straightforward account in the system argued for here is Radford's (1988, 1990, 1996) observation that children show substantial limitation in the early acquisition of *wh*-questions. Radford reports

early utterances of questions by children acquiring English with *wh-phrases* being regularly omitted. Some sample of the data from Radford (1990:123) is given below:

(29) a. *Bow-wow go?* (=Where did the bow-wow go?)

b. Mummy doing? (=What is mummy doing?)

c. Doing there? (=What are you doing there?)

A similar observation that children under the age 2 do not show much evidence of having acquired productive *wh*-fronting was in fact given before Radford, mainly in Klima & Bellugi (1966), Brown (1968), Bowerman (1973), and Wells (1985). Even in the contexts where children imitate adult's utterance the *wh*-phrase is omitted (together with the auxiliary):

(30) a. Adult: *Where does Daddy go*?

Child's imitation: *Daddy go*?

b. Adult: *Where does it go?*

Child's imitation: *Go?* (Radford 1990 after Brown & Fraser 1963)

Radford gives examples like (31), indicating that the imitation of adult *wh*-fronting is still unsuccessful at this age even when the child attempts to use *wh*-phrases.

(31) Adult: *What* are they doing there?

Child: Doing *what* there?

Child's utterance in (31) is therefore *wh-in-situ* with *what* appearing in its base-generated complement position inside the VP. Furthermore, Roeper & De Villers (1989) note that *know what*, with the non-fronted *wh*-phrase, is very commonly observed to be one of the first *wh*-expressions

in child English, appearing around 22 months. Then, a series of phrases containing initial *wh*-phrases appear, but these are argued to be strictly formulaic expressions ('routines'), which are not productive but instead they are acquired as unanalyzed wholes. Hence, movement does not occur in such examples (also Brown 1968:279).

The lack of *wh*-fronting at the early stages of acquisition is explained by Radford as a consequence of the lack of the functional projection of C (i.e., CP) at that stage. Since the complementizer is missing, the specifier of this projection that would host the moved *wh*-phrase is missing as well. Radford argues that additional evidence for the claim regarding the lack of the C-system at the early stage of language development is provided by the lack of preposed auxiliaries, as well as apparent problems that children face with correct parsing of *wh*-questions with fronted *wh*-phrases (children's early responses to *wh*-questions are not consistent with the categories of *wh*-phrases in questions being asked (based on Klima & Bellugi 1966)).

Radford's finding can be interpreted in two different ways in my system: one possibility is that C is missing in early production since C requires a Spec and the children are trying to avoid this dispreferred configuration. Alternatively, if $C_{[+wh]}$ must be present for question interpretation, then C is present for the children too, since they are interpreting the relevant utterances as questions. But they are avoiding creating a specifier by moving a *wh*-phrase there, hence the attested lack of *wh*-movement in their speech.

Returning to Radford, the author actually makes a more general claim that children go through a stage in their acquisition in which they lack all functional categories. This initial stage (up to around age 1;8) is proposed to be strictly lexical-thematic. Children lack all of the major functional categories and only later the relevant functional projections start emerging.

Going beyond CP, Radford thus argues for the lack of DP. According to him, the lack of DP is reflected in the data as the lack of referential determiners (*a*, *the*), problems with the use of the possessive determiner '*s*, problems with the morphosyntax of case-marked pronominal determiners, and the lack of the person/binding properties of (pro)nominals.

(32) a. *Open door* (Radford 1990:84)

b. *Mummy car* (Radford 1990:88)

c. *me talk* (=I talk) (Radford 1996:50)

d. *Kendall see Kendall* (=I can see myself) (Radford 1990:97)

The lack of I-system in child English is postulated based on the lack of early evidence for modal auxiliaries, finite verb inflections, *do*-support, copula and progressive *be*, perfective *have*, nominative case marking, empty categories, or acquisition of infinitival *to* as in (33) below:

(33) *Want* [*read*] (Radford 1990:141)

Analogically with the way the lack of *wh*-questions was interpreted in the current system, these data can be interpreted in two ways: in some of these cases the relevant projections are missing because they would trigger movement and involve creation of specifiers, and this is exactly what the children are trying to avoid in light of the labeling-induced dispreference for phrase-phrase configurations. Thus, the lack of I would be connected to the lack of subject movement, whereas the lack of D would be related to the lack of possessor's movement (for arguments that possessors are generated in NP and move in adult grammar, see Munn 1995, Radford 2000, Alexiadou 2005).

Alternatively, these projections are present, but they fail to trigger movement, both phrasal and head movement.[8]

However, there is another possibility, namely, that head-labeling is not just a preference, but the only available labeling option in children's grammar. In other words, as the labeling by the head is the most economical option to satisfy this requirement (which is really a need for a structure to have its syntactic identity determined), there is a possibility that this way of labeling emerges first, while feature-sharing (and as a result of this, movement) would indeed need to mature in the grammar. This would make head-labeling not a preference, but in fact the only option in the early grammar.

Interestingly, Murasugi (2018) suggests that since the way labeling works in {XP,YP} configurations is subject to cross-linguistic variation (unlike labeling by the head; see the discussion below), it has to take some time and sufficient amount of experience for children to figure out which way of labeling such symmetrical configurations is employed in the language they are acquiring. She illustrates this with a difference in labeling in phrase-phrase merger configuration between English and Japanese.

Chomsky (2013) argues that the labeling of the {DP,TP} merger in finite clauses is possible due to the presence of the shared φ-feature. Hence these two elements enter a feature-sharing relation which results in assigning a $<\varphi, \varphi>$ label for the created structure. For this to take place the subject DP has to be in a Spec-head configuration with the Tense projection (and for which to take place, movement has to be invoked as well in a language like English). This is the role the φ-

[8] Regarding DP, Bošković (2008, 2012) argues that the DP projection is not universal and that languages without definite articles lack it. Koulidobrova (2012) argues that all children, even English children, go through the NP stage initially. This can explain all these DP-related patterns noticed in Radford (1990). Koulidobrova shows that all these problems disappear when children acquire definite articles, which is the trigger to acquire DP.

features play in syntax, and this is how they are utilized. In a language like Japanese, on the other hand, the φ-agreement is absent (see e.g. Kuroda 1988), and hence it cannot resort to it to label the syntactic configuration. Saito (2014a, 2016) explores the possibility that Japanese is indeed devoid of φ-agreement and argues that in Japanese, it is the suffixal case marking that plays a role in labeling. Case markers enable labeling by serving as an 'anti-labeling' device, namely elements that have a case-particle are ignored for labeling in a {XP-case, YP} merger and thus enable the other constituent (in the case just discussed, that would be the head of YP), to be chosen as the label of thus created syntactic objects. Hence in a configuration {XP-NOM, TP}, only the head T can provide the label for this whole merger, case particles being ignored for labeling analogically to traces in Chomsky (2013).

This indicates that φ-feature-labeling may not be a universal property and there could be a variation between languages as to the availability of this option.

Murasugi (2018) argues that we have some indication of this in language acquisition. Children go through the Root Infinitive (RI) stage both in languages like English (e.g., Wexler 1998), i.e., with φ-features, but also in languages like Japanese or Korean, i.e., without φ-features (but with case-particles).[9]

(34) a. *My go in there* [Child English]

b. *Paula play with ball* (Murasugi 2018:6)

[9] See also Hoekstra and Hyams (1998) for evidence from Dutch, or Pierce (1992) for French. See also Snyder (2007) for a more general discussion of this phenomenon.

(35) a. tii si-ta (adult form: *si-ta-i*) [Child Japanese]

pee(mimetics) do-PAST

Int: I want to pee.' (Murasugi 2018:7)

b. mek-e emma (adult form: *mek-ca* (Propositive)) [Child Korean]

eat-DECL mommy

'Let's eat, Mommy' (Murasugi 2018:9)

Examples in (34) indicate that in a language like English, children go through an early stage producing caseless nominals and non-finite verbs in root contexts (see also Wexler 1996, 1998). Interestingly, Japanese or Korean children go through a corresponding stage, producing non-finite verbs in finite contexts and these 'Surrogate' Root-Infinitives as in (35) are used to express events both in the present and the past or future, as well as in modal contexts (Murasugi 2018).

Murasugi (2018) argues that we can treat those RI stages in both types of languages as the evidence that children are actively trying to discover the labeling mechanism for {XP,YP} merger in their target language. Her conclusion is that Labeling Algorithm is still under construction during the early acquisitional period. This is especially visible in errors made in configurations involving labeling in {XP,YP}mergers. This also confirms the proposal endorsed here, as to the 'marked' or dispreferred status of labeling in phrase-phrase and head-head configuration. We can expect that less economical and marked, hence dispreferred options may indeed emerge later in the acquisition. If this is the case, the diachronic reanalysis occurs because children utilize the only option that they actually have at this stage (i.e., a head-phrase merger). This could mean that no ambiguity arises in terms of the possible interpretation of the input that children receive (PLD), as children analyze the structure using their only available tools (this mirrors the suggestion in Walkden 2014 discussed earlier in section 4.2.).

This supports the analysis based on labeling considerations.

The analysis based on labeling considerations is also supported by situations where when faced with the problematic {XP,YP} structure that they may not have the resources to label just yet, children resort to other means. Returning to *wh*-questions, Radford (1990, 1996, among others) argues that the first appearance of what looks like productive fronted *wh*-questions around the age 2.5 years (36) actually involves adjunction.

(36) *What kitty doing?*

Radford postulates that at this stage, children already know that a *wh*-expression needs to take the wide scope for the question interpretation (Radford refers to this knowledge as Scope Principle). He also argues that children treat *wh*-phrases more like quantifiers than operators, hence they undergo adjunction.

If those *wh*-questions involve adjunction, we would like to know what the actual adjunction site for these early *wh*-phrases is. We have already discussed the two possibilities to account for children's non-adult production of *wh*-questions at this stage: either they lack the functional projections in question, or they just lack the specifiers, with the relevant functional heads (C,T, or D) being already in place in children's grammar. Radford, who assumes the former, argues for VP-adjunction for *wh*-fronting in (36). The alternative could be CP-adjunction, if CP is present at this stage. Under the latter analysis, one argument for the adjunction rather than movement to Spec,CP can be the lack of subject-auxiliary inversion at this stage, as in (37), assuming the inversion occurs when a *wh*-phrase moves to the specifier of CP. ((37) illustrates that before the age 2;5 their grammar does not involve subject-auxiliary inversion in *wh*-questions):

(37) *a. Where the other Joe will drive?*

b. What he can ride it?

c. Where I should put it when I make up (Radford 1996:69 based on Bellugi & Klima 1966:205)

Regardless of the actual height of the *wh*-phrase, i.e., whether it is adjoined to CP or to VP, what is important for us here is that children use adjunction in the early *wh*-questions. If Radford is correct, I suggest that children may resort to adjunction to solve a labeling problem.

While discussing adjunction, Chomsky suggests that these structures should be created via *Pair-Merge* operation (i.e. creation of an ordered set <X,Y>), in contrast to *Set-Merge* (i.e. unordered set {X,Y}), operating in the contexts of complementation. The notation of the ordered set suggests that there is some intrinsic asymmetry already built into the adjunction operation, similarly to head-phrase merger. This may suggest that this structure may have already pre-determined label during Merge (it is always the projection of the host of adjunction, not the adjunct itself). In other words, the label here is automatically determined as with head-phrase merger, hence preferred to non-adjunction phrase-phrase merger.

There are also proposals to dispense with labeling for adjuncts (Chametzky 2000, Hornstein and Nunes 2008, Hunter 2010, Bošković 2016a), where what we have traditionally referred to as adjunction involves unlabeled objects.[10] This hypothesis opens up an interesting possibility that if adjunction does not require labeling for interpretation, or well-formedness, this could be the way for children to escape the early {XP,YP} labeling limitation or dispreference for this merger.[11]

[10] Under this view, I hypothesize that this holds for phrasal adjunction, but not head-head adjunction, which we have seen earlier is subject to labeling pressures.

[11] We may also speculate that diachronic change from specifiers to adjuncts could be another strategy for grammar to get rid of disfavored feature-sharing configuration. It can be argued that this is attested with Slavic possessors, e.g., in Serbo-Croatian, where the fact that possessives bind out of the nominal phrase they modify (causing e.g. Condition

At any rate, the next change in children's grammar occurs around the age 2;5 with the emergence of inversion. This is accompanied by an overgeneralization of this process (as in 38a-b), as well as with the phenomenon of auxiliary-copying (38c-d).

(38) a. *don't wait me go on it* (Ruth 2;6)

b. *is I can do that ?* (a girl; 3,0 based on Akmajian & Heny 1975:17)

c. *does it doesn't move?* (Nina 2;10)

d. *Is the clock is working?* (Shem 2;5) (Radford 1996:72)

This indicates the beginning of the third stage in the acquisition, which soon will lead to the adult-like mastery of *wh*-questions due to the abundance of the data that children hear around them indicating the obligatoriness of this operation in English and the availability of feature-sharing to label the {WH,CP} mergers as <Q,Q>.

