

CYCLES OF AGREEMENT: ROMANCE CLITICS IN DIACHRONY

BY

MATTHEW LEROY MADDOX

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Spanish
with concentrations in Romance Linguistics and Medieval Studies
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor José Ignacio Hualde, Chair
Associate Professor Jonathan E. MacDonald, Director of Research
Professor James Yoon
Professor Martin Joseph Camargo
Professor Elly van Gelderen, Arizona State University

ABSTRACT

In this dissertation I investigate two linguistic cycles in Romance per van Gelderen's (2011) framework: the Subject Agreement Cycle (SAC) and the Object Agreement Cycle (OAC). These grammaticalization cycles turn pronouns into agreement morphology on the verb. Both cycles are comprised of three stages. At stage (a) the pronoun is a full DP. At stage (b) the pronoun is reanalyzed as a D-head and at stage (c) it is reanalyzed as a T-head in the SAC or a *v*-head in the OAC. I extend the SAC to account for the grammaticalization of impersonal pronouns. I show that Modern French *on* is at stage (c) of the Impersonal Subject Cycle (ISC). Old Spanish (OldS) *omne* was on this cycle but it disappeared only to be replaced by Modern Spanish (ModS) *uno*. I propose that the reason for this disappearance was that impersonal subject pronouns will only be reanalyzed if personal subject pronouns are being reanalyzed first via the SAC.

I also build upon van Gelderen (2011) by examining the OAC in Spanish in more detail. I show that in OldS, object clitics were full DPs and thus OldS was at stage (a) of the OAC. Based on diagnostics of coordination, modification, and movement, ModS object clitics are more deficient than OldS object clitics. Patterns of clitic doubling are evidence that standard ModS is at stage (b) of the OAC while Rioplatense Spanish is at stage (c). I adopt an analysis of accusative clitic doubling (ACD) based on Harizanov (2014) and Kramer (2014) whereby the object merges and moves to Spec,*v* (object shift) as a DP. I also show how object movement feeds the OAC. Object movement results in the object pronoun being in Spec,*v*, where it is associated with the *v*-head. Since ACD depends upon object shift to Spec,*v*, we expect

languages that have developed ACD to have had object movement at an earlier period. This is the case for Romance.

I show that the reflexive clitic *se* in ModS has been subjected to a type of OAC which I label the “Reflexive Object Cycle” (ROC). The ROC is a grammaticalization cycle that takes a reflexive object pronoun and turns it into a valency-marking morpheme, a Voice or *v*-head. I present evidence that in Latin and OldS, the reflexive pronoun was a full DP. It was later reanalyzed as a D-head and then a Voice head. This reanalysis is supported by diagnostics of interpolation, modification, coordination, and doubling.

I demonstrate that null subjects and null objects relate to the stages of the SAC and the OAC. Null arguments are allowed in a language only if that language has reached stage (c) of the relevant cycle. I extend a D-feature and topic-identification type of analysis based on Holmberg (2005, 2010) and Holmberg et al (2009) to the licensing of null objects; i.e., null objects are licensed by a D-feature in *v*. I argue that this D-feature is only present on *v* in some varieties of Spanish because the clitic’s D-feature has been reanalyzed as a feature of *v*, which is a result of how the OAC works. I show how clitic left-dislocation and ACD are tied to the stages of the OAC, which accounts for their cross-linguistic distribution. I propose a typology of null object languages based on Holmberg’s (2005, 2010) typology of null subject languages. As for null subjects, I illustrate how the SAC and the ROC have interacted in the history of Spanish to give rise to passive *se* (Pass_{se}). In order to develop passive *se*, two elements are need: null subjects and a reflexive Voice head. These elements are present due to the SAC and the ROC, respectively. This also accounts for the presence or absence of passive reflexive constructions crosslinguistically. Pass_{se} develops when a language has subject *pro* and *se* as a Voice head.

Acknowledgements

I first travelled to Urbana-Champaign in April 2012 to visit with faculty and find an apartment. I am now putting the finishing touches on this dissertation while living in Lincoln, Nebraska, and it is currently April 2019. How can I describe the seven years in between? Intellectually stimulating, sure. Personally revelatory, of course. However, the adjective that seems appropriate as this phase of my life draws to a close is: busy. Now that I have successfully defended, I actually have some time to write a few words of gratitude to those people who have helped me along the way.

First, I am grateful to the audiences at the conferences and other venues where this material has been presented. I thank specifically the organizers and participants at LSRL 45 in Campinas, Brazil, and LSRL 46 at Stony Brook University. I also thank the audiences at HLS from 2015-2017. I received very helpful feedback during the Worskhop on Romance Se-Si at the University of Wisconsin-Madison in 2016. I am grateful to Francisco Ordóñez, Paula Kempchinsky, Grant Armstrong, Anne Wolfsgruber, Javier Ormazabal, Juan Romero, and Alexandru Nicolae. I also thank the audiences at ICHL 23 at the University of Texas at San Antonio and DiGS 18 at Ghent University.

I thank the members of my dissertation committee: José Ignacio Hualde, James Yoon, and Martin Camargo. I am grateful to Elly van Gelderen for participating as an external member. Her insight was crucial since this research has been carried out in the grammaticalization framework she formulated. I also thank my advisor Jonathan E. MacDonald. His support and collaboration was critical during the dissertation. Also, if it hadn't been for one fateful conversation he and I had on the Gold line bus early in the Fall semester of 2012, I doubt I would have chosen to write roughly 250 pages about theoretical syntax.

I have received a lot of support from colleagues at my former institution as well. At the University of Nebraska-Lincoln my thanks goes to Errapel Mejías-Vicandi, Catherine Johnson, and Jordan Stump. And of course I must thank Isabel Velázquez who was not visibly disappointed (I think) when I told her I was switching from sociolinguistics to diachronic syntax.

I am also grateful to my parents and other family members for their encouragement over the years. Finally, I thank my partner Sean for his love and support. Not only was he patient during my three-and-a-half year residence in Champaign, but he was also a major source of motivation during the writing process. And, naturally, I have to thank my writing buddies, Bowie and Kihei (the dogs). They were present on the chair in my office for almost every word I wrote. Without their “help” I would likely have finished the first draft much more quickly.

The critics who

Can 't break you

They somehow help to make you.

The critics who

Can 't break you

Unwittingly they make you.

(Morrissey)

TABLE OF CONTENTS

Chapter 1: Introduction.....	1
Chapter 2: Impersonal Pronouns and the Subject Agreement Cycle.....	32
Chapter 3: Object Agreement and Object Movement in Spanish.....	69
Chapter 4: The Reflexive Object Cycle from Latin to Spanish.....	126
Chapter 5: Agreement Cycles and the Licensing of Null Arguments.....	158
Chapter 6: Conclusion.....	201
References.....	205

Chapter 1: Introduction

Preliminary remarks

In this dissertation my primary goal is to illustrate how van Gelderen's (2011) formal theory of linguistic cycles provides a principled explanation of clitic-related phenomena in the Romance language family, focusing on Spanish and French pronominal clitics. In doing so I show how these agreement cycles interact with other parts of morphology, syntax, and information structure. As a first step toward this goal, it necessary to start with some preliminary discussion. This introductory chapter is formatted as follows. In Section 1.1, I briefly review the central questions addressed in previous literature on clitics and discuss why clitics should be of interest to theories of language change. Next, in Section 1.2, I present some of the diachronic work that has been carried out from a generative perspective. I also present van Gelderen's (2011) theory of linguistic cycles which is the framework that I adopt throughout this dissertation. Section 1.3 is an overview of the two specific cycles that will be relevant for what follows: the Subject Agreement Cycle and the Object Agreement Cycle. The format of the dissertation as a whole is summarized in Section 1.4.

1.1 Overview of clitics

1.1.1 *Definition and typology*

Clitics are unstressed linguistic elements, often pronominal, that share properties with full words and inflection. In a sense, clitics are somewhere in between syntax and morphophonology. As an example, consider the English pronoun *him*. This pronoun can appear in reduced or contracted form <'em> in a phrase like the following: "we're gonna get'em!"¹ This

¹ Actually there may be more than just one clitic element in this example, given the presence of two other contracted forms: *'re* for *are* and *gonna* for *going to*.

type of reduction is typical of clitics and the fact that in that example the contracted form is attached to the verb *get* suggests that '*em*' is not a full "word" in the conventional sense. Formulating an adequate definition of clitics is just one of the many issues that have fascinated linguists working in this area. Clitics first start receiving serious scientific attention in the nineteenth century with the philological work of Wackernagel (1892), Tobler (1875), and Mussafia (1886). Wackernagel's contribution was his observation that certain types of clitics always occur after the second constituent in a sentence in Indo-European languages like Ancient Greek. Tobler and Mussafia made a related observation for Romance clitics, which in the medieval period could not appear sentence initially but rather attached to whatever the first constituent was. Both studies share in common their claim that the placement of clitics was regulated systematically and not by chance.

More recent work has attempted to determine whether there is a typology of clitics. Kayne's (1975) study of clitics in French, for example, is the earliest attempt in the generative tradition to propose diagnostics that formally differentiate these elements from regular pronouns. According to Kayne, French clitics appear in special positions in comparison with full syntactic phrases like Determiner Phrases (DPs). For example, they cannot be stressed, they cannot be separated from the verb except by other clitics, and they cannot be modified or coordinated. Spanish clitics follow the same patterns as shown in the Spanish adaptation of Kayne's data in (1) below.

(1) a. María los conoce.

María them knows

‘María knows them.’

b. María los va a conocer. / María va a conocerlos.

María them goes to know María goes to know-them

‘María is going to meet them.’

In (1a), the direct object clitic *los* precedes a finite verb or may attach to a nonfinite form, in which case it forms a single orthographic word, *conocerlos*, as in (1b). When *los* precedes the verb it also forms a single prosodic word, but this is not reflected in the written language. When a clitic precedes its host it is "proclitic" and when it follows its host it is "enclitic."² Now consider how full lexical DPs pattern in (2) below.

(2) a. María *(a tus amigos) conoce a tus amigos.

María DOM³ your friends knows DOM your friends

‘María knows your friends.’

b. María *(a tus amigos) va a conocer a tus amigos.

María DOM your friends goes to know DOM your friends

‘María is going to meet your friends.’

In (2a), however, the lexical DP *a tus amigos* cannot precede the finite verb and when it follows the verb it does not form a single word (2b). Another distinction is that full DPs like *María* (3a) or *ella* (3b) can be stressed (3a) while clitic pronouns like *la* cannot (3c).

² There is also the possibility of incorporating into its host, "mesoclisis," which I do not discuss here.

³ DOM indicates the differential object marker *a*.

(3) a. Vi a MARÍA ayer.
 I-saw DOM María yesterday
 ‘I saw MARY yesterday.’

b. La_i vi a ELLA_i ayer.
 her I-saw DOM her yesterday
 ‘I saw HER yesterday.’

c. *LA vi ayer.
 her I-saw yesterday
 Intended: ‘I saw HER yesterday.’

Also, the clitic must be adjacent to the verb (4a). It cannot be separated from the verb by negation (4b) but full DPs can (4c).

(4) a. No la comió.
 not it he-ate
 ‘He didn’t eat it.’

b. *La no comió.
 c. El caballo no comió la manzana.
 the horse not ate the apple
 ‘The horse did not eat the apple.’

Finally, full DPs can be coordinated (5a) while clitic pronouns cannot (5b).

(5) a. Tú y yo somos amigos.
 you and I are friends
 ‘You and I are good friends.’

b. *Te y me vieron ayer.

you and me they-saw yesterday

Intended: ‘They saw you and me yesterday.’

The work of Zwicky (1977) and Zwicky & Pullum (1983), focuses on distinguishing between clitics and affixes. They propose six criteria: 1) clitics are much less selective with respect to their host while affixes are highly selective, 2) clitics are less likely to have “arbitrary gaps” in their set of combinations while affixes often have these gaps, 3) affixes tend to be morphophonologically idiosyncratic while clitics do not, 4) affixes tend to be semantically idiosyncratic while clitics do not, 5) affixes can be affected by syntactic rules while clitics cannot, and 6) clitics can attach to a host that already has other clitics on it while affixes cannot. Out of their study comes a proposed typology of clitics according to which there are two classes of clitics: “simple” and “special.” Simple clitics are weakened versions of a full form that have the same syntactic distribution as the full counterpart. Special clitics do not have a full counterpart and display “special” morphosyntactic characteristics. Anderson (2005) adopts this bipartite scheme but relabels simple clitics as “phonological” clitics and special clitics as “morphosyntactic” clitics.

A more recent typology of clitics has been proposed by Cardinaletti & Starke (1999). They expand the traditional two-way division and propose a tripartite scheme according to which there are strong pronouns, and two types of deficient pronouns, weak and clitic. Strong pronouns can appear in peripheral positions, can be modified and coordinated. The main difference between clitic and weak pronouns is that the former are heads, subject to head

movement like verbal adjunction, while the latter are phrases, so they can occur in the same position as other types of maximal projection phrases.⁴

1.1.2 *Clitic placement*

A major question that has been of interest to researchers of clitics is: what principles regulate the positioning of clitics within the clause and with respect to the verb? Clitics in European Portuguese, for example, can be proclitic (6a) or enclitic (6b).⁵

(6) a. Só o Pedro o viu.

Only the Pedro him saw

“Only Pedro saw him.”

b. Viu o só o Pedro.

saw him only the Pedro

Enclisis with finite verbs has been lost in Modern Spanish (ModS), but it was active in Old Spanish (OldS) as in the data in (7) from Fontana (1993).

(7) Respondió les el que lo no farie.

answered them he that it not would-do

“He answered them that he would not do it.”

In French, clitics can precede non-finite verbs (8a) while in ModS clitics can never be immediately left-adjacent to an infinitive (8b), though this was possible in OldS (8c).

⁴ Recent studies such as Roberts (2010), building on work by Muysken (1982), Chomsky (1995), and Belletti (1999), have proposed that clitics are both heads and phrases simultaneously.

⁵ Data from Barbosa (1996)

(8) a. Elle va le voir. (Modern French)
 she goes him see
 “She is going to see him”

b. *Ella quiere lo ver. (Modern Spanish)
 she wants him see
 Intended: 'She wants to see him'

c. a lo menos devian lo ver...⁶ (Old Spanish)
 at the least they-should him see
 “At least they should see him...”

While this description may suggest that clitics move rather unrestrictedly, this is not the case.

Clitics are much more constrained in how they can move when compared to other constituents like DP arguments and PP adjuncts. This has led to the development of three major types of accounts attempting to deal with clitic placement: 1) the movement account, 2) the base-generation approach, and 3) a mixed or hybrid of the two. In order to provide context for the research proposed here, each of the three types of accounts is briefly summarized below.⁷

The movement account assumes that clitics are merged in argument position to receive a theta-role and then move to a different structural position, often for reasons having to do with Case assignment in the Government & Binding framework or feature checking in the Minimalist Program. Kayne (1975, 1989, 1991) analyzes clitics as left-adjoining to a functional head. He explains the different behavior of clitic placement with respect to infinitives as being the result of distinct types of verbal movement rather than clitic movement. For example, in French, where

⁶ Anonymous, *Petición de traslado [Libro del Concejo y documentos del Archivo Municipal de Castro Urdiales]*, para. 85; 1483.

⁷ This summary is based in part on the overview given in González (2008).

the order [clitic + Infinitive] occurs, the verb moves out of VP to adjoin to an Infn (infinitive) functional head. The clitic then left adjoins to the [V + Infn] complex. In Italian, however, which displays [Infinitive + clitic] order, the verb moves to Infn, as in French, but then the [V + Infn] complex moves again to adjoin to T' while the clitic left-adjoins to T. A different movement analysis is given by Belletti (1999, 2001). Under this approach the clitics are merged to receive a theta-role and then move to an Agr projection. Thus, subject clitics move to AgrS while object clitics move to AgrO.

Alternatively, base-generation approaches such as Borer (1984) take clitics to be base-generated in their verbally adjacent position and to be coindexed with an XP in argument position, which may be null. One of the explanations for the clitic's location is that it binds its coindexed argument from that position, as in Aoun & Sportiche (1981). Suñer (1988) builds on earlier analyses by further proposing that clitics are agreement inflection base-generated on the verb. Thus, on her analysis, object clitics in Spanish are object agreement morphemes. Suñer focuses on clitic doubling structures wherein the object clitic and its argument cooccur in the same clause. This establishes an important precedent; namely, the correlation of clitics to agreement inflection. In fact, this relation is an integral part of the historical analysis that I develop throughout this dissertation; i.e., clitics becoming agreement.

The mixed or hybrid approach to clitic placement incorporates aspects of both the base-generation and movement analyses and is best represented in the work of Uriagereka (1995) and Sportiche (1996). According to Uriagereka, clitics encode referentiality and license *pro* in a functional projection F. Clitics move to F as part of the licensing of *pro*. The role of F is to assign reference to speakers in the discourse. Cross-linguistically F may be “morphological,” in which case clitics move there overtly, or it may only be active at LF, in which case clitics do not

move to F but rather affix to AgrO. Uriagereka further proposes that there are two types of special clitics: weak determiner clitics and strong phrasal clitics. Determiner clitics head DP and in clitic doubling constructions their coreferential constituent is in the specifier position of that DP. Strong clitics, on the other hand, are not D heads and do not permit doubling. They can behave either as heads or phrases.

Sportiche (1996) takes clitics to be the heads of inflectional Voice projections that are in a special relationship with an associated XP. This XP moves to specifier position in the Clitic Voice phrase since the clitic head licenses special properties on that XP; i.e., agreement. In order to account further for what triggers clitic movement, Sportiche proposes a Clitic Criterion and a Generalized Licensing Criterion, according to which functional heads, which includes clitics on his analysis, must be in a spec-head relationship with an XP at LF.

1.1.3 *Clitics and language change*

A third question of interest to clitic research is a diachronic one, and it is formulated by Salvesen (2013:12) as “where do clitics come from?” In other words, how do clitics come to be clitics, given that we can trace their historical sources back to what were once independent, strong pronouns? The bulk of the work in this area has been done from the functionalist perspective (Tauli 1956, Givón 1976, Mithun 1991, *inter alios*) and has been couched within the framework of “grammaticalization.” Grammaticalization is the historical process that takes lexical material and turns it into functional elements; i.e., demonstratives becoming complementizers, pronouns becoming copulas, etc. Examples of theories of grammaticalization within the generative perspective include Roberts & Roussou (1999), Simpson & Wu (2002), Tse (2016). Extensive work by van Gelderen (2004ab, 2007, 2008, 2011, 2016) is focused on the grammaticalization of subject agreement affixes from subject pronouns and clitics

specifically. She also proposes a Minimalist model of language change which is discussed in detail in Section 1.3 below.

Clitics have gained attention within historical linguistics because they are elements that apparently interact with all levels of language. Phonologists, for example, are interested in determining how a prosodically independent word comes to be destressed. Morphologists, on the other hand, might study how a lexical item is reanalyzed as inflection. Since clitics have this ability to interact with different parts of language, investigation into their historical development may inform what we know about how morphology and phonology interact diachronically, for example. Thus, by studying clitics linguists can learn more about how language change takes place generally.

1.2 Language change and grammaticalization in Generative Theory

Grammaticalization is the diachronic process by which lexical items are reanalyzed as functional elements, or functional elements are reanalyzed as different functional items. This process has been the subject of extensive study among functional linguists (Traugott & Heine 1991, Hopper & Traugott 2003, *inter alios*). Within generative linguistics grammaticalization and language change generally has received less attention, though this has changed in recent years. In what follows I present two different generative approaches to serve as a point of comparison with van Gelderen's (2011) model, which I adopt and present in Section 1.2.2.

1.2.1 Lightfoot (1979), Roberts & Roussou (1999)

While not the first generativist work of diachronic syntax, Lightfoot's (1979) study is one of the earliest and most influential.⁸ Lightfoot's model of language change is developed within Extended Standard Theory (Chomsky 1973). Lightfoot's study focuses primarily on changes in

⁸ Lightfoot (1979:21) provides a thorough review of earlier generative historical work such as Klima (1964) and Kiparsky (1968).

the verbal paradigm in English, but he also discusses changes affecting verbal mood in Greek, and the development of serial verbs in Kwa.⁹ He proposes a "Transparency Principle" that motivates language change in general and reanalysis in particular. This principle dictates that initial linguistic structures should be derivationally close to their surface structure. In other words, once a certain part of the grammar becomes too complex "therapeutic restructuring" will take place. One of Lightfoot's goals is to identify the point of maximum tolerance a grammar can reach before the restructuring is forced.

In his analysis of the development of modal verbs from fully lexical verbs in the history of English, Lightfoot illustrates his Transparency Principle at work.¹⁰ He describes a cluster of changes that took place affecting these verbs. For example, they ceased to have infinitive forms and they no longer took direct objects. Additionally, when V-to-T movement was lost, the verbs that would become modals continued to follow the older patterns of negation and subject-verb inversion. Eventually the modals-to-be were reanalyzed as T, while lexical verbs merged lower as V and did not move up. This "category reanalysis" was forced due to the increased "opacity" displayed by modals. In other words, they developed a series of "exception features" that made it so that learners no longer analyzed them as lexical verbs. The exception features involved relate to their morphology, syntax, and semantics, and the reader is referred to Lightfoot (1979) for the details. In sum, once the level of opacity passed a certain threshold, the Transparency Principle took effect and the verbs in question were reanalyzed as modals (T) rather than lexical verbs (V).

A significant early attempt at formalizing grammaticalization within the Minimalist Program is Roberts & Roussou (1999). These authors modify Chomsky's (1995) checking

⁹ Kwa is a group of Niger-Congo languages spoken in parts of the Ivory Coast, Ghana, and Togo.

¹⁰ See also Roberts (2007) for a discussion of Lightfoot's (1979) analysis.

theory by assuming that the same set of formal features are available to all languages, but there is variation as to which ones are expressed phonetically. In their system, a phonetically overt functional head is designated F^* . An F^* element can be realized at PF by move, merge, or both and economy conditions dictate which will apply. Roberts & Roussou (R&R) formalize parametric variation as in (9) below:

(9) a. F^* ? Yes/NO
b. If F^* , is * satisfied by move or merge?

Thus, variation is reduced to whether a functional item is pronounced and whether its PF realization is accomplished via move or merge.

Adapting this to language change, R&R propose that whether an item is phonetically overt varies diachronically. They follow Lightfoot (1979) in assuming that the language acquirer is naturally conservative and thus extra complexity, as in syntactic movement, is the marked option. The least marked option is for a functional item to be phonetically null. If in the Primary Linguistic Data (PLD) that the learner is exposed to a given functional head is not overt, they can reanalyze it as involving merge or move, but the null option is the preferred one since it is the simplest, involving the least amount of operations. Within this system, grammaticalization happens when a functional head is reanalyzed by learners as involving merge instead of move. Learners may also reanalyze F as being null, but since this is not reanalysis, R&R do not consider it a case of grammaticalization.

R&R also provide an analysis for the development of modals in English, building on Lightfoot (1979). At an early stage, when a modal-to-be cooccurred with another verb, the structure was biclausal and the modal-to-be raised to T since V-to-T movement was still available. V-to-T movement was lost because, as discussed above, learners prefer the most

economical option which is merge rather than move. As a consequence of this loss there was no evidence in the PLD for a biclausal structure with modals-to-be and thus they were reanalyzed as T. As for the causation of this change, R&R attribute it to the loss of infinitive morphology ending in *-en*. This suffix was evidence to learners for two T projections. Once it was lost, reanalysis took place and there was now a clear class of fully modal verbs in English.

1.2.2 *Van Gelderen's (2011) model of linguistic cycles*

1.2.2.1 Economy Principles

I adopt van Gelderen's (2011) model of grammaticalization which is couched within the framework of the Minimalist Program (Chomsky 1995). Van Gelderen argues that grammaticalization is cyclical and is the result of three economy principles: the Head Preference Principle, the Late Merge Principle, and the Feature Economy Principle. Each of these principles is given below according to their formulation in van Gelderen (2011).

(10) Head Preference Principle (HPP)

Be a head rather than a phrase.

(11) Late Merge Principle (LMP)¹¹

Merge as late as possible.

(12) Feature Economy Principle (FEP)

Minimize the semantic and interpretable features in the derivation, for example:

Adjunct	Specifier	Head	Affix
Semantic feature	> interpretable	> uninterpretable	> uninterpretable

The HPP in (10) motivates the reanalysis of phrases into heads while the LMP in (11) drives grammaticalization of a lexical head to a functional head or a functional head to a higher

¹¹ This principle originated in Chomsky (1995).

functional head. The FEP in (12) is what causes renewal. Semantic features are originally interpretable, but when the element being grammaticalized is reanalyzed as a head, those features become uninterpretable. Thus, in order for feature checking to take place a new element must merge that has interpretable features to contribute to the derivation. The change from interpretable to uninterpretable can also be accompanied by loss of specific features. For example, an item might lose a Case feature but maintain φ -features when reanalyzed by learners. The intuition behind these principles is related to language acquisition. Based on previous studies (Bloom 1970, 1973; Diessel 2004; *inter alios*), van Gelderen argues that children exhibit a preference for heads over phrases during the acquisition process. Thus, when possible they will reanalyze phrases as heads.

Additionally, grammaticalization is viewed as cyclical because after an item is reanalyzed as a result of the aforementioned principles, "renewal" may take place, subsequently restarting the cycle. As a concrete example, take the well-known case of Jespersen's (1917) Negative Cycle. This is the cycle whereby a language expresses negation preverbally at an early stage, later pre- and post-verbally (negative concord), and finally post-verbally, as shown in the changes in the history of French in (13), adapted from Roberts & Roussou (1999:1029).

(13) a. **jeo ne dis** (Old French)
b. **je ne dis pas** ("Standard" Modern French)
c. **je dis pas** (Colloquial Modern French)

This change is cyclic because eventually a new item will merge that "renews" the cycle. In the case of the negative cycle, van Gelderen (2011:292) claims that indefinite phrases and certain types of verbs tend to be reanalyzed as negative heads due to the HPP.

1.2.2.2 Diagnostics

A significant part of van Gelderen's model relies on distinguishing between heads and phrases, given the principles she has adopted. To that end, she proposes several tests to help tease apart phrases from heads, based in part on earlier studies such as Zwicky & Pullum (1983). It should be noted that van Gelderen avoids the label "clitic" because it is descriptively ambiguous. On her analysis, a clitic may exhibit head-like or phrasal properties depending on what stage of the cycle it is at. The tests she adopts to diagnose head or phrase status are modification and coordination. Full XP phrases can undergo both, while heads can undergo neither. An argument is an XP while agreement is a head. This is illustrated in the Malagasy data below:¹²

(14) a. Hita-**ny** tany an-tokotany i-Koto (-**ny** = Agreement)

see-3 there DET-garden Koto

'She/he/they saw Koto in the garden.'

b. Hitan' **izy** sy ny zaza tany an-tokotany i-Koto. (**izy** = Argument)

see 3S and the child there DET-garden Koto

'She/he and the child saw Koto in the garden.'

In (14a) the *-ny* suffix is agreement. In (14b) there is coordination with *ny zaza* and so the full argument pronoun *izy* appears. Following Mithun (2003), van Gelderen also assumes that agreement has a fixed-position while arguments may be subject to movement. However, the distinction between fixed-position and movement are not necessarily binary, which is why theta-role and φ -features are also to be taken in consideration. What looks like a full pronoun can be an XP or an X, but if it still has interpretable φ -features it is indeed a pronoun that receives a

¹² These data are from Pearson (2001:43) cited in van Gelderen (2011:39).

theta-role. Additionally, agreement is always a head but may still have interpretable φ -features or it may have uninterpretable φ -features, as in polysynthetic languages. This is summarized in Table 1 below, which is an adaption of van Gelderen's (2011:40) Table 2.1.

	Theta-role	XP or X	Fixed Position	$i\varphi$	Language
Full pronoun	Yes	XP	No	Yes	Hindi/Urdu, Japanese
Head pro-Noun	Yes	X	No	Yes	French, (English)
Polysynthetic Agreement	Yes	X	Yes	Yes	Navajo, Old English
Agreement	No	X	Yes	No	Hind/Urdu, etc.

Table 1. Pronouns vs. Agreement.

1.3 Agreement cycles and clitics

1.3.1 *Subject Agreement Cycle*

The first cycle presented in van Gelderen (2011) is the Subject Agreement Cycle (SAC). This is the cycle that takes subject pronouns and turns them into subject agreement morphology on the verb. As an example of how the SAC works within van Gelderen's model, let us start by examining subject pronouns in the history of French.

1.3.1.1 Historical developments from Latin to Modern Colloquial French

That French subject pronouns are becoming prefixed subject agreement morphemes has been studied extensively and in Modern Colloquial French (MCF) this cycle is nearly complete (Lambrecht 1981, van Gelderen 2011, Bahtchevanova & van Gelderen 2016, *inter alios*). There are a couple of developments that took place in the history of French that illustrate the cycle's progress. It began when subject agreement affixes that were inherited from Latin underwent morphophonological attrition. There is no consensus as to when the distinct endings ceased to be realized phonetically in French, but Foulet (1936) points to before the thirteenth-century while

others like Franzen (1939) and Herman (1954) claim this happened after the thirteenth-century.¹³

The attrition from Latin to "Standard" Modern French (ModF) is summarized in (15):

(15)	Latin	Old French	Modern French¹⁴
	PORTŌ	port – [pɔrt]	porte – [pɔrt]
	PORTĀS	portes – [pɔrtəs]	portes – [pɔrt]
	PORTĀT	porte – [pɔrtə]	porte – [pɔrt]
	PORTĀMUS	portons – [pɔrtɔnz]	portons – [pɔrtɔ]
	PORTĀTIS	portez – [pɔrtets]	portez – [pɔrte]
	PORTĀNT	portent – [pɔrtənt]	portent – [pɔrt]

In (15), the personal endings that were fully realized in Latin start to lose some phonological features in Old French (OldF). For example, the third-singular final /t/ in PORTĀT is dropped and in OldF it is pronounced [pɔrtə]. Nevertheless, even in the OldF period, each personal ending was distinct so that first-singular and second-singular endings sounded different. In ModF, however, there are only two endings with distinct phonetic realizations: first- and second-person plural. This is relevant to the SAC because it is only when agreement affixes cease to unambiguously pick out the subject that an overt DP needs to be introduced in order to disambiguate. In other words, once verbal morphology is no longer enough to identify the subject, a topic subject DP will merge which renews the cycle; i.e., a subject pronoun.

An additional change that took place in the transition from OldF to ModF is that subject pronouns began to show different distributional patterns with respect to interpolation, elision, and doubling.¹⁵ In OldF, subject pronouns were independent DPs. They were used for emphasis and could be separated from the verb by other DPs; i.e., interpolation.¹⁶

¹³ See also Harris (1978), Nyrop (1904), Revol (2005), Wartburg (1963) for alternative dates.

¹⁴ Adapted from Alkire & Rosen (2010)

¹⁵ Data in this section taken from Kaiser (1992) and van Gelderen (2011, 2016). I avoid discussion of third-person subject clitics since they are lagging behind in the cycle.

¹⁶ Similar data can be found in Gardner & Greene (1958). Adams (1987:41) also claims that in OldF pronouns were "full pronouns, not clitics."

(16) Si con tu meismes le preves.¹⁷
 if when you self it prove
 'If you prove it yourself.'

(17) Se je meïsme ne li di.¹⁸
 if I self not him tell
 'If I don't tell him myself.'

In MCF, interpolation of subject pronouns is not allowed:

(18) *si tu même le prouves

(19) *si je même ne lui dis

Additionally, in OldF subject pronouns could be elided when two verbs are coordinated, as in (20), while in MCF they must be repeated, as in (21):

(20) Mais je doi plorer et do-1 faire.¹⁹ (21) Mais je dois pleurer et *(je) dois le faire.
 but I must cry and must-it do but I must cry and I must it do
 'But I must cry and I must do it.' 'But I must cry and I must do it.'

Finally, in MCF subject pronouns can be doubled, as in that data below:

(22) euh moi je trouve ce qui en souffre le plus... (23) moi je suis un bloggeur
 uh me I find that who of-it suffers the most me I am a blogger
 'I find that the one who suffers most...' 'I am a blogger.'

In (22) and (23), the subject pronoun *je* is doubled by *moi*. Thus, the SAC appears to be complete in MCF.

1.3.1.2 Analysis

In formal terms, how does all this work? Van Gelderen proposes three stages that revolve around the categorial status of the pronoun and the features at play. At stage (a), the pronoun is a full DP that merges in Spec, *v* to contribute interpretable φ -features. The DP receives its theta-role and then moves to Spec, *T* to check *T*'s uninterpretable φ -features. The

¹⁷ Guillaume de Lorris, *Roman de la Rose*, c. 1230

¹⁸ Chréten de Troyes, *Cligès*, c. 1176

¹⁹ Anonymous, *Aucassin et Nicolette*, XXIV, 42; 12th century

subject pronoun also has an uninterpretable Tense feature which is valued by the T-head. At stage (b), feature loss begins. The DP merges as an argument since it still has interpretable φ -features but at this point it may lack the Tense feature and possibly other features as well. The D-head of the DP then moves to T. At stage (c), the pronoun only has uninterpretable φ -features and so it is now agreement. Since T needs to probe a goal to have its φ -features valued, it probes another constituent, such as another DP, which may be a topic and/or an emphatic pronoun. This last stage is renewal since the topic/emphatic pronoun will be reanalyzed as a subject pronoun, at which point the cycle goes back to stage (a). The SAC as described above is summarized in (24) below.

(24) Stages of the Subject Agreement Cycle

Stage (a): Pronoun is DP with $i\varphi$ $\rightarrow [_{TP} DP T [_{VP} <DP> v [_{VP} V]]]$

Stage (b): Pronoun is DP/D with $i\varphi$ $\rightarrow [_{TP} T + v + D [_{VP} DP v [_{VP} V]]]$

Stage (c): Pronoun is $u\varphi$ on T $\rightarrow [_{TopP} DP Top [_{TP} <DP> T_{u\varphi} [_{VP} <DP_{i\varphi}> v [_{VP} V]]]$

Van Gelderen observes that first- and second-person pronouns are typically reanalyzed ahead of third-person. In English, for example, first- and second-person pronouns behave more like agreement rather than DP pronouns. Additionally, first- and second-person pronouns become clitics ahead of third-person pronouns. In order to account for these differences, van Gelderen proposes that third-person pronouns have more features than first- and second; i.e., third-person pronouns have gender and deictic features. Since they have more features, it takes longer for them to be reanalyzed.

As mentioned above, after stage (c) the cycle may start over again. This renewal takes place because there is no goal providing $i\varphi$ features to check T's $u\varphi$ features. According to van Gelderen, the most common sources of renewal are pronouns, nouns, and demonstratives. These

are the types of constituents that can check T's $u\varphi$ -features. In French, the cycle was renewed by emphatic pronouns for first- and second-person and demonstratives for third-person. Van Gelderen (2011:78) observes that in OldF, *je* and *tu* were used as both emphatic and regular subject pronouns. At this time, *moi* and *toi* were emphatic object pronouns while the regular object pronouns were *me* and *te*. After *je* and *tu* were reanalyzed as agreement, *moi* and *toi* renew the cycle since they still have $i\varphi$ features. For the third-person, the OldF subject pronouns are etymologically derived from Latin demonstratives rather than Latin subject pronouns.

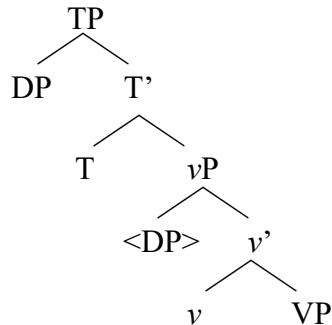
Van Gelderen also discussed the role that topics play in renewing the cycle. The emphatic pronouns discussed above often undergo topicalization to the left periphery. Renewal will be triggered if the emphatic pronoun no longer moves to CP but rather to Spec,T. If the pronoun occupies Spec,T, it is now associated with the T-head and the cycle is back at stage (a). Thus, at stage (c), overt subject pronouns are actually topics. This is the case for standard Modern Spanish, as has been argued by Ordóñez & Treviño (1999). Thus, in a sense, DP movement to Spec,T feeds the SAC. In Chapter 3, I develop a similar claim for object movement and the Object Agreement Cycle. The stages of the SAC are represented derivationally in Section 1.3.1.4 below.

1.3.1.3 Stages of the SAC

Stage (a) – Old French

Subject pronoun = DP [iφ, uT]

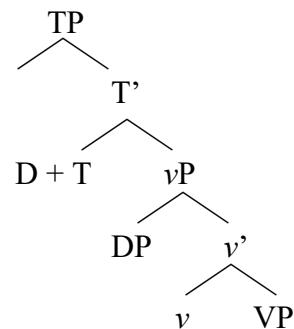
T [uφ, iT]



Stage (b) – Standard Modern French

Subject pronoun/clitic = DP/D [iφ, uT]

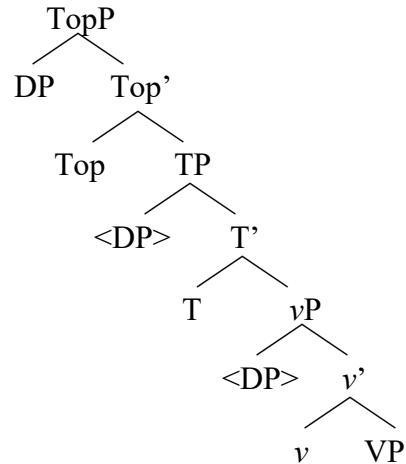
T [uφ, iT]



Stage (c) – Modern Colloquial French

Subject clitic = T [uφ, iT]

DP or *pro* merges as argument; DP can be topic



1.3.1.4 Null subjects and the SAC

One more important aspect of the SAC to discuss is its relevance to null subjects. I have presented data from OldF and MCF to show that the SAC has gone from stage (a) to stage (c). However, an additional change that took place in the history of French was its status as a null subject language. Standard Modern French is a non-null subject language but OldF allowed null

subjects (Adams 1987). Are these two developments, the grammaticalization of subject pronouns and the loss of null subjects connected or is the latter an independent development? Van Gelderen (2011:44) mentions that whether a language allows null subjects or not is “connected” the SAC because it depends on how T’s φ -features are being checked. Depending upon which stage of the SAC a language is in, this will take place in different ways. Of course, in radical pro-drop languages like Chinese this is not the case since those types of languages lack φ -features altogether.

While van Gelderen (2011) assumes that null subjects are related to stages of the SAC, she does not develop the point fully. Based on her analysis, stage (c) is when one would expect null subjects. This is because at stage (c), the external argument position is open and another element must merge there as a goal for the T probe. Van Gelderen does point out that this goal could be overt or phonetically null. If the latter, we have a case of null subjects. Thus, null subject languages should be at stage (c) of the SAC. Spanish, one of the most commonly discussed null subject languages, is consistent with this. In standard ModS, the verb has overt subject agreement ($u\varphi$) and overt subjects are topics, as argued by Ordóñez & Treviño (1999). Throughout this dissertation I will develop the connection between null arguments and cycles more explicitly, adopting an analysis based on Holmberg (2005, 2010), whereby null arguments are licensed by a D-feature on T for subjects and v for objects. I argue that the D-feature ends up on the relevant head as a result of the linguistic cycle.

1.3.2 *Object Agreement Cycle*²⁰

The Object Agreement Cycle (OAC) is similar to the SAC since, in both cases, a pronoun is being reanalyzed as verbal morphology. However, while the principles behind each cycle

²⁰ The data in this section are from van Gelderen (2011) and references therein.

remain the same, they differ with respect to the syntactic domain concerned. The SAC takes place primarily in the T-domain while the OAC is relevant to vP . Before seeing examples of the OAC at work, let us consider in more detail what exactly constitutes “object agreement,” since it is typologically less frequent.

1.3.2.1 Defining object agreement

Object agreement can be defined as a morphosyntactic matching relationship between a verb and its direct object complement. In Corbett’s (2006) terminology the direct object is the “controller” and the verb is the “target.” The “domain of agreement” is a transitive clause. The verb reflects morphological features of its object complement. Within Generative theory, agreement is formalized as the realization of φ -features on the verbal head. During the Government & Binding era, object agreement was analyzed as follows. The object DP merges as complement to verb where it is assigned theta-role. The object then moves to $Spec, AgrO$ to check accusative Case (Pollock 1989, Chomsky 1991). This analysis has been revised as part of the Minimalist program (Chomsky 2000, 2001). Agreement is a relationship between a “probe” and its “goal.” Direct object agreement involves the v head as a probe, since it has unvalued φ -features, and a lexical DP with valued φ -features as the goal. Thus, the DP object values v ’s unvalued φ -features. The φ -features appear as an agreement marker at PF, though overt realization of agreement is subject to cross-linguistic variation. The GB analysis depends on Agr heads which have since been dispensed with, even though $AgrO$ was useful for accounts of object shift.²¹ Within the Minimalist approach, Case assignment has been separated from movement but an additional feature has been proposed as a trigger for object shift; i.e., EPP.

²¹ Chomsky (1995) dispenses with Agr projections because they are uninterpretable at the interfaces and thus do not contribute anything semantically.

According to van Gelderen (2011), object agreement is found in many different languages and language families including Afro-Asiatic, Bantu, Dravidian, Indo-European, Austronesian, Athabascan, and Uto-Aztecán.²² Since it is not isolated to a single group of languages it is worthy of study from the perspective of Universal Grammar. Furthermore, object agreement parallels subject agreement. They both occur via the same operation (Agree) but differ with respect to domain: *T* for subjects, *v* for objects.

1.3.2.2 Analysis and cross-linguistic variation

The OAC resembles the SAC since they both involve reanalysis of pronouns into agreement. The OAC takes object pronouns and turns them into object agreement morphology on the verb. Example (25) illustrates object agreement in Taqbaylit Berber.

(25) zri-x-t **umcic**

saw-I-OBJ the-cat

‘I saw the cat.’

The OAC is divided into three stages. At stage (a) the object pronoun is a full DP that merges as complement to the verb. It contributes *i*φ features that check *v*’s *u*φ features. The DP pronoun has its *u*Case feature checked by *v*’s *i*Case feature. At stage (b) the pronoun is reanalyzed as a clitic/head. It merges as a DP but then moves as a D-head with the verb to *v*.²³ The clitic pronoun still has *i*φ features. Clitic doubling is not allowed at this stage since the clitic pronoun occupied the object position upon first merge. The clitic is reanalyzed as the spell-out of *u*φ features on *v* at stage (c). The clitic/agreement morphology may now be doubled by an object

²² For more examples of languages with object agreement, see the Siewierska (2008). Object agreement is grouped under feature 102A (Verbal Person Marking) in the *World Atlas of Language Structures* (<https://wals.info>).

²³ I depart from van Gelderen’s original analysis of stage (b) in Section 3.3.4 below.

DP or *pro*. Thus, clitic doubling is agreement once the clitic is *v*. Renewal by a lexical object can take the cycle back to stage (a).

The same diagnostics that are used to determine the categorial status of subject pronouns can be applied to object pronouns: coordination, modification, movement.²⁴ Van Gelderen (2011) provides data from the following languages and identifies the stage they are in. Consider first languages in stage (a) such as Urdu/Hindi, Japanese, and Malayalam. In the Urdu/Hindi in (26), the object pronoun *us-ko* is a full DP since it can be separated from the verb by the PP *gher me*.

(26) m̄ey nee **us-ko** gher me dekhaa
I ERG him/her-OBL house in saw
'I saw him/her in the house.'

An object pronoun that can be separated from the verb has likely moved out of its A-position and thus it is a full DP. Stage (b) languages include English, some Semitic languages such as Arabic, Hebrew, and Coptic. English object pronouns may be contracted as in (27) and doubling is only acceptable if the object DP is a topic in the left periphery as in (28b). In (28a) the doubling is unacceptable because the object pronoun, while contracted phonologically, still occupies argument position. If *that guy* were separated by a prosodic break, it would be acceptable as a hanging right-dislocated topic. Object pronouns are in complementary distribution with lexical objects.

(27) a. Hear how I salute **him**. (28) a. You hate (*em) *that guy*
b. Hear how I salute 'em. b. That guy, you hate'm.

²⁴ Clitic doubling is an additional diagnostic that is dealt with in detail in Section 3.2 below.

Finally, languages in stage (c) include Southern Slavic, Bantu, and some Austronesian languages. In Kambera (Central-Malayo-Polynesian group of Austronesian), for example, definite objects must also be marked on the verb as in (29a) but indefinite objects do not have to be marked (29b).

(29) a. Mbàda manahu-da-*(nyà)-ka **na** uhu
 already cook-3P-3S-PF the rice
 ‘They have already cooked the rice.’

b. Mbàda manahu-da-ka **uhu**
 already cook-3P-PF rice
 ‘They have already cooked some rice.’

This object marking is object agreement and so Kambera is in stage (c) of the OAC.

An additional phenomenon related to the OAC is Accusative Clitic Doubling (ACD), as in Bulgarian (30) below.

(30) Vidjah (go_i) Ivan_i.
 I-saw him Ivan
 ‘I saw Ivan.’

ACD occurs when a clitic and a coreferential DP occur in the same clause. In (30), *go* and *Ivan* both refer to the same individual. On van Gelderen’s analysis, ACD becomes possible at stage (c) because it is at this point that the clitic is a *v*-head and thus, nothing is occupying the verbal complement position. Consequently, a coreferential object may merge which results in ACD.²⁵

According to van Gelderen, the further advanced a language is at stage (c), the less restrictive

²⁵ In Chapter 3, I challenge the claim that ACD is only possible at stage (c). I show that, at least in Spanish, ACD occurs at both stage (b) and stage (c), but they are derivationally distinct.

ACD is. In the Bulgarian data in (30) above, the clitic optionally doubles the object. Now consider the Macedonian data below:

(31) a. Daniela go_i poznavava nego_i.
Daniela him knows him
'Daniela knows him.'

b. Daniela ja_i kupi kniga-ta_i.
Daniela it bought book-the
'Daniela bought the book.'

c. Daniela kupi edna kniga / knigi.
Daniela bought one book books
'Daniela bought a book / books.'

In (31a), *go* obligatorily doubles the full pronoun *nego*. In (31b) an inanimate object is doubled. However, ACD is not completely unrestricted in Macedonian, because indefinite objects as in (31c) cannot be doubled. Van Gelderen (2011) takes these different patterns of ACD to indicate that both Bulgarian and Macedonian are either at stage (c) of the OAC or "moving towards it." Bulgarian is behind Macedonian since it has more restricted ACD.

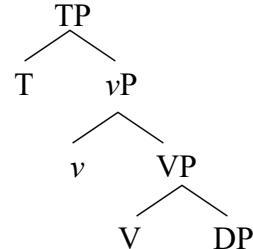
Recall that the SAC can be renewed by pronouns, nouns, or determiners, and it typically starts in the first- and second-person rather than the third. The OAC differs from the SAC in this respect. Van Gelderen observes that renewal in the OAC can have a variety of sources and does not necessarily start at first- and second-person. One common denominator in this cycle is the importance of definiteness and animacy, which conditions doubling. Due to this, van Gelderen proposes an Aspect feature [Asp] on *v* involved in the OAC, in addition to the regular φ -features.

The stages of the OAC with their corresponding languages are summarized in Section 1.3.2.3 below.

1.3.2.3 Stages of the OAC

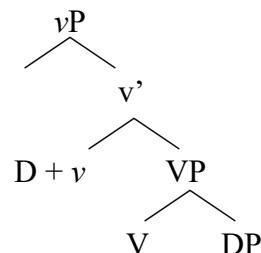
Stage (a) – Urdu, Hindi²⁶

Object pronoun = DP [iφ, uAsp]
 v [uφ, iAsp]



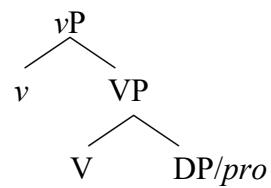
Stage (b) – Modern English, Arabic

Object pronoun/clitic = DP/D [iφ, uAsp]
 v [uφ, iAsp]



Stage (c) – Kambera, Southern Slavic

Object clitic = v [uφ, iAsp]
DP or *pro* merges as argument



1.3.2.4 Null objects and the OAC

As with the SAC, van Gelderen does not elaborate on null objects and their relation to the OAC. She does point out that, on her analysis, object *pro* becomes available at stage (c) of the OAC because now that the object clitic is a v -head, complement position is open for either a lexical DP object or *pro* to merge. Above I discussed French as an example of a language in

²⁶ Hindi is an SOV language. Van Gelderen (2011:90) places the object pronoun in Spec, V but does not go into detail on how the object lands in a preverbal position. She does state explicitly that she does not assume an EPP-feature in v , which would trigger object shift to Spec, v . I assume the object pronoun can undergo further movement but for simplicity sake I leave it as in the verbal complement in stage (a) above.

which the null subject status has changed. OldF allowed referential null subjects while ModF does not. Throughout this dissertation I will pursue the notion that null referential objects are licensed in a language as a result of the OAC. The language of focus in this regard will be Spanish rather than French. This may seem an odd choice since Spanish is not conventionally considered a null object language, like Pashto. Nevertheless, I show in Section 5.2 that some varieties of Spanish like Rioplatense do indeed allow null referential objects. I argue that this is the result of the OAC, which has resulted in a D-feature being present in *v*, which is how null referential objects are licensed on a Holmberg (2005, 2010) type of analysis.

1.4 Dissertation synopsis

Chapter 2: Impersonal Pronouns and the Subject Agreement Cycle

In Chapter 2, I extend the Subject Agreement Cycle (SAC) to impersonal subject pronouns. I adopt Egerland's (2003) typology of impersonal pronouns and apply it to Spanish. After diagnosing to what extent impersonal pronouns in Spanish are deficient I introduce a new diagnostic, cross-clausal coreferentiality, to distinguish among the different types of impersonal pronouns in Spanish and French. I argue the Modern French impersonal *on* has undergone grammaticalization via the SAC. I also explain how the Old Spanish impersonal pronoun *omne* was lost diachronically and I propose that impersonal pronouns participate in a cycle that is a sub-type of the SAC, which I label the "Impersonal Subject Cycle." My analysis leads to the prediction that impersonal pronouns are grammaticalized after personal subject pronouns. Both types of pronouns may serve as input to the SAC, but impersonal pronouns are less likely to be reanalyzed than are personal pronouns.

Chapter 3: Object Agreement and Object Movement in Spanish

Chapter 3 is focused on the Object Agreement Cycle (OAC). I start by discussing the categorial status (phrase or head) of object clitics in Old Spanish. Building on van Gelderen (2011) I show based on data from Latin, Old Spanish, and Modern Spanish that object clitics are undergoing grammaticalization via the OAC based on diagnostics discussed in Chapter 1. I also develop clitic doubling as a diagnostic further than was done in van Gelderen (2011). I discuss patterns of accusative clitic doubling (ACD) in “standard” and regional varieties of Spanish such as Rioplatense which indicate that some varieties of Spanish are further ahead in the OAC. I claim *contra* van Gelderen that ACD is actually possible at both stages (b) and (c) of the OAC, though the derivation is not the same. After adopting an analysis of doubling based on Harizanov (2014) and Kramer (2014), whereby ACD involves object movement to *Spec,v*, I show how object movement has fed the OAC historically in Spanish.

Chapter 4: The Reflexive Object Cycle from Latin to Spanish

In Chapter 4, I turn to reflexive clitics and show that they are participating in a cycle similar to the OAC which I refer to as the "Reflexive Object Cycle" (ROC). I start by showing based on distributional patterns and phonological attrition that the reflexive pronoun in Latin and Old Spanish was a full DP that underwent reanalysis. It is now a Voice head in Modern Spanish that marks valency. I analyze Old Spanish (1200-1400) as stage (a) moving to stage (b) of the ROC, Middle Spanish (1400-1600) as stage (b), and Early Modern and Contemporary Modern Spanish (1600-present) as stage (c) of the ROC. In so doing, I use the same diagnostics that were used previously but I also take advantage of patterns of auxiliary selection to show that the Old Spanish reflexive clitic was still an argument and, consequently, a DP. I also discuss how doubling with the reflexive clitic has become obligatory in Modern Spanish where it was

optional in Old Spanish. This is further evidence for stage (c) of the ROC. While Chapter 4 is focused on “true” reflexive *se* constructions in which the external and internal argument are coreferential, I briefly consider patterns of doubling with other types of *se* such as passive and impersonal *se*.

Chapter 5: Agreement Cycles and the Licensing of Null Arguments

In Chapter 5 I pursue some of the consequences of the analysis of cycles I have given throughout the dissertation and how they connect with the development of null subjects and null objects. I adopt a theory of the licensing of null subjects based on Holmberg (2005) and Holmberg et al (2009) and I extend it to null objects. I claim that null objects are the result of reanalysis of a D-feature on *v* which was encouraged by patterns of clitic-left dislocation (CLDT) and accusative clitic doubling (ACD). Interestingly, CLDT occurs historically prior to ACD in the languages under consideration. I argue that this pattern is tied to what stage of the Object Agreement Cycle (OAC) a language is in. I then propose a typology of null object languages similar to Holmberg's (2005) typology of null subject languages. After this I discuss null subjects and show how they have developed in French as a result of the SAC. I argue following MacDonald & Maddox (2018) that passive *se* constructions can only arise if a language has null subjects and if the language has reached the stage of the ROC where the reflexive is reanalyzed as a Voice head. Since this development involves two cycles, the SAC and the ROC, I frame this as a case of "cyclic interaction" following work by Bahtchevanova & van Gelderen (2016).

Chapter 6: Conclusion

Chapter 6 is where I conclude and summarize the findings of this dissertation. I also consider how to take this line of research further and what areas to pursue for future work.

Chapter 2: Impersonal Pronouns and the Subject Agreement Cycle

Introduction

Various impersonal pronouns (Imp) in Romance are derived from the Latin noun HOMŌ, “man” in English; e.g., French (1) and Catalan (2).

(1) On ne parle pas anglais ici.	(2) Hom no parla anglès aquí.
Imp Neg speaks Neg English here	Imp Neg speaks English here
‘One does not speak English here.’	‘One does not speak English here.’

The focus of this chapter is the corresponding Old Spanish (OldS) *omne*, as in (3) and (4).

(3) E los veniales pecados, quando omne más come e beve de lo que deve...¹

and the venial sins when Imp more eats and drinks of it that should

‘And the venial sins: when one eats and drinks more than they should...’