4.5.2. Head labeling Preference and the Acquisition of Optional *Wh*-movement Languages

In this context, it is interesting to examine the acquisition of *wh*-fronting in languages that have the option of *not* moving the *wh*-phrase (such as Modern Romance languages discussed in chapter 2), hence avoiding the feature-sharing option in *wh*-questions by leaving the *wh in-situ*, where it can be head-labeled. There are many studies that indicate that this is exactly what is happening in the early stages of acquisition, where children prefer *wh-in-situ* over *wh*-movement, as predicted by the analysis advocated here.

B/C violations) has been taken to indicate that they are adjuncts (hence the phrase they modify does not confine their binding domain; see Bošković 2008, 2012, Despić 2011, 2013).

Zuckerman (2001) observes that French-acquiring children prefer *wh-in-situ* in matrix clauses despite the fact that this option has a lower frequency in their input. French speaking children are faced with several overt *wh*-movement strategies, but their initial stage is *wh*-in-situ, i.e., this is the first option that they acquire, only later followed by fronted *wh*-phrases without inversion (no head movement of the verb to C), then *wh* in clefts, and finally fronted *wh* accompanied by inversion, acquired as last.

Zuckerman also observes that older children still prefer to use more of *wh-in*-situ compared to their parents. She interprets this as essentially indicating that children avoid less economical alternants, which they are capable of, when more economical options are available (with the current proposal providing an explantion for the move vs less economical options). Children thus produce more *in-situ* questions than their parents. It is easy to see that this could increase the usage of *wh-in-situ* in adult language, eventually leading to French becoming a *wh-in-situ* language.

This is confirmed by an experimental study in Oiry (2011), where French children (mean age: 5.07) exceeded the adults in producing Long-Distance (LD) *wh-in-situ* (over Long Distance movement) not only showing that the *wh-in-situ* in French does not have to be limited to matrix contexts (as claimed in Bošković 1998), but also showing that children actively produce this option, even more than the adult speakers (who vary in the availability of this option). Thus, children produced LD *wh-in-situ* in 12.75% of the presuppositional context (which satisfies felicity conditions for both Scope Marking (which are also shown to be used by French children during the acquisition; see also Thornton (1990) for English) and long distance movement), whereas adults produced such a construction only in 4.75% of cases. In non-presuppositional context (where only long distance movement is felicitous) the situation is similar, with children using LD *in-situ* option in 18% of the situations, and adults only in 8.5%.

Oiry notes that the children who produced Long-Distance *wh-in-situ* in her experiment are those who also produce the fronted version, hence it is not the case that the *in-situ* option is the only availability they have, but rather a choice they make, again, in more cases than adults. Thus, again, we can see that children favor the *in-situ* option which eliminates an instance of dispreferred {XP,YP} labeling.[12]

There are other studies also reporting the preference for the initial *wh-in-situ* in French children (Hamann 2000, Plunket & DeCat 2001, Plunkett 2004). Hamann (2001), using the data from both spontaneous speech (Geneva corpus) and elicitation experiments, looks at the distribution of *wh-in-situ* versus *wh*-fronted questions in French in both typical children (age 1;8-2;10) and children with Specific Language Impairment (SLI) (age 3;10-9;1). She finds that children in both groups substantially prefer *wh-in-situ* over the movement option in 'free choice' contexts (i.e., where both strategies are in principle available).[13]

Table 2. Free choice contexts (Hamann 2001)

Child	**Wh-in-situ**	**Fronted wh**
Augustin	90.7 (49/54)	9.3 (5/54)
Louis	73.3 (22/30)	26.7 (8/30)
Marie	76.1 (35/46)	23.9 (11/46)

[12] Oiry suggests that what drives this difference between the use of *wh-in-situ* vs. wh-movement in these contexts is the fact that children are free from the prescriptive pressures urging the speakers to avoid *wh-in-situ* (social indesirability of *wh-in-situ*). If this is indeed the case, we can argue that children's performance and their lack of the 'social filter' connected to the prescriptive language norms shows which option is more natural in those contexts. Since these children have both options in place, their choice does suggest that there is a preferred variant, which may not be the case for adults, as for them other considerations than purely grammatical preference may play some role.

[13] There is one exception in the corpus: Philippe (also reported in Crisma 1992) in the first period produces only fronted wh-questions. In the second period, only one in-situ is attested. Finally, the boy uses both types with the preference for movement again. This exception is very interesting, as all other children reported (also based on the York corpus) start with *wh-in-situ*.

If fronted *wh*-questions do appear in the data, they tend not to be inverted as in Louis below (similarly to observations in Zuckerman 2001 reported above; recall that inversion is also dispreferred from the labeling point of view).

(39) **Quoi** ça *fait* le robe
what that does the dress
'What does the dress do?' (Louis 2;2.4; Hamann 2001:163)

While typical children show a strong preference for the in-situ variant of *wh*-questions, SLI children's asymmetry in their preference for the *in-situ* questions is even more prominent. 3 out of 6 children *never* produce fronted *wh*-questions spontaneously in a context where a choice (movement vs. *in-situ*) is possible.

Table 3. Questions in Free choice contexts; SLI children (Hamann 2001: 169-170)

Child	**Age**	**Mean MLU**	**Wh-in-situ**	**Fronted Wh**
Younger Group				
Rafaelle	3.10-5.1	2.75	50 (9/18)	50.0 (9/18)
Aurélie	4.2-5.6	3.24	82.5 (47/57)	17.5 (10/57)
Loris	4.7-6.0	3.68	84.2 (16/19)	15.8 (3/19)
Corentin	4.9-5.8	2.09	100.1 (1/1)	0 (0/1)
Didier	4.9-5.8	3.46	88.9 (8/9)	11.1 (1/9)
Martin	5.0-6.1	3.71	78.6 (22/28)	21.4 (6/28)
Older Group				
Fabrice	5.7-6.11	3.40	100.0 (7/7)	0
Noelle	6.9-7.11	4.93	52.6 (10/19)	47.4 (9/19)
Lea	7.7-8.9	3.31	33.3 (1/3)	66.7 (2/3)
Candide	7.10-9.1	4.64	100.0 (3/3)	0
Noa	7.11-9.1	3.99	33.3 (1/3)	66.7 (2/3)

The elicitation data reflect the results for the spontaneous speech in both groups, with SLI children producing almost twice as many in-situ questions than the control group (40.4% vs 24.2 %) (with the lead-in in one of the conditions including the bias towards the fronting wh-word). Hence it is clear that as both groups avoid *wh*-fronting (also inversion is reported to be avoided to a great extent), SLI children have even greater tendency to use *wh-in-situ* and have many more difficulties with movement than typical population. This confirms other SLI-based studies reporting similar problems with the complexity of wh-movement in this population (Hamann et al. 1998, Van der

Lely 1998). Again, this supports postulated complexity of the labeling in the Spec-head configuration, which is even more striking in populations having language difficulties like SLI children here.

The tendency to acquire *wh-in-situ* before the fronted version in a language that allows both options is also reported to hold in other languages apart from French. Lessa-de-Oliveira (2003) provides evidence to this effect from the Bahia dialect of Brazilian Portuguese. The author investigates the spontaneous speech of two children acquiring this dialect and confirms that their acquisition reflects the French pattern, *wh-in-situ* being the preferred strategy in the early acquisition. The placement of *wh-in-situ* has actually become the preferred strategy in this dialect, even in adult question formation. In fact, even aggressively non-D-linked expressions like *que diabo* 'what the hell' (which must be *wh*-moved in English) stay *in-situ* (even in adult speech).

4.5.3. Head labeling Preference and the Bilingual Acquisition

In this context, it is interesting to look at bilingual acquisition, especially when a child acquires languages that differ in terms of the obligatoriness of *wh*-fronting. Yip & Matthews (2007) (Y&M) show that Cantonese-English bilinguals first go through a *wh-in-situ* stage during their acquisition of English. *Wh*-movement in *wh-questions* is acquired later, and for some time both *in-situ* and movement exist as two options for a child.

Yip & Matthews report that the first English interrogative occurrences of their son Timmy lacked *wh*-fronting and were analogous to Cantonese *wh*-questions, which is a *wh-in-situ* language.

(40) a. *This **what** colour?* (Timmy 2;10;01)

b. *The snail **why** live in the water?* (Timmy 3;03;08) (Y&M 2007:94)

(41) a. Li1 go3 **mat1je5** (ngaan4)sik1 aa3?

This CL what colour SFP

'What colour is this?' (adult Cantonese)

b. Zek3 wo1ngau4 **dim2gaai2** zyu6 hai2 seoi2 dou6 aa3?

CL snail why live in water there SFP

'Why does the snail live in the water?' (adult Cantonese; Y&M 2007:94)

The preference for the *in-situ* option is reported to have been very strong, as the child responds using this strategy even when prompted by questions involving fronting:

(42) INV: Look, what do they want?

CHI: It's a what? (Timmy 2;07; Y&M 2007: 95)

Looking at the longitudinal data of this bilingual child, we can see that the amount of *wh-in-situ* grows steadily from its initial occurrence at the age 2;01. It is ranging between 25%-80% up to a stage where it predominates, reaching up to 100% of object *in-situ* questions (ages 2;05-2;08). After that, *wh-in-situ* gradually recedes, appearing in 33%-67% occurrences. Yip & Matthews observe that the peaks in Timmy's *wh-in-situ* correspond to peaks in the boy's Mean Length of Utterance (MLU)[14], which occur between ages 2;01 and 2;08, reaching 100% at age 2;07. Since the boy's MLU for Cantonese exceeds that for English, the authors postulate that the *wh*-in-situ in English is a result of a linguistic transfer from Cantonese, which is Timmy's dominant language.

[14] MLU is the average number of words or morphemes per utterance, which the authors take to be the most objective indicator of a child's linguistic development.

As Yip & Matthews extend the scope of the study, including five more Cantonese-English bilinguals, they find the same pattern as in Timmy, i.e., early prevalence of *wh-in-situ* in their English, which for some children is actually at 100% (Sophie and Alicia, in *where* used *wh-situ*). One girl (Alicia) is reported to be still using non-fronted *what* at age 5.

(43) *Daddy, Lulu birthday you give to Lulu* ***what*** ? (Alicia 4;04;11; Y&M 2007:103)

Starting from *wh-in-situ*, as predicted by our analysis, the acquisition of *wh*-movement in bilinguals involves going through some intermediate stages. One of them is partial *wh*-movement.

(44) a. *You t'ink where do you put it?* (Timmy 3;11;20) (Y&M 2007:115)

b. Father: What are you making, Sophie?

Child: *You want* ***what*** *do I make*? (Sophie 5;04;19) (Y&M 2007:115)

As these questions involve movement, albeit only within the embedded clause, the authors take them as an intermediate step between initial *wh-in-situ* and full adult *wh*-movement of the English type. From the current perspective, cases like (44) involve creation of one specifier (in the embedded clause). If the example were to involve full wh-movement, two Specs would be created. The embedded Spec, CP would still be created due to the PIC, but the matrix Spec,CP would also be created (I hypothesize that for these children labeling does occur in the embedded CP).

Another interesting strategy that bilinguals adopt while developing *wh*-fronting is the inversion without the movement of the *wh*-phrase.

(45) *Daddy, are you having* ***what****?* (Alicia 3;09;11) (Y&M 2007:119)

The construction in (45) is reported to be very rarely attested in acquisition (as well as in adult grammars in general). Its appearance in bilingual data can be attributed to special circumstances where a child has access to two grammars, one employing more economical and thus preferred strategy, the other using more marked option. Such intermediate stages could then be indicating the child's leaning towards the preferred system while surrounded by the PLD indicating the presence of a more marked option, involving movement and feature-sharing. Examples like (45) are also important since they cannot be treated in terms of Radford's proposal that children who do not have *wh*-movement lack CP (CP is present here, given inversion).

4.5.4. Head labeling Preference and Contact Languages

Having shown that the labeling-induced preference for a head-complement merger and dispreference for Spec-head configurations is attested in bilingual acquisition and results in the initial adhering to more economical labeling-wise structures, we will show that the same tendency is attested in contact languages.

Singapore Colloquial English (SCE) displays analogical trajectory to the bilingual acquisition of Hong Kong children. As a result, initially only *wh-in-situ* is the option for children acquiring SCE.