(4) Cuando omne non puede dormir...²

when Imp not can sleep

‘When one cannot sleep....’

OldS *omne* was in use through the medieval period but then disappears in the sixteenth century (Brown 1931).³ The history of this pronoun poses a potential problem for a cyclic theory of grammaticalization because it does not follow the typical pattern of phonological weakening, reanalysis to a head, and renewal.

¹ Pedro de Cuéllar, *Catecismo*; 1325

² Juan Manuel, *El Conde Lucanor*; 1328-1335

³ There is orthographic variation the Old and Middle Spanish periods but for simplicity sake I use <*omne*>. Brown (1931:265) lists (*h*)*ome*, (*h*)*omme*, and (*h*)*ombre*, but there may be other variants as well.

While *omne* does undergo some phonological weakening, it is not reanalyzed before it falls out of use. Thus, its cycle may be characterized as broken or incomplete.⁴ Previous studies like Brown (1931), Company Company & Pozas Loyo (2009) have attributed its loss to competition from other constructions such as impersonal *se* or generic *uno*. Giacalone Ramat & Sansò (2007) offer an areal account in which they refer to HOMÓ-derived pronouns as a “recessive” feature that spread throughout European languages due to contact. The areal study does not provide an explanatorily adequate reason for their loss but rather a description of the process. Competition accounts are problematic because they do not define clearly what is meant by “competition.” Must the relevant constructions be the same categorially (both being pronouns, for example) or is semantic equivalence sufficient for them to be competing?

The goal of this chapter is to explain more satisfactorily from a formal point-of-view why the impersonal pronoun was lost in at least one language, Spanish. My main claim is that OldS *omne* was undergoing a grammaticalization cycle similar to the Subject Agreement Cycle (SAC), which was discussed in Chapter 1 above. I will refer to the cycle *omne* was on as the “Impersonal Subject Cycle” (ISC). In fact, the ISC is just the SAC with a HOMÓ-derived pronoun as its input rather than person subject pronouns. I argue that the ISC stalled in Spanish because there was no SAC taking place; i.e., subject pronouns were not being grammaticalized. *Omne* is used most frequently as an impersonal subject pronoun, and thus there was no reanalysis of *omne* as a head, contrary to what took place in French to the corresponding pronoun impersonal *on* (Imp_{on}). In Modern Spanish (ModS), generic *uno* has replaced *omne* on the ISC. This is not the cause but rather the consequence of the loss of *omne*.

⁴ Or “gramaticalización trunca,” in the words of Company Company & Pozas Loyo (2009).

As part of the analysis I make a comparison with Modern French Imp_{on} which is a completed example of the ISC. I take Imp_{on} to be a generic T head that licenses *pro* in Spec, v . The ultimate end of the ISC is to take an impersonal pronoun and turn it into an impersonal affix.⁵ Historical data show that Old French Imp_{on} was at an earlier stage in the cycle than OldS *omne*. Furthermore, in French the subject pronouns have been undergoing grammaticalization since the earliest documents. Old French (OldF) was a null subject language while standard Modern French (ModF) is a non-null subject language. Modern Colloquial French (MCF), on the other hand, is going back to being a null subject language with clitic-pronouns acting as agreement T-heads that license *pro*.⁶ Since the other subject pronouns like *je* and *tu* were being reanalyzed as heads in French, the reanalysis of *on*, also a subject pronoun, was stimulated. In Spanish, however, subject pronouns have not been undergoing grammaticalization and thus there has been no push to reanalyze *omne*.

The format of this chapter is as follows. In Section 2.1, I apply the diagnostics to determine the categorial status (head or XP) of the impersonal pronouns under consideration at their different historical stages. I show that while ModF Imp_{on} is a head, OldS *omne* was a full phrase. I also present Egerland's (2003) typology of impersonal pronouns and apply his classification to ModS and OldS. Section 2.2 introduces an additional diagnostic of cross-clausal coreferentiality to show that OldS *omne* patterns like ModS generic *uno*, but ModS Imp_{se} does not. I also compare ModF Imp_{on} and OldF_{on} to show that the latter was less grammaticalized. In

⁵ Some non-Indo-European languages that have impersonal clitic pronouns or affixes (not necessarily derived from a MAN-pronoun) include Hausa (Jagger 2001, Pawlak 2009), Northern Tepehuan (Bascon 1982), and Somali (Cabredo-Hofherr 2004).

⁶ By “Modern Colloquial French” I mean the spoken continental variety as opposed to the written one. I adopt this term from previous literature. See van Gelderen (2011:50) and references therein. For my purposes, MCF excludes Canadian and Swiss varieties of French.

Section 2.3 I discuss the stages of the ISC and discuss how OldS *omne* was lost and replaced by generic *uno*. This chapter is summarized in Section 2.4.

2.1 Diagnosing degrees of deficiency

In this section, using the same diagnostics as discussed in Chapter 1, I show that ModF Imp_{on} is a head while OldS *omne* was a full DP. I then introduce Egerland's (2003) typology of impersonal pronouns, which distinguishes between two different classes based on their level of grammaticalization. Using Egerland's diagnostics I show that ModF Imp_{on} belongs to the more advanced class while OldS *omne* exhibits properties of the less grammaticalized class of impersonal pronouns.

2.1.1 The status of ModF Imp_{on} and OldS omne: head or phrase?

Recall from Chapter 1 that tests of coordination, modification, and separation from the verb serve to distinguish a head from a phrase (XP). With respect to ModF Imp_{on}, consider the following data:

(5) *On et elle parlent beaucoup.

Imp and she speak a-lot

(6) *On intelligent parle français.

Imp intelligent speaks French

(7) *On souvent arrive tard.

Imp often arrives late

In (5), Imp_{on} cannot be coordinated with the subject pronoun *elle*. The data in (6) show that Imp_{on} cannot be modified by the adjective *intelligent* and in (7) Imp_{on} cannot be separated from the verb *arrive*. This leads to the conclusion that in ModF, Imp_{on} is indeed a head.

Now consider the following data with OldS *omne*:

(8) si omne bien non cata...⁷ (9) todo omne que deva deuda a otro...⁸
 if Imp well not observe every Imp that owes debt to another
 ‘If one does not observe well...’ ‘every one that owes debt to another...’

(10) como Cristo, que es Dios e omne...⁹
 like Christ who is God and man
 ‘like Christ, who is God and man...’

In (8) *omne* is separated from the verb by the adverb *bien* and negation. *Omne* is modified by a relative clause in (9) and in (10) *omne* is coordinated with the noun *Dios*. These diagnostics indicate that in the OldS period, *omne* was a full DP.

2.1.2 Egerland’s (2003) typology of impersonal pronouns¹⁰

Egerland (2003) argues that the variation exhibited by impersonal pronouns cross-linguistically is due to the distinct featural makeup of the pronoun and the stage of grammaticalization it is at. He identifies two classes based on diagnostics of variable agreement, semantic reading (generic or arbitrary), and syntactic function. Class 1 impersonal pronouns display variable agreement, can be either generic or arbitrary, and cannot be objects. Class 2 impersonal pronouns, however, do not display variable agreement, can only be generic, and can be subjects and objects.¹¹

Egerland arrives at these diagnostics by comparing the impersonal pronouns, some of which are derived from HOMŌ or MAN, in Swedish (*man*), Italian (*si*), Icelandic (*maður*), and

⁷ Gonzalo de Berceo, *Vida de San Millán de Cogolla*; c. 1230

⁸ *Fuero de Burgos*, Philadelphia Ems.245; 1290-1300

⁹ Alfonso X, *General Estoria, Primera Parte*; c. 1275

¹⁰ The data in this section are from Egerland (2003).

¹¹ The terminology “Class 1” and “Class 2” is my own. Egerland refers to Class 1 as “arbitrary impersonal pronouns” and Class 2 as “generic impersonal pronouns.” Egerland’s terminology can be confusing since arbitrary impersonal pronouns can be both arbitrary and generic. I have introduced the new labels to help avoid confusion.

French (*on*).¹² To see how these diagnostics work, starting with variable agreement, consider the data below for the first three languages. French Imp_{on} and Old Spanish *omne* are treated in Section 2.1.3 below.

(11) Om man inte är gift / gifta måste man ha skilda rum på detta hotell.

if Imp not is married.S / married.Pl must Imp have separate rooms in this hotel

‘If one is not married, one must have separate rooms in this hotel.’

(12) Quando si è donna / donne, si è disposta / disposte a rinunciare a molte cose
when Imp is woman women Imp is ready.S ready.Pl to renounce to many things
per i propri figli.

for the own children

‘When one is a woman, one is ready to renounce many things for their children.’

(13) Þrátt fyrir sannanirnar var maður ekki alveg sannfærður / *sannfærðir um
despite for the-evidence was Imp not completely convinced.S / convinced.P about
sekt hans.

guilt his

‘Despite the evidence, one was not completely convinced about his guilt.’

(14) a. *Om du är gifta... Swedish

if you are married.Pl

b. *Quando tu sei giovani... Italian

when you are.2S young.Pl

c. *Quand tu es belles... French

when you are.2S lovely.Pl

Generic *you*

¹² Italian *si*, like Spanish *se*, is derived from the Latin reflexive pronoun.

Swedish *man* in (11) can be modified by an adjective in the singular, *gift*, or the plural, *gifta*.

Italian impersonal *si* (Imp_{si}) in (12) shows the same pattern while Icelandic *maður* in (13) cannot be modified by a plural adjective.¹³ As for generic *you* in (14), it is unacceptable with a plural adjective in all three languages. Thus, vis-à-vis variable agreement, Swedish *man* and Italian Imp_{si} pattern together in one group (Class 1), while Icelandic *maður* and generic *you* are in the other group (Class 2).

The second diagnostic Egerland uses is what kind of semantic reading the impersonal pronoun can have; i.e., generic or arbitrary. In this context “generic” refers to a quasi-universal set of individuals that can potentially include the speaker. It may be paraphrased by the English *one* or *people*. An arbitrary reading, on the other hand, refers to a non-specific set of individuals within a specific time reference, paraphrased by English *some people* or *they*. This reading always excludes the speaker. With this terminology in mind, now consider the following data:

(15) a. Man måste arbeta till 65. Swedish

‘One must work until the age of 65.’

b. Man arbetade i två månader för lösa problemet.

‘They worked for two months to solve the problem.’

(16) a. Si deve lavorare fino all’età di 65 anni. Italian

‘One must work until the age of 65.’

b. Si è lavorato per due mesi per risolvere il problema.

‘They worked for two months to solve the problem.’

¹³ For the purpose of this chapter I leave aside the issue of whether Italian Imp_{si} is a pronoun or not.

(17) a. Maður vinnur til 65 ára aldurs. Icelandic
 ‘One works until the age of 65.’

 b. *Maður hefur unnið að því í two mánuði að leysa vandamálið.
 ‘They worked for two months to solve the problem.’

(18) a. Du arbetar till 65. Swedish (generic *you*)
 ‘One works until the age of 65.’

 b. *Du har arbetat i två månader för att lösa problemet.
 Intended: ‘They worked for two months to solve the problem.’

In (15a), Swedish *man* is generic while in (15b) it is arbitrary. Italian Imp_{si} in (16) patterns with Swedish *man* again while in (17a) Icelandic *maður* is generic but it cannot be arbitrary as in (17b). The generic *you* in Swedish in (18) patterns with Icelandic *maður*. Just as was seen above for variable agreement, *man* and Imp_{si} are in one group (Class 1) with respect to the type of reading allowed while *maður* and generic *you* are in the other group (Class 2).

Egerland refers to his third diagnostic as “syntactic function,” by which is meant whether the impersonal pronoun can be both a subject and an object or just a subject. Now consider the Swedish and Italian data below:

(19) *De har sett man. Swedish
 they have seen Imp

(20) *Loro si hanno visto. Italian
 they Imp have seen.

The data in (19) and (20) show that *man* and Imp_{si} cannot be objects. How do Icelandic *maður* and generic *you* pattern in this respect?

(21) Svona tölur segja *manni* að eitthvað sé í ólagi. Icelandic
 ‘Such figures tell one that something is wrong.’

(22) Om de litar på *dig*, får du inte göra dem besvikna. Swedish
 ‘If they rely on you, you must not make them disappointed.’

In (21) the accusative form of the Icelandic impersonal pronoun, *manni*, is used as an object. Generic *you* in (22) also has an accusative form and can be an object. Thus, once again, Swedish *man* and Italian Imp_{si} pattern into one group while Icelandic *maður* and generic *you* pattern together in their own group. Swedish *man* and Italian Imp_{si} may be categorized as Class 1 impersonal pronouns. They allow variable agreement, can have either generic or arbitrary readings, and are restricted to being subject. Icelandic *maður* and generic *you* belong to Class 2. They do not allow variable agreement, they can only have generic readings, and they can be either subjects or objects.

Thus, there are two different classes of impersonal pronouns cross-linguistically. In his analysis Egerland (2003), following Chomsky (1995) and Marantz (1993, 1997), assumes that there are two sets of φ-features: 1) abstract features involved in the derivation and 2) lexically specified features inherent to pronominals. These sets of features must match post-syntactically. Impersonal pronouns, however, lack the second set of lexical features (van Gelderen 1997, Rivero 2000). Egerland proposes that the different patterns exhibited by the two classes of impersonal pronouns are a result of their featural makeup. He proposes the following grammaticalization cline for impersonal pronouns:

(23)	Early Stage	Middle Stage - “Class 2”	Late Stage - “Class 1”
	Lexical DP	> Impersonal generic pronoun	> Impersonal arbitrary pronoun
		Icelandic <i>maður</i> , generic <i>you</i>	Swedish <i>man</i> , Italian Imp _{si}

At the middle stage of the cline in (23), lexical φ-features are maintained; they are lost at the late stage. Class 2 pronouns like Icelandic *maður* and generic *you* are at the middle stage while Class 1 pronouns like Swedish *man* and Italian Imp_{si} are at the late stage; i.e., they are more grammaticalized and lack lexical φ-features. In Section 2.3 below I adopt Egerland's analysis, in part, and extend it to ModF Imp_{on} and OldS *omne*.

2.1.3 Egerland's (2003) typology applied to Modern French and Old Spanish

Now that we have distinguished the two classes of impersonal pronouns according to Egerland's (2003) typology, these diagnostics can be applied to ModF Imp_{on} and OldS *omne*.

Starting with ModF Imp_{on}, consider the following data:

(24) Quand on est belle / belles...¹⁴

when Imp is.3S beautiful.F.S beautiful.F.Pl.

‘When one is beautiful...’

(25) a. On doit travailler jusq’à l’age de 65 ans.

Imp must work until the-age of 65 years

‘One must work until the age of 65.’

b. On a travaillé pendant deux mois pour résoudre le problème.

Imp has worked during two months to resolve the problem

‘Someone worked for two months to fix the problem.’

(26) Cela (*on) conduit (*on) / pro à la conclusion suivante.

this Imp leads Imp to the conclusion following

‘This leads one to the following conclusion.’

¹⁴ From Rey & Rey-Debove (1984), cited in Egerland (2003:79)

In (24), Imp_{on} can have a singular or a plural form of the adjective. In (25a), Imp_{on} is generic while it is arbitrary in (25b). That Imp_{on} cannot be an object is shown in (26), where it is unacceptable as either a proclitic or an enclitic. Thus, $\text{ModF Imp}_{\text{on}}$ is a Class 1 impersonal pronoun. However, it is not a full DP but rather a head, as was shown above in Section 2.1.1. Thus, $\text{ModF Imp}_{\text{on}}$ is actually a Class 1 impersonal affix.

How does OldS *omne* pattern with respect to these diagnostics? First, as concerning variable agreement, in the CORDE *omne* always occurs with a masculine singular adjective as in the data below:

(27) algunas veces omne es avaro en guardar sus cosas.¹⁵

some times Imp is greedy in guarding their things

‘Sometimes one is greedy about guarding their things.’

(28) de aquesta cosa tal, omne es maravillado.¹⁶

of this thing such Imp is amazed

‘By such a thing one is amazed.’

Additionally, I have found no examples in the CORDE where OldS *omne* receives an arbitrary reading. This observation is backed up by Kärde (1943). In his corpus he noted that *omne* contrasts with $\text{ModF Imp}_{\text{on}}$ in this way.

As for syntactic function, OldS *omne* can be a subject, as has been shown in many of the examples above. However, it can also be an object as below:

¹⁵ *Libro del Tesoro*, Girona, Catedral 20a5; 1400-1425

¹⁶ Pero López de Ayala, *Rimado de Palacio*; c. 1378-1406

(29) ca non ensuzia a omne comer con las manos non lavadas.¹⁷

for not soils DOM Imp eat with the hands not washed

‘For it does not soil a person to eat with their hands unwashed.’

(30) Dios guarde a omne de fazer fecho malo.¹⁸

God prevent DOM Imp from doing deed evil

‘May God prevent one from doing an evil deed.’

Thus, these diagnostics indicate that OldS *omne* was a class 2 impersonal pronoun. It was a full DP, as shown in Section 2.1.1, that was at the middle stage of grammaticalization. It appears to never have progressed to the last stage of Egerland’s cline before it was lost.

2.2 Cross-clausal coreferentiality

In Section 2.1 above various diagnostics showed that ModF Imp_{on} patterns like a functional head which is at the end stage of grammaticalization. OldS *omne*, however, was always a full DP at the earlier stage of the Impersonal Subject Cycle (ISC). In this section I provide an additional diagnostic of cross-clausal coreferentiality based in part on Frascarelli (2007); i.e., whether an element can license a coreferential null subject in subsequent clauses. Here again ModF Imp_{on} and OldS *omne* pattern differently, which I take as further evidence of distinct stages of grammaticalization. There is also a brief discussion how ModS impersonal *se* (Imp_{se}) patterns in this respect.¹⁹ Finally, I draw parallels between ModS generic *uno* and OldS *omne* and I propose that the former replaced the latter on the ISC.

¹⁷ Alfonso X, *Setenario*; c. 1252-1270

¹⁸ Juan Manuel, *El Conde Lucanor*; 1328-1335

¹⁹ This aspect of ModS Imp_{se} is also relevant to Chapter 4 below, where I argue that Imp_{se} is a valency marking head that licenses *pro*.

2.2.1 Licensing of *pro* via topic-identification

2.2.1.1 Referential *pro* in Italian

Frascarelli's (2007) study focuses on how referential third-person *pro* is assigned interpretation in Italian via the syntax-discourse interface. She shows that a topic, or its null counterpart, in the left periphery values the features of *pro* and thus, the topic and *pro* are coreferential. Her account adopts a cartographic approach to the topic-domain based on Frascarelli & Hinterhöltzl (2007) according to which there are three types of topics: 1) the Aboutness-shift Topic (A-topic), 2) the Contrastive Topic, and 3) the Familiar Topic.²⁰ Frascarelli (2007) provides intonational evidence from a corpus of spoken Italian in order to justify this fine-grained Topic-domain. Each type of topic is characterized by a different tonal event. The A-topic is the topic that values features on *pro*. She adopts the following hierarchy for the C-domain, based in part on Rizzi (1997) and on her previous work:

(31) [ForceP [ShiftP [GP [ContrP [FocP [FamP [FinP

The topic positions in (31) are ShiftP for the A-topic, ContrP for the Contrastive-topic, and FamP for the Familiar topic. As noted by Frascarelli & Hinterhöltzl (2007:88), these different types of topics do different things. The A-Topic is what introduces a new topic into the discourse, the Contrastive Topic "induces alternatives which have no impact on the focus value and creates oppositional pairs with respect to other topics," and the Familiar Topic "is a D-linked constituent, used to refer to background information, for topic continuity, or in the right periphery as an afterthought" (Frascarelli 2007:699).²¹

²⁰ See also Reinhart (1981) and Lambrecht (1994) for the Aboutness Topic.

²¹ FocP is for Focus Phrase. GP is a Ground Phrase projection based on Poletto & Pollock (2004), which is a functional projection in C-domain that is "targeted by presupposed information." Since neither of these latter projections qualify as topics they are disregarded in what follows.

As an example of how the A-topic and *pro* relation works, consider the following data:²²

(32) **Il mio capo**_i ... *pro_i* è un exreporter...*pro_i* è stato in giro per il mondo...*pro_i* mi ha preso in simpatia solo che siccome *pro_i* è mostruosamente lunatico è apace che domani non gli sto più simpatica e *pro_i* mi sbatte fuori... -- poi c'è **M.F._k** che è questo che appunto sta facendo tipo praticantato per poi andare a fare l'esame da giornalista/fra un anno e mezzo quindi **lui_k** c'ha quanto meno la garanzia che *pro_k* può rimanere lì finché *pro_k* non farà l'esame cioè ehm **lui_i** poi gli deve fare/scrivere le referenze...

'**My boss**_i ... *he_i* is a former reporter ... and *he_i* has been all over the world...and *he_i* likes me, however, as *he_i* is extremely moody, maybe tomorrow *he_i* does not like me any longer and *he_i* fires me ... -- then there is **M.F._k**, who is practicing for his exam as a journalist/in one and a half years, so at least **he_k** has a guarantee that *he_k* will stay there till *he_k* has made the exam because **he_i** then must make/write a report...'

In (32), the first overt subject, *il mio capo*, is the topic that the following instances of *pro* refer to and so they are coindexed. That it qualifies as an A-topic is supported by the intonational evidence; i.e. it has an L*+H tone. When a new A-topic is introduced, *M.F.*, it could identify *pro* but instead the speaker uses another overt pronoun, *lui* and the subsequent *pros* are linked to it and, by extension to *M.F.* Both the overt *M.F.* and *lui* also have the appropriate tone needed to qualify as A-topics. Finally, in the last sentence another overt *lui* is introduced. This is new A-topic referring once again to *il mio capo*. In this case, the *lui* is required. If a *pro* is merged instead, it will be identified with *M.F.* Frascarelli takes this as evidence that *pro* is always interpreted in relation to the closest A-topic. These data also show that the interpretation of *pro* does not depend upon the agreement feature of the head under which it is licensed but rather by a "matching" relation with the topic. In other words, since all the subjects in question in these data

²² Adapted from Frascarelli (2007:703). The bold-faced words are overt nouns and pronouns. Italian *pro* is *he* in italics in the English gloss. Her study is based on a corpus of spoken Italian, which were analyzed for intonation patterns.

have the same φ -features (masculine third-person singular) it is not φ -feature ambiguities that are the issue but rather topic ambiguity, which will be encoded via an "aboutness" feature, that needs to be checked.

Following Roberts (2004) and Holmberg (2005), Frascarelli assumes a non-pronominal account of Agr. Only the topic-antecedent relation is responsible for the interpretation of referential *pro*. This takes place through a "matching" operation; i.e. Agree. Importantly, she claims that this type of topic-antecedent identification of *pro* is part of the core grammar and it applies under specific structural conditions. Given that the topic is located in the left-periphery and *pro* is merged in Spec,*v*, this might pose an obstacle to the theory, since Agree is an operation that takes place within phases per the Phase Impenetrability Condition (Chomsky 2001, 2004). In order to deal with this, Frascarelli assumes *pro* is at the edge of the phase, Spec,*v*, where it is accessible to agree with the topic in ShiftP. She also assumes that after an A-topic is merged, its null counterpart will be merged in later clauses, until a new A-topic is introduced. For example, consider the following extract from (32) above:

(33) Il mio capo_i è un exreporter ... *pro_i* è stato in giro per il mondo...

the my boss is a former-reporter he has been on tour through the world

'My boss is a former-reporter... He has been all over the world...'

In (33) the overt subject *il mio capo* in the first clause is coindexed with the *pro* in the second clause. In this case, its null copy is present in the second clause, valuing the φ -features on *pro*. Additionally, a [Person] feature and an [Aboutness] feature are valued on *pro* by the topic or its null copy.²³ The structure of the null subject clause in (33) is represented in (34) below.²⁴

²³ "Person" in the logophoric sense of Sigurðsson (2004).

²⁴ I follow Frascarelli's annotation, where Aboutness = α ; Person = Pn. While Frascarelli maintains Agr projections, I dispense with them here and throughout the rest of this chapter. In this analysis Agr plays no role in the identification of *pro*.

(34) [ShiftP *Il mio capo*_[φ, α, Pn] Shift' [TP [vP *pro*_[φ, α, Pn] [VP]]]] ...

In (34), the line going from the topic in Spec,Shift and *pro* in Spec,v represents Agree; i.e., *pro*'s Person and Aboutness features are valued by the topic. *Il mio capo* is struck through because this is a null copy of the overt topic first introduced in the prior clause, now base-generated in Spec,Shift. Notice that *pro* is at the phase edge in Spec,v, allowing agreement to take place across phases. The same derivation takes place with each subsequent *pro* that follows, as in (33), until a new A-topic is introduced.

Frascarelli's analysis requires that every predicational sentence have an Aboutness topic. This is akin to the Extended Projection Principle (EPP), whereby every sentence has a projected subject position. Based on Rizzi's (2006) Subject Criterion, Frascarelli proposes a Topic Criterion to account for this, given below.

(35) Topic Criterion

- a) [+Aboutness] is connected with an EPP feature in the high Topic field that yields a specific discourse-related property, namely "Aboutness."
- b) The [+aboutness] Topic matches with an argument in the main clause through Agree.
- c) When continuous, the [+aboutness] Topic can be null (i.e., silent).²⁵

(35a) ensures that all predicational sentences have a Topic; (35b) explains the nature of the relationship between the topic and the argument; and (35c) allows for null subjects. Frascarelli also observes that the proposed Topic Criterion explains why null subjects are more frequent cross-linguistically than null objects; i.e., in order for the null argument to occur it must be at the phase edge (Spec,v) where it can agree with the topic.

²⁵ A "continuous" topic refers to the topic going unchanged from sentence to sentence.

Importantly, only preverbal subjects qualify as A-topics. Postverbal subjects are not topics but rather associated with Focus (Frascarelli 2007:726), which may be coreferential with the null A-topic present in the C-domain. This suggests that *by*-phrases, which typically occur postverbally, also are disqualified as topics that can identify *pro*. That this is indeed the case is shown by the following data from Samek-Lodovici (1996), also cited in Holmberg et al (2009):

(36) a. Questa mattina, la mostra è stata visitata da Gianni_i. Più tardi egli_i/lui_i/**pro*_i
this morning the exhibit is been visited by Gianni more late he he he
ha visitato l'università.

has visited the-university

‘This morning, the exhibit was visited by Gianni. Later, he visited the university.’

b. Questa mattina, Gianni_i ha visitato la mostra. Più tardi *pro*_i ha visitato
this morning Gianni has visited the exhibit more late he has visited
l'università.

the-university

‘This morning, Gianni visited the exhibit. Later, he visited the university.’

In (36a), the preverbal subject of the first sentence is *la mostra* and not *Gianni*, which occurs in a *by*-phrase. As a result, a *pro* that would refer to *Gianni* in the following sentence is ungrammatical; only an overt pronoun may be used. In (36b), however, *pro* is acceptable in the second sentence because *Gianni* is now a preverbal subject that can now serve as a topic identifying *pro*'s content.

2.2.1.2 Referential *pro* in Spanish

Frascarelli's (2007) analysis of Italian can be extended to the licensing of referential *pro* in Spanish as well. Consider the following data, a Spanish adaptation of (36) above.