(46) a. You are doing what? (Elvoo 3;06) (Y&P 2007:123)

b. We are going to eat where? (Elvoo 3;09) (Y&P 2007:123)

However, the situation in Singapore Colloquial English is different from the Hong Kong English in the fact that adult Singaporean English (or Singlish) involves many more instances of *wh-in-situ*:

(47) a. Doll lie down where? (Y&P 2007:124)

b. John must have dinner with who? (Chow 1995:32)

c. After that, you went to what? (*International Corpus of English*; Roberts 2017:430)

Singlish is therefore analogical to Modern Romance as it is undergoing a change towards the loss of *wh*-movement, along the lines discussed in chapter 2. Hence, children acquiring this variety of English may be influenced both by the transfer from Chinese dialects (Hokkien, Teochew, Cantonese), Baba/Bazaar Malay (see Sato 2011) and Tamil, as well as the English input they are exposed to, already containing a strong indication of possibility of leaving *wh*-phrase *in-situ*.

Besides the absence of the obligatory *wh*-movement, Singlish also lacks V-to-C (or T-to-C) head movement. Hence, questions do not involve subject-auxiliary inversion and instead, the language uses a particle system borrowed from Chinese varieties.

(48) This one **what** *hah*? (Y&P 2007:124)

(49) a. **Who** you think buying John's car *ah*?

b. You think **who** buying John's car *ah*? (Sato 2011; 6)

Singlish thus can be taken here to be an interesting example of a mixed system, where in a situation where one language employs a more costly labeling option than the other contact language(s), the more costly labeling option is eliminated in the target language. This is expected under the theory of labeling playing important role in language change.

Similar kind of evidence is present in creoles, which as DeGraff (2005) argues, emerge via ordinary processes of language change (which Roberts 2007 refers to as a case of *extreme language change*). From the current perspective, the labeling pressures result in many similarities shared by these languages.

That such pressures can be quite strong is shown by the lack of V-to-T movement even in Romance-based creoles (possibly as a consequence of the morphological simplicity of these languages, given the role of morphology in labeling discussed in chapter 2). The lack of V-to-T movement in a French creole is illustrated in (50) from Haitan, where the verb cannot occur before the adverb 'already'. The movement to this higher position is not allowed.

(50) a. *Bouki **pase** *deja* rad yo
Bouki iron already cloth their
b. Bouki *deja* **pase** rad yo
Bouki already iron cloth their
'Bouki has already ironed their clothes.' (Haitan; Roberts 2007:410)

Similarly, certain Dutch-based creoles display SVO word order, despite Dutch being characterized as an SOV word order language. This is illustrated with Berbice Dutch (spoken in Guyana), which is derived from Dutch but directly related to the Ijo language, both of which are SOV languages. Nevertheless, this Caribbean creole language is SVO, as shown below in (51):

(51) ɛk wa jefi-a kali kali
I ANT eat DUR little little
'I was eating very little.' (Berbice Dutch; Muysken 1988:290)

We have already seen how the labeling-based system argued for here explains the loss of V-to-T movement and the OV-to-VO word order change. Creoles therefore may serve as an important

indication of the existence of disfavoring movement and feature-sharing structures, which can be explained by labeling considerations argued for here.[15]

4.6.Conclusion

This chapter has shown that postulated connection between language acquisition and language change can be explained from the perspective of labeling-induced preference for head-complement structures. We have seen in the previous chapters that historical changes affect the Spec-head configurations, reducing them to head-complement configurations (by a loss of movement and specifiers). In this chapter, we have seen evidence for similar forces being at work in language acquisition. Just like languages diachronically reduce the dispreferred {XP,YP} and {X,Y} configurations, children avoid them during early acquisition, opting instead for mergers that can be head-labeled.

[15] Analogically, many creoles do not have *wh*-movement:

(i) a. Yutupela sutim husat tru ?
You shot who really
'Who did you really shoot?' (Tok Pisin; Woolford 1979: 43)

Chapter 5

Conclusion

In this dissertation I have investigated a number of cases of a diachronic loss of syntactic movement occurring cross-linguistically. I offered a unified analysis for all these cases based on Chomsky's (2013, 2015) labeling system. I argued that all the cases discussed in the present work can be deduced from a preference for a configuration in which a head and a phrase are merged, where labeling occurs straightforwardly. A merger of two phrases {XP,YP}, and a merger two heads {X,Y} are dispreferred from the standpoint of labeling, hence they are diachronically fragile and subject to reanalysis.

Chapter 2 showed that there is a broad diachronic trend to lose obligatory *wh*-movement, which was also deduced from Chomsky's (2013, 2015) labeling system. The argument was made that the loss of movement in this case actually involves the loss of a specifier, which can be captured by the labeling system. I have shown that the cases of the loss of wh-movement, which occurred (or is occurring) in Japanese, Chinese, Indic, Romance, and Basque, can be unified with other diachronic phenomena, such as reanalysis of specifiers as heads (traditional grammaticalization), as in Early Germanic *wh*-exclamative *hwæt* or complemenizers in Welsh, the loss of head movement (as in English or Mainland Scandinavian), the change from the OV word order to the VO word order, and the loss post-verbal adjunction: I have shown that all these cases involve the loss of configurations that are dispreferred from the labeling perspective.

I have also discussed specifiers which are more resistant to diachronic change, in particular cases involving multiple movement to the same position in the structure (in particular, multiple wh-fronting). I have shown that the loss of multiple *wh*-fronting proceeds through a single *wh*-fronting stage, not going directly to the option without any movement at all.

Chapter 3 discussed the Verb-Second phenomenon. I have shown that this phenomenon involves (at least) two distinct configurations with distinct motivations, syntactic mechanisms, and licensing conditions. These two types of V2 are the discourse related V2 (V2D) and the illocutionary-force related V2 (V2I). I have shown that the latter type of V2 is more stable diachronically than the former type of V2. I have also shown that the different behavior of these two types of V2 can be explained in the labeling-based system, where V2I is derived via self-attaching head-movement, and the considerations of prosody. Under the proposed analysis, V2I is not subject to the pressures of the labeling algorithm, unlike V2D, which is labeled in a feature-sharing configuration, hence the labeling pressures are operative in this construction, leading to its diachronic loss.

Chapter 4 explored acquisitional consequences of the working of the labeling algorithm and showed a connection between language acquisition and diachronic change, both of them being affected by labeling-induced pressures. I have shown that the acquisitional errors made by Polish children acquiring SE-reflexives offer us a valuable insight into the labeling pressures operating within UG, favoring head-complement structures and head-labeling. Furthermore, I have shown that the same forces that lead to the omission of SE-reflexives in language acquisition are behind a diachronic reanalysis of SE in Russian. Additionally, I have shown that the labeling-related preference for the head-complement structure (head-labeling) and dispreference for symmetrical mergers that need some extra syntactic machinery is visible in language acquisition more generally, the main test case being acquisition of *wh*-questions by both monolingual and bilingual children.

Overall, the account offered here explains many diachronic facts, which can now be seen as part of a much broader phenomenon and accounted for in terms of a structure-building mechanism with labeling playing a central role.

BIBLIOGRAPHY

Adams, Marianne. 1987. From Old French to the Theory of Pro-Drop. *Natural Language & Linguistic Theory* 5(1): 1–32.

Adger, D. 2006. Post-syntactic movement and the Old Irish Verb. Natural Language & Linguistic Theory 24, 605–654.

Adli, A. 2006. French wh-in-situ questions and syntactic optionality: Evidence from three datatypes. *Zeitschrift für Sprachwissentschaft*, 25, 163-203.

Aldridge, Edith. 2009.Short *wh*-movement in Old Japanese. *Japanese/Korean Linguistics* 17:549-563.

Aldridge, Edith. 2010.Clause internal *wh*-movement in Archaic Chinese. *Journal of East Asian Linguistics* 19:1-36.

Aldridge, Edith. 2011. Neg-to-Q: Historical development of one clause-final particle in Chinese. *The Linguistic Review* 28.4: 411-447.

Aldridge, Edith. 2018. C-T inheritance and the left periphery in Old Japanese. *Glossa* 3(1)26: 1-22.

Andersen, Henning. 1973. Abductive and deductive change. *Language* 49:765-93.

Anderson, Stephen. 2000. Towards an Optimal Account of Second Position Phenomena. Available at https://pdfs.semanticscholar.org/01c5/7a1c26538fffc842dbd3ac3292ff8568912f.pdf (accessed April 2019)

Anderson, Stephen. 2005. *Aspects of the Theory of Clitics.* Oxford: Oxford University Press.

Andrews, Avery. 1983. A note on the constituent structure of modifiers. *Linguistic Inquiry* 14: 695–697.

Aoun, J. and A. Li. 2003. *Essays on the Representational and Derivational Nature of Grammar: The Diversity of WH-constructions*. Cambridge MA, MIT Press.

Ashby, William J. 1977. Interrogative forms in Parisian French. *Semasia* 4: 35-52.

Axel-Tober, Katrin. 2017. The Development of the Declarative Complementizers in German. *Language* 93(2): e39-e65.

Bammesberger, Alfred. 2006. The syntactic analysis of the opening verses in Beowulf. *ANQ: A Quarterly Journal of Short Articles, Notes, and Reviews* (19): 3–7.

Benincà, Paola. 2006. A detailed map of the left periphery of medieval Romance. In Raffaella Zanuttini (ed.), *Crosslinguistic research in syntax and semantics: Negation, tense and clausal architecture*, 53–86. Georgetown: Georgetown University Press.

Benincà, Paola, and Cecilia Poletto. 2004. Topic, focus, and V2: Defining the sublayers. In by Luigi Rizzi (ed.), *The Cartography of Syntactic Structures, vol. 2: The Structure of CP and IP*. New York and Oxford: Oxford University Press, 52−75.

Bhatt, Rakesh.1999. *Verb Movement and the Syntax ofKashmiri.* Dordrecht: Kluwer.

Bhatt, Rakesh & James Yoon. 1992. On the Composition ofComp and Parameters ofV2. In *Proceedings ofWCCFL* X., D. Bates (ed). Stanford: CLSI, 41- 52.

Bhatt, Rajesh, and Sadaf Munshi. 2009. Two locations for negation: evidence from Kashmiri. *Linguistic Variation Yearbook* 9: 205-240.

Biberauer, Theresa. 2002. Verb Second in Afrikaans: Is This a Unitary Phenomenon? *Stellenbosch Papers in Linguistics* 34 (SPIL 34): 19-68.

Biberauer, Theresa, Glenda Newton, and Michelle Sheehan. 2009a. Limiting Synchronic and Diachronic Variation and Change: The Final-over-Final Constraint. *Language and Linguistics 10.4: 701-743.*

Biberauer, Theresa, Glenda Newton, and Michelle Sheehan. 2009b. The Final-over-Final Constraint and predictions for diachronic change. *Toronto Working Papers in Linguistics (TWPL) 31:*1-17.

Biberauer, Theresa, Michelle Sheehan, and Glenda Newton. 2010. Impossible Changes and Impossible Borrowings. In Anne Breitbarth, Chris Lucas, Sheila Watts, and David Willis (eds.), *Continuity and Change in Grammar*. Amsterdam: John Benjamins, 35–60.

Biberauer, Theresa, Anders Holmberg, and Ian Roberts. 2014. A Syntactic Universal and Its Consequences." *Linguistic Inquiry* 45:169–225.

Biberauer, Theresa, and Ian Roberts. 2006. Loss of residual 'head final' orders and remnant fronting in Late Middle English: causes and consequences. In Jutta Haartmann and Laszlo Molnarfi (eds.), *Comparative studies in Germanic syntax*. Amsterdam: John Benjamins, 263-297.

Biberauer, Theresa, and Ian Roberts. 2008.Cascading parameter changes: internally-driven change in Middle and Early Modern English', in Þórhallur Eyþórsson (ed.), *Grammatical change and linguistic theory: the Rosendal papers*. Amsterdam: John Benjamins, 79–114.

Bennet, Charles E. 1966. *Syntax of Early Latin I, II.* Hildesheim, Olms.

Behnstedt, Peter.1973. *Viens-tu? Est-ce que tu viens? Tu viens? Formen und Strukturen des direkten Fragesatzes im Französischen*. Tübingen: Tübinger Beiträge zur Linguistik.

Bentzen, Kristine, Gunnar Hrafn Hrafnbjargarson, Þorbjörg Hróarsdóttir, and Anna-Lena Wiklund. 2007. The Tromsø guide to the force behind V2. Working Papers in Scandinavian Syntax 79: 93–118.

Blümel, Andreas. 2017. Exocentric root declaratives: Evidence from V2. In *Labels and roots*, ed. by Leah Bauke and Andreas Blümel, 263-290. Berlin: de Gruyter Mouton.

Behnstedt, Peter. 1973. *Viens-tu? Est-ce que tu viens? Tu viens? Formen und Strukturen des direkten Fragesatzes im Französischen*. Tübingen: Tübinger Beiträge zur Linguistik.

Benett, Charles E. 1966. *Syntax of Early Latin.* Hildesheim: Georg Olms.

den Besten, H., 1977. On the interaction of root transformations and lexical deletive rules. Ms., MIT and University of Amsterdam.

Boeckx, Cedric. 1998. A minimalist view on the passive. In: *UConn Working Papers Occasional Papers in Linguistics 2.* Department of Linguistics, University of Connecticut, Storrs. [Distributed by MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.]

Borsley, R.D., Kathol, A., 2000. Breton as a V2 language. Linguistics 38, 665–710.

Bošković, Željko. 1997. On certain violations of the Superiority Condition, AgrO, and economy of derivation. *Journal of Linguistics* 33:227-254. In Ángel J. Gallego and Dennis Ott (eds.) *50 Years Later: Reflecitons on Chomsky's Aspects,* MIT Working Papers on Linguistics (77),13-32.

Bošković, Željko. 1999. On multiple feature checking: multiple wh-fronting and multiple head-movement. In S. Epstein and N.Hornstein (eds.), *Working Minimalism*, 159-187. MIT Press.

Bošković, Željko. 2000. Sometimes in SpecCP, sometimes in-situ, in R. Martin, D. Michaels, and J. Uriagereka (eds.) Step by step: Essays on minimalism in honor of Howard Lasnik, 53-87. MIT Press, Cambridge, Mass.

Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena.* Amsterdam: Elsevier Science.

Bošković, Željko. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33: 351-383.

Bošković, Željko. 2003. On the interpretation of multiple questions. *Linguistic Variation Yearbook* 1:1-15.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38: 589-644.

Bošković, Željko. 2008.What will you have, DP or NP? In *NELS 37*, In Emily Elfner and Martin Walkow (eds), Amherst: University of Massachusetts, Graduate Linguistic Student Association, 101–114.

Bošković, Željko. 2015a. Phrasal movement: Wh-movement. In A Fábregas, J. Mateu, and M. Putnam (eds.) *The Handbook of Parameters.* Bloomsbury.

Bošković, Željko. 2015b. On prosodic boundaries. In *Formal Studies in Slavic Linguistics: Proceedings of Formal Description of Slavic Lanaguges* 10.

Bošković, Željko. 2016a. On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling. *The Linguistic Review: Special Issue on Labels* 33:17-66.

Bošković, Željko. 2016b. Introduction. *The Linguistic Review: Special Issue on Labels* 33: 1-16.

Bošković, Željko. 2018. On Movement out of Moved Elements, Labels, and Phases. *Linguistic Inquiry* 49: 247-282.

Bošković, Željko. 2019a. On the Coordinate Structure constraint and Labeling. In R. Stockwell, M. O'Leary, Z.Xu, and Z.L.Zhou (eds.), *Proceedings of the 36th West Coast Conference on Formal Linguistics.* Cascadilla Proeceedings Project, 71-80.

Bošković, Željko. 2019b. Generalized Asymmetry. Ms., University of Connecticut.

Bošković, Željko. In press. On the impossibility of moving IPs and V-2 clauses and labeling. In Peter Kosta, Teodora Radeva-Bork (eds.). Frankfurt am Main: Peter Lang.

Bošković, Željko. To appear (a). On the Syntax and Prosody of Verb Second and Clitic Second. In T. Biberauer, S. Wolfe, R.Woods (eds.), *Rethinking Verb Second.* Oxford: Oxford University Press.

Bošković, Željko. To appear (b). On the coordinate structure constraint, Across-the-Board-Movement, Phases, and Labeling. Ms, University of Connecticut.

Bouchard, D. 1984. *On the Content of Empty Categories*. Dordrecht: Foris.

Bowerman, M. 1973. *Early Syntactic Development.* Cambridge: Cambridge University Press.

Brandner, Ellen. 2004. Head-movement in minimalism, and V2 as Force-marking. In Lohnstein, H. and S.TRissler (eds.) *The Syntax and Semantics of the Left Periphery.* Mounton de Gruyer.

Brattico, Pauli, and Taija Saikkonen. 2010.Grammatical movement and functional heads: Evidence from Child Finnish. Nordic Journal of Linguistics 33 (1): 5–29.

Brinton, Laurel J. 1996. *Pragmatic markers in English: grammaticalization and discourse functions*. Berlin: Mouton de Gruyter.

Brown, R. 1968. The development of wh-questions in child speech. *Journal of Verbal Learning and Verbal Behaviour* 7: 279-290.

Brown, R. and C. Frazer. 1963. The Acquisition of Syntax. In C. Cofer and B. Musgrave (eds), *Verbal Behavior and Learning Problems and Processes.* McGraw-Hill;158-201.

Brown, Paul, Joseph, Brian, and Wallace, Rex. 2009. Questions and answers. In Baldi and Cuzzolin (eds.), *New Perspectives on Historical Latin Syntax,* 489-530, Berlin/New York: Mouton De Gruyter.

Bury, D., 2002. A reinterpretation of the loss of Verb-Second in Welsh. In: Lightfoot, D.W. (Ed.), Syntactic effects of Morphological Change. Oxford University Press, pp. 215–231.

Bury, D. 2010. Verb-second, particles, and flexible verb-initial orders. *Lingua* 120:303-314.

Burzio, L.1986.*Italian Syntax: A Government-Binding Approach*. Studies in Natural Language and linguistic theory. D. Reidel Publishing Company.

Cardoso, Adriana. 2012. Extraposition of restrictive relative clauses in the history of Portuguese. In Galves, Charlotte, Sonia Cyrino, Ruth Lopes, Fiomena Sandalo, and Juanito Avelar (eds.) *Parameter theory and linguistic change*. Oxford University Press, 77-96.

Carnie, Andrew, Heidi Harley and Elizabeth Pyatt. 2000. 'VSO Order as Rais- ing out of IP: Some evidence from Old Irish', in A. Carnie and E. Guilfoyle (eds.), *The Syntax of Verb Initial Languages*, Oxford University Press, New York, 39-59.

Carnie, A., Harley, H. (Eds.), 2005. On the Syntax of Verb Initial Languages. John Benjamins Publishing Company, Amsterdam/Philadelphia.

Cecchetto, Carlo & Caterina Donati. 2015. *(Re)labeling*. Cambridge: MIT Press.

Chametzky, Robert. A. 2000. *Phrase structure: From GB to Minimalism.* Blackwell.

Cheng, Lisa. 1991. *On the Typology of Wh-Questions*. Doctoral dissertation, MIT.

Cheng, Lisa Lai-Shen, and Johan Rooryck. 2000. Licensing wh-in-situ. *Syntax* 3.1:1-19.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries. In R. Martin, D. Michaels, and J. Uriagereka, eds., *Step by step: Essays on Minimalism in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In M. Kenstowicz, ed., *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36, 1–22.

Chomsky, Noam. 2008. On phases. In: Freidin, R., Otero, C.P., Zubizarreta, M.L. (Eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud* (Current Studies in Linguistics 47). MIT Press, Cambridge, Massachusetts, pp. 133–166.

Chomsky, Noam. 2013. Problems of Projection. *Lingua* 130:33-49.

Chomsky, Noam. 2015. Problems of Projection: Extensions. In Elisa Di Domenico, Cornela Hamann, and Simona Matteini, (eds.), *Structures, Strategies, and Beyond-Studies in Honor of Adrianna Belletti*, 3-16. Amsterdam: John Benjamins.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.

Citko, Barbara. 2008. Missing labels. *Lingua* 118. 907–944.

Clark, Eve V. 1987. The principle of contrast: A constraint on language acquisition. In Brian Mac Whinney (ed.), *Mechanisms of language acquisition: the 20th annual Carnegie Symposium on Cognition*. Mahway, NJ: Lawrence Erlbaum, 1-34.

Clark, Eve V. 1990. On the pragmatics of contrast. *Journal of Child Language* 17: 417-31.

Clark, R., and I. Roberts. (1993). 'A Computational Approach to Language Learnability and Language Change'. Linguistic Inquiry 24: 299–345.

Collins, Chris. 1991. Why and how come. *More Papers on Wh-Movement*, 31-45. MIT Working Papers in Linguistics 15. Cambridge: MIT.

Collins, Chris. 2002. Eliminating labels. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and explanation in minimalist syntax*. New York: Wiley, , 42–64.

Collins, Chris. 2014. Merge(X.Y) = {X,Y}. Ms., New York University.

Corver, Norbert. 2017. Freezing effects. In Blackwell Companion to Syntax.

Coveney, Aidan. 1995. The use of the QU- final interrogative structure in spoken French. *Journal of French Language Studies* 5: 143-171.

Coveney, Aidan. 2002. *Variability in Spoken French: a sociolinguistic study of Interrogation and negation*. Bristol/Portland: Elm Bank. Reprint from 1996 original.

Crisma, Paola.1992. On the acquisition of Wh in French. *Geneva Generative Papers* 0 (1-2): 155-122.

Croft, W. 2003. *Typology and Universals*. Cambridge: Cambridge University Press.

Dadan, Marcin. 2015. Reflexiva Tantum acquire differently? Splitting the voice domain in reflexive clitics in child Polish. Ms., University of Connecticut.

Dadan, Marcin. 2019a. Loss of movement and labeling: Grammatical pressure and diachronic change. In *Proceedings of the 36th West Coast Conference on Formal Linguistics*.

Dadan, Marcin. 2019b. Movement that is kept and movement that is lost: Movement, adjunction, and labeling in diachronic perspective. Paper presented at LSA 93, NYC, January 3-6, 2019.

Dadan, Marcin. 2019c. Loss of Specifiers in Slavic reflexives: Acquisitional structure reanalysis and diachronic change. Paper presented at FASL 28, Stony Brook University, May 3-5, 2019.

Danckaert, Lieven. 2012. *Latin Embedded Clauses: The Left Periphery.* Amsterdam: John Benjamins.

Danckaert, Lieven. 2017. *The development of latin clause structure: a study of the extended verb phrase*. (Oxford Studies in Diachronic and Historical Linguistics). New York, NY: Oxford University Press.

DeGraff, M. 2005. Morphology and Word Order in ‘‘Creolization’’ and Beyond’, in G. Cinque and R. Kayne (eds.), The Oxford Handbook of Comparative Syntax. Oxford: Oxford University Press, pp. 293–372.

Dehé, Nicole. 2009. Clausal parentheticals, intonational phrasing, and prosodic theory. *Journal of Linguistics* 45:569-615.

Devine, Andrew, and Stephens, Laurence. 2006. *Latin Word Order: Structured Meaning and Information* .Oxford: OUP.

Despić, Miloje. 2011. *Syntax in the absence of determiner phrase*. Doctoral dissertation, University of Connecticut, Storrs.

Despić, Miloje. 2013. Binding and the structure of NP in Serbo-Croatian". *Linguistic Inquiry* 44: 239–70.

Diesing, M. 1990. Verb-second in Yiddish and the nature of the subject position". *Natural Language and Linguistic Theory* 8: 41-79.

Djamouri, R.,Waltraud, P. ,Whitman, J. 2013. Syntactic change in Chinese and the argument - adjunct asymmetry. *Breaking Down the Barriers*, 577-594.

Doherty, C., 2000. Residual verb second in Early Irish: on the nature of Bergin’s construction. In: *Diachronic*a, vol. 17 (1), John Benjamins Publishing Company, 5–38.

Dryer, Matthew S. 1992. The Greenbergian Word Order Correlations. *Language* 68. 81-138.

Duguine, Maia and Aritz Irurtzun. 2014. From obligatory WH-movement to optional WH-in-situ in Labourdin Basque. *Language* 90: 1-30.

Ehala, M., 2006. The Word Order of Estonian: Implications to Universal Language. Journal of Universal Language 7, 49–89.

Eide, Kristine. 2006. *Word order structures and unaccusative verbs in Classical and Modern Portuguese*. Doctoral dissertation, University of Oslo.

Embick, David. 2004. Unaccusative syntax and verbal alternations. In A. Alexiadou, E. Anagnostopoulou & M. Everaert (eds), *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*,Oxford: OUP, 137–158.

Etter, A. 1985. *Die Fragesätze im Rgveda.* Berlin.

Eythórsson, Thórhallur. 1995. Verbal syntax in the early Germanic languages. PhD dissertation. Cornell University.

Faarlund, J. T. 1990. *Syntactic Change: Toward a Theory of Historical Syntax*. Berlin: Mouton de Gruyter.

Feng, Shengli. 1996. Prosodically Constrained Syntactic Changes in Early Archaic Chinese. *Journal of East Asian Linguistics* 4: 235-260.

Ferreira, M. 2009. Null subjects and finite control in Brazilian Portuguese. In J. Nunes (ed.), *Minimalist Essays on Brazilian Portuguese Syntax*. Amsterdam: John Benjamins, 17–49.