(37) a. Esta mañana, la exhibición fue visitada por Juan_i. Luego, él_i/*pro_i fue a la
 this morning the exhibit was visited by John later he went to the
 universidad.
 university

‘This morning, the library was visited by John. Later, he went to the university.’

b. Esta mañana, Juan_i visitó la exhibición. Luego, pro_i fue a la universidad.
 this morning John visited the exhibit later he went to the university
 ‘This morning, John visited the library. Later, he went to the university.’

In (37a), the preverbal subject is *la exhibición*. Following Frascarelli, it is also the A-topic and its null copy will occur in later clauses, identifying *pro*. In the second sentence of (37a), *pro* is unacceptable, but *él* is acceptable. The null subject is not acceptable because it can only be identified with *la exhibición*, and this is not the intended meaning of the sentence. Thus, (37a) shows the same pattern as Italian (36a) above. In (37b), *pro* is acceptable in the second sentence because *Juan*, the preverbal subject of the preceding sentence, is also the topic. Consequently, its null copy is present in later clauses and can identify *pro*. (37b) shows the same patterns as Italian (36b) above. The same type of agreement takes place as shown below.

(38) [ShiftP Juan_[φ, α, Pn] Shift' [TP [vP pro_[φ, α, Pn] [VP *fue a la universidad*]]]]

In (38), the null topic *Juan* is in Spec,Shift while *pro* is merged in Spec,v, at phase-edge, allowing the two to Agree. The A-topic values the Person and Aboutness features on *pro*, thereby ensuring that they receive the same interpretation. In conclusion, Spanish referential *pro* is identified by A-topic, just like Italian referential *pro* following Frascarelli's (2007) analysis.

2.2.2 Impersonal pronouns in Spanish

It was seen in the preceding section that there is a syntactic relationship between an overt subject in one clause and its coreferential null counterpart in the later clauses. Specifically, overt subjects in Italian and Spanish license a coreferential *pro* in subsequent clauses via topic-identification and agreement. The overt subjects under consideration previously were nouns like *il mio capo*, definite pronouns, and proper names. These are lexical DPs that pattern together. In the following subsections I extend this diagnostic to impersonal pronouns like Spanish generic *uno* and impersonal *se*. These elements do not pattern together which suggests that they are categorially distinct. Generic *uno* patterns more like lexical DPs while impersonal *se* does not. Hence, *se* is not a DP but rather a deficient pronoun.

2.2.2.1 Generic *uno*

The licensing of generic subject *pro* in Spanish takes place via the same operation as seen above for referential subject *pro*; i.e., topic identification and Agree (Maddox 2018). Consider the data below:

(39) Uno_i duerme bien, cuando *pro*_i duerme en ese cuarto.

one sleeps well when one sleeps in that bedroom

‘One_i sleeps well when one_i sleeps in that bedroom.’

In (39), the null subject in the second clause is coreferential with the generic *uno* in the first clause. In previous work (Maddox 2018), I argued that this pattern poses a problem for Holmberg’s (2005, 2010) Null Generic Subject Generalization, according to which consistent null subject languages like Spanish do not allow generic null subjects like the one in the second clause; i.e., bare generic third-person singular *pro*. The same analysis that was given for referential *pro* applies to generic *pro* as in (40) below.

(40) a. Matrix: [ShiftP *uno* [TP <*uno*> [vP <*uno*> [VP *duerme bien*]]]]

b. Adjunct: [CP *Cuando* ShiftP *uno* [TP [vP *pro_i* [VP *duerme en ese cuarto*]]]]

In (40), *uno* is introduced as an A-topic in the first clause. Consequently there is a null copy of it present in Spec,Shift in the second clause which enters into Agree with *pro*, thereby identifying its content.²⁶ Thus, the impersonal pronoun *uno* in Spanish licenses a coreferential *pro* in subsequent clauses. Note that the full DP status of generic *uno* is indicated by diagnostics of modification and separation from the verb:

(41) “Un minuto de rezo intenso; con eso basta”. Lo decía uno que nunca rezaba.²⁷

one minute of prayer intense with that is-enough it said one who never prayed

‘One minute of intense prayer; that’s enough. So said one who never prayed.’

(42) uno siempre debe lavar-se las manos.

Imp always should wash-*Refl_{se}* the hands

‘One should always wash one’s hands.’

In (41), *uno* is modified by a relative clause and in (42) *uno* is separated from the verb by an adverb, *siempre*.

2.2.2.2 Impersonal *se*

The pattern of cross-clausal coreferentiality becomes more informative when we look at another impersonal element in Spanish, impersonal *se* (Imp_{se}), as in (43).

(43) En ese cuarto, se *pro* duerme bien.

in that room Imp_{se} one sleeps well

‘In that room, one sleeps well.’

²⁶ For additional diagnostics and the issue of whether indefinite pronouns can be topics, see Maddox (2018).

²⁷ Escrivá de Balaguer, José María, *Surco*, p. 216; 1986

Imp_{se} is uncontroversially a clitic and taken by some authors to be the spell-out of a *v* or Voice head that licenses *pro* in its specifier (Otero 1986, Mendikoetxea 2008, MacDonald 2017).

Generic *uno* and Imp_{se} can both be interpreted generically. Nevertheless, vis-à-vis cross-clausal coreferentiality, these elements pattern differently, as seen in the data below:

(44) a. Se *pro* duerme bien cuando se *pro* duerme en ese cuarto.

‘One sleeps well when one sleeps in that room.’

b. *Se *pro_i* duerme bien cuando *pro_i* duerme en ese cuarto.

In (44a), Imp_{se} can only license a *pro* within the same vP in which it is introduced, hence the unacceptability of the *pro* in (44b). Compare these data with (39) above, repeated as (45) below:

(45) Uno_i duerme bien, cuando (uno_i) *pro_i* duerme en ese cuarto.

one sleeps well when one sleeps in that bedroom

‘One_i sleeps well when one_i sleeps in that bedroom.’

In (45), the *uno* in the second clause may be repeated, but it is optional. Maddox (2018) proposes that this difference is due to the generic *uno* being a full DP and Imp_{se} being a functional head. Since Imp_{se} is a deficient clitic pronoun, it lacks sufficient features to be able to serve as a topic, which is a requirement for it to license *pro* in later clauses following the Frascarelli (2007) type analysis. Generic *uno*, on the other hand, is a full DP pronoun that can serve as a topic, hence it can license *pro* non-locally (outside its own clause). This leads to the following generalization:

(46) Structurally deficient elements can only license a coreferential *pro* clause-internally while full DP pronouns can license a coreferential *pro* in subsequent clauses.

The generalization in (46) is relevant to diagnosing the categorial status of the other pronouns to be considered below.

2.2.2.3 Old Spanish *omne*

OldS *omne* patterns with ModS generic *uno* with respect to cross-clausal coreferentiality, as shown by the following data:

(47) ...porque asi commo por fazer limosna o ayuno o guardar castidad e en
because thus as by doing alms or fast or guarding chastity and in
otros actos virtuosos omne_i ha merito, asy *pro_i* lo ha en la oração.²⁸
other acts virtuous Imp has merit thus Imp it has in the prayer
'...because just as by giving alms or fasting or guarding chastity and in other virtuous
acts one has merit, thus also one has it in prayer.'

(48) ...si omne_i caye de casa o de algun edificio o de arbol o de muro
if Imp fall from house or from some building or from tree or from wall
o de otra cosa qualquier ... & *pro_i* moriere daquella cayda...²⁹
or from other thing what ever and Imp die from-that fall
'If one falls from a house or some building or a tree or a wall or anything else ... and one
dies from that fall...'

(49) Ca si omne_i ayuna commo en guisa que *pro_i* non aya fanbre, yo non digo
for if Imp fasts as in way that Imp not have hunger I not say
que tal ayuno sea malo.³⁰
that such fasting be bad
'For if one fasts in such a way that one is not hungry, I do not say that such fasting is
bad.'

²⁸ Alfonso de Cartagena, *El Oracional*; 1456

²⁹ *Fueros de Aragón*, para. 2; 1247

³⁰ Juan Manuel, *Libro de estados*; 1327-1332

In (47) to (49), an overt impersonal pronoun, OldS *omne*, is introduced in the first clause. The coreferential *pros* that occur in later clauses are licensed by a null local copy of *omne* in Spec,Shift, as in the analysis given above for generic *uno*. Since generic *uno* is a full DP impersonal pronoun and OldS *omne* patterns with it, we may conclude that OldS *omne* was also a full DP impersonal pronoun, which is consistent with what was observed for OldS *omne* in Section 2.1.1 above.

2.2.3 Impersonal pronouns in French

In the preceding subsection it was shown that generic *uno* and OldS *omne* both license a coreferential null subject in subsequent clauses while Spanish Imp_{se} does not. This difference is due to the status of Imp_{se} as a deficient head while generic *uno* and OldS *omne* are full DP pronouns. In what follows the same test is applied to impersonal pronouns in French.

2.2.3.1 Modern French impersonal *on*

Recall that ModF Imp_{on} is an example of a completed Impersonal Subject Cycle. If this is the case then ModF Imp_{on} is a deficient head which should pattern like ModS Imp_{se}. Consider the following data:

(50) a. Si on_i ne fait pas attention, on_i peut s'enrhummer.³¹
 if Imp not makes not attention Imp can catch-cold
 'If one is not careful one can catch a cold.'
 b. *Si on_i ne fait pas attention, *pro_i* peut s'enrhummer.

In (50a), Imp_{on} can have a coreferential Imp_{on} in the second clause but in (50b) Imp_{on} must be repeated or the sentence is ungrammatical. This is the same pattern that was seen for Spanish Imp_{se} in (44) above, which is to be expected if both elements are functional heads.

³¹ Adapted from Cabredo-Hofherr (2010)

2.2.3.2 Old French impersonal *on*

If ModF Imp_{on} is the result of a grammaticalization cycle, we should expect to see evidence of it behaving more like a full DP impersonal pronoun at earlier stages of the language. Evidence that this is indeed the case comes from the following data:³²

(51) Sainz Boneface que l'um martir apelet.³³

Saint Boniface whom the-Imp martyr call

‘Saint Boniface, whom people call martyr.’

(52) Quant l'en en la meson Dieu entre.³⁴

when the-Imp in the house God enters

‘When one enters the house of God.’

In (51) Imp_{on} is separated from the verb by the noun *martir* while in (52) it is separated from the verb by a PP and a noun. This suggests that Imp_{on} in Old French (OldF) was a full DP. The diagnostic of cross-clausal coreferentiality lends further credence to this claim, as shown in the data below:

(53) ne se peüst **on_i** porpenser de richece ...en cel leu trover

not Pron_{se} could Imp imagine of richness in that place find

ne **pro_i** peüst.³⁵

not Imp could

‘One could not imagine any richness...in that place that one could not find.’

³² The orthography was not fixed at this point in the language, hence the variant spellings of Imp_{on}.

³³ Anonymous, *La Vie de Saint Alexis*; 1040-50

³⁴ Rutebeuf, *La Miracle de Théophile*; c. 1261

³⁵ Anonymous, *Roman d'Énéas*; c. 1160

(54) Mais laquele jagunce que **hom** portet u en deit u al col pendue,
 but that jacinth that Imp carries or on finger or on-the neck hung
 seurement puet **pro** aler en autre terre senz pour de engrutement...³⁶
 surely can Imp go in other land without fear of illness
 'But that jacinth that one carries either on his finger or hanging on his neck, surely he can go into another land without fear of illness...'

In both (53) and (54) Imp_{on} occurs in the first clause with null coreferential *pro* in a later clause; i.e., the same pattern displayed by ModS generic *uno* and OldS *omne*. Hence, OldF Imp_{on} was less grammaticalized than its ModF counterpart. It was a full DP and had sufficient features to license *pro* via topic identification.

2.3 Impersonal Subject Cycle: loss and replacement

In this section I give a formal analysis of ModF Imp_{on} and OldS *omne*. I discuss the kind of renewal that can be observed with ModF Imp_{on} and I give a summary of the ISC. I then return to the Subject Agreement Cycle (SAC), which was introduced in Chapter 1, and argue that the ISC is a subtype of the SAC. That the ISC was not completed in Spanish but was completed in French is a result of the status of the SAC in each language. I then show that in Spanish, generic *uno* has replaced OldS *omne* on the ISC. The last subsection deals with a prediction that falls out of the relationship between the SAC and the ISC.

2.3.1 *The Impersonal Subject Cycle (ISC)*

2.3.1.1 Analysis of ModF Imp_{on} and OldS *omne*

For his analysis of impersonal pronouns, Egerland (2003) assumes that there are two types of φ -features involved: lexical features inherent to pronouns and abstract features that are

³⁶ Anonymous, *Lapidaire en prose*; c. 1250-1300

relevant to the syntax. Following Chomsky (1995) and Marantz (1993, 1997), these two sets of φ -features must match after the syntax when a pronoun is inserted by the morphology. Egerland proposes that impersonal pronouns of the Class 1 type like Swedish *man* and Italian Imp_{si} are deficient elements that do not have any lexical φ -features. Egerland assumes that they do have a [+human] semantic feature.

I follow Egerland in assuming that Class 1 impersonals are deficient and lack lexical φ -features. I further propose that Class 1 impersonals, which have reached the end of the ISC, are the spell-out of a T head. Recall from the discussion in Chapter 1 that reanalysis as a T head is the expected outcome of the SAC. Since the ISC is a subtype of SAC, this makes sense. This differs from previous work (Maddox 2017), in which I claimed that impersonal pronouns like $\text{ModF Imp}_{\text{on}}$ are the spell out of a Voice or *v* head. However, since the impersonal pronouns that have completed the ISC are actually subject pronouns that only occur in the nominative, they are relevant to the T rather than the *v* domain.³⁷ Furthermore, Imp_{on} does not change the valency of the verb. On the other hand, Spanish Pass_{se} does change valency, so it is fitting that it be reanalyzed as a Voice head.

Thus, Class 1 impersonals like $\text{ModF Imp}_{\text{on}}$ are T heads and, following Mendikoetxea (2008), they license a *pro* in Spec, v , where it receives the agent theta-role. Imp_{on} has interpretable T-features and uninterpretable φ -features, which are checked when a *pro* or other overt DP merges. A generic subject reading is induced if a generic operator is present. When the generic operator is absent, an arbitrary or first-person interpretation may arise. If there is a [+human] feature involved, I assume it is merely lexical and, consequently, not involved in the derivation. The structure of $\text{ModF Imp}_{\text{on}}$ is given in (55) below.

³⁷ I thank Elly van Gelderen for pointing this out to me.

(55) ModF Imp_{on}: [TP T_{on} [vP *pro v* [VP V DP]]]

OldS *omne* is structurally distinct from ModF Imp_{on} since it has not completed the ISC; i.e., it is a full DP. OldS *omne* has interpretable φ -features. It merges in Spec,*v* and then may be subjected to further movement to Spec,*T* or other specifier positions. This is how it gets separated from the verb. Similar to Imp_{on}, if a generic operator is present, OldS *omne* may receive a generic interpretation, but it is never arbitrary. The structure in (56) below shows the full DP *omne* moving through specifiers to be a topic in CP.

(56) OldS *omne*: [CP *omne* C [TP <*omne*> T [vP <*omne*> *v* [VP V DP]]]]

2.3.1.2 Renewal

Once an impersonal pronoun has been reanalyzed as a T head, another element may merge where the original full DP pronoun originally merged, in Spec,*v*; i.e., renewal. In Chapter 1 it was shown that clitic doubling is a type of renewal. Recall, for example, that Modern Colloquial French (MCF) is at stage (c) of the SAC, and so subject clitics are agreement. As such they can be doubled by another pronoun without an intonational break, as in (57) below from van Gelderen (2011:53).

(57) Moi je suis un bloggeur.

me I am a blogger

‘I am a blogger.’

This doubling is allowed because *je* is fully grammaticalized as a T-head in MCF. If ModF Imp_{on} is also fully grammaticalized, there should be instances where it too may be doubled. In MCF, doubling of Imp_{on} does occur.³⁸ Consider the following data taken from the *Corpus d’entretiens spontanés*:

³⁸ Bahtchevanova & van Gelderen (2016:119) also discuss doubling of *on* by *nous*, concluding that *nous* is the “real pronoun” and *on* is agreement.

(58) **Nous on** a quand même plus de liberté.

we Imp has really more of freedom

‘We really have more freedom.’

(59) et (ils) partent euh l’été euh au bord des plages comme comme **nous on** fait.

and they start uh the-summer uh at-the side of-some beaches like like we Imp does

‘And they uh start the summer uh next to some beaches like we do.’

(60) Oui c'est ça nous, **nous on** s'entend bien.

yes it-is that we we Imp Pron_{se}-get-on well

‘Yes that's it we, we get along well.’

In (58) through (60), the *on* is doubled by *nous* with no prosodic pause and inside the same clause. This suggests that *on* is functioning as agreement rather than as an argument. However, there are also instances where a pause follows *nous* and *on* occurs afterward as in the data below:

(61) Mais **nous, on** fabrique tout nous-même.

but we Imp makes everything we-self

‘But we, we make everything ourselves.’

(62) c'est le travail que que, que **nous, on** ferait pas quoi.

it-is the work that that we Imp would-do not what

‘It's the work that that, that we, we wouldn't do at all.’

The data (61) and (62) then are instances of clitic-left dislocation rather than agreement.

Doubling may be even more advanced in Québécois French. The data in (63) below is a lyric by French-Canadian singer Renée Martel where, again, there is no pause in between *nous* and *on*.

(63) **Nous on** aime la musique country.

we Imp love the music country

‘We love country music.’

Many additional examples of this kind of doubling can be found in the *Corpus de français parlé au Québec*.

One might object that *nous* is a clitic and that doubling by another clitic does not qualify.

However, *nous* may actually be less grammaticalized than the first- and second-person clitics in French. For example, it can be separated from the verb as in (64) below. In (65), *nous* is separated from the verb and modified by a relative clause.

(64) Nous, jeunes et étudiants, sommes les oubliés de cette présidentielle.³⁹

we young and students are the forgotten of this presidential-election

‘We, the young and the students, are the forgotten ones of this presidential election.’

(65) C'est nous qui avons déclaré la guerre!⁴⁰

it-is we who have declared the war

‘It's is we who declared the war! ’

Furthermore, *nous* can be coordinated as in (66) and (67).

(66) Nous et nos amis sommes intelligents.

we and our friends are intelligent

‘We and are friends are intelligent.’

³⁹ http://www.lemonde.fr/campus/article/2017/04/21/nous-jeunes-et-etudiants-sommes-les-oubliés-de-cette-présidentielle_5115356_4401467.html

⁴⁰ <https://www.les-crisés.fr/pierre-conesa-cest-nous-qui-avons-declare-la-guerre/>

(67) a. Nous et vous sommes amis.

we and you are friends

b. *Je et tu sommes amis.

I and you are friends

Thus, Imp_{on} , although it does not receive a generic interpretation in these data, can be doubled by *nous* and so renewal is taking place. *Nous* is contributing the person features that have been lost by *on*. Additionally, the licensing of *pro* in Spec, v by *on* as represented in (55) can be viewed as another type of renewal, albeit with an unpronounced element.⁴¹

2.3.1.3 Summary of the Stages of the ISC

The three stages of the Impersonal Subject Cycle are summarized below, with *homō* representing any impersonal pronoun derived from a *man*-related lexical item; i.e., *omne* in Spanish, *on* in French.

(68) Impersonal Subject Cycle (ISC)

Stage (a): $\text{Imp} = \text{DP}, [\text{TP } homō \text{ T} [\text{vP } <homō> \text{ v} [\text{VP}]]] \rightarrow$ (Old Spanish, Old French)

Stage (b): $\text{Imp} = \text{D-clitic}, [\text{TP } \text{DP } homō + \text{T} [\text{vP } <homō> \text{ v} [\text{VP}]]] \rightarrow$ (Modern French)

Stage (c): $\text{Imp} = \text{T}, [\text{TP } \text{DP } T_{homō} [\text{vP } <\text{DP}> \text{ v} [\text{VP}]]] \rightarrow$ (Modern Colloquial French)

The stages of the ISC are essentially identical to those of the SAC as discussed in Chapter 1. At stage (a), the impersonal pronoun is a full DP that merges in argument position, Spec, v , to receive the theta-role. It may undergo other movement afterwards such as to Spec, T for EPP. At stage (b), the impersonal clitic merges as a DP but moves as a D-head. Thus, it satisfies theta-role assignment but moves as a head to T. In the third stage, the impersonal clitic is reanalyzed as the spell-out of T.

⁴¹ This type of renewal by *pro* is similar to that proposed by MacDonald & Maddox (2018) for *Pass_{se}* in Spanish and Romanian.

2.3.2 The Subject Agreement Cycle and the Impersonal Subject Cycle

In previous sections I have referred to the ISC as a sub-type of SAC, but what exactly is the relationship between them? In this section I consider more directly how these two cycles are connected. As was seen in Section 1.3.2, the SAC is advanced in ModF and nearly complete in MCF. However, I show below that in Spanish the SAC is still at stage (a). As for impersonal pronouns, in ModF Imp_{on} is fully grammaticalized while in Spanish the grammaticalization of the HOMŌ-pronoun equivalent *omne* froze and *omne* has been lost in ModS. These historical patterns lead to the generalization that the ISC can only advance in a given language if the SAC has moved past stage (a). In other words, the ISC piggy backs off the SAC. If subject pronouns in a language are not being reanalyzed, the HOMŌ-derived pronoun will not be reanalyzed either. This is why in ModF Imp_{on} is at stage (c) in the ISC and OldS *omne* disappeared from the lexicon.

2.3.2.1 SAC in Spanish

If we compare the developments that took place from Latin to OldF to ModF with the history of “standard” Spanish with respect to the SAC, it appears that not much has happened.⁴² Consider the following paradigm:

(69)	Latin	Old Spanish	Modern Spanish
	CANTŌ	canto	canto
	CANTĀS	cantas	cantas
	CANTĀT	canta	canta
	CANTĀMUS	cantamos	cantamos
	CANTĀTIS	cantades	cantáis
	CANTĀNT	cantan	cantan

While there has been some attrition from Latin to OldS such as the loss of final /t/ in the third-persons singular and plural, OldS retained distinct endings. ModS also has different endings for

⁴² There is evidence that some Caribbean varieties of Spanish are undergoing change in this respect; v. Gupton & Lowman (2013).

every person. Thus, there are no major morphophonological changes pointing to the progress of the SAC in Spanish as there are in French.⁴³

As for distributional changes, again OldS and ModS are very similar. For example, in OldS the subject pronouns were full DPs, similar to OldF, as shown in the following data:

(70) mas que tu mismo que guardas la ley...⁴⁴ (71) senor si yo esto todo les di...⁴⁵
 but that you self that you-guard the law sir if I this all them give
 ‘but that you yourself guard the law...’ ‘Sir, I if give them all of this...’

In (70), the subject pronoun *tu* is separated from the verb by the intensifier *mismo* and the complementizer *que*, while in (71), *yo* is separated from the verb by a quantifier and a DP. The situation in ModS is the same as shown in the following data:

(72) Tú, mi amor, no tienes razón. (73) yo también quiero ir.
 you my love not you-have reason I also I-want to-go
 ‘You, my love, are not right.’ ‘I want to go too.’

Additionally, in both OldS and ModS subject pronouns can be coordinated and modified:

(74) que mientras yo & tu fueremos vivos...⁴⁶ (75) e yo que mandase...⁴⁷
 that while you and I may-be living and I who ordered
 ‘that while you and I may be alive...’ ‘And I who ordered...’

(76) cuando tú y yo viajamos (77) y yo, que andaba por allí
 when you and I travel and I who walked by there
 ‘When you and I travel...’ ‘And I, who was walking there...’

⁴³ The development of the first- and second-person plural pronouns, *nosotros* and *vosotros*, may be a type of renewal, but whether this is due to the SAC merits further study. See Penny (2002:138).

⁴⁴ Anonymous, *El Nuevo Testamento según el manuscrito escurialense I-J-g*; 1260

⁴⁵ Alfonso X, *Estoria de España, II*; 1270-1284

⁴⁶ Alfonso X, *General Estoria, Segunda parte*; 1275

⁴⁷ Anonymous, *Leyes nuevas*; 1255-1280

The OldS in (74) and ModS in (76) show coordination while OldS (75) and ModS (77) show modification by a relative clause.

OldS and ModS also pattern together with respect to omission of subject pronouns:

(78) con quanta devoción yo puedo, suplico a Vuestra excelencia...⁴⁸

with as-much devotion I can beg to your excellency

‘with as much devotion as I am able, I beg Your Excellency...’

(79) yo debo y puedo escribir...⁴⁹

I should and can write

“I can and I should write...”

The sentence in (78) shows that in OldS the subject pronoun *yo* could be left out in the second clause. (79) is acceptable in both OldS and ModS.

Finally, whereas in MCF the subject pronouns could be doubled, in OldS I find no evidence of this, nor is it acceptable in ModS, except when prosodically marked.

(80) *Mí yo pienso que...

me I think that

(81) Yo, Juan, pienso que...

I Juan think that

In (80) the stressed pronoun *mí* is unacceptable. The stressed pronoun is used since this is the kind of doubling that would be expected if the SAC were at stage (c), as with ModF *moi* doubling *je*. Thus, throughout its history, the subject pronouns in Spanish have remained stable as full DPs. This is surprising when compared to French where abundant change has taken place. The primary difference here is that in French, the SAC has progressed quite far, turning subject pronouns into subject agreement morphemes. Spanish, on the other hand, is still at stage (a) of this cycle.

⁴⁸ Anonymous, *Cancionero de Juan Fernández de Ízar*; 1424-1520

⁴⁹ Alonso de Santa Cruz, *Crónica del Emperador Carlos V*; 1550

Returning now to the status of impersonal pronouns in each language, a pattern emerges. Recall that in OldS *omne* was always a full DP prior to its disappearance. In French, however, Imp_{on} was reanalyzed from a full DP to a head. The central question this chapter is attempting to address is: why did grammaticalization continue in French but freeze in Spanish? Based on the above discussion a likely explanation is that since other personal subject pronouns like *je* and *tu* in French were undergoing the SAC, the impersonal subject pronoun *on* could also participate in this cycle. In fact, it appears that personal pronouns have to be reanalyzed before impersonal pronouns can be. While it is true that third-person pronouns lag behind, as pointed out by van Gelderen (2011) and Bahtchevanova & van Gelderen (2016), they are still undergoing grammaticalization. On the other hand, in Spanish the SAC has remained at stage (a) since the earliest documentation. All subject pronouns in ModS are still full DPs. Consequently, it is no surprise that *omne* was not reanalyzed as a head because none of the other subject pronouns were reanalyzed either.

2.3.2.2 Generic *uno* and the ISC⁵⁰

Generic *uno*, which was discussed above in Section 2.2.2.1, first starts to appear in Spanish during the 16th century. Prior to then it could occur as a numeral, a pronoun, and an indefinite article. During the same period OldS *omne* disappears and impersonal *se*, which was on its own cycle, as discussed in Chapter 4, becomes fully grammaticalized. Generic *uno* displays properties similar to OldS *omne*. For example, it is a full DP that can license *pro* in subsequent clauses as in the data below:

⁵⁰ Some data in this section adapted from Company Company & Pozas Loyo (2009).

(82) *uno_i* se govierna en su manera de bevir conforme al estado y

Imp Refl_{se} governs in his way of living according to-the state and
condición que *pro_i* tiene...⁵¹

condition that Imp has

‘One governs himself in his way of living according to the state and condition that he
has...’

(83) porque más facilmente menospresia *uno_i* lo que *pro_i* vee con los ojos...⁵²

because more easily despises Imp it that Imp sees with the eyes

‘because one more easily despises that which he sees with his eyes...’

Since generic *uno* patterns like the now defunct OldS *omne*, it appears that *uno* has replaced *omne* on the ISC. In fact, generic *uno* is still undergoing grammaticalization. One development that has occurred is that it can now have a first-person reference as in (84) below:

(84) “Voy a arreglar la casa. Hay cosas que debe uno hacer.”⁵³

I-go to fix the house there-are things that should one do

‘I’m going to fix up the house. There are things one must do.’

Company Company & Pozas Loyo (2009) claim that this is the latest step in the evolution of this impersonal pronoun. They point out that this same development took place with OldS *omne* as well, but only toward the end of its existence and only in the genre of rustic drama. In (85) below *el hombre* is the equivalent of *omne*.

⁵¹ Juan de Valdés, *Diálogo de le Lengua*; 1535

⁵² Antonio de Guevara, *Menosprecio de corte y alabanza de Aldea*; 1539

⁵³ Diario, *Reforma*; 15-11-2006

(85) ¡Ah cuerpo de San Antón, Cómo stá el hombre acosado!⁵⁴

oh body of Saint Anthony how is the man harassed

‘Oh Saint Anthony’s body! How one is harassed!’

Thus, generic *uno* is now going through the same changes that OldS *omne* went through and at least one sociolinguistic study by Flores-Ferrán (2009) shows that there is additional dialectal variation vis-à-vis the interpretation of *uno*. All this suggests that generic *uno* has now replaced OldS *omne* on the ISC. However, given the generalization formulated in the preceding section, generic *uno* will not be reanalyzed because Spanish subject pronouns have not been reanalyzed.

2.3.2.3 A prediction

In French, the HOMÔ-derived impersonal pronoun went from being a full DP, to a head, to the spell-out of T. This is the expected outcome of the ISC. In Spanish, however, *omne* was never reanalyzed and its grammaticalization was cut off, with generic *uno* now replacing it on the ISC. Additionally, in French subject pronouns are being turned into subject agreement morphemes on the verb as a result of the SAC. This cycle has not progressed past stage (a) in Spanish, where the subject pronouns are still full DPs. These patterns point to the prediction in (86) below:

(86) If the personal subject pronouns in a language are undergoing the SAC, the impersonal subject pronoun(s) may also participate in it. If the personal subject pronouns in a language are not subject to the SAC, neither will the impersonal pronoun(s) in that language be. The SAC affects personal subject pronouns prior to impersonal ones.

This prediction holds for the two languages that have been discussed throughout this chapter up to this point, French and Spanish.

⁵⁴ Juan del Encina, *Auto del repelón*; c. 1529

2.4 Chapter Summary

In this chapter I focused on the Old Spanish impersonal pronoun *omne* derived from Latin HOMŌ. I compared it with French impersonal *on* by means of diagnostics adopted from Egerland (2003), van Gelderen (2016), and an original diagnostic of cross-clausal licensing to *pro* via topic-identification to show that while in French the Impersonal Subject Cycle is nearly complete in Spanish it is frozen and OldS *omne* was ultimately replaced by Spanish generic *uno*. OldS *omne* remained a full DP throughout its history while Imp_{on} was reanalyzed as a T head. I proposed three stages of the Impersonal Subject Cycle which is a linguistic cycle (a subtype of the Subject Agreement Cycle) that takes impersonal pronouns and turns them into impersonal subject marking affixes. The ISC failed to complete in Spanish because the Subject Agreement Cycle did not target subject pronouns for reanalysis, which is what would also have led to the reanalysis of OldS *omne*. In French, on the other hand, the SAC is quite advanced and so Imp_{on} has been reanalyzed. This leads to the prediction that if the SAC is underway in a language, the ISC may also take place; i.e., if subject pronouns are reanalyzed, so too will the impersonal pronoun. Alternatively, if a language has an impersonal subject marking affix, it should also have subject agreement morphology derived from subject pronouns. This prediction holds in French and Spanish, though other languages need to be examined as well to determine its accuracy.

Chapter 3: Object Agreement and Object Movement in Spanish

Introduction

In Chapter 1, subject clitics in Modern Colloquial French (MCF) were shown to be an example of a stage (c) of the Subject Agreement Cycle (SAC). Chapter 3 is dedicated to object clitics in Spanish, specifically direct object (DO) clitics as in (1); additional data in (2) and (3):¹

(1)	Singular	Plural
1	<i>me</i>	<i>nos</i>
2	<i>te</i>	<i>os</i>
3	<i>lo</i> (M) / <i>la</i> (F)	<i>los</i> (M) / <i>las</i> (F)
(2)	Juan te ama.	(3) Juan la ama.
	Juan you loves	Juan her loves
	'Juan loves you.'	'Juan loves her.'

Object clitics in Spanish are an example of another linguistic cycle, the Object Agreement Cycle (OAC). While the DO clitics in all varieties of “standard” Modern Spanish (ModS) are more grammaticalized than their Latin ancestors, some varieties of Modern Spanish like Rioplatense, are closer to full grammaticalization than the standard, as is discussed below.

In this chapter I show how DO object clitics in Spanish have been reanalyzed from DP to D to v -heads as part of the OAC. I argue further that this process has been encouraged by patterns of object movement which were frequent in Latin and Old Spanish (OldS) but are highly constrained in ModS; i.e., object movement feeds the OAC. This chapter is organized as follows: In Section 3.1 I discuss patterns of object clitic placement in OldS that suggest OldS

¹ The clitic system in (1) is the “etymological” one present in “standard” Spanish. Other systems such as *leísta*, *laísta*, etc., are not considered here. See Ormazabal & Romero (2013) for a discussion of different varieties of Peninsular Spanish.

object clitics were full DPs. I then apply van Gelderen's (2011) diagnostics in order to further bolster this claim. Accusative clitic doubling is discussed in Section 3.2 and, following van Gelderen (2011), I adopt it as a diagnostic for evaluating which stage in the OAC a language is at. I adopt a mixed-analysis following Kramer (2014) and I propose that clitic doubling can actually occur at different stages of the OAC, but that early doubling is derivationally distinct from late doubling. The focus of Section 3.3 is patterns of object movement such as scrambling and object shift and how they feed the reanalysis of clitics. First I give some background on different types of object movement and then analyze data from OldS to determine whether object movement in OldS was scrambling or object shift. A prediction comes out of the analysis whereby if a language allows object movement and has object clitics, those clitics will start to occur in clitic doubling constructions. This chapter is summarized in Section 3.4

3.1 Categorial status of object clitics in Spanish

In this section I show that object clitics have changed their categorial status in the history of Spanish. In OldS they were full phrases (DPs) while in ModS they are heads. Evidence for this comes from patterns of clitic placement. In OldS object clitics could be separated from the verb by other maximal projections (interpolation), they had the same distribution as full phrases, and accusative clitic doubling was rare and optional. These patterns no longer hold in ModS. After going through these data I present previous analyses that bear on the status of Spanish object clitics. I then apply van Gelderen's (2011) diagnostics for distinguishing between phrases and heads to Latin and Spanish, adding further support for the DP status of object clitics in Latin and OldS and the head status of object clitics in ModS.

3.1.1 *Object clitic placement in Old Spanish*

3.1.1.1 Interpolation

One of the most unique properties of object clitics in OldS when compared to ModS is that in OldS they could be separated from the verb, as in (4) below.²

(4) si **lo** non fiziere, non erede.³

if it not he-does not inherits

‘If he does not do it, he does not inherit.’

(5) pero que-**lo** non fallamos en toda la estoria que auemos contada.

but that it not we-find in all the story that we-have told

‘But we do not find it in all the story that we have told.’

(6) Ca si-**la** tu non amparas.

because if-it you not protect

‘Because if you do not protect it...’

While interpolation by negation is the most frequent pattern, other elements may also intervene between the clitic and verb, as the following data show.

(7) quien **te** algo prometiere...

who you something would-promise

‘the one who would promise something to you...’

(8) Busca todas buenas vias...para **lo_i** mejor complir que **lo_i** ella non mando.

searches all good ways to it better accomplish than it she not ordered

‘He looks for all the good ways...in order to accomplish it better than she ordered it.’

² Examples (5), (6), and (9) are from Fontana (1993); (7) and (8) are from Rivero (1986).

³ Anonymous, *Fuero de Cuenca*, para. 100; 1284-1295

(9) porque-**te** assi encerreste?

why-you thus locked-up

‘Why did you lock yourself up like this?’

In (7), the pronoun *algo* separates *te* from the verb, while in (8) the clitic is separated by a pronoun and negation. An adverb is the intervening element in (9).

3.1.1.2 Parallel distribution with DPs

OldS object clitics also share the same distribution as full DPs. Consider the following data.⁴

(10) a. El infante ovo respuesta del rrey.

the prince had answer from-the king

‘The prince had an answer from the king.’

b. El infante **esta rrepuesta** ovo del rrey.

the prince this answer had from-the king

‘The prince had this answer from the king.’

(11) a. El rrey recibio-lo muy bien.

the king received-him very well

‘The king received him very well.’

b. Ellos **lo** entendieron.

they it understood

‘They understood it.’

In (10) a full DP occurs preverbally and postverbally. The object clitic in (11) does the same.

Furthermore, both full DPs and object clitics can climb as in (12) below.

⁴ Data in (10) to (12) are from Rivero (1986).

(12) a. quando...Jesuchristo **las sus divinales bodas** quisyere celebrar.
 when Jesus-Christ the his divine wedding would-wish to-celebrate
 'When Jesus Christ would wish to celebrate his divine wedding.'

b. antes que la queria complir.
 before that it he-wanted to-fulfill
 'Before he wanted to fulfill it.'

3.1.1.3 Accusative clitic doubling

OldS accusative clitic doubling (ACD), as in (13) and (14) below, was rare.⁵

(13) Priso lo_i al conde_i.
 he-took him DOM-the count
 'He took the Count.'

(14) sy de otra guisa la_i matare a ella_i...⁶
 if of other fashion her he-kill DOM she
 'if in another way he should kill her...'

When ACD does occur it appears to be optional since pronominal objects occur most often without a clitic double as in (15) and (16) below.

(15) Otrosy, sy matare a él & non a la muger...⁷
 however if he-kill DOM he and not DOM the woman
 'However, if he shall kill him and not the woman...'

⁵ Example (13) is from Rivero (1986).

⁶ Anonymous, *Fuero de Úbeda*; 1251-1285

⁷ Anonymous, *Fuero de Úbeda*; 1251-1285

(16) que quando mataron ami padre que llagaron a él de muerte.⁸

that when they-killed DOM-my father that they-wounded DOM he of death

‘that when they killed my father that they injured him mortally.’

The datum in (17) below is particularly interesting since it shows that the same object can be realized as a stressed pronoun in the first clause and as an enclitic in the second clause.

(17) y tomó a mí_i en la boca & llevó-me_i al monte.⁹

and she-took DOM me in the mouth and took-me to-the mountain

‘and she (the lioness) took me in her mouth and took me to the mountain.’

In ModS, on the other hand, ACD is obligatory with pronominal objects, as in (18) below.

(18) a. Si *(*lo_i*) mata a él_i... (Modern Spanish)

if him kills DOM he

‘If he kills him...’

b. Juan *(*me_i*) llevó a mí_i al aeropuerto...

Juan me took DOM me to-the airport

‘Juan took me to the airport.’

3.1.2 Previous analyses

As part of the OAC, pronouns change from full DPs to D-heads to *v*. In this section I review the following studies bearing on the categorial status of Spanish object clitics historically: Rivero (1986, 1991), Fontana (1993), Fischer et al (2016, 2018), Navarro et al (2017), and Vilanova et al (2018), and Franco (1993).

⁸ Anonymous, *Cuento muy fermoso de Otas de Roma*; c. 1300-1325

⁹ Anonymous, *Libro del cavallero Cifar*; 1300-1305

3.1.2.1 Rivero (1986, 1991)

Rivero argues that OldS object clitics are full XPs that occur in A-positions. They merge as complement to V and do not license *pro*. Syntax treats them as XPs; they cliticize at PF. In ModS, object clitics are base-generated heads.¹⁰ Rivero reaches these conclusions based on patterns of interpolation and parallel distribution with other XPs, as discussed in Section 3.1.1 above. On Rivero's analysis, in interpolation structures the XP clitic is moved from its first-merge position and adjoined to the complement of CP. This is a type of focalization which also affects NPs and PPs in OldS. The complement of CP is IP in affirmative sentences and NegP in negative sentences. ACD in OldS differs significantly from ModS, though they can appear identical on the surface. The clitic is the complement of V and the doubling XP is base-generated in an A'-position and adjoined to another maximal projection. Rivero analyzes the doubling phrase as a topic- or focus-type constituent that is licensed via predication or quantification.

Rivero proposes a “Critic Parameter” to account for differences between OldS and ModS. In OldS the parameter has the effect such that clitics are dependent on a host only at PF. In ModS the parameter applies in the morphological component, by which Case is transferred from the verb to the clitic. A possible gap in Rivero's analysis is that she does not explain adequately how the clitic changed from an XP in OldS to a head in ModS. She attributes the difference to a parameter that is set differently in either variety of Spanish. In OldS it is only effective at PF while in ModS it operates at word-formation, i.e., morphology. This is an unusual approach to parameters especially given the current understanding of them in the Minimalist Program, sometimes referred to as the “Borer-Chomsky Conjecture” as in (19) below.

¹⁰ Data in this section adapted from Rivero (1986) and references therein.

(19) Borer-Chomsky Conjecture (Baker 2008)¹¹

All parameters of variation are attributable to differences in the features of particular items (e.g. the functional heads) in the Lexicon.

3.1.2.2 Fontana (1993)

Fontana builds on Rivero (1986, 1991) and offers a more explanatorily adequate account for the different categorial status of clitics in OldS and ModS (particularly from a diachronic perspective). On Fontana's account, reanalysis of clitics is the result of a conspiracy of two independent developments: a) loss of V2 syntax and b) loss of V1 declaratives. Fontana argues that OldS clitics are second place (2P) clitics since they parallel Homeric Greek (HG) clitics (Taylor 1990). Fontana, like Rivero, also takes patterns of interpolation and optional/rare doubling as support for OldS clitics as phrases. Many analyze the obligatory co-occurrence of object clitics and pronouns to indicate that the clitic itself is not the argument but rather an affix-like element (Jaeggli 1982, Borer 1984, Suñer 1988, *inter alios*).

Fontana's analysis of how clitics changed status from OldS to ModS relies on the loss of V2 syntax and V1-declaratives. The datum in (20) below is a putative example of OldS V2 and (21) is a V1-declarative.

(20) esto fazien por razon de los ricos omnes.

this they-did for reason of the rich men

'They did this because of the rich men.'

(21) & fizo el papa penitencia & dixo Sant Antidio la missa en su lugar...

and did the pope penitence and said Saint Antidius the mass in his place

'And the Pope did penance and Saint Antidius said the mass in his place...'

¹¹ The BCC and its relation to the Null Subject Parameter is also discussed in Roberts & Holmberg (2010:32).

Fontana's analysis is as follows. First, OldS clitics in interpolation structures are NP/DP-arguments that adjoin to the right or left of the first XP position dominated by IP; or they can move into Spec,I. Fontana also provides data showing that OldS was a symmetric V2 language.¹² V moves to I, the subject remains VP-internal, and the first position element is topicalized via movement out of VP (if a direct object) to Spec,I. In V1-declaratives/Narrative Inversion, V moves to C, the subject moves to Spec,I and the object remains in situ. Some verb-initial structures can occur in ModS, but this is not the same as narrative inversion in OldS. In ModS, V moves to I and not to C as in OldS.

Fontana notes that IO doubling clearly increases before DO doubling, but there are examples of DO doubling even in the 15th century which must be explained. Fontana (1993:268) posits a dual system in which some speakers may analyze the clitic as either a head or a full DP; i.e., the two coexist in the speakers' grammar. Such overlap is not unexpected in linguistic change (Kroch 1989, Santorini 1989, Pintzuk 1991). OldS clitics are second place (2P) clitics that cannot occur in first position. V2 syntax ensures that a different XP occurs in first-position. V1 declaratives also keep clitics from occurring in first-position. Loss of V2 syntax entails loss of topicalization to Spec,I. There is also a steady decrease of V_[+finite]-Cl configurations from the 14th century on that coincides with the increase in Cl-V_[+finite] structures, which means I-to-C movement is lost. With these developments, clitics end up in first position more regularly with nothing preceding them other than perhaps a conjunction. Consequently, learners reanalyze them as being proclitic on the tensed verb rather than enclitic on whatever occupies first-position. This tight association between clitic and verb then encourages reanalysis of the clitic as the realization of the verb's inflectional features.

¹² See Fontana (1993) Chapter 3 for evidence in favor of Old Spanish as a V2 language.

Fontana's (1993) study is groundbreaking since it is the first to offer a formal diachronic explanation for the grammaticalization of clitics in Spanish. Nevertheless, it is not without its problems. First, Fontana's analysis rests on the status of OldS as a V2 language but this has been challenged (Sitaridou 2012, 2016). Also, an implication is made that loss of V2 and V1 declaratives will have to happen before clitics are reanalyzed as heads. This should be examined in other varieties of Romance as well as other languages to see if it holds true. While Fontana does observe that indirect object doubling develops prior to ACD, he tends to lump these clitics together in other respects. Since at least Suñer (1988), Spanish IO clitics have been shown to pattern differently than DO clitics with respect to doubling. Additionally, authors such as Ormazabal & Romero (2013) have shown that even within the DO clitic paradigm, first- and second-person are much less restricted than are third-person clitics in doubling constructions.

3.1.2.3 Fischer et al (2016, 2018), Navarro et al (2017), Vilanova et al (2018)

Fischer, Navarro, and Vilanova investigate the grammaticalization of object clitics and Spanish dialectal variation based on quantitative/corpus data on doubling. They carry out a comparative analysis with Catalan and connect reanalysis of clitics to verb movement.¹³ Vilanova et al's (2018) study focuses on general word order change as a contributing factor to the rise of clitic doubling. Their data show that doubling of pronominal objects was optional in OldS (up until the 15th century). They show that obligatory doubling of pronominal objects develops from the 16th to the 20th century.

These authors also analyze data from non-peninsular varieties of Spanish. In Buenos Aires Spanish pronominal DOs are obligatorily doubled and animate nominal DOs are optionally doubled (22).

¹³ Data in this section adapted from Fischer et al (2016, 2018) and Navarro et al (2017) and references therein. These authors also provide data from Catalan which will not be discussed here.

(22) *Las_i* saludé a las maestras_i del jardín.

them I-greeted DOM the teachers from-the garden

‘I greeted the teachers from the kindergarten.’

Fischer et al (2016, 2018) is the first study I am aware of to incorporate data from Judeo-Spanish, where animate definite DOs are doubled (23).

(23) La hija hazina *la_i* yamó a la madre_i, después di kuarenta días.

the daughter sick her called DOM the mother after of forty days

‘The sick daughter called her mother after forty days.’

In Lima Spanish and Andean Spanish indefinite (24) and inanimate DOs (25) can be doubled.

(24) *Lo_i* saludé a un estudiante_i que conozco.

him I-greeted DOM a student that I-know

‘I greeted a student that I know.’

(25) a. *Lo_i* vendo toditos los carros_i.

him I-sell all the cars

‘I sell all the cars.’

b. Eso también *lo_i* mata las plantas_i.

that also him kills the plants

‘That also kills the plants.’

In (32) the clitic does not agree in number or gender with its coindexed DP. Additionally,

Fischer et al do not point this out but in (25a) a quantified DP is doubled by a clitic which, per van Gelderen (2011), is evidence for the final stage of grammaticalization.

Fischer et al propose that clitic doubling is a cycle comprised of five stages ranging from no doubling at all, as in Latin and Proto-Romance, to generalized doubling including animate

DOs as in Lima and Andean Spanish. The categorial status of the clitic is related to the doubling possibilities, but so too is the availability of verb movement because it provides a position for A'-movement. As doubling becomes less restricted, verb movement becomes more restricted. Word-order also becomes much less flexible and “does not convey discourse information” (Fischer et al 2018:15). In sum, doubling takes over for verb movement as an influencer of information structure (the authors do not go into detail as to what specific information structure they are referring; i.e, topic, focus, something else?). Vilanova et al (2018) build on Navarro et al (2017) and Fischer et al (2018) using mostly the same data. Here they propose that changes in word order and grammaticalization status of the clitic motivate the development of doubling. In Old Romance, word order was more flexible as it was tied to information structure; as word order becomes less flexible, doubling develops as a replacement.

Fischer et al (2016, 2018), Navarro et al (2017), and Vilanova et al (2018) analyze data from a corpus that is wide-ranging both diachronically and synchronically. However, their claim about datives contradicts what most other studies have observed with respect to chronology. They also group DO and IO clitics as if they are both participating in the same cycle, though the same authors are currently looking into this in more detail as part of a larger project (S. Fischer, p.c.; March 5, 2018). The claim that doubling replaces verb-movement for information structure rests on the assumption that these two phenomena are semantically equivalent for discourse purposes. The same can be said for Vilanova et al’s (2018) claim about word order. Unfortunately, the authors have not shown this.

3.1.2.4 Franco (1993)

In Franco’s (1993) dissertation Spanish object clitics are analyzed as agreement morphemes rather than pronominal arguments. He maintains AgrO in his analysis and takes

object clitics to be AgrO heads that project AgrO phrases. Franco provides morphological evidence for clitics as agreement: fixed order, strict adjacency to host, variation in agreeing features, etc. He argues that the doubled XP originates in argument position, so the clitic must be a head. Franco's analysis is synchronic, but it makes sense diachronically as well when situated within the OAC.

Let us examine some of Franco's morphological diagnostics. First, inflectional affixes, unlike lexical items, are very restricted with respect to the types of constituents they can be adjacent to.¹⁴ Consider the following data:

(26) a. Acogen-*se-le* omnes de todas partes. (Old Spanish)
b. *Se-* *le-* acogen hombres de todas partes. (Modern Spanish)

Refl_{se} Dat.3S join men from all parts
‘Men join him from everywhere.’

(27) a. Juan *se-* *lo-* mandó. (28) a. Juan ya *lo* vio.
Juan Dat.3S Acc.M.3S sent Juan already Acc.3S saw
‘Juan sent it to him.’ ‘Juan already saw it.’

b. *Juan *lo se* mandó. b. *Juan *lo ya* vio.

(29) a. Alfredo *la-* trajo del Perú.
Alfredo Acc.F.3S-brought from-the Peru
‘Alfredo brought her/it from Peru.’

b. *¿La-* trajo Alfredo del Perú de verdad?
c. **¿Trajo* Alfredo *la* del Perú de verdad?

¹⁴ Data in this section adapted from Franco (1993) and references therein.

Example (26) shows that the Tobler-Mussafia Law is no longer in effect in ModS; i.e., clitics cannot appear in second position which was an option in OldS. ModS clitics follow a fixed order with respect to the verb and other clitics as in (27). Furthermore, strict adjacency to host is obligatory in ModS (28). And finally, (29) shows that clitics form a single syntactic item with the verb. In Franco's (1993) analysis accusative clitics head the AgrO projection.¹⁵ AgrO has strong NP features which force clitic-doubled NPs to raise overtly to Spec,AgrO, where they check Case. The strong NP features on AgrO are related to the Animacy Hierarchy: [+animate/specific/referential, etc.]. Thus only NPs bearing these features will move to Spec,AgrO. Franco proposes that Spanish also has an AgrO with weak NP features headed by a phonological null clitic. In this case the object NP raises covertly to Spec,AgrO.

3.1.3 *Application of van Gelderen's (2011) diagnostics in Latin and Spanish*

Having summarized earlier studies on the categorial status of OldS clitics in the preceding section, we are now in a position to build upon that by applying van Gelderen's (2011) diagnostics for determining at which stage of the OAC a clitic is at. First- and second-person object clitics in Spanish are derived from the accusative forms of the first- and second-person "determinative pronouns" in Latin. Third-person object clitics in Spanish are derived from the accusative forms of the third person demonstrative pronouns in Latin. Demonstrative pronouns in Latin can be coordinated. In (30) the third-person masculine singular and the first-person singular are coordinated. In (31) the third-person feminine singular and a DP are coordinated.

¹⁵ Franco (1993) actually labels it AgroDO to contrast it with an AgroS head for subjects and an AgroIO for indirect objects.

(30) et **illum et me** vehementer ignoras.¹⁶

both him and me vehemently not-know

"Both him and me you vehemently do not know."

(31) Vidi, **et illam et hospitem**, complexam atque osculantem.¹⁷

I-saw and her and guest embraced.FS and kissing.MS

'I saw both her and the guest; her embraced and him kissing.'

In (31) above *illam* is modified by the participle *complexam*. In (32) below *illum* is modified by a relative clause headed by *quem*.

(32) meus vero discipulus valde amat illum quem Brutus noster sauciavit.¹⁸

my but disciple greatly loves him whom.Acc.MS Brutus our has-wounded

'But my disciple...greatly loves him whom our Brutus has wounded.'

In (33) *illum* has moved as the complement of *eripui* past the subject *ego*. In (34) *me* has moved past the subject *ambitio* in the first clause and *te* has moved into a topicalized position probably in the CP domain.

(33) **illum** ego per flammas et mille sequentia tela eripui his umeris...¹⁹

him I through flames and thousand falling spears rescued these shoulders

'I rescued him on these shoulders through flames and a thousand falling spears...'

¹⁶ Cicero, *Pro Rabirio Postumo* 33.2; 54 BCE

¹⁷ Plautus, *Miles Gloriosus* 533; c. 254-184 BCE

¹⁸ Cicero, *Epistulæ ad Atticum* 14.22.1.8; 68-44 BCE

¹⁹ Virgil, *Aeneis*, 6.110; 29-19 BCE

(34) quod **me** ambitio quaedam ad honorum studium, **te** autem alia minime reprehendenda
that me ambition certain to honors zeal you however other least restraining
ratio ad honestum otium duxit.²⁰

reason to honest leisure led

‘...that a certain ambition led me to zeal of honors while a not in the least blame-worthy
reason led you to honest leisure.’

Based on van Gelderen’s diagnostics, Latin pronouns are DPs and at most might be considered weak pronouns but not clitic pronouns per Cardinaletti & Starke’s (1999) typology. Thus, they are DPs and not heads.²¹

Moving onto OldS, I have found no examples of coordination of object clitics in the CORDE. As for modification, you might expect something like (35) below, similar to Latin (32) above, but I have not found it.

(35) *María ama-*lo* que está a la derecha.

María loves-him that is to the right

Intended: ‘María loves him/the one on the right.’

Movement of the clitic is still less fixed than in ModS and may be movement to Spec positions.

The data in (36) to (47) below are examples of possible clitic movement in OldS.

(36) et el juez non *lo* quisiere recibir...²²

and the judge not him wanted to-receive

‘And the judge did not want to receive him...’

²⁰ Cicero, *Epistulae ad Atticum*, 1.17.5.7; 68-44 BCE

²¹ I am aware of the line of research represented by Bošković (2008), according to which languages without articles, like Latin, do not have DPs. I do not believe there is consensus on this, so I follow other linguists of Latin like Gianollo (2007) in assuming Latin did have DPs.