Figueiredo Silva, Maria Cristina, and Elaine Grolla. 2016. Some syntactic and pragmatic aspects of WH-in-situ in Brazilian Portuguese. In Mary A. Kato and Francisco Ordóñez (eds.) *The Morphosyntax of Portuguese and Spanish in Latin America*, 259-285. New York City: OUP.

Fischer, Olga, Ans van Kemenade, Willem Koopman, and Wim van der Wurff. 2001/2004. *The syntax of early English*. Cambridge: Cambridge University Press.

Fontana, Josep. 1993. *Phrase Structure and the Syntax of Clitics in the History of Spanish*. Doctoral Dissertation, University of Pennsylvania.

Fontana, Josep. 1997. On the integration of second position phenomena. In Ans van Kemenade & Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 207–249. Cambridge: Cambridge University Press.

Fortson IV, Benjamin W. 2004, *Indo-European Language and Culture: An Introduction.* Blackwell Publishing.

Fuss, Eric. 2003. On the historical core of V2 in Germanic. *Nordic Journal of Linguistics* 26(02). 195–231.

Galves, Charlotte. & Kroch, Anthony. 2016. Main syntactic changes from a Principles-and - Parameters View. In W. Wetzels et al. (eds.), *The handbook of Portuguese Linguistics*- Wiley-Blackwell, 487-503.

Gelderen, Elly, van. 2004. *Grammaticalization as Economy.* Amsterdam: John Benjamins.

Gelderen, Elly. van. 2009. Renewal in the left periphery: economy and the complementizer layer. *Transactions of Philological Society* 107 (2): 131-195.

Gelderen, E. van. 2011. *The Linguistic Cycle.* Oxford: OUP.

Gelderen, E. van. 2014. Changes in Pysch-Verbs: A reanalysis of little v. *Catalan Journal of Linguistics* 13:99-122.

Gelderen, E. van. 2015. Problems of Projection: The role of language change in labeling paradoxes. *Studia Linguistica*: 1-15.

Gelderen, E. van. 2018. Cyclical Change and Problems of Projection. Lingbuzz/003815 [v1].

Grolla, Elaine. 2009. Speculations about the Aquisition of Wh-Questions in Brazilian Portuguese *In* Acrísio Pires & Jason Rothman (eds.), *Minimalist Inquiries into Child and Adult Language Acquisition: Case Studies across Portuguese,* Mouton de Gruyter.

Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and composition of the left periphery* (The Cartography of Syntactic Structures). Oxford: Oxford University.

Hagstrom, P. 1998. *Decomposing questions*. PhD Dissertation, MIT.

Haeberli, Eric. 2002. Inflectional morphology and the loss of verb-second in English. In *Syntactic effects of morphological change*, ed. David Lightfoot, 88–106. Oxford: Oxford University Press.

Haeberli, Eric. 2002. Observations on the loss of verb second in the history of English. In Jan-Wouter Zwart & Werner Abraham (eds.), *Studies in Comparative Germanic Syntax*, 245–272. Amsterdam: John Benjamins.

Haeberli, Eric and Tabea Ihsane. 2016. Revisiting the loss of verb movement in the history of English. *Natural Language and Linguistic Theory 34: 497-542.*

Haegeman, L. 1996. Verb Second, the Split CP and Null Subjects in Early Dutch Finite Clauses." *GenGenP* 4.2, 13 3 -17 5.

Hale, Mark. 1987. *Studies in the Comparative Syntax of the Oldest Indo-Iranian Languages.* Doctoral dissertation, Harvard University.

Hale, Mark. 1995. Wackernagel's Law in the Rigveda. Ms., University of Concordia.

Hale, Mark. 1996. Theory and Method in Historical Linguistics. Ms., University of Concordia.

Hale, Mark. 1998. Diachronic Syntax. *Syntax* 1: 1–18.

Hallman, P., 2000. On the derivation of V-final and its relation to V2. In: Hirotani, et al. (Eds.), Proceedings of NELS 30. GLSA.

Hamann, Cornelia. 2000. The acquisition of constituent questions and the requirements of interpretation. In: Friedemann, Marc-Ariel; Rizzi, Luigi (eds.). *The Acquisition of Syntax*. London: Longman, pp. 170-201.

Hamann, Cornelia. 2001. Speculations about early syntax: The production of wh-questions by normally developing French children and French children with SLI. Catalan Journal of Linguistics 5, 143-189.

Hamann, Cornelia; Penner, Zvi; Lindner, Katrin. 1998. German impaired grammar: The clause structure revisited». Sonderheft: 'Specific Language Impairment in Children. *Language Acquisition* 7: 193-246.

Harris, Alice C, and Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 1989. From Purposive to Infinitive – a Universal Path of Grammaticization. *Folia Linguistica Historica* 10: 287–310.

Haspelmath, Martin. 1997. *Indefinite Pronouns*. Oxford: Oxford University Press.

Haspelmath, Martin. 2001. Non-Canonical Marking of Core Arguments in European Languages. In Aikhenvald, Alexandra Y.,Dixon, Robert M. W.,Onishi, Masayuki (eds), *Non-Canonical Marking of Subjects and Objects*. Amsterdam: John Benjamins, 53-83.

Hayes, Bruce. 1989. The prosodic hierarchy in meter. In *Rhythm and Meter*, ed. by Paul Kiparsky and Gilbert Youmans, 201–260. Orlando: Academic Press.

Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of Grammaticalization.* Cambridge University Press.

Hendriksen, Hans. 1986. *Himachali Studies, III: Grammar.* København: Munksgaard

Hendriksen, Hans. 1990. Sentence Position of the Verb in Himachali. *Acta Linguistica Hafniensia X. 159-172.*

Heycock, Caroline, Antonella Sorace, Zakaris Svabo Hansen, Frances Wilson, and Sten Vikner. 2012. Detecting the late stages of syntactic change: The loss of V-to-T in Faroese. *Language* 88: 558–600.

Heycock, Caroline and Joel Wallenberg. 2013. Explaining the loss of verb movement: How embedded V2 and V-in-situ conspired against V-to-T. *Journal of Comparative Germanic Linguistics* 16, 2-3:127-157.

Hoekstra, T., and N. Hyams. 1998. Aspects of Root Infinitives. Lingua 106:81-112

Hock, Hans Henrich. 1991. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.

Hock, Hans Henrich and B. Joseph. 1996. *Language History, Language Change, and Language Relationship: An Introduction to Historical and Comparative Linguistics*. Berlin: Mouton de Gruyter.

Holmberg, Anders. 1986. *Word Order and Syntactic Features in the Scandinavian Languages and English*. Dissertation, University of Stockholm.

Holmberg, Anders. 2000. Deriving OV order in Finnish. In Peter Svenonius (ed.). *The Derivation of OV and VO*, 123-152. Amsterdam: John Benjamins.

Holmberg, Anders. 2005. Stylistic fronting. In: Everaert, M., van Riemsdijk, H. (Eds.), The Blackwell Companion to Syntax. Blackwell Publishing, Blackwell.

Holmbert, Anders. 2010. Verb Second. Available at: https://ling.auf.net/lingbuzz/001087 (Accessed August 2018).

Holmberg, Anders. 2015. Verb second. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: an international handbook*, 342–382. Berlin: de Gruyter.

Holmberg, Anders. 2016. Non-V2 in V2 languages. Presented at the workshop Rethinking Verb Second: Assessing the Theory and Data. Cambridge, March 22-24, 2016.

Holmberg, Anders and C. Platzack. 1995. *The role of inflection in Scandinavian syntax.* Oxford: Oxford University Press.

Hooper, J & S Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4. 465–497.

Hornstein, N. 2009. *A theory of syntax: Minimal operations and Universal Grammar*. Cambridge: Cambridge University Press.

Hornstein, Norbert & Jairo Nunes. 2008. Adjunction, labeling and Bare Phrase Structure. *Biolinguistics* 2. 57–86.

Horvath, Julia. 2010. "Discurse Features", syntactic displacement and the status of contrast. *Lingua* 120:1346-1369.

Hróarsdóttir, T. 2000. *Word Order Change in Icelandic: From OV to VO*. Amsterdam: John Benjamins.

Huang, C.–T. J. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. Dissertation. MIT.

Huang, C.-T. James. 2006. The macro-history of Chinese and the theory of change. Paper presented at the Chinese Linguistics Workshop, December 1-2, 2006 at the University of Chicago.

Hulk, Aafke, & Ans van Kemenade. 1995. V2, pro-drop, functional projections and language change. In Allan Battye & Ian Roberts (eds.), *Clause Structure and Language Change*. Oxford: Oxford University Press, 227–256.

Hunter, Tim. 2010. *Relating Movement and Adjunction in Syntax and Semantics*. Doctoral Dissertation, University of Maryland.

Jayaseelan, Karattuparambil. 2010. Stacking, stranding and pied-piping: A proposal about word order. *Syntax* 13: 298–330.

Jespersen, Otto. 1917. *Negation in English and Other Languages*. Copenhagen: A.F. Høst.

Jiménez, Maria Luisa. 1997. *Semantic and Pragmatic Conditions on Word Order in Spanish.* Doctoral Dissertation, Georgetown University.

Jonsson, Johannes Gisli. 1991. Stylistic fronting in Icelandic. In *Working papers in Scandinavian syntax* 48: 1-43. Department of Scandinavian Languages, University of Lund.

Jouitteau, M., 2007. The Brythonic reconciliation. In: Craenenbroek, J., Rooryck, J. (Eds.), *The Linguistics Variation Yearbook* 2007. John Benjamins, Netherland.

Jouitteau, Maria Luisa. 2010. A typology of V2 with regard to V2 and second position phenomena: An introduction to the V1/V2 volume. *Lingua* 120; 197-209.

Jouitteau, Maria Luisa. to appear. Left Edge Filling Trigger. Wolfe, S. and R. Woods (eds.) *Rethinking V2.* Oxford: Oxford University Press.

Julien, Marit. 2007. *Embedded V2 in Norwegian and Swedish*. Working papers in Scandinavian Syntax. http://ling.auf.net/lingbuzz/000475/v1.pdf (August 2017).

Julien, Marit. 2015. The Force of V2 Revisited. *Journal of Comparative Germanic Linguistics* 18:139-181

Kaiser, Georg A. 2002. *Verbstellung Und Verbstellungswandel in Den Romanischen Sprachen*. Max Niemeyer Verlag.

Kaiser, Georg A. 2009. Losing the null subject. A contrastive study of (Brazilian) Portuguese and (Medieval) French. In Georg A. Kaiser & Eva Maria Remberger (eds.), *Proceedings of the Workshop "Null Subjects, expletives and locatives in Romance,"*

Kaiser, Georg A. & Michael Zimmermann. 2015. Exploring verb order differences between (Medieval) Romance and (Modern) Germanic. Paper presented at the Traces of History Project: Workshop on Verb Second. 9th-10th March 2015., Oslo, Norway.

Kato, Mary A. 1993. Word order change: the case of Brazilian Portuguese wh-questions. Paper presented at the *XI International Congress of Historical Linguistics.* Los Angeles.

Kato, Mary A. 2004. Two types of wh-in-situ in Brazilian Portuguese. Paper presented at *GURT* (2004), Washington.

Kato, Mary A. 2013. Wh-movement in the history of Brazilian Portuguese. Accessible at http://www.diacronia.ro/ro/indexing/details/A22713/pdf. Accessed on October 19, 2017.

Kato, Mary A. and M. E. L. Duarte. 2002. A Diachronic Analysis of Brazilian Portuguese wh-questions. *Santa Barbara Portuguese Studies,* vol. VI, University of California at Santa Barbara, Center for Portuguese Studies: 326-339.

Kayne, Richard.1975. *French Syntax: The Transformational Cycle*. Current Studies in Linguistics Series, 6. The MIT Press, Cambridge, Massachussetts and London, England.

Kayne, Richard. 1986. Participles, Agreement, Auxiliaries, Se/Si and pro. Handout to talk at Princeton University.

Kayne, Richard. 1988.Romance se/si", GLOW newsletter 20.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, Richard. 2010. Why are there no directionality parameters? In *Proceedings of the 28th West Coast Conference on Formal Linguistics*, ed. by Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer, and Barbara Tomaszewicz, 1-23. Somerville: Cascadilla Press.

Kayne, Richard. 2014. Why are there no directionality parameters? In Theresa Biberauer and Michelle Sheehan (eds.), *Theoretical Approaches to Disharmonic Word Order,* 219-244. Oxford: Oxford University Press.

Kemenade, Ans van. 1987. Syntactic case and morphological case in the history of English. Dordrecht: Foris.

Kemenade, Ans van. 1997. V2 and embedded topicalization in Old and Middle English. In *Parameters of morphosyntactic change*, eds. Ans van Kemenade and Nigel Vincent, 326–352. Cambridge: Cam-bridge University Press.