²² Anonymous, *Fuero de Zorita de los Canes*, p. 53; 1218-1250

(37) Et que *lo* non podades vender nin camiar...²³

and that it not you-can sell nor change

‘And that you cannot sell it nor change it...’

(38) E priso *lo* el rei Nabuchodonosor...²⁴

and took him the king Nabuchodonosor

‘And King Nabuchodonosor took him...’

(39) ...puede *lo* muy bien guardar.²⁵

he-can him very well protect

‘he can protect him very well.’

(40) Et asi todos los vienes ... son tesoro de santa Eglesia et puede *lo* partir el papa.²⁶

and thus all the goods are treasure of holy church and can it distribute the pope

‘And thus all the goods...are the treasure of holy Church and the Pope can distribute it.’

(41) Ca atal commo este_i puede *lo_i* el por derecho alçar entre los otros de su linaie...²⁷

for to-such as this can him he by right raise among the others of his line

For to such a one as this he can by right raise him among the others in his line.

Example (36) looks like standard clitic climbing, which is possible and frequent in ModS as well. However, Rivero (1986) points out that lexical DPs also climb in OldS, so this is not necessarily evidence against the DP status of the clitic. The data in (37) to (41) are not possible in ModS. Example (37) is an interpolation structure which Fontana (1993) would analyze as XP-adjunction or scrambling. Example (38) is a V1-declarative where, if we adopt Fontana’s

²³ Anonymous, *Documentos [Documentos del Monasterio de Santa María de Trianos]*, para. 1; 1299

²⁴ Anonymous, *Liber Regum [Documentos Lingüísticos Navarros]*, para. 27; 1194-1211

²⁵ Juan Manuel, *Libro de estados*, p. 304; 1327-1332

²⁶ Juan Manuel, *Libro de estados*, p. 472; 1327-1332

²⁷ Anonymous, *Siete Partidas de Alfonso X. BNM I 766*, p. II,64R; 1491

analysis, the verb has moved to C. The subject may be in situ but the object clitic clearly is not since it precedes the subject. While some VOS structures can occur in ModS, they must involve a lexical object since enclisis to finite verbs is no longer possible. The data in (39) to (41) are examples of clitic climbing to a position after the finite verb and preceding the infinitive, an order no longer possible in ModS, though still found in French. In (39) the adverbial *muy bien* separates the clitic from the infinitive. The datum in (47) may be an instance of clitic-left dislocation and VOS, since the clitic and *este* corefer. Different analyses can be applied to these data, but it is clear that the object clitic has much less freedom in ModS than it had in OldS, which may be taken as further evidence supporting its DP status in OldS.

In ModS, object clitics cannot be coordinated, as in (42).

(42) *No lo y me abrazas.

not him and me hug

"You do not hug him and me."

ModS clitics only move (if that is the analysis) to positions immediately adjacent to a verb. In (43) below the clitic cannot be separated from the verb by negation. The same sentence shows that while proclisis to a finite verb and enclisis to an infinitive is acceptable, enclisis to a finite verb is not, even in climbing constructions. Example (44) shows the same property in a simple clause. This is contrary to what was shown for OldS above.

(43) María (**lo*) no *lo* quiere (**lo*) abrazar (*lo*).

Maria him not him wants him to-hug him

'Maria does not want to hug him.'

(44) María *lo* abrazó(**lo*).

‘Maria hugged him.’

An additional diagnostic not adopted by van Gelderen that demonstrates the non-DP status of object clitics is repetition in conjuncts, a.k.a., wide scope in conjoined phrases, group inflection, suspended affixation.²⁸ Culbertson (2010) shows that subject clitics in Modern Colloquial French (MCF) must be repeated in coordinated VPs. She takes this to be evidence that subject clitics are becoming subject agreement morphemes since, like other inflectional morphology they cannot be omitted. Lexical DPs clearly pattern differently. Applying this diagnostic to subjects in ModS, consider the following:

(45) a. Juan canta y (Juan) baila. *Juan* = lexical DP

‘Juan sings and dances.’

b. Juan cant-a y bail-*(a). *-a* = subject agreement morpheme

c. Él canta y (él) baila. *él* = full DP pronoun (not a clitic)

In (45) the full DPs like *Juan* and *él* can be omitted in the second conjunct but the subject agreement morpheme *-a* must be repeated. This diagnostic can be extended to object pronouns and clitics as well. Consider first the following Latin data, where the underline indicates that an element has not been repeated but it is being interpreted in that position:

(46) Refige illum et mitte ____ in senatum...²⁹

unloose him and send to senate

‘Make him whole again and send him back to the senate...’

²⁸ Spencer & Luís (2012) give an overview of this phenomenon as it relates to the clitic vs. affix distinction. See Kayne (1975) for French; Uriagereka (1995:104) for Galician.

²⁹ Seneca the Younger, *Dialogi* 1.3.9.7; c. 64 CE. Translation by John W. Basore, *Seneca’s Essays*, vol. I, Loeb.

(47) quoi(u)s nunc pudet me et miseret ____ ...³⁰

of-whom now shames me and grieves

‘on whose account it now shames me and grieves me...’

In Latin (46) and (47), the object pronouns *illum* and *me* are elided in the second conjunct.³¹

Now consider similar data in OldS:

(48) assi como *lo* auemos & *lo* deuemos hauer...³²

thus as it we-have and it we-ought to-have

‘Thus as we have it and we ought to have it...’

(49) e cortaron a Nicanor la cabeza e la mano diestra, que tendiera ssobergiosa

and they-cut DOM Nicanor the head and the hand right which lay arrogant

mente en el templo, e levaron *lo*, e colgaron ____ contra Iherusalem.³³

-ly in the temple and they-took him and hung facing Jerusalem

‘And they cut off Nicanor’s head and his right hand, which lay arrogantly in the temple

and they took him and hung him facing Jerusalem.’

(50) E algunos dellos conosçieron *lo* e quisieron *lo* matar o ____ tomar preso...³⁴

and some of-them knew him and wanted him to-kill or to-take prisoner

‘And some of them knew him and wanted to kill him or take him prisoner.’

³⁰ Terence, *Heauton Timorumenos*, l. 260; 163 BCE

³¹ In fact, in the corpus I used I did not find any examples of *illum* being repeated in the second conjunct. This may be a result of the size of the corpus or there could be an actual preference in Latin to delete redundant object pronouns. Also, Latin has another third-person object pronoun, *eum* in the masculine, which is not related to the Spanish forms etymologically. A search for this yielded no results of the targeted coordinated clauses.

³² Anonymous, *Carta de donación [Documentos del Reino de Castilla]*, para. 349; 1225-1228

³³ Anonymous, *Biblia ladinada I-i-3*, p. 502RB; c. 1400

³⁴ Pero López de Ayala, *Crónica del Rey Don Pedro*, para. 277; c. 1400

(51) Et desto son testigos que lo vieron & ___ odieron...³⁵

and of-this they-are witnesses who it saw and heard

‘And these are the witnesses of this who saw it and heard it.’

The object clitic *lo* may be repeated as in (48) or elided as in (49) to (51). In (50) the second out of three is repeated, but the third is elided. How does ModS pattern? For most of my informants, repeating the clitic in the second conjunct is obligatory, as in (52).³⁶

(52) a. María <i>me</i> vio y *(<i>me</i>) abrazó.	(53) a. María <i>lo</i> besó y <i>lo</i> abrazó.
Maria me saw and me hugged	b. María <i>lo</i> besó y ___ abrazó.
‘Maria saw me and hugged me.’	c. *María ___ besó y <i>lo</i> abrazó.
b. María <i>lo</i> vio y *(<i>lo</i>) abrazó.	

One informant observed that you might get an elided clitic in formal written texts, as in (53b), but for all speakers (53c), where the first clitic is elided and the second maintained, is unacceptable. That (53b) might occur in a written text is not surprising nor does it count as evidence against the claim that it is highly grammaticalized, since formal written texts are by their nature linguistically conservative. Additionally, (53b) might be analyzed as a type of topic-licensing of object *pro* similar to subject *pro* as in Frascarelli (2007) and Maddox (2018).

In this section I have shown that Latin object pronouns can be coordinated and modified, and they move rather freely. They also can be omitted in conjoined clauses. This is evidence that they are full object DPs. In OldS, object clitics cannot be coordinated nor modified, and their movement is more restricted than that of Latin object pronouns. OldS object clitics can be omitted in conjoined clauses. This suggests they were undergoing reanalysis from full DPs to D-heads. In ModS, object clitics cannot be coordinated nor modified and they move as heads.

³⁵ Anonymous, *Carta de venta [Documentos del Reino de Castilla]*, para. 371; 1224

³⁶ These informants are speakers of Mexican, Peninsular, and Colombian Spanish.

They cannot be omitted in conjoined clauses. ModS object clitics are D- or *v*-heads (more on this below), per van Gelderen's model of the OAC.

3.2 Clitic doubling

3.2.1 Preliminary remarks

Critic doubling (CLD) is a well-studied phenomenon not only in Romance but also in Slavic, Greek, and other languages. Spanish exhibits doubling of direct objects (accusative critic doubling/ACD), with some variation, as in (54) below, and indirect objects (55).

(54) Juan <i>la_i</i> abrazó a <i>ella_i</i> .	(55) Juan <i>le_i</i> mandó un regalo a <i>María_i</i> .
Juan her hugged DOM her	Juan to-her sent a gift to Maria
‘Juan hugged her.’	‘Juan sent Maria a gift.’

Since some studies have analyzed ACD as object agreement, van Gelderen (2011) adopts it as a diagnostic that can be used to determine to what extent object pronouns/clitics are grammaticalized; i.e., if a clitic can occur in doubling structures, it is at a later stage in the OAC.

In this section I first present some influential analyses of cliticization and critic doubling.³⁷ I then discuss how van Gelderen makes use of doubling as a diagnostic of grammaticalization and discuss patterns of variation in Spanish. While van Gelderen identifies stage (c) as the only stage in which CLD can occur, I propose that CLD is possible at stage (b) and stage (c), though they differ derivationally. I extend Kramer's (2014) analysis to stage (b). Stage (c) is agreement rather than “true” CLD, though they appear identical superficially.³⁸

³⁷ For more in depth reviews see Anagnostopoulou (2006), Ordóñez (2012), Spencer & Luís (2012).

³⁸ Critic doubling can occur with all types of clitics and so the general phenomenon itself is abbreviated as CLD while doubling of specifically direct object clitics is abbreviated as ACD. This distinction is made to avoid confusion with another doubling pattern, clitic-left dislocation, which is abbreviated below as CLDT rather than CLLD.

3.2.2 Previous analyses of clitic doubling

The main problem posed for syntactic theory by CLD is that a single argument is represented in two structural positions. This is problematic if one assumes that Case and theta-role are assigned or checked only once per the Case Filter and the Theta Criterion. If there are two elements in a structure that are both relevant to accusative Case and theme, such as the clitic *la* and the pronoun *ella* in (61) above, do both receive Case? Perhaps one gets Case and the other the theta role? This is one aspect of CLD that has made it a popular area of research. Another area of research is how CLD varies cross-linguistically. While in Spanish there is dialectal variation, French and Italian lack ACD altogether, though French and some varieties of Italian do display subject clitic doubling.

One of the main types of analyses proposed for cliticization generally is the movement approach as in Kayne (1975). This type of account assumes that clitics are merged in argument position to receive a theta-role and then move to a different structural position, often for reasons having to do with Case assignment in the Government & Binding framework or feature checking in the Minimalist Program. Kayne (1975, 1989, 1991) analyzes clitics as left-adjoining to a functional head. He explains the different behavior of clitic placement with respect to infinitives as being the result of distinct types of verbal movement rather than clitic movement. For example, in French, where the order clitic+Infinitive occurs, the verb moves out of VP to adjoin to an Infn (infinitive) functional head. The clitic then left adjoins to the V+Infn complex. In Italian, however, which displays Infinitive+clitic order, the verb moves to Infn, as in French, but then the V+Infn complex moves again to adjoin to T' while the clitic left-adjoins to T.

A different movement analysis is given by Belletti (1999, 2001). Under this approach the clitics are merged to receive a theta-role and then move to an Agr projection. With CLD, the

doubled DP and the clitic both receive the single theta-role since they are both members of a larger constituent. Subject clitics move to AgrS while object clitics move to AgrO, and then further up to adjoin to the verb. While Kayne's general approach works well for French and Italian, it does not account for CLD which is absent in those languages. Belletti's analysis is an improvement since it can account for CLD, but it relies on Agr projections which have since been dispensed with in the most recent theoretic work within the Minimalist Program.

Other authors have pursued a base-generation approach. Early non-movement analyses such as Jaeggli (1982) and Borer (1984), for example, proposed that clitics are not moved from the complement position of the verb but rather are base-generated at their surface position, to the left of a finite verb. This leaves the original object position open for an additional DP in cases of ACD. Suñer (1988) expands on these previous analyses and argues that object clitics are object agreement morphemes.³⁹ Suñer (1988) argues that both IO and DO clitics are inflectional heads generated on V that license an empty category/*pro*; when there is a doubled constituent it is in argument position. DO clitics have more lexical features than IO clitics; i.e., they are lexical/inherently specified as [+specific]. Otherwise IO and DO clitics behave the same with respect to extractability, weak crossover, and scope. On Suñer's analysis, clitics do not absorb Case or receive a theta-role and specificity is the relevant feature that constrains ACD. Much of her data is taken from Rioplatense Spanish, which patterns differently from "standard" Spanish.

A third type of approach to CLD is the mixed analysis. For example, Sportiche (1996) focuses on object clitics taking them to be the heads of inflectional Voice projections that are in a special relationship with an associated DP. This DP moves to specifier position in the Clitic Voice phrase since the clitic head licenses special properties on that DP; i.e., agreement. In order

³⁹ Franco's (1993, 2000) investigations continue this line of research.

to account further for what triggers clitic movement, Sportiche proposes a Clitic Criterion and a Generalized Licensing Criterion, according to which functional heads, which includes clitics on his analysis, must be in a spec-head relationship with a DP at LF. In languages that allow CLD the movement of the DP may be overt, whereas in languages without CLD the movement is covert. Thus, both doubled and non-doubled clitic constructions receive the same analysis which is a mixture of base-generation (clitic voices) and movement (of the DP). Sportiche (1999) is an extension of the same analysis to subject clitics.

An alternative approach is taken by Uriagereka (1995), who argues that there are two types of clitics: weak determiner clitics and strong phrasal/pronominal clitics. According to his analysis, clitics encode referentiality and license *pro* in a functional projection F. Clitics move to F as part of the licensing of *pro*. The role of F is to assign reference to speakers in the discourse. Cross-linguistically F may be “morphological,” in which case clitics move there overtly, or it may only be active at LF, in which case clitics do not move to F but rather affix to AgrO. Uriagereka further proposes that determiner clitics head DP and in clitic doubling constructions their coreferential constituent is in the specifier position of that DP. Hence, this approach is sometimes referred to as the “Big DP” analysis. Strong clitics, on the other hand, are not D heads but phrases that undergo adjunction scrambling. Nevertheless, in languages lacking F strong clitics can end up in head positions.

A couple more recent studies also bear mentioning, since they have taken the movement approach in a new direction. Harizanov (2014) studies CLD in Bulgarian and argues that it is a type of A-movement based on diagnostics of binding and quantifier stranding. Harizanov’s analysis of CLD is as follows. First, the verbal complement is a KP with unvalued Case and φ -features. When the KP merges as complement it probes for a goal to have its features valued;

i.e., Agree. The v -head has an optional EPP-feature which can trigger movement of the KP object to Spec, v . In order to account for the double representation of a single argument, Harizanov adopts and modifies Matushansky's (2006) notion of "m-merger" whereby the head of a phrase in the specifier of some head can adjoin to the higher head. Harizanov extends m-merger so that it applies to both branching and non-branching maximal projections. Thus, the K-head m-merges with v and the $K+v$ compound head is realized as a clitic. Both the clitic and its associate are pronounced in CLD via multiple spell-out.

Kramer's (2014) study of the object marker in Amharic is similar to Harizanov (2014) since they both rely on m-merger and movement to Spec, v in their formal analysis. Kramer applies a battery of diagnostics such as optionality, number of marker per clause, presence/absence of an obligatory default, etc., to illustrate that the object marker in ACD is a doubled clitic, which means it is not object agreement in the true sense. She then reviews the morphological properties of the Amharic object marker which also suggest it is D rather than v , which means instances of ACD are just that, a doubled clitic construction and not object agreement. Since the object marker affects binding relationships, it is subject to A-movement. Kramer argues that the "doubled" object merges as DP verbal complement where it is probed by v (Agree) and then the DP object moves to Spec, v , where it undergoes m-merger with v , as in Harizanov (2014). In ACD, there are two copies of the DP object: one in argument position and the other in Spec, v . Following work by Kandybowicz (2007) and Nunes (2004), *inter alios*, Kramer proposes that both copies are pronounced because they are distinct at PF. Importantly for what I discuss below in Section 3.4, Kramer points out that movement of the object DP to Spec, v is basically object shift. Since on her account ACD relies on this movement, "object shift

feeds clitic doubling, causing a doubled DP to be close enough to v for v to Agree with it” (Kramer 2014:622).

3.2.3 *Clitic doubling in van Gelderen (2011)*

As previously discussed, there are three stages of the OAC. On van Gelderen’s (2011) account, ACD only becomes available at stage (c) where the clitic is a v -head because it is only then that a coreferential full DP can merge as complement to the verb.⁴⁰ Van Gelderen does not formally adopt a specific analysis of ACD, but one can see how a movement/copy analysis makes sense for stage (b) and base-generation for stage (c). I propose that depending on the grammaticalization status of the clitic, a different analysis applies. First, consider how ACD varies dialectally in Spanish. Van Gelderen, building on previous work (Jaeggli 1982, Suñer 1988, Franco 1993) argues that the more diverse types of constituents a clitic can double, the more grammaticalized it has become. In other words, the less restricted ACD is, the further along in the cycle a language has gone. Unrestricted doubling is characteristic of agreement, hence stage (c) in the OAC.⁴¹ Standard Spanish shows restricted doubling while other varieties such as Rioplatense and Malinche are unrestricted.

⁴⁰ Van Gelderen (2011) does not discuss indirect object CLD since indirect object clitics are presumably on a different cycle than direct object clitics. See the discussion on the Applicative Cycle in van Gelderen (2011:118).

⁴¹ The following data are adapted from van Gelderen (2011:102ff) and references therein, except where otherwise indicated.

(56) a. (*La_i) vimos la casa_i de Maria. “Standard” Spanish
it we-saw the house of Maria
‘We saw Maria’s house.

b. *(Lo_i) vimos a él_i.
him we-saw DOM he
‘We saw him.’

c. Pedro (*lo_i) vio a Juan_i.
Pedro him saw DOM Juan
‘Pedro saw Juan.’

d. ¿A quién_i (*lo_i) viste?
DOM whom him you-saw
‘Whom did you see?’

In “standard” Spanish doubling of inanimates is ungrammatical (56a), doubling of pronominal objects is obligatory (56b), doubling of proper names is ungrammatical (56c), and doubling of quantifiers is ungrammatical (56d). In non-standard varieties names can be doubled (57ab), inanimates can be doubled (57c), and quantifiers can be doubled (57d).

(57) a. Pedro lo_i vio a Juan_i. Rioplatense
 Pedro him saw DOM Juan
 ‘Pedro saw Juan.’

b. De repente la_i vio a Grimanesa_i ... Limeño
 of sudden her saw DOM Grimanesa
 ‘Suddenly s/he saw Grimanesa...’

c. Lo_i trae un chiquihuite. Malinche
 it he-brings a basket
 ‘He brings a basket.’

d. ¿A quién_i le_i viste? Basque Spanish/Argentinean
 DOM whom him you-saw
 ‘Whom did you see?’

Consequently, van Gelderen concludes that the non-standard varieties of Spanish are already at stage (c) in the OAC and clitics are object agreement. “Standard” Spanish, on the other hand, is transitioning from stage (b) to stage (c). Van Gelderen (2011:103), based on previous work by Tomić (2008) and Kalluli & Tasmowski (2008), notes that similar patterns of ACD occur in South Slavic and Balkan languages. While Bulgarian ACD is restricted to “pragmatically marked circumstances” such as topicalization, in which case it might be an instance of clitic left dislocation, Macedonian ACD is less restricted. For example, inanimate objects and pronominal objects can be doubled, though the latter is not obligatory for all speakers. Importantly, the clitics that can double object DPs are derived etymologically from pronouns, and thus this is another example of the OAC, where object pronouns are becoming object agreement.

3.2.4 Analysis: ACD at stage (b) versus ACD at stage (c)

As previously mentioned, van Gelderen does not fully develop a specific analysis for ACD. She assumes stage (b) of the OAC involves the clitic merging as DP and then moving to the V+v complex as a D-head. On van Gelderen's account, ACD is only possible at stage (c), when the clitic is the spell-out of u φ features of v. She does allude to the possibility of ACD when a language is moving from stage (b) to stage (c), as was discussed in the previous section, but she does not explain how ACD would work derivationally when the clitic is a D-head (stage b). Moreover, van Gelderen does not provide a formal explanation of why a wider range of lexical object types are allowed in ACD when the clitic is more grammaticalized.

In order to address these issues, I propose extending the Kramer/Harizanov-type analysis, summarized above in Section 3.2.2, to ACD at stage (b). This analysis seems appropriate for several reasons. First, it is a way of dealing with ACD while the clitic is still a D-head, which is what Kramer argues the Amharic object marker to be. Second, it involves object movement which is significant to the argument I make in Section 3.3 below that object movement feeds the OAC. Third, the m-merger operation provides a way to formalize what is meant by “reanalysis” of D to v.

With this in mind, let us consider how ACD can occur at different stages of the OAC. Recall first, however, that at stage (a), ACD cannot occur because the object pronoun is a full DP that merges as verbal complement. It may move after that, but on its own it satisfies theta-role, Case, and feature-checking requirements. ACD at stage (b) is where Kramer's (2014) analysis, following Harizanov (2014), comes in. Let us examine a typical case of ACD in Spanish, as the following:

(58) María lo_i vio a él_i

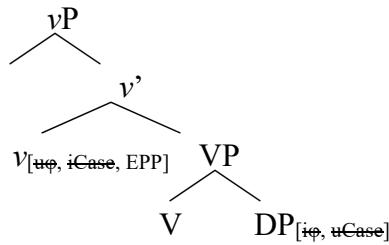
‘Mary saw him.’

In (58), the object DP *él* merges as verbal complement. The object DP is then probed by *v* which has up features that are checked by the ip features on the object and the object has its Case feature valued by *v*.⁴² Per Kramer’s analysis, an optional EPP on *v* triggers movement of the DP to Spec,*v*. There are now two copies of the object DP present: a low copy in first-merge verbal complement position and a higher copy in Spec,*v*. M-merger takes place between the higher copy of the DP and *v* which causes the higher copy to be realized as a clitic, as in Matushansky (2006). In ACD both copies are phonetically realized, but they are pronounced differently because only the higher copy of the object DP has undergone m-merger. The step-by-step derivation of (58) is represented below, based in part on Kramer (2014:620):

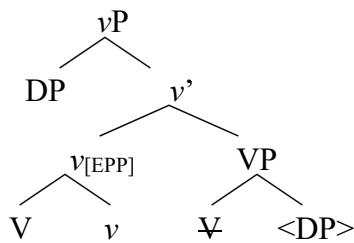
⁴² For simplicity sake I assume the differential object marker *a* is a morphological realization of accusative Case, though I am aware that there is a body of research that likely argues against this.

(59) ACD at Stage (b)

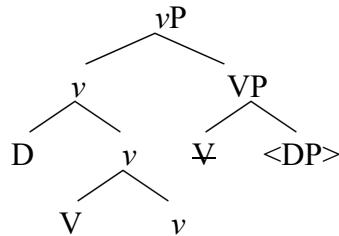
Step 1 → Object DP merges and Agrees with *v*.



Step 2 → V moves to *v*; Object DP moves to Spec,*v*.



Step 3 → M-merger between object DP and *v*.



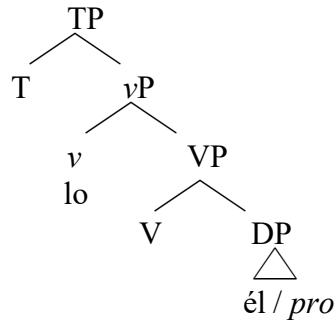
My analysis requires some modification to how stage (b) of the OAC was previously characterized by van Gelderen (2011). Rather than the clitic moving as a D-head to *v*, the whole DP moves to Spec,*v*. M-merger is how D becomes associated with *v*.

As discussed previously, stage (b) exhibits variation with respect to the type of object that can be doubled in ACD. In standard Spanish only pronominal objects can be doubled, but in other varieties like Rioplatense this restriction does not hold and thus Rioplatense is at stage (c). In a sense, ACD starts with pronominal objects and becomes possible only later with lexical objects. In order to account for this, let us consider the nature of m-merger. Matushansky

(2006:84) discusses Romance cliticization as an example of m-merger with phrasal movement. Following Kayne (1975) she assumes subject clitics are pronominal argument DPs that adjoin to T. Matushansky also takes the clitic to be simultaneously a maximal and a minimal projection (DP/D), as in Chomsky (1995). On her account, m-merger only applies to elements that can be both an XP and a head. Matushansky (2006:86) also states that m-merger “allows us to capture the diachronic transition from what looks like scrambling to cliticization.” This suggests to me that scrambling is what happens when pronouns are full DPs while cliticization and m-merger is what happens when they have lost some structure and become DP/D. This is consistent with van Gelderen’s original formulation of stage (b). How does this relate to ACD being restricted to pronominal objects at stage (b)? It is at this stage where the clitic has been reanalyzed from full DP to DP/D, and so it is only at this stage that m-merger can take place. M-merger is the synchronic version of reanalysis of clitics into morphology. Consequently, ACD starts with pronominal objects because they are deficient enough to undergo m-merger. Full lexical DPs, on the other hand, cannot be targeted by m-merger because they are not DP/Ds.

After reanalysis of the clitic from DP/D to *v* has taken place, the OAC is at stage (c) and the derivation of ACD is “simpler” than at stage (b) since there is at least one less operation given the lack of m-merger. Now the clitic is the spell-out of *u_φ* features on *v*; i.e., it is base-generated as *v*. The argument position is subsequently available for either *pro* or a lexical direct object, which merges to value *v*’s *u_φ* features. ACD is unrestricted at this stage because m-merger is not involved. Thus, at stage (c) we have agreement rather than “true” ACD, as in (60) below.

(60) ACD/Agreement at stage (c); clitic = *v*-head



Rioplatense is one of the varieties of Spanish that is at stage (c) of the OAC since it displays unrestricted ACD. Independent evidence that Rioplatense is at stage (c) comes from patterns of clitic left-dislocation (CLDT). Standard ModS allows a topicalized object in the left periphery to be coreferential with a clitic as in (61) below.

(61) Las flores_i las_i compré ayer.

the flowers them I-bought yesterday

‘The flowers, I bought them yesterday.’

Rioplatense also allows (61) but exhibits an additional pattern of CLDT with epithets. An example of this is (62) below, adapted from Suñer (2006).

(62) [A mi mejor amiga]_i, la_i vi [a esa loca linda]_i el jueves.

DOM my best friend her I-saw DOM that crazy beautiful the Thursday

‘I saw my best friend, that crazy beautiful girl, on Thursday.’

In (62), there are three coreferential elements: the topic, the object clitic, and the epithet directly after the verb. Importantly, (62) is unacceptable in ModS. Given my analysis, this is to be expected. At stage (c), the clitic is a *v*-head that never occupies complement position while at stage (b) the clitic is the realization of a copy of the object DP which merged as complement.

Since Rioplatense is at stage (c), the clitic is v , and the complement position is open for the epithet to merge.⁴³

3.3 Object movement (object shift/scrambling) and the OAC

3.3.1 *Preliminary remarks*

In the previous section I adopted an analysis of Spanish ACD at stage (b) whereby the object DP moves to Spec, v . This object movement is triggered by an optional EPP-feature per Kramer (2014)'s analysis of the Amharic object marker. The postulation of an EPP-feature in v is consistent with the word order of Amharic. Amharic is an SOV language (Leslau 1995) and, assuming SVO to be basic word order following Kayne's (1994) Linear Correspondence Axiom, OV in Amharic must involve object movement. However, the primary language this dissertation is focused on is Spanish, which is clearly SVO. This raises the question: why would we expect Spanish to pattern like Amharic in this respect? Object movement is a regular feature of Amharic, but in Spanish it is rather constrained, occurring in structures such as topicalization and focus fronting, which are conventionally considered distinct from scrambling and object shift.⁴⁴

The aim of this section is to investigate the presence of object movement in Spanish so as to justify the EPP-feature in v , which is a crucial component of the Kramer-type analysis. Per her analysis of ACD, EPP in v triggers object movement to Spec, v , where the object DP can undergo m-merger with v . If the object does not move, m-merger does not take place. I have proposed that m-merger is a synchronic version of diachronic reanalysis, which is a necessary aspect of the OAC.⁴⁵ Consequently, in what follows I examine patterns of object movement in Old Spanish. The main claim is that standard Modern Spanish (ModS) has an optional EPP-in- v

⁴³ See Chapter 5 for my full analysis of CLDT and its relation to ACD and null referential objects.

⁴⁴ See Ordóñez (1997, 1998, 2000) and Gallego (2012) for analyses of VOS in Spanish as object shift or scrambling.

⁴⁵ In this I am pursuing Matushansky's (2006:86) observation that m-merger "allows us to capture the diachronic transition from what looks like scrambling to cliticization."

as a historical remnant inherited from Old Spanish (OldS), which, in turn, inherited it from Latin. Previous authors such as MacKenzie & van der Wurff (2012) and Mensching (2012) have also claimed that OldS had EPP-in-*v*. In this section I build on this work by connecting it to ACD and the OAC. Others have proposed synchronic analyses in which object movement and ACD and/or clitic-left dislocation are derivationally related. I contribute to this line of research by taking a diachronic approach. In other words, I argue that object movement and ACD are also related historically. This section is organized as follows: In Section 3.3.2, I briefly review the literature on object shift and scrambling, identifying the main properties of each, and I present a couple of the synchronic studies that have attempted to link object movement and ACD. Section 3.3.3 is comprised of OldS data that show to what extent object movement was productive. Finally, in Section 3.3.4 I give an analysis of how object movement relates to CLD, how it feeds the OAC, and the prediction that falls out of this.

3.3.2 *Background*

3.3.2.1 Object shift

Object shift (OS) is a syntactic operation in which an object is moved to the left of the verb. Various analyses have identified either Spec,AgrO or Spec,*v* as the position targeted by object shift.⁴⁶ OS has been investigated primarily in Scandinavian languages like Icelandic (63), Danish (64), and Swedish (65).

(63) a. Af hverju las Pétur aldrei [VP <las> [DP þessa bók]] ?
why read Peter never this book

b. Af hverju las Pétur [DP þessa bók] aldrei [VP <las> [DP <þessa bók>]] ?
'Why did Peter never read this book?

⁴⁶ Unless otherwise indicated, the data in this section are taken from Vikner (2006) and references therein.

(64) a. *Hvorfor læste Peter aldrig [VP <læste> [DP den]] ?
 why read Peter never it
 b. Hvorfro læste Peter den aldrig [VP <læste> [DP <den>]] ?
 ‘Why did Peter never read it?’

(65) a. Damen såg inte [VP <såg> [DP den]] .
 lady-the saw not it
 b. Damen såg den inte [VP <såg> [DP <den>]] .
 lady-the saw it not
 ‘The lady did not see it.’

Object movement in Germanic languages is typically analyzed as scrambling rather than OS. In Icelandic both lexical and pronominal objects can shift. In Danish only pronominal objects may shift and do so obligatorily. In Swedish, OS is limited to pronominal objects and is optional.⁴⁷

Various constraints on OS have been identified in previous literature.⁴⁸ One constraint is Holmberg’s (1986) Generalization, according to which the verb must move out of VP before the object is allowed to shift. Thus, in auxiliary constructions OS is blocked because the verb does not move. A number of other constraints have also been identified. For example, OS is clause-bound. OS of lexical DPs optional but it is obligatory with pronominal objects. OS can be blocked by prepositions, particles, and indirect objects. OS is A-movement to a clause-internal specifier position, and OS cannot target a position between two adverbials.

Most studies take OS to be A-movement and various authors have offered different explanations for why it takes place. The two main explanations have to do with Case assignment

⁴⁷ Object raising in verb particle constructions has also been analyzed as object shift, but for the sake of brevity I do not discuss it here.

⁴⁸ Some of these constraints have been challenged. See Thráinsson (2001) and Vikner (2006) for data and discussion.

or information structure. According to the Case-assignment approach (Holmberg 1986, Holmberg & Platzack 1995), if an object is not assigned Case by V because V has moved, the object may move to a higher projection such as AgrO for Case. Diesing's (1996) information structure-based analysis argues that a shifted object receives a non-specific interpretation while an object in situ is interpreted as specific. OS is never fully optional but rather is dependent on which interpretation the speaker intends. Given that within the Minimalist Program, Case is checked when the object DP is probed by v , there is no need to relate OS to Case anymore. An account based on information structure may be a more promising approach, though there appears to not yet be consensus on this matter.

3.3.2.2 Scrambling⁴⁹

Scrambling can be defined as movement of a constituent (not limited to objects) to the left of the verb. It may superficially resemble OS, but the two operations differ in various ways, some of which are discussed below. Investigation has focused primarily on Germanic languages, with German being the “classic” scrambling language. However, scrambling has also been studied in some Asian languages, notably Japanese, and even Romance languages like Old Portuguese.⁵⁰ In German (66) and Dutch (67) below, the object has been scrambled from VP-internal position past negation.

⁴⁹ Most of the data in this section are taken from Thráinsson (2001) and Vikner (2006) and references therein except where otherwise indicated.

⁵⁰ Scrambling in Japanese differs significantly from Germanic. For example, Japanese scrambling is not clause bound while Germanic scrambling is. See Sato & Goto (2014) for an overview.

(66) ...dass Jens die Bücher nicht <die Bücher> kauft. (German)

that John the books not buys

(67) ...dat Jan de boeken niet <de boeken> koopt. (Dutch)

that John the books not buys

'...that John does not buy the books.'

While both scrambling and OS can move DPs (both lexical and pronominal), only scrambling can also move PPs. Neither one can move adverbs. Various constraints on scrambling have been identified. For example, scrambling is clause-bound (except for the Japanese/Russian-type) and it does not require verb movement, unlike OS. Thus, auxiliary constructions do not block scrambling. Scrambling is obligatory with pronominal objects. It is not blocked by prepositions, particles, or indirect objects and can target a position between two adverbials

Most take scrambling to be A'-movement. The main questions addressed in the literature on scrambling focus on determining what the structure of scrambling is (movement or base-generation) and identifying the trigger for scrambling. See Haider (2006:249) for a thorough overview. Studies aimed at determining the structure of scrambling include base-generation accounts such as Haider (1984), Kiss (1994), Bošković & Takahashi (1998). A movement approach is adopted in De Hoop & Kosmeijer (1995) and Müller & Sternefeld (1994). Regarding the trigger for scrambling, Haider & Rosengren (1998) argue there is not one and that scrambling is completely optional. De Hoop (1992) claims Case is a trigger while Müller (1995) argues it is a topic-feature. On Bošković & Takahashi's (1998) account, scrambling is triggered by a theta-feature.

3.3.2.3 Summary: Object shift compared with scrambling

Based on the preceding discussion, the properties summarized in Table 2 below are characteristic of object shift and scrambling.

Object shift	Scrambling
Obeys Holmberg's (1986) Generalization: the verb must move out of VP before the object is allowed to shift. Thus, in auxiliary constructions OS is blocked because the lexical verb does not move.	Does not require verb movement, Thus, auxiliary constructions do not block scrambling.
Clause-bound	Clause-bound (except for Japanese/Russian-type)
Optional with lexical DPs. Obligatory with pronominal objects.	Optional for most constituents. Obligatory with pronominal objects.
Can be blocked by prepositions, particles, and indirect objects.	Is not blocked by prepositions, particles, or indirect objects.
Cannot target a position between two adverbials.	Can move a constituent to a position between two adverbials
Moves DPs/NPs only.	Can move DPs/NPs, PPs, etc., but not APs.

Table 2. Object shift vs. Scrambling.

3.3.2.4 Unification approaches

Before moving on, it should be noted that some authors have linked object shift and/or scrambling to clitic doubling and clitic-left dislocation. Let us review a couple of studies that take this approach since it is relevant to the argument made below in Section 3.3.4. Suñer (2000), for example, identifies some similarities between clitic doubling (CLD) and OS. For example, CLD of strong pronouns is obligatory even when quantified, coordinated, or non-contrastive. Mainland Scandinavian uses OS to move the same type of pronouns obligatorily while in Icelandic it is optional. On her analysis, OS and CLD are two ways to comply with Diesing's (1992) Mapping Hypothesis and Heim's (1982) Novelty Condition according to which definite (specific, referential, existential) items must move out of VP since it is the domain of existential closure. In Spanish CLD the clitic is base-generated outside of VP where it acts as an

operator which allows the strong pronoun to be interpreted as specific. In languages with true OS the object pronoun actually moves out of VP so it can have a specific interpretation.

Alexiadou & Anagnostopoulou (1997) note some of the similarities between scrambling and CLD. Both phenomena are similar with respect to binding effects, semantic interpretation, the ability to de-stress the object intonationally. Their analysis focuses on CLD in Greek and German/Dutch scrambling. Scrambling is A-movement to Spec,AgrO. In CLD constructions, the clitic is a verbal agreement morpheme that doubles the object, forming a chain. They adopt the EPP condition for subjects; i.e., a D-feature in the subject domain must be checked by either moving or merging an XP. In null subject languages, EPP is satisfied by moving the verb to AgrS. They argue that, similar to the EPP for subjects, objects must also be licensed in AgrO. In scrambling languages this is accomplished by moving an XP to AgrO while in CLD languages a head moves there. The only difference between feature checking for subjects as compared to objects is that the former is obligatory while the latter is optional. While one might simply adopt this analysis and extend it diachronically, since I have chosen not to adopt Agr projections this is not the best option here, especially since these authors explicitly state that their analysis “crucially relies on the existence of Agr” (p. 158). Another option would be, as they also suggest, to modify the analysis to involve T for subjects and *v* for objects.

Harizanov’s (2014) analysis which, following Kramer (2014), I have adopted for ModS ACD, makes explicit the connection between object shift and ACD. He points out that since both involve movement of an object to Spec,*v*, the semantic and binding parallels exhibited by both phenomena are to be expected. I discuss in more detail Harizanov’s perspective on this connection in Section 3.3.4 below, since it forms the basis of my own analysis of how object movement feeds the OAC.

3.3.3 Patterns of object movement in Old Spanish

Given the analysis I have proposed following Kramer (2014), ModS ACD involves object movement to *Spec,v*. If this movement is triggered by an EPP-feature, we should expect further evidence to support this, such as other types of object movement not involving clitics. In this section I present data from previous studies and my own additional data to show that in OldS object movement was quite productive, much more so than in ModS. These patterns of object movement lend support to the postulation of EPP-in-*v* in OldS. Since object movement is highly constrained in ModS, its occurrence in ACD is a kind of historical remnant left over from the previous grammar.⁵¹

3.3.3.1 Previous studies

Object movement has been the focus of less attention in Romance than in Germanic and Scandinavian languages. Nevertheless, I review some of the relevant studies in what follows. Let us start with Parodi (1995), which is the sole work I am aware of that focuses on object shift specifically in OldS. Parodi takes object shift to be an instantiation of object agreement where the object moves to *Spec,AgrO*. Her main concern is showing that OS existed in compound tenses in OldS as in (68) below and that it takes place overtly without violating the Shortest Movement and Strict Cycle Conditions proposed in Chomsky (1993).

(68) Dixol cuemo avia su obra acabada.⁵²

he-told-them how had his work finished

‘He told them how he had finished his work.’

⁵¹ Old Spanish inherited object movement from Latin, though it is much more restricted in Old Spanish since the language was changing from Latin SOV to Romance SVO. I do not discuss the Latin facts here but see Devine & Stephens (2006), who argue that Latin objects undergo obligatory movement past the verb to topic and focus projections. As for object movement in Modern Spanish, some authors such as Ordóñez (1997, 1998, 2000) and Gallego (2013) have argued that Spanish VOS order is object shift. Consequently, it would not be surprising to find object shift in OldS.

⁵² Alfonso X, *Primera Crónica General de España*, 12.28^a; late 13th cent.

According to Parodi, objects can undergo further scrambling after OS. Thus, on her account, both phenomena existed in OS. Following Sportiche (1992), she takes cliticization to be an example of OS. Parodi attributes the loss of OS to it being too costly of an operation for language acquisition.

Sitaridou's (2011) article is focused on word order in OldS while her (2012) study investigates word order throughout Old Romance. Sitaridou (2011) argues that OldS was not V2 and that V2 syntax arises as an epiphenomenon of how discourse information is conveyed. While ModS expresses focus to the right and topic to the left of the verb, in OldS both discourse relations were expressed on the left. Sitaridou does not point this out but this parallels Latin which had obligatory movement of DOs and IOs to preverbal topic and focus projections (Devine & Stephens 2006). In Sitaridou's (2012) study she makes the same argument as Sitaridou (2011) but she extends the analysis to other varieties of Old Romance such as Old Portuguese, Old French, and Old Occitan, all of which have been analyzed as having V2 syntax. On Sitaridou's account, object movement in OldS is motivated by information structure rather than Case or other relations.

Another study, MacKenzie & van der Wurff (2012), is a comparative account of object movement in Middle English and Medieval Spanish. These authors identify four types of objects that can appear preverbally in both languages: i) quantified DPs (69), ii) negative DPs (70), iii) objects that are “given” information (71) and v) bare nouns in common V + N_{Obj} collocations (72):

(69) e si los huerfanos algund pleyto le fizieren.⁵³

and if the orphans some action to-him made

‘And if the orphans bring some action against him.’

(70) e no le quiso mostrar que ninguna sospecha havia d-el.⁵⁴

and not to-him he-wanted show that not-any suspicion he-had of-he

‘And did not want to show him that he had any suspicion about him.’

(71) Quando la Inffante donna Urraca estas razones avie con los de Çamora.⁵⁵

when the Infanta lady Urraca these reasons had with the-ones of Zamora

‘When the Infanta doña Urraca had these words with the Zamorans.’

(72) si tu verdat dizes vengan sobre ti todas las bendiciones que...⁵⁶

if you truth speak may-come upon you all the blessings that

‘If you speak true may all the blessings come upon you that...’

These authors do not opine as to whether the movement exhibited in OldS is object shift or scrambling. However, they do analyze it as A'-movement and note that preverbal objects in compound tenses usually precede the auxiliary, which is typical of scrambling rather than OS. They analyze object movement as being triggered by an EPP feature that moves an object DP to Spec,v*P, where it may then move higher to CP. Object movement is lost diachronically because the EPP-feature is lost. A further complication, they claim, is that preverbal objects were reanalyzed as being at the edge of v*P in Middle English but at the edge of CP in OldS. This was due to Spec,T being filled overtly in Middle English since it was a non-null subject

⁵³ Alfonso X, *Siete partidas*; 1256-1265. (109) to (112) are MacKenzie & van der Wurff's data and so I cite them as they do; i.e., without line or page numbers.

⁵⁴ Anonymous, *Gran conquista de Ultramar*; 1291-1295

⁵⁵ Alfonso X, *Estoria de España II*; 1260-1284

⁵⁶ Alfonso X, *Ordenamiento de las tafurerías*; 1276

language whereas Spanish has always had null subjects, so there was no overt material present to act a guidepost for reanalysis. In other words, the setting of the null subject parameter interfered with the reanalysis of OV word order.

Menschling's (2012) study investigates XP-V-S and Aux-XP-Participle structures in Old Romance generally. One might consider the former pattern to be object shift when the XP is a direct object. He mentions that this differs from ModS focus fronting OVS because in OldS the XP is usually topic and not focus, though focus fronting does occur in OldS. This makes sense diachronically since topics feed grammaticalization (Givón 1976, Gelderen 2011). Mensching also analyzes object movement to be triggered by an EPP feature on *v*. This feature moves the object to *Spec,v* where it can then move higher afterward. Mensching claims that ModS lacks this EPP feature and so a clitic has to merge to get the type of topicalization structures that were found in Old Romance.

3.3.3.2 Additional data

In an attempt to determine whether object movement in OldS is either object shift or scrambling, I have gathered more data from manual searches as well as searches in the CORDE to which I now turn. Starting first with the diagnostics given in Table 1 above, we need to address the question of whether verb movement is required or not. Recall that OS is blocked by auxiliaries but scrambling is not. MacKenzie & van der Wurff (2012) observe that preverbal objects in OldS with compound tenses almost always precede the auxiliary, which would be evidence of scrambling.⁵⁷ Parodi (1995), on the other hand, considers even those examples of objects preceding the auxiliary to be OS and not scrambling. While I have not carried out a quantitative analysis, the data I found concur with MacKenzie & van der Wurff's in that

⁵⁷ See MacKenzie & van der Wurff's (2012) footnote 10 about the rare examples in their corpus of Aux-O-V order.

preverbal objects following the auxiliary is certainly a minority pattern, since I found zero examples of it. This does not mean Aux-O-Participle order did not occur. Below is an example from Mensching (2012):

(73) porque ella non avia las cartas resçebidas.⁵⁸

because she not had the letters received

‘Because she had not received the letters.’

While (73) could be considered an example of OS, it should be pointed out that the source, *El libro de buen amor*, is a poem, and so it is not the optimal text to use given the well-known possibility for poetic registers to bend the rules when it comes to syntax.

As for preverbal objects with simple tense verbs, the prototypical context for OS, consider the following data:

(74) Tod esto cuenta Moysen en este sobredicho libro.⁵⁹

all this relates Moses in this aforementioned book

‘Moses relates all of this in this aforementioned book.’

(75) cuando esto sopo Berzebuey buscó aquellas escripturas.⁶⁰

when this found-out Berzebuey he-looked those scriptures

‘When Berzebuey found this out he looked at those scriptures.’

In (74) the object *tod esto* precedes the verb *cuenta*. The verb must have moved since it precedes the subject *Moysen*. The same pattern occurs in (75). This is a very common pattern and is highly reminiscent of Scandinavian OS. Nevertheless, verb movement also occurs in sentences where the object appears to remain in situ, as in the following data:

⁵⁸ Juan Ruiz, *El libro de buen amor*, I 191a; 1330

⁵⁹ Alfonso X, *Primera Crónica General*, p. 4; 1260-1284

⁶⁰ Anonymous, *Calila e Dimna*; 1251

(76) poníale el padre una corona de oro...⁶¹

he-put-on-him the father a crown of gold

‘His father put a golden crown on him.’

(77) Non conosce el hombre la grandeza de Dios...⁶²

not knows the man the greatness of God

‘Man does not know God’s greatness...’

Thus, verb movement may be necessary but it may not be the only condition under which OS is licensed in OldS.

Another relevant datum I found in this respect is given below:

(78) E este poder ovo Sant Pedro mientras que fue vivo e ovieron lo cuantos

and this power had Saint Peter while that he-was alive and they-had it as-many

apostoligos fueron despues del.⁶³

apostles were after of-he

‘And Saint Peter had this power while he was alive and as many apostles had it that went after him.’

In (78) both the first and second clauses have preverbal objects. However, in the second clause the object is the clitic *lo*. Thus, if the analysis of OVS as object shift in OldS is correct, we can state that it applies to both lexical DPs such as *este poder* and pronominal clitics like *lo*.

Nevertheless, in typical Scandinavian OS, the object does not move past the subject, so one might still contend that the above data are scrambling and not OS.

⁶¹ Anonymous, *Fazienda de Ultramar*; mid-13th century

⁶² Anonymous, *Fazienda de Ultramar*; mid-13th century

⁶³ Alfonso X, *El Espéculo*, p. 11; 1255-1260

Recall that OS is clause-bound while scrambling may move a constituent outside of the clause in some languages. I have found no evidence of object movement to a clause-external position in OldS. The furthest an object can move is up to CP, where one often finds an object clitic encliticized to a verb (79) or complementizer (80).

(79) e matolo el rey Artur ante la cibdad de Paris.⁶⁴

and he-killed-him the king Arthur before the city of Paris

‘And King Arthur killed him before the city of Paris.’

(80) e los Obispos pusieron sentencia de descomulgamiento sobre todos aquellos quelo non
and the bishops placed sentence of excommunication on all those that-it not
tovieren.⁶⁵
held

‘And the bishops placed an excommunication sentence on all those who did not hold it
(their mandate).’

With respect to optionality, object movement of lexical DPs does appear to be optional since the majority pattern as identified by MacKenzie & van der Wurff (2012) is for objects to be postverbal, even when they belong to the four classes identified by these authors. However, if we take cliticization to involve OS, it should be noted that object clitics most frequently occur preverbally in simple clauses. Thus, in a sense, OS of object clitics is obligatory. Recall that in OldS, clitics are full DPs. When they occur preverbally they have moved to Spec,v. When object clitics are postverbal, the verb has moved, as in (79) above, and the clitic has scrambled up past the subject.

⁶⁴ Anonymous, *La demanda del Sancto Grial*, p. 260; c. 1470

⁶⁵ Anonymous, *Ordenamiento de las cortes celebradas en Valladolid*, p. 55; 1258

Another pattern that distinguishes OS from scrambling is that the former is blocked by an indirect object while the latter is not. Consider the following data:

(81) *todas estas razones te cuento yo.*⁶⁶

all these reasons you tell I

‘I tell you all these reasons.’

(82) *si acaeciese que otra desonra les feziesen...*⁶⁷

if should-happen that other dishonor them they-made

‘if it should happen that they made them another dishonor...’

In (81) the direct object *todas estas razones* has moved up past subject and past the indirect object *te*. In (82) the movement of *otra desonra* is not blocked by *les*. Thus, these data are instances of scrambling rather than OS.

Recall that while the position targeted by OS may not be between two adverbials, this constraint does not hold for scrambling (Vikner 2006). The only relevant data I found for adverbs is the following:

(83) *E pues que tamaño bien puso Dios en este sentido...*⁶⁸

and since that amount large placed God in this sense

‘And since God indeed placed a large amount in this sense...’

(84) *E quien estas cosas asi non guardase...*⁶⁹

and who these things thus not keep

‘And whoever does not keep these things in this manner...’

⁶⁶ Abu al-Wafa' al-Mubashshir ibn Fatik, *Bocados de Oro*; mid-13th century

⁶⁷ Alfonso X, *El Espéculo*, p. 20; 1255-1260

⁶⁸ Abu al-Wafa' al-Mubashshir ibn Fatik, *Bocados de Oro*; mid-13th century

⁶⁹ Alfonso X, *El Espéculo*; p. 17; 1255-1260

(85) E qui esto asi non feziese...⁷⁰

and who this thus not did

‘And he who did not do it this way...’

In (83) the preverbal object *tamaño* is separated from the verb by a single adverb. Examples (84) and (85) are more interesting since in both cases the displaced object is separated from the verb by *asi* and *non*. The former might be considered an adverb of manner while the latter is negation. Vikner (2006:408) takes negation to count as an adverb in his data for German and Icelandic and shows that both OS and scrambling can move an element past two adverbs. It is only scrambling that moves an element between two adverbs. Consequently, the data in (84) and (85) are ambiguous in this regard.

Given the tests applied in this section it appears that OldS had both OS and scrambling. The latter may be more common on the surface, but as mentioned above, if OS is movement to *Spec,v*, then this will have to happen first before an object can undergo further movement (scrambling) since it cannot skip a specifier position. Thus, it is not surprising that both phenomena existed in OldS. There are more diagnostics that can be applied but for now I leave that for future work. Nevertheless, the extensive object movement exhibited in OldS lends support to the presence of an EPP feature in *v*, which is necessary to trigger the movement.

3.3.4 Object movement feeds the OAC

In this section I show how patterns of object movement feed the reanalysis of full object pronouns into object agreement; i.e., the Object Agreement Cycle (OAC). Building on van Gelderen (2011), Harizanov (2014), and Kramer (2014), I argue that object movement feeds the OAC in the sense that it moves the object to a position (*Spec,v*) where it is associated with the *v*

⁷⁰ Alfonso X, *El Espéculo*; p. 17; 1255-1260

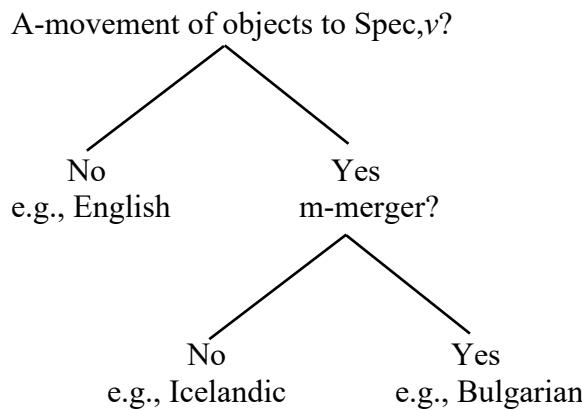
head. OldS compared with ModS is an example of this relationship. Recall that in Section 3.2.4 above I proposed that ACD at stage (b) of the OAC is derivationally distinct from ACD at stage (c). Stage (b) ACD involves object shift and m-merger. At stage (c) ACD is agreement. A prediction falls out of this analysis whereby if a language has object movement, full object pronouns will be reanalyzed as clitics and later agreement. Additionally, ACD will develop after object movement, which may later become constrained or absent altogether. These predictions hold throughout the Romance family. My analysis aims to contribute to our understanding of what separates object shift/scrambling languages from clitic doubling languages as has been discussed by previous authors such as Sportiche (1996), Alexiadou & Anagnostopoulou (1999), and Chomsky (2001). While formal synchronic mechanisms may account for some of the differences, I propose, following van Gelderen (2011), that the categorial status of the clitic must also be considered; i.e., what stage of the OAC a language is at.