Kemenade, Ans van. 2007. Formal syntax and language change: developments and outlook, *Diachronica* 24: 155–69.

Kemenade, Ans van. 2012. Rethinking the loss of verb second. In Terttu Nevalainen and Elizabeth Closs Traugott (eds.) *The Oxford handbook of the history of English*, 823–834. Oxford: Oxford University Press.

Kemenade, Ans van, and Marit Westergaard. 2012. Syntax and information structure: Verb-second variation in Middle English. In Anneli Meurman-Solin, María José Löpez-Couso, and Bettelou Los (eds.), *Information structure and syntactic change in the history of English*, Oxford: Oxford University Press, 87–118.

Kiparsky, Paul. 1995. Indo-European Origins of Germanic Syntax. In Adrian Battye and Ian Roberts (eds.) *Clause Structure and Language change.* New York City: Oxford University Press, 140-170.

Kiparsky, Paul. 1996. The Shift to Head-Initial VP in Germanic, in H. Thráinsson, S. Epstein, and S. Peter (eds.), Studies in Comparative Germanic Syntax II. Dordrecht: Kluwer, pp. 140–79.

Kiparsky, Paul, and Leo Condoravdi. 2006. Tracking Jespersen's Cycle. In Mark Janse, Brian D. Joseph, and Angela Ralli (eds.), Proceedings of the Second International Conference on Modern Greek Dialects and Linguistic Theory, Mytilene: Doukas, 172–197.

Klima, E.S, and U.Bellugi. 1996. Syntactic regularities in the speech of children. In J.Lyons and R.Wales (eds.), *Pyscholinguistic Papers,* Edinburgh University Press; 183-207.

Ko, Heejeong. 2011. Predication and edge effects. Natural Language & Linguistic Theory 29. 725–778.

Koeneman, Olaf. 1995. Flexible phrase structure: on the role of verbal projections in Dutch, English and Yiddish. Unpublished MA dissertation. Utrecht University.

Koeneman, Olaf. 2000. The flexible nature of verb movement. Doctoral dissertation., Utrecht University.

Koeneman, Olaf. 2002. The distribution of declarative verb second in Germanic. In C. Jan-Wouter Zwart & Werner Abraham (eds.), *Studies in Comparative Germanic Syntax: Proceedings from the 15th Workshop on Comparative Germanic Syntax (Groningen, May 26-27, 2000)*, 175–201. Amsterdam ; Philadelphia: John Benjamins Publishing.

Koizumi, M. (1994). Layered specifiers. In: *Proceedings of NELS 24* (M. Gonzalez, ed.), pp. 255- 269. GLSA, University of Massachusetts, Amherst.

Kotek, Hadas. 2014. *Composing Questions*. PhD Dissertation. MIT.

Kroch, Anthony.1989. Reflexes of Grammar in Patterns of Language Change, *Journal of Language Variation and Change* 1**,** 199-244.

Kroch, Anthony. 1994. Morphosyntactic variation. *Chicago Linguistic Society.*

Kroch, A., A. Taylor, D. Ringe. 1995. The Middle English verb-second constraint: a case study in language contact and language change. In: Herring, S., van Reenen, P., Schoesler, L. (Eds.), Textual Parameters. John Benjamins, Philadelphia.

Kroch, A., and A. Taylor. 1997. Verb movement in Old and Middle English: Dialect variation and language contact. In A. van Kemenade and N. Vincent (eds.), *Parameters of Morphosyntactic Change.* Cambridge: Cambridge University Press, 297-325.

Kroch, Anthony & Taylor, Ann. 2000. Verb-object order in early Middle English. In *Diachronic Syntax: Models and Mechanisms*, Susan Pintzuk, George Tsoulas &AnthonyWarner (eds), 132–163. Oxford: OUP.

Kuroda, Shige.-Yuki. 1988. Whether we agree or not: A comparative syntax of English and Japanese. *Linguisticae Investigationes* 12. 1–47.

Kuroda, Shige-Yuki. 2007. On the syntax of Old Japanese. In *Current issues in the history and structure of Japanese*, ed. Bjarke Frellesvig, J.C. Smith, and Masayoshi Shibatani, 263–318. Tokyo: Kurosio.

Kuznetsova, Julia, and Tore Nesset. 2015.In Which Case are Russians Afraid? *Bojat'sja* with Genitive and Accusative Objects. *Journal of Slavic Linguistics* 23(2):255-83.

Labelle, Marie & Paul Hirschbühler. 2012. Topic and Focus in Old French V1 and V2 Structures. Paper presented at the DiGS 14, Ottowa.

Labelle, Marie & Paul Hirschbühler. 2017. Leftward Stylistic Displacement in Medieval French. In Eric Mathieu & Robert Truswell (eds.), *Micro-change and Macro-change in Diachronic Syntax*. Oxford: Oxford University Press.

Ladd, D. Robert.1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

L'Arrivée, Pierre. 2010. The pragmatic motifs of the Jespersen Cycle: Default, activation, and the history of negation in French. *Lingua* 120 (9):2240–2258.

L'Arrivée, Pierre. 2019. To be or not to be informational: Preverbal complements in Medieval French V2 configurations. *Glossa* 4 (1):85;1-19.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19.3: 335-391.

Lasnik, Howard. 1999. *Minimalist analysis*. Oxford: Blackwell.

Lasnik, Howard. 2001. Subjects, objects, and the EPP. In William D. Davies and Stanley Dubinsky (eds.), *Objects and Other Subjects: Grammatical Functions, Functional Categories, and Configurationality*. Dordrecht: Kluwer, 103–121.

Lebeaux, David. 1988. *Language acquisition and the form of the grammar*. Doctoral dissertation, University of Massachusetts, Amherst.

Ledgeway, Adam. 2008. Satisfying V2 in early Romance: Merge vs. Move. *Journal of Linguistics* 44(02). 437–470.

Ledgeway, Adam. 2010. Introduction: The clausal domain: CP structure and the left periphery. In Roberta D'Alessandro, Adam Ledgeway & Ian Roberts (eds.), *Syntactic Variation: The Dialects of Italy*.Cambridge: Cambridge University Press, 38–52.

Ledgeway, Adam. 2012. *From Latin to Romance. Morphosyntactic Typology and Change.* New York: Oxford University Press.

Ledgeway, Adam. 2017. Late Latin Verb Second: The Sentential Word Order of the Itinerarium Egeriae. *Catalan Journal of Linguistics* (Special Issue: Generative Approaches to Latin Syntax),163–216.

Ledgeway, Adam, and Ian Roberts. 2017. Principles and Parameters. In Ledgeway, A., and I. Roberts, *The Cambridge Handbook of Historical Syntax*, Cambridge University Press, 581-628.

Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese*. Cambridge, MA: MIT Press.

Lehmann, W. 1973. A Structural Principle of Language and its Implications. *Language* 49: 47–66.

Lehmann, C. 1995. *Thoughts on Grammaticalization.* Munich: Lincom.

Lessa-de-Oliveira, Adriana. 2003. *Aquisição de constituintes-QU em dois dialetos do português brasileiro* [Acquisition of wh-constituents in two dialects of Brazilian Portuguese]. Master Thesis, Unicamp, Brazil.

Lewis, Charlton T., and Charles Short (1879). A Latin dictionary. Oxford: Clarendon.

Lightfoot, David, 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.

Lightfoot, David. 1991. *How to Set Parameters: Arguments from Language Change.* Cambridge, Mass.: MIT Press.

Lightfoot, David. 1995. Why UG needs a learning theory: Triggering verb movement. In Adrian Batye and Ian Roberts (eds.), *Clause Structure and Language Change.*Oxford University Presss, 31-52.

Lobeck, Anne. 1995. *Ellipsis: Functional heads, licensing and identification*. New York: Oxford University Press.

Lopes-Rossi, Maria Aparecida. (1996). *As Orações Interrogativas-q no Português de Brasil: Um Estudo Diacrônico*. Doctoral dissertation, Universidade Estadual de Campinos, Brazil.

Lunt, Horace,G. 2001. *Old Church Slavonic Grammar.*Mouton de Gruyter.

Lyons. Christopher. 1999. *Definiteness.* Cambridge: Cambridge University Press.

MacDonell, Arthur, A. 1916. *A Vedic Grammar for Students*. Oxford University Pres

MacWhinney, Brian (2000). *The CHILDES Database: Tools for Analyzing Talk*. Mahwah, NJ: Lawrence Erlbaum Associates.

Madariaga, Nerea. 2017. Understanding Grammars through Diachronic Change. *Frontiers in Psychology*(8):1-16.

Mahajan, Anoop. 2003. Word order and (remnant) VP movement. In Simin Karimi (ed.), *Word Order and Scrambling*, 217–237. Malden MA: Blackwell.

Manetta, Emily.2011. *Peripheries in Kashmiri and Hindi-Urdu*. John Benjamins.

Manzini, M. R. 1986. On Italian SI. In Borer, H. (ed.) *Syntax and Semantics 19: The Syntax of Pronominal Clitics*. Academic Press, New York, 241–262.

Marantz, A. P. 1984. *On the Nature of Grammatical Relations*. Cambridge: MIT Press.

Martins, Ana Maria. 2002. The loss of IP-scrambling in Portuguese: Clause structure, word order variation and change. In David Lightfoot (ed.), *Syntactic Effects of Morphological Change*, 232–248. Oxford: Oxford University Press.

Massam, D. 2010. V1 or V2?: On the left in Niuean, Jouitteau (ed.), *Lingua* 120: 284-302.

Mathieu, E. 2004. The Mapping of Form and Interpretation: The Case of Optional WH movement in French. *Lingua* 114:1090-1132.

Mathieu, Éric. 2012. The left periphery in Old French. In Deborah Arteaga (ed.), *Research in Old French: The state of the art*, 327–350. (Studies in Natural Language and Linguistic Theory). Dordrecht: Springer.

Mathieu, Éric. 2016. The wh parameter and radical externalization. In Luis Eguren, Olga Fernández-Soriano, and Amaya Mendikoexea (eds.), *Rethinking Parameters, 252-290* Oxford: Oxford University Press.

McFadden, Thomas. 2004. *The Position of Morphological Case in the Derivation: A Study on the Syntax-morphology Interface*. Doctoral dissertation, University of Pennsylvania.

McFadden, Thomas. 2005. OV-VO in English and the role of case-marking in word order. *English Language and Linguistics* 9(1): 63–82.

McGinnis, Martha. 2004. Lethal Ambiguity. *Linguistic Inquiry* 35:47-95.

Meelen, Marieke. To appear. Reconstructing the rise of V2 in Welsh. In Theresa Biberauer, Sam Wolfe, and R. Woods (eds.) *Rethinking Verb Second.* Oxford: Oxford University Press.

Medova, L. 2009. *Reflexive Clitics in the Slavic and Romance Languages. A Comparative View from an Antipassive Perspective*. Doctoral Dissertation, Princeton University.

Meyer, Roland. 2017. The C System and Complement Clauses in the History of Slavic Languages. *Language* 93(2): e97-e113.

Meyers, Lindsy Lee. *Wh-interrogatives in Spoken French: A Corpus-Based Analysis of their form and function.* Doctoral dissertation, The University of Texas, Austin.

Migdalski, Krzysztof. 2016. *Second Position Effects in the Syntax of Germanic and Slavic Languages.*Wydawnictwo Universytetu Wrocławskiego.

Moro, Andrea. 2000. *Dynamic Antisymmetry*. Cambridge, MA: MIT Press.

Mossé, Fernand. 1968. *Manual of Middle English*, Johns Hopkins Press.

Möhlig-Falke, Ruth (2012). *The Early English Impersonal Construction*. Oxford University Press.

Munaro, Nicola and Hans-Georg Obenauer. 1999. On underspecified wh-elements in pseudo-interrogatives. *University of Venice Working Papers in Linguistics* 9:181-253.

Munaro, Nicola, Cecilia Poletto, and Jean-Yves Pollock. 2001. Eppur si muove! On comparing French, Portuguese, and Bellunese wh-movement. Ms., University of Padua and Université de Picardie à Amiens.

Murasugi, Keiko. 2018. Acquisition of Labeling: Evidence from a comparative study of English and East-Asian child languages. Talk given at 50 years of Linguistics at UConn (8 September 2018, University of Connecticut, Storrs).

Muysken, P. 1988. Are Creoles a Special Type of Language?, in F. Newmeyer (ed.), *Linguistics: The Cambridge Survey*, Volume 2. Cambridge: Cambridge University Press, pp. 285–302.

Muysken, P. 2000. *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.

Narrog, Heiko. 2019. Origin and structure of focus concord constructions in Old Japanese- a synthesis. *Glossa* 4(1):108;1-41.

Neelman, Ad, Elena Titov, Hans van de Koot, and Reiko Vermeulen. 2009. A syntactic typology of topic, focus and contrast. In Craenenbroeck, Jeroen van (ed.), *Alternatives to Cartography.* De Gruyter Mouton, 15-52.

Nespor, Marina, and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Ning, C.–Y. 1993. *The Overt Syntax of Relativization and Topicalization in Chinese*. Ph.D. Dissertation, University of California, Irvine

Nunes, Jairo, 2016. Edge Features on Phase Heads or Moving Elements. Available at: https://www.researchgate.net/publication/309488171_Edge_features_on_phase_heads_or_moving_elements (Accessed: January 2016).

Nomura, Takashi. 1993. Jodaigo no ‘no’ to ‘ga’ nituite [OJ ‘no’ and ‘ga’]. *Kokugo Kokubun* 62(2): 1-17.

Ochi, Masao. 2004. *How come* and other adjunct *wh*-phrases: A cross-linguistic perspective. *Language and Linguistics* 5.1:29-57.

Ogawa, Kunihiko., 1976. Where Diachronic and Synchronic Rules Meet: A Case Study from Japanese Interrogatives and Kakari-Musubi. Papers in Japanese Linguistics 5-77, 193-242.

Oiry, Magda. 2011. A case of true optionality: Wh-in-situ patterns like long movement in French. In Anna Roussou and Christos Vlachos (eds.) *The Optionality of Wh-movement. Linguistic Analysis* 37: 115-142.

Oiry, Magda. 2015. French embedded cleft questions. Talk at SyMiLa Toulouse, June 12, 2015.

Otaki, Koichi. 2011. Argument ellipsis arising from non-fusional case morphology. In *Online Proceedings of GLOW in Asia Workshop for Young Scholars 2011*, ed. by Koichi Otaki, Hajime Takeyasu, and Shin-ichi Tanigawa, 247–261. Mie: Mie University.

Panhuis, Dirk G.J. 1982. *The Communicative Perspective in the Sentence. A Study of Latin Word Order* [Studies in Language Companion Series 11], Amsterdam: John Benjamins.

Pesetsky, David. 1995. *Zero Syntax. Experiencers and Cascades.* Cambridge, MA: MIT Press.

Pesetsky, David. 2000. *Phrasal Movement and Its Kin*. Cambridge, MA: MIT Press.

Pierce, A. 1992. Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars. Dodrecht: Kluwer

Pinkster, Harm. 1990. *Latin Syntax and Semantics*. London: Routledge.

Pintzuk, S. 1991. *Phrase Structure in Competition: Variation and Change in Old English.* Doctoral Dissertation, University of Pennsylvania.

Pintzuk, S., 1993. Verb seconding in Old English: verb movement to Infl. The Linguistic Review 10, 5–35.

Pires, Acrisio, and Heather Taylor. 2007. The Syntax of wh-in-situ and Common Ground. In *Romance Languages: Structure, Interfaces, and Microparametric Variation. LSRL 37,* Amherst: GLSA.

Platzack, Christer. 1995. The Loss of Verb Second in English and French. In Ian Roberts & Adrian Battye (eds.), *Clause Structure and Language Change*. Oxford: Oxford University Press, 200–226.

Plunkett, Bernadette. 2004.Early peripheries in the absence of C. In: Adger, David; De Cat, Cecile; Tsoulas, George (eds.). *Peripheries. Syntactic Edges and Their Effect.* Dordrecht: Kluwer, 383-405.

Plunkett, Bernadette, and Cécile De Cat (2001). Root specifiers and null subjects revisited. In Do, Anna H.-J., Dominguez, Laura; Johansen, Aimee (Eds.). *Proceedings of the 25th Annual Boston University Conference on Language Development*. Somerville, Mass.: Cascadilla Pres, 611-622.

Poletto, Cecilia. 2002. The left-periphery of V2-Rhaetoromance dialects: a new view on V2 and V3. In Sjef Barbiers, Leonie Cornips & Susanne van der Kleij (eds.), *Syntactic Microvariation*, 214–242. (Meertens Institute Electronic Publications in Linguistics 2). Amsterdam: Meertens Institute.

Poletto, Cecilia. 2013. On V2 Types. In Silvia Luraghi & Claudia Parodi (eds.), *The Bloomsbury Companion to Syntax*, 154–164. London: Bloomsbury.

Poletto, Cecilia. 2014. *Word Order in Old Italian*. Oxford: Oxford University Press.

Poole, Geoffrey. 1996. Optional movement in the Minimalist Program. In Wernham, Samuel D. Epstein, Hoskuldur Thrainsson, and Jan-Wouter Zwart (eds.), *Minimal ideas.* 199-216. Amsterdam Benjamins.

Quinn, Charles. 1997. On the origins of Japanese sentence particles *ka* and *so*. In Ho-min Sohn & John Haig (eds.), *Japanese/Korean linguistics* 6. 61–89. Stanford: Center for the Study of Language and Information.

Rackowski, Andrea and Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. Linguistic Inquiry 36: 565-99

Radford, Andrew. 1988. Small children's small clauses. *Transactions of Philological Society* 86(1): 1-43.

Radford, Andrew. 1990. *Syntactic Theory and the Acquisition of English Syntax: The Nature of Early Child Grammar of English.* Oxford: Blackwell.

Radford, Andrew. 1994. The syntax of questions in child English. *Journal of Child Language* 21: 211-236.

Radford, Andrew. 1996. Towards a structure-building model of acquisition. In Harald Clahsen (ed.), *Generative perspectives on language acquisition*. Amsterdam: John Benjamins, 43-89

Reglero, Lara. 2004. *A'-Dependencies in Spanish and Basque*. Doctoral dissertation, University of Connecticut.

Reglero, Lara. 2007. Wh-in-Situ Interrogatives in Spanish. *Probus* 19:267-297.

Reglero, Lara, and Emma Ticio. 2013. A unified analysis of wh-in-situ in Spanish. *The Linguistic Review* 30(4): 501-546.

Reinhart, Tanya. 1998. Wh-in-situ in the framework of the minimalist program. Natural Language Semantics 6. 29–56.

Reis, M. 1997. Zum syntaktischen Status unselbständiger Verbzweit-Sätze. In C. Dürscheid, K. Ramers, & M. Schwarz (eds.), *Sprache im Fokus*, 121–144. Tübingen: Niemeyer.

Reuland, Eric. 2001. Primitives of binding. *Linguistic Inquiry* 32: 439–492.

Ribeiro, I. 1995. *A sintaxe da ordem no português arcaico; o efeito V*2. Doctoral dissertation, University of Campinas.

Rice, C., and P. Svenonius. 1998. Prosodic V2 in Northern Norwegian. Paper presented at the West Coast Conference on Formal Linguistics 17, University of British Columbia, Vancouver.

Richards, Norvin. 2001. *Movement in language: Interactions and architectures*. Oxford: Oxford University Press.

Richards, Norvin. 2010. *Uttering Trees*. Cambridge, MA: MIT Press.

Richards, Norvin. 2016. *Contiguity theory*. Cambridge, MA: MIT Press.

Rinke, Esther. 2009. Verb placement in Old Portuguese. In Andreas Dufter & Daniel Jacob (eds.), *Focus and Background in the Romance Languages*, 309–332. Amsterdam ; Philadelphia: John Benjamins.

Rivero, Maria Luisa. 1999. Stylistic verb-movement in yes-no Questions in Bulgarian and Breton, In Kenesei, Istvan (ed.), *Crossing Boundaries, advances in the theory of Central and Eastern European languages.* Amsterdam, John Benjamins, 67-90.

Rivero, Maria Luisa and Magda Golędzinowska.2002. The Acquisition of Constructions withReflexive Clitics in Polish. *Cahiers Linguistiques d'Ottawa*, Vol 30, 60-90.

Rivero, María Luisa and Milena Milojević Sheppard. 2003. Indefinite Reflexive Clitics in Slavic: Polish and Slovenian. *Natural Language and Linguistic Theory* 21:89-155.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In L. Haegeman (ed.), *Elements of Grammar.* Dordrecht: Kluwer, 281-337.

Rizzi, Luigi. 2016. Labeling, maximality and the head – phrase distinction. *The Linguistic Review: Special Issue on Labels* 33(1): 103–127

Roberts, Ian. 1991. Excorporation and Minimality, Linguistic Inquiry, 22:1, 209-218.

Roberts, Ian. 1993. *Verbs and Diachronic Syntax: A Comparative History of English and French*. Dordrecht: Kluwer.

Roberts, Ian. 1996. Remarks on the Old English C-system and the Diachrony of V2. *Language Change and Generative Grammar* 7. (Linguistiche Berichte). 154–167.

Roberts, Ian. 1997. Directionality and Word Order Change in the History of English. In A. van Kamenade and N. Vincent (eds.), *Parameters of Morphosyntactic Change.* Cambridge University Press, 397-426.

Roberts, Ian. 2004. The C-System in Brythonic Celtic Languages, V2, and the EPP. *The Structure of CP and IP*, vol. 2, 297–328. (The Cartography of Syntactic Structures). Oxford: Oxford University Press.

Roberts, Ian. 2005. *Principles and parameters in a VSO language: a case study in Welsh*. Oxford; New York: Oxford University Press.

Roberts, Ian. 2007. *Diachronic Syntax*. New York: Oxford University Press.

Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge: MIT Press.

Roberts, Ian. 2012a. Phases, head movement and second-position effects. In Ángel J Gallego (ed.), *Phases developing the framework*, 385–440. Berlin; Boston: De Gruyter Mouton

Roberts, Ian. 2012b. Diachrony and Cartography: Paths of Grammaticalization and the Clausal Hierarchy. In Laura Brugé, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro & Cecilia Poletto (eds.), *Functional Heads: The Cartography of Syntactic Structures, Volume 7*, 351–367. Oxford University Press.

Roberts, Ian. 2017. Inertia. In Ledgeway, A., and I. Roberts, *The Cambridge Handbook of Historical Syntax*, Cambridge University Press, 425-445.

Roberts, Ian and Anna Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization.* Cambridge University Press.

Roberts, Ian, Anders Holmberg, Theresa Biberauer & Michelle Sheehan. 2010. *Parametric variation: null subjects in minimalist theory*. Cambridge: Cambridge University Press.

Roeper, Tom., and Jill De Villiers. 1989. Ordered parameters in the acquisition of wh-questions. Ms, University of Massachusets.

Roeper, Tom., and Jill De Villiers. 2011. The acquisition path for wh-questions. In J. De Villiers and T. Roeper (eds.), *Handbook of Generative Approaches to Language Acquisition*, Springer; 189-246.

Rohrbacher, Bernhard W.1999. *Morphology-Driven Syntax.*John Benjamins.

Rudin, Catherine. 1988. On multiple questions and multiple *wh*-fronting. *Natural Language and Linguistic Theory* 6:445-501.

Saito, Mamoru. 1989. Scrambling as semantically vacuous A'-movement. In Mark R. Baltin & Anthony S. Kroch (eds.), Alternative conceptions of phrase structure, 182–200. Chicago: University of Chicago Press.

Saito, Mamoru. 2012. Case checking/valuation in Japanese: Move, Agree, or Merge? *Nanzan Linguistics* 8:109-127.

Saito, Mamoru. 2014a. Case and labeling in a language without Φ-feature agreement. In Anna Cardinaletti, Guglielmo Cinque & Yoshio Endo (eds.), On peripheries: Exploring clause initial and clause final positions, 269–297. Tokyo: Hituzi Syobo Publishing.

Saito, Mamoru. 2014b. Selection and incorporation in complex predicate formation. In Audrey Li, Andrew Simpson & Wei-Tien Dylan Tsai (eds.), *Chinese syntax in a cross-linguistic perspective*, 251–269. New York: Oxford University Press.

Saito, Mamoru. 2016. (A) case for labeling: labeling in languages without φ-feature agreement. *The Linguistic Review* 33(1): 129-175.

Saito, Mamoru & Keiko Murasugi. 1990. N'-deletion in Japanese: A preliminary study. *Japanese/ Korean Linguistics* 1. 285–301.

Santorini, Beatrice. 1994. Some similarities and differences between Icelandic and Yiddish. In Adrian Battye and Ian Roberts, eds., *Clause structure and language change.* Oxford: Oxford University Press, 53-79.

Sapir, Edward. 1921. *Language.*New York: Harcourt Brace and Co.

Sato, Yosuke. 2011. Wh-Questions in Colloguial Singapore English: Syntactic evidence for the Malay Substrate hypothesis. lingbuzz/001234.(Available at: https://ling.auf.net/lingbuzz/001234; Accessed: May 2017.)

Schäfer, F. 2008. *The syntax of (anti-) causatives*. Amsterdam: John Benjamins.