3.3.4.1 Analysis

As was mentioned in Section 3.3.2.3 above, object movement and CLD are similar with respect to interpretational effects and binding patterns, which has led some to formally connect object movement and CLD. Kramer (2014), whose analysis I have adopted here, explicitly states that “optional object shift feeds clitic doubling, causing a doubled DP to be close enough to v for v to Agree with it” (p. 622). Harizanov (2014) also considers the relation between ACD and object shift. He points out that since object shift is A-movement to Spec, v and, on his analysis, ACD also relies on A-movement of the object DP to Spec, v followed by m-merger, this explains the similarities exhibited by ACD and object shift. Harizanov goes on to state that what distinguishes an object shift language from an ACD language is morphophonological since in the former there is no m-merger. Additionally, there are two crucial aspects that separate different

types of languages: the presence/absence of EPP in *v* and the possibility of m-merger. Thus, in a language like English which lacks object shift, there is no EPP in *v*. Icelandic has EPP in *v* and so it has object shift but m-merger is not possible. Bulgarian has EPP in *v* and m-merger is allowed, so it has ACD rather than object shift alone. The schema in (86) below is Harizanov's (2014:1075) representation of how languages with object shift differ parametrically from languages with CLD.

(86) Parametric Variation of Object Shift and Clitic Doubling (Harizanov 2014)



Harizanov's analysis is synchronic, but I propose that this variation appears diachronically as well. Recall that, depending on the stage, the object pronoun/clitic can be either a full DP, pronounced as a D-clitic head, or *v* (Gelderken 2011). And based on the discussion in Harizanov, languages may or may not have EPP in *v* and m-merger may or may not be allowed. EPP is satisfied by A-movement to Spec,*v*, so it targets XPs rather than heads. Consequently, when the object pronoun/clitic is still a DP, it can move to Spec,*v* for EPP (object shift). EPP-triggered movement to Spec,*v* is object shift and subsequent movement higher is scrambling. If a language has EPP and m-merger is allowed, the DP object in Spec,*v* will form a complex head with *v* and will be spelled out as a D-clitic. It is languages like this that can have ACD on the Harizanov/Kramer-type analysis.

What accounts for these patterns? For example, why do some languages allow A-movement to Spec, v ? Why do only some languages license m-merger? Harizanov attributes this to different parameters, as in (86) above. If we adopt the Borer-Chomsky Conjecture according to which parameters can be reduced to the presence or absence of a feature on a functional head, this makes sense for the EPP issue; i.e., does v have an EPP-feature? However, what is the critical feature that licenses m-merger? Furthermore, since m-merger is a morphological operation rather than syntactic, it does not depend on formal syntactic features and thus is not parametric in the conventional sense.⁷¹ In order to resolve this conflict, I propose that m-merger is always possible, but can only operate when the clitic/pronoun is already undergoing grammaticalization. In other words, if the object is a full pronominal DP it can move to Spec, v (if the language has EPP in v) but it cannot undergo m-merger. Only a deficient pronominal object, perhaps in the sense of Cardinaletti & Starke (1999), can undergo m-merger, but it has to still have enough structure to satisfy EPP. This is stage (b) of the OAC. The language still has EPP-in v and the clitic/pronoun can still satisfy EPP but it is also deficient enough to be targeted by m-merger with v . Thus, m-merger is essentially synchronic reanalysis of the clitic as v . For one generation of speakers, m-merger is a regular synchronic operation in ACD. The next generation of language learners may analyze the string of ACD as signaling that the clitic is object agreement. This is then stage (c) of the OAC. Naturally, some of the details remain to be worked out here, but historical patterns support the general idea, which I discuss below.

3.3.4.2 Predictions and patterns

Let us consider how the above analysis applies to Old and Modern Spanish since these are the primary languages that have been discussed in this chapter as characterizing the different

⁷¹ Of course, this relies on the assumption that morphology and syntax are distinct domains that do not derivationally interact.

stages of the OAC. One prediction resulting from my analysis is that in order to develop ACD a language will have to have object movement at an earlier stage. Recall that, as discussed above in Section 3.2.4, ACD can be superficially identical but derivationally distinct. At stage (b), it is movement of the clitic to Spec, v followed by m-merger. Both copies of the clitic are pronounced, but the higher one is realized as a clitic. At stage (c) ACD, the clitic is the spell out of v and the doubled DP is in argument position; i.e., object agreement. Stage (b) cannot happen unless object shift is licensed, which we are attributing to an EPP-feature in v . Thus, since ModS has ACD, it must have had object shift at an earlier stage. As was shown in Section 3.3.3, OldS did have object shift and scrambling triggered by EPP in v and so the prediction holds at least for Spanish.

Nevertheless, ModS no longer allows these types of object movement on most conventional analyses and so we need to account for there being an EPP in v to still license stage (b) type ACD, which is what I have proposed ModS ACD is. One possibility is to follow Harizanov and stipulate that EPP is optional since it has effects on interpretation. Alternatively, we can consider the EPP in v to be a historical remnant from OldS, which allowed a wider variety of object movement than ModS does. Assuming that ModS lacks true scrambling and object shift, it now only has object shift in ACD constructions. In fact, these two alternatives are not necessarily mutually exclusive. One might postulate that the EPP is not optional in some languages but can be in others. Latin might be an example where EPP in v is obligatory since it is an SOV language and allowed object shift and scrambling. In OldS EPP in v was optional in a variety of contexts while in ModS EPP is optional only in stage (b) ACD.

How does this prediction fare in other languages? If we stay within the confines of the Romance family we see a consistent pattern: less constrained object movement in the older

variety and more restricted object movement in the modern variety. Sitaridou (2012) analyzes word order in Old French, Old Portuguese, Old Spanish, and Old Occitan. She concludes that while these varieties of Old Romance were not V2 in the conventional sense (except possibly Old French), they did have frequent V2 strings due to a conspiracy of factors, one of which is allowing preverbal objects when they are topics or contrastive focus. Thus, discourse-motivated object movement in modern Romance is more constrained than in Old Romance due to a change in information structure. Other studies that demonstrate the possibilities of object movement in Old Romance include Martins (2002) for scrambling in Old Portuguese and Poletto (2014) for scrambling in Old Italian. Nicolae (2019) investigates scrambling in Old Romanian and Dindelegan (2016:141) states that preverbal objects were much more frequent in Old Romanian than in Modern Romanian, which is similar to what I have shown above for Spanish. For Old French, both Zaring (1998) and Mathieu (2009) argue that object shift was available. In sum, all the major varieties of Old Romance had object movement, though it was more constrained than in Latin and becomes even more constrained in the modern varieties. Thus, the first part of the prediction holds.

The second part of the prediction relates to the grammaticalization of clitics. If object movement feeds reanalysis of clitics then we should expect that those languages that had object movement should, at least to some extent, have clitics that are undergoing grammaticalization. On the face of it, this is true because all modern varieties of Romance have object clitics which differ from Latin object pronouns in their distribution and relation to the verb. However, not all varieties of Romance are at the same stage of the OAC. Evidence for this comes from the possibilities of ACD. Recall that at stage (b) of the OAC, restricted ACD becomes possible and at stage (c) it is unrestricted. What do we see across Romance? Standard French and Italian do

not allow ACD, Spanish and Romanian do allow ACD, with varieties of Spanish like Rioplatense being unrestricted. Portuguese is notable in that the European variety (EP) differs from the Brazilian (BP) in this respect. EP lacks ACD (Barrie 2000) while in BP first- and second-person object clitics can be doubled while third-person object clitics tend to be omitted (Machado Rocha & Ramos 2016). This suggests that EP is very conservative, more like Italian and French, while BP is even more innovative than Rioplatense Spanish, since object clitics are being lost. Crucially, all of these varieties allowed extensive object movement during the medieval period and thus the prediction holds. Since object pronouns could move first to Spec, v and then higher up, they came to be associated with v , which fits with how the OAC works. Nevertheless, the situation is even more complicated than it appears. In Chapter 5, I revisit object clitics and I argue that the diachronic and synchronic distribution of object clitic-left dislocation, ACD, and null objects in Spanish and beyond are tied to the stages of the OAC.

3.5 Chapter summary

This chapter covered the Object Agreement Cycle (OAC), which takes object pronouns and turns them into object agreement on the verb. I showed that Spanish object clitics are undergoing grammaticalization via this cycle. Building on van Gelderen (2011), I adopted clitic doubling as a diagnostic of grammaticalization. The less restricted doubling is, the more advanced in the cycle a language has gone. I proposed that clitic doubling is derivationally distinct at different stages of the OAC. At an earlier stage, the object DP moves to Spec, v and forms a complex head with v , which is phonetically realized as a clitic. Later, the clitic is the spell out of v and a lexical DP or *pro* can merge in argument position, which renews the OAC. Object movement shares certain characteristics with clitic doubling. I proposed that this is because the two phenomena are diachronically related; i.e., object movement feeds clitic

doubling (reanalysis of object clitics). This led to the prediction that if a language has productive object movement and develops clitics, it will later develop clitic doubling. This prediction holds within the Romance language family.

Chapter 4: The Reflexive Object Cycle from Latin to Spanish

Introduction

Reflexive pronouns are commonly grammaticalized cross-linguistically into valency-marking morphemes (Cennamo 1993, Haspelmath 1990, Geniušienė 1987, Heine & Kuteva 2002). In most varieties of modern Romance the morpheme is *se* or *si*. It can have different uses such as reflexive, anticausative, and passives as in (1) below for Spanish and (2) for French.

(1) a. Juan se lava.	(2) a. Jean se lave.
John Refl _{se} washes	John Refl _{se} washes
‘John washes himself.’	‘John washes himself.’
b. Se abrió la puerta.	b. La porte s’est ouverte.
AntiC _{se} opened the door	the door AntiC _{se} -is opened
‘The door opened.’	‘The door opened.’
c. Se vendieron los pisos.	c. Les appartements se sont vendus
Pass _{se} sold the apartments	the apartments Pass _{se} were sold
‘Apartments were sold.’	‘Apartments were sold.’

The focus of this chapter is the grammaticalization of the reflexive pronoun from Latin to Spanish, which I refer to as the Reflexive Object Cycle (ROC). This is a subtype of the Object Agreement Cycle (OAC) that was discussed in Chapter 3. The primary difference between these cycles is that whereas in the OAC the input to the cycle is the direct object pronoun, in this case it is the direct object reflexive pronoun.

Like the OAC the ROC is comprised of three cycles. My main claim is that Latin and Old Spanish were at stage (a), Middle Spanish was at stage (b), and Modern Spanish is at stage

(c) of the ROC. Evidence for each stage comes from patterns of clitic placement, auxiliary selection, and phonological attrition. The obligatory doubling of the stressed reflexive *a sí mismo* by the reflexive clitic in Modern Spanish also supports this claim. This type of doubling was absent in OldS because *se* was still a full DP complement. Consequently, the object position was not open for *a sí mismo* to merge. At stage (a) of the ROC, *se* is a full DP argument. At stage (b) *se* is a DP that moves as a D-head. *Se* is reanalyzed as a Voice or *v* head at stage (c).

This chapter is organized as follows. In Section 4.1, an example of the ROC in Scandinavian is presented, based on the discussion in van Gelderen (2011:120) and data from Faarland (2004) and Ottosson (2004). The ROC in Scandinavian serves as a model for the ROC in Romance. Section 4.2 focuses on determining the categorial status (phrase or head) of the reflexive pronoun during different diachronic stages based on diagnostics of interpolation, coordination, modification, etc. In Section 4.3, I examine the extent to which clitic doubling takes place with the reflexive pronoun throughout the history of Spanish since this pattern also bears on the categorial status of *se*. This chapter is summarized in Section 4.4.¹

4.1 ROC in Scandinavian

The Object Agreement Cycle is not restricted to definite object pronouns; it can also affect reflexive pronouns. Van Gelderen (2011:120), based on data in Faarland (2004) and Ottosson (2004), outlines how this cycle has taken place in the Scandinavian and Slavic languages.² In Old Norse, the ancestor of Modern Swedish and other Scandinavian languages, reflexive *sik*, as in (3), is a DP pronoun that can be modified by *sjalfa*, as shown in (4).

¹ Reflexive clitics are also discussed in Chapter 5, Section 5.3, which is based on Maddox & MacDonald (in contract).

² See also Cennamo (1993) for a functionalist analysis of this cycle in Romance and Russian. The grammaticalization of reflexives into valency markers is well-attested cross-linguistically; v. Geniušienė (1987), Haspelmath (1990), and references in Heine & Kuteva (2002). A reflexive cycle has also been proposed by Faltz (2008).

(3) Hann nefndi sik Ola. Old Norse

he called REFL Ola

‘He called himself Ola.’

(4) Sumir hofðu sik sjalfa deydda. Old Norse

some had REFL.Acc self.Acc killed

‘Some had themselves killed.’

As (5) below shows, *sik* can also be a suffix that gets a reflexive or passive reading.

(5) Kalla-sk. Old Norse

calls-REFL

‘He calls himself/He is called.’

In Modern Swedish, *sik* has become the valency marking suffix *-s*, as in the anticausative in (6).

(6) Dörren öppnades. Swedish

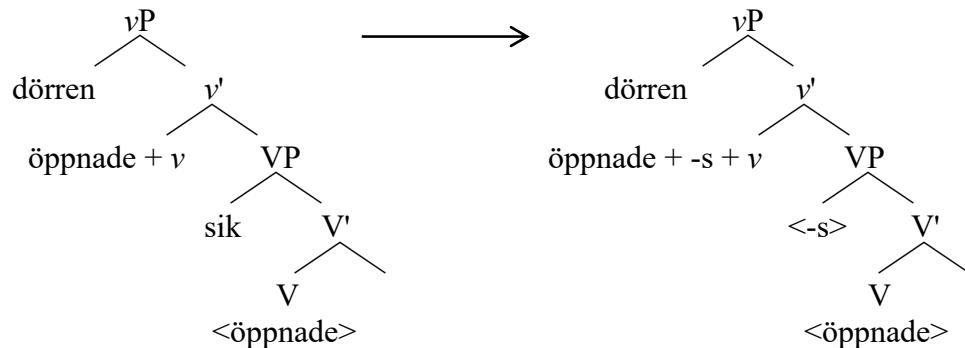
door-the opens-REFL

‘The door opens.’

Thus, the full DP reflexive pronoun *sik* was weakened and is now a functional head *v* or Voice.

Two stages of this cycle are represented below in (7), adapted from van Gelderen (2011:121).

(7) Stage (a) Stage (b)



On van Gelderen's analysis, in stage (a) in (7), the DP *sik* merges in object position, and may then be subjected to further XP movement. At stage (b), it has been reanalyzed as a D head that can be "picked up" by the verb on the way to *v*. If the reflexive suffix can be doubled, Swedish would be at stage (c) of the cycle which would be followed by renewal. My central proposal is that this is the type of cycle that took place in Spanish.

4.2 Categorial status of *se* diachronically

In this section I argue that *se* changed its status diachronically from a DP in Latin, to a D head in Middle Spanish (MidS), and finally a Voice head in Modern Spanish (ModS). Originally a pronominal argument, *se* is now an inflectional element, a marker of valency. Some of the evidence for this kind of change comes from distributional differences with respect to modification, coordination, and movement/interpolation. Phonological attrition and auxiliary selection also point to a change in status. In ModS, *se* displays the properties of inflection and is undergoing paradigmatization in some varieties. The changes from DP to D head to Voice head each represent a stage of the reflexive object cycle. I use evidence similar to what was seen for Scandinavian in Section 4.1 above but I introduce additional diagnostics such as clitic doubling and auxiliary selection.

4.2.1 Stage (a) - Latin

4.2.1.1 Background

Latin had a multifunctional *-r* suffix parallel to ModS *se*; i.e., it could have a reflexive (8), anticausative (9), passive (10), or impersonal (11) reading.

(8) Excepit Seleucus fabulae partem et 'ego' inquit 'non cotidie **lavor**.'³
 took-hold Seleucus conversation part and I said not daily wash
 'Seleucus took up part of the conversation and 'I,' he said, 'do not wash myself daily.'

(9) Omnis liquor vapore **solvitur** ac frigoribus magnis conficitur.⁴
 all liquid vapor dissolve and great cold congeal
 'All the liquid is thinned by the heat and congealed by great cold.'

(10) **Mittitur** ad eos C. Arpinius eques Romanus.⁵
 sent to them C. Arpinius cavalryman Roman
 'The Roman cavalryman, C. Arpinius, was sent to them.'

(11) **Vivitur** ex rapto: non hospes ab hospite tutus, non socer a genero.⁶
 lives from plunder not guest from host safe nor father from child
 'One lived off plunder: a guest was not safe from the host, nor a father from his child.'

Nevertheless, Latin also had a reflexive pronoun *sē*, the ancestor of ModS *se*, which could occur in reflexive (12) and anticausative constructions (13).⁷

(12) similī tālem sē vidit in aurō. ⁸ likewise such Refl _{se} sees in gold 'Likewise he sees himself in the gold.'	(13) dum calor sē frangat. ⁹ while heat AntiC _{se} breaks '... while the heat breaks.'
---	---

³ Petronius, *Satyricon*, 42.2.1; 66 CE

⁴ Columella, *De re rustica*, 1.6.18.7; 70 CE

⁵ Caesar, *De bello gallico*, 5.27.1.1; 58-49 BCE

⁶ Ovid, *Metamorphoses*, 1.144; 8 CE

⁷ There is no consensus in the literature as to whether there was a passive *se* in Latin; v. Muller (1924), Kärde (1943), Monge (1954), Cennamo (1999), Adams (2013). There is no evidence of an impersonal *se*. See Green (1991) for an analysis (among many others) of how the *-r* form is lost due to competition with *se*.

⁸ Statius, *Achilleid*, 1.865; 94-96 CE

⁹ Cicero, *De Oratore* 1.265; 55 BCE

4.2.1.2 Distribution

At this stage, *sē* has the distribution of a full pronominal DP rather than that of a functional head.¹⁰ For example, Latin *sē* can be coordinated, as in (14a), while ModS *se* cannot (14b).

(14) a. mē et sē hīsce impedīvit nuptiīs!¹¹ Latin
 me and Refl_{se} this shackled marriage
 ‘He shackled me and himself in this marriage!’

b. *Me y se aprisionó en este matrimonio. Modern Spanish
 me and himself imprisoned in this marriage

Additionally, Latin *sē* can be modified, as in (15a) by the intensifier *ipse*, while its ModS counterpart cannot (15b).

(15) a. sē ipse sine mūnītiōne dēfenderet. Latin
 Refl_{se} very.M.S without fortification defended
 ‘He defended his very self without fortification.’

b. *se mismo defendió sin municiones. Modern Spanish
 Refl_{se} veryself defended without fortification

Finally, a particularly striking difference between Latin *sē* and ModS *se* is that the former can be separated from the verb via scrambling, as seen in (16).

¹⁰ Faltz (2008) observes that Latin *se* was not a clitic, but makes no claims as to its status as a head or phrase.

¹¹ Terence, *Phormio*, 2.4; 161 BCE

¹² Caesar, *de Bello Gallico*, 20.5; 58-49 BCE

(16) apud Platonem Socrates in caelum effert laudibus Protagoram Hippiam Prodicum
 with Plato Socrates in heaven brings praises Protagoras Hippias Prodicus
 ceteros, **sē** autem omnium rerum inscium **fingit** et rudem.¹³
 others Refl_{se} but all things ignorant represents and coarse
 ‘Along with Plato Socrates praises to the heavens Protagoras, Hippias, Prodicus, and
 others, but himself he represents as coarse and ignorant of all things.’

Having an intervener between the verb and *se* other than another clitic is not possible in ModS, as (17) below shows.

(17) *Juan *se* no lava regularmente.

John Refl_{se} not washes regularly

Intended: ‘Juan does not wash himself regularly.’

The Latin datum in (16) also shows that *sē* could be contrastively focused, a property especially strange to speakers of ModS. In (16), Cicero states that Socrates praises other philosophers exceedingly, but considers himself ignorant; i.e., he contrasts himself with these other philosophers. In ModS, *se* cannot be contrastively focused, as (18) shows.

(18) *Juan_i alaba a sus amigos mucho pero SE_i considera ignorante.

John praises DOM his friends a-lot but Refl_{se} considers ignorant

Intended: ‘Juan praises his friends a lot but he considers himself to be ignorant.’

Thus, Latin *sē* differs from its ModS descendant in at least four ways: it can be coordinated, modified, separated from the verb, and contrastively focused. This suggests that Latin *sē* was a fully phrasal DP rather than a type of head.

¹³ Cicero, *Brutus*, 292.14; 46 BCE

4.2.1.3 Phonological attrition

Further evidence that Latin *sē* was undergoing grammaticalization comes from "attrition," a well-known property of elements undergoing this process (Lehmann 1985). This is the tendency for lexical items to lose phonological and semantic properties. Hopper & Traugott (2003:10) discuss the case of the imperative *let's* in English as in (19a), which is reduced phonologically from *let us* (19b).

(19) a. Let's go to the circus tonight.
b. Let us go.

The reduced (19a) has a first-plural "adhortative" reading while the full form in (19b) is a second-person imperative, asking one's captor to release them. Hopper & Traugott claim that in some varieties it has been further reduced from [lets] to [les] to a clitic that prefixes to the following verb as in *s-go*, for *let's go* or *s-fight*, for *let's fight*. Another example of reduction in English includes the lexical *will* meaning want or desire being reduced to the affix *-ll* as a future marker as in "I'll eat at nine o'clock."

In the case of the Latin reflexive pronoun, attrition begins already in the classical language, where the archaic *sēd* loses the final consonant and becomes *sē* (Lindsay 1894:424). An example from Old Latin is given in (20), where *sēd* is a prepositional complement.

(20) Neve post hac inter sed coniourase...¹⁴
and-not after this among SED plot
'And they shall not plot among themselves...'

While this was the older form, early playwrights like Plautus (254-184 BCE) already use the reduced form of *sē*. The next reduction takes place in the passage from Classical to Vulgar Latin

¹⁴ *Senatus consultum de Bacchanalibus*, 13-14; 186 BCE

when vowel quantity is lost, resulting in Proto-Romance *sě*, with a short rather than a long vowel. Thus, attrition as a symptom of grammaticalization is clear in the case of *se*. This attrition process is represented below in (21).

(21)	Old Latin	Classical Latin	Proto-Romance
	(prior to 75 BCE)	(75 BCE - 500 CE)	(500 - 1000 CE)
	<i>sēd</i>	<i>sē</i>	<i>sě</i>

In conclusion, I have shown that Latin *sē* was at stage (a) in the ROC. As a DP it could be separated from the verb with few constraints. It merged as the internal argument and was thereafter subject to scrambling or other types of XP movement. The reflexive pronoun also underwent phonological attrition from Latin to Romance.

4.2.2 Stage (a) moving to stage (b) - Old Spanish (1200-1400)

4.2.2.1. Distribution

From the earliest documents, Old Spanish (OldS) had *Refl_{se}* (22), *AntiC_{se}* (23), and *Pass_{se}* (24).¹⁵

(22) Es semejante al puerco, que dexa el agua clara y se baña en el cieno.¹⁶
 is similar to-the pig which leaves the water clear and *Refl_{se}* bathes in the mud
 'He is similar to the pig, which leaves the clear water and bathes itself in the mud.'

(23) E cuenta aquí la Biblia que se abrió entonces la tierra bien allí.¹⁷
 and tells here the Bible that *AntiC_{se}* opened then the earth well there
 'And here the Bible states that the earth indeed opened up there.'

¹⁵ OldS also had pronominal/inherent *se* which I do not discuss here. It follows the same patterns with respect to interpolation as other types of *se*. Impersonal *se* is not clearly distinguishable from passive *se* until the sixteenth century.

¹⁶ Ferrand Martínez, *Libro del cavallero Cifar*, fol. 42r; 1300

¹⁷ Alfonso X, *General Estoria*; 1270

(24) ¡Con tal cum esto se vençen moros del campo!¹⁸

with such with this Pass_{se} conquers Moors of-the field

‘In this way Moors are conquered in the field.’

Crucially, however, OldS *se* differs from ModS *se* in several respects. First, it can be separated from the verb by interveners such as DPs, PPs, and adverbs, though negation is by far the most frequent. This phenomenon, illustrated in (25) through (26) below, is typically referred to in the literature as "interpolation;" see Chenery (1905), Ramsden (1963), Poole (2007), *inter alios*.

(25) Desí mando que se non rasiessen.¹⁹

thus order that Refl_{se} not shave

‘Thus I order that they not shave (themselves).’

(26) este algodon es atal que se non quema por fuego.²⁰

this cotton is such that AntiC_{se} not burns by fire

‘This cotton is such that it does not burn from fire.’

(27) No hay guisa por que se esto diga.²¹

not is fashion by which Pass_{se} this say

‘There is no way by which this is said.’

Recall that the same behavior was displayed by Latin *sē* as discussed above in Section 4.2.1.2, which I take as evidence supporting *se* as a DP and not a head.

Additionally, OldS *se* can occupy the same position as stressed pronouns and DPs.

Consider the following data.

¹⁸ Anonymous, *Cid*, line 1753; 1207

¹⁹ Alfonso X, *General Estoria, Primera Parte*, fol. 277V; c. 1275

²⁰ Anonymous, *Lapidario*, fol. 17R; c. 1250

²¹ Anonymous, *Calila e Dimna*, Chap. 3; 1251

(28) Et los sabios dizen que quién **a sí non guarda** a otri non fará pro.²²
 and the wise say that who DOM self not defend DOM other not do good
 ‘And the wise men say that he who does not defend himself will not do good to another.’

(29) si el enperador **todas estas cosas non guarda** et yerra en todas...²³
 if the emperor all these things not defends and errs in everything
 ‘If the emperor does not defend all these things and erres in everything...’

(30) si **se non guarda** delo bever tanto quel pueda del venir danno.²⁴
 if Refl_{se} not guards from-it drink so-much that-to-him can from-it come damage
 ‘If he does not protect himself from drinking so much of it, there may come to him
 damage from it.’

In (28), the stressed pronoun *sí* precedes negation as a complement of the preposition *a*. Since this is a stressed pronoun it can be assumed that it is a full DP per Cardinaletti & Starke (1999). A quantified DP occurs in the same position in (29), followed by negation and then the verb. Finally, in (30), *se* also occurs in this position, suggesting that it is also a DP rather than a head.

Additional evidence for the DP status of OldS *se* comes from climbing environments. As Rivero (1986:783) notes, climbing is not restricted to clitics alone in OldS; lexical DPs can also climb.²⁵ The data she provides are related to direct object clitics, but the same observation applies to *se*, as shown in the following data.

²² Anonymous, *Calila e Dimna*, para. 9; 1251

²³ Juan Manuel, *Libro de estados*, para. 12; 1327-1332

²⁴ Juan Manuel, *Libro de estados*, para. 13; 1327-1332

²⁵ Though in the case of lexical DPs it is typically labelled “scrambling.” Rivero (1986) uses “climbing” for both.

(31) et quisiere meterse so otro sennorío.²⁶

and should-wish place-*Refl_{se}* under other lordship

‘And should he wish to place himself under another lord's authority...’

(32) e quisiere quitar la heredat o el huerto.²⁷

and should-wish take-away the estate or the garden

‘And should he wish to take away the estate or the garden...’

(33) qual quier que sus cosas quisiere vender.²⁸

which ever that his things should-wish sell

‘Whosoever should wish to sell his things...’

(34) Et todo omne que se quisiere saluar d'estas calonnyas.²⁹

and every man who *Refl_{se}* should