Seely, T. Daniel.2006. Merge, derivational c-command, and subcategorization in a label-free syntax. In Cedric Boeckx (e.d), *Minimalist Essays,* 182-220. Amsterdam: John Benjamins.

Selkirk, Elisabeth. 1986. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Selkirk, Elisabeth. 2011. The Syntax-Phonology interface. In Goldsmith, J, Riggle, J, and A.C.L.YU(eds.) *The Handbook of Phonological Theory.*Blackwell; 435-484.

Sheehan, Michelle, Theresa Biberauer, Ian Roberts, and Anders Holmberg. 2017. *The Final-over-Final Condition: A syntactic Universal*. Cambridge: MIT Press.

Shimada, J. 2007. Head Movement, Binding Theory, and Phrase Structure. Ms, MIT.

Shinzato, Rumiko & Leon A. Serafim. 2013.*Synchrony and Diachrony of Okinawan Kakari Musubi in Comparative Perspective with Premodern Japanese.* Leiden: Brill.

Shlonsky, Ur. 2012. Notes on wh in situ in French. In *Functional heads: The Cartography of Syntactic Structures, Volume 7,* ed. by Laura Brugè, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro and Cecilia Poletto, 242-252. New York: Oxford University Press.

Sigurðsson, Halldór Ármann. 2010. On EPP effects. *Studia Linguistica* 64: 158–189.

Simons, Mandy. 2007. Observations on embedding verbs, evidentiality, and presupposition. *Lingua* 117: 1034–1056.

Sitaridou, Ioanna. 2012. A comparative study of word order in Old Romance. In Anne Carlier, Béatrice Lamiroy & Walter De Mulder (eds.). *Folia Linguistica* 46(2 "The pace of Grammaticalisation in Romance"). 1–51.

Spevak, Olga. 2010. *Constituent Order in Classical Latin Prose.* Amsterdan: Benjamins.

Speyer, Augustin. 2010. *Topicalization and Stress Clash Avoidance in the History of English*. Walter de Gruyter.

Sportiche, Dominique. 1998. *Partitions and Atoms of Clause Structure. Subject, agreement, case and clitics*. Routledge Leading Linguists 2. Routledge, London and New York.

Sportiche, Dominique. 2014. Assessing Unaccusativity and Reflexivity: Using Focus Alternatives to Decide What Gets Which θ-Role. *Linguistic Inquiry* 45(2): 305-321.

Stepanov, Arthur. 1998. On *wh*-fronting in Russian. In *Proceedings of the North Eastern Linguistic Society 28,* eds. Pius Tamanji and Kiyomi Kusumoto. 453–467. Amherst: GLSA.

Stepanov, Arthur. 2001. *Cyclic domains in syntactic theory*. Doctoral dissertation, University of Connecticut, Storrs.

Takahashi, Daiko. 1994. Minimality of movement. Doctoral dissertation, University of Connecticut, Storrs.

Takita, Kensuke, Nobu Goto & Yoshiyuki Shibata. 2016. Labeling through Spell-Out. *The Linguistic Review 33(1); 177-198.*

Thornton, Rosalind. 1990. *Adventures in Long-distance Moving: The Acquisition of Complex Wh-questions*. Doctoral dissertation, University of Connecticut.

Thráinsson, Höskuldur. 2003. Syntactic Variation, Historical Development and Minimalism. In R. Hendrick (ed.), *Handbook of Minimalist Syntax.*Blackwell, Oxford, 152-191.

Tonoike, S. 2009. "Minimarisuto Puroguramu [Minimalist Program]." In H. Nakajima, ed., *Gengogaku-no Ryooiki* [Subfields of Linguistics] 1, Asakura Shoten, Tokyo, 135–168.

Travis, Lisa de Mena. 1984. Parameters and effects of word order variation. Doctoral Dissertation, MIT.

Truckenbrodt, Hubert. 2006. On the semantic motivation of Syntactic Verb movement to C in German. *Theoretical Linguistics* 32-3: 257-306.

Uriagereka, Juan. 1988. *On government.* Doctoral dissertation, University of Connecticut, Storrs.

Uribe-Etxebarria, Myriam. 2002. In situ questions and masked movement. *Linguistic Variation Yearbook* 2. 259–303.

Van der Lely, Heather. 1998. SLI in children: movement, economy and deficits in the computational-syntactic system». *Language Acquisition* 7: 161-192.

Vance, Barbara. 1995. On the decline of verb movement to Comp in Old and Middle French. In Adrian Battye & Ian Roberts (eds.), *Clause structure and language change.*Oxford: Oxford University Press, 173–199.

Vance, Barbara. 1997. *Syntactic Change in Medieval French*. Dordrecht: Kluwer.

Vikner, S., 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford University Press, New York.

Villa-García, Julio. 2012. Recomplementation and locality of movement in Spanish. *Probus* 24(2). 257–314.

Villa-García, Julio. 2019. Recomplementation in English and Spanish: Delineating the CP space. *Glossa* 4(1):56; 1-44.

Vovin, A. 2009. *A descriptive and Comparative Grammar of Western Old Japanese. Part 2.* Folkestone, Kent: Global Oriental.

Wackernagel, Jacob. 1892. Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1: 333–436.

Waldmann, Christian. 2012. Moving in small steps towards verb second: A case study. *Nordic Journal of Linguistics* 34(3). 331–359.

Walkden, George. 2014. *Syntactic reconstruction and Proto-Germanic*. New York, NY: Oxford University Press.

Walkden, George. 2017. Language contact and V3 in Germanic varieties new and old. *Journal of Comparative Germanic Linguistics* 20: 49-81.

Wallenberg, Joel, C. 2016. Extraposition is disappearing. *Language* 92 (4): e237-e256.

Watanabe, Akira. 2002. Loss of Overt Wh-Movement in Old Japanese. In D. Lightfoot (ed.) *Syntactic Effects of Morphological Change,* 179-195. Oxford: Oxford University Press.

Watanabe, Akira. 2004. The genesis of Negative Concord: Syntax and morphology of negative doubling. *Linguistic Inquiry* 35: 559-612.

Warner, A. 1997. The structure of parametric change, and V movement in the history of English. In A. van Kemenade and N.Vincents (eds.) *Parameters of Morphosyntactic Change.* Cambridge: Cambridge University Press, 203-225.

Webelhuth, G. (1992). *Principles and parameters of syntactic saturarion.* Oxford: Oxford University Press.

Wechsler, Stephen. 1991. Verb second and illocutionary force. In Katherine Leffel and Denis Bouchard *(eds.), Views on Phrase Structure*. Dordrecht: Kluwer, 177–191.

Wells, G. 1985. *Language Development in Pre-school years.*Cambridge University Press.

West, Martin L.2011.*Old Avestan Syntax and Stylistics.* De Gruyter.

Westergaard, Marid. 2007. English as a mixed V2 grammar: synchronic word order inconsistencies from the perspective of first language acquisition. *Poznań Studies in Contemporary Linguistics* 43(2), 2007, pp. 107–131

Westergaard, Marit. 2009. *The Acquisition of Word Order. Micro-Cues, Information Structure and Economy*. Amsterdam/Philadelphia: John Benjamins.

Westergaard, Marit**,** Vangsnes, Øystein A. 2005. WH-questions, V2, and the Left Periphery of Three Norwegian Dialect Types. *Journal of Comparative Germanic Linguistics* 8 (1-2): 119-160.

Wexler, K. 1998. Very early parameter setting and the Unique Checking Constraint: A new explanation of the Optional Infinitive stage. *Lingua* 106:23-79.

Wexler, Kenneth and Peter Culicover. 1980. *Formal Principles of Language Acquisition*. Cambridge, Mass.: MIT Press.

Wiklund, Anna-Lena.2010. In Search of the Force of Dependent Verb Second. *Nordic Journal of Linguistics 33*; 81-91.

Wiklund, Anna-Lena, Kristine Bentzen, Gunnar Hrafn Hrafnbjargarson, and Þorbjörg Hróarsdóttir. 2009. On the distribution and illocution of V2 in Scandinavian that-clauses. Lingua 119: 1914–1938.

Whitman, John. 1997. Kakarimusubi from a Comparative Perspective. In Ho-min Sohn and John Haig (eds.) *Japanese/Korean Linguistics* 6: 161-178.

Whitman, John. 2001. Kayne 1994: p.143, fn.3. In Galina Alexandrova and Olga Arnaudova (eds.), *The Minimalist Parameter: Selected Papers from the Open Linguistics Forum*, Ottawa, 12-23 March 1997, 77-100. Amsterdam: John Benjamins.

Willis, David. 1998. *Syntactic change in Welsh: a study of the loss of the verb-second*. Oxford: Oxford University Press.

Willis, David. 2007. Specifier-to-Head Reanalyses in the Complementizer domain: evidence from Welsh. *Transactions of the Philological Society* 105 (3):432–480.

Willis, David, Christopher Lucas, and Anne Breitbarth (eds.). 2013. *The History of Negation in the Languages of Europe and the Mediterranean. Vol.1*. Oxford: Oxford University Press.

Wolfe, Sam. 2016. On the Left Periphery of V2 Languages. *Rivista di Grammatica Generativa: Selected Papers from the 41st Incontro di Grammatica Generativa*(38). 287–310.

Wolfe, Sam. 2019. *Verb Second in Medieval Romance.* Oxford: Oxford University Press.

Woolford, Ellen, B. 1979. *Aspects of Tok Pisin Grammar.Pacific Linguistics* (B)66.

Wood, Jim. 2015. *Icelandic Morphosyntax and Argument Structure*. Dordrecht: Springer.

Wurmbrand, Susi. 2014. The Merge Condition. In *Minimalism and Beyond*, ed. by Peter Kosta, Steven L. Franks, Teodora Radeva-Bork, and Lilia Schürcks, 139–177. Amsterdam: John Benjamins.

Wurff, W. Van Der. 1997. Deriving object-verb order in Late Middle English. *Journal of Linguistics* 33:485-509.

Wurff, W. Van Der. 1999. Objects and Verbs in Modern Icelandic and Fifteenth-century English: A word order parallel and its causes. *Lingua* 109: 237-65.

Yanagida, Yuko. 2005. *The syntax of FOCUS and WH-questions in Japanese: A cross-linguistic perspective*. Tokyo: Hituzi Syoboo.

Yanagida, Yuko. 2006. Word order and clause structure in Early OJ. *Journal of East Asian Linguistics* 15. 37–67.

Yanagida, Yuko. 2012. The syntactic reconstruction of alignment and word order: The case of Old Japanese. Paper presented at the 19th International Conference on Historical Linguistics held at Radboud University, Nijmegen, on August 14, 2009

Yanagida, Yuko. 2017. Genitive/active to nominative case in Japanese: the role of complex experiencer constructions. Talk at ICHL 23 Workshop: Alignment Typology in Diachronic Perspective, The University of Texas at San Antonio, on August 4, 2017.

Yanagida, Yuko, and John Whitman. 2009. Alignment and word order in Old Japanese. *Journal of East Asian Linguistics* 18:101-144.

Yang, Charles. 2002. *Knowledge and Learning in Natural Language.* Oxford: OUP.

Yang, C. 2002. Grammar competition and language change, in *Syntactic Effects on Morphological Change*, ed D. Lightfoot (Oxford: Oxford University Press), 367–380.

Yip, Virginia and S. Matthews. 2007. *The Bilingual Child.*Cambridge University Press.

Zaliznjak, A. A. (2008). *Drevnerusskie Ènklitiki*. Moskva: Jazyki slavjanskoj kul'tury.

Zimmermann, Michael. 2014. *Expletive and Referential Subject Pronouns in Medieval French*. (Linguistiche Arbeiten 556). Berlin ; New York: Mouton de Gruyter.

Zocca DeRoma, Cynthia L. 2011. *Divide et Impera- Separating Operators from their Variables.* Doctoral dissertation, University of Connecticut.

Zombolou, K. and A. Alexiadou .2012. Early Voice Acquisition in German. Paper presented at the Workshop "Acquisition at the Interface", University of Tromsø, Norway, 22-23 October.

Zuckerman, Shalom.2001. *The acquisition of optional movement*. Groningen Dissertations in Linguistics 34

Zwart, C. Jan Wouter. 1993. *Dutch Syntax: A minimalist approach*. Doctoral dissertation, Groningen.

Zyman, Erik. 2018. On the Driving Force for Syntactic Movement. Doctoral dissertation, University of California, Santa Cruz.

Primary Sources

Polish: *Biblia Królowej Zofii (Queen's Sophia's Bible).*

(Available at: https://ijp.pan.pl/publikacje-i-materialy/zasoby/korpus-tekstow-staropolskich/)

English: *Corpus of Contemporary American English*

Latin: *Opera Latina* (Laboratoire d'Analyse Statistique des Langues Anciennes (LASLA); Centre Informatique de Philosophie & Lettres (CIPL)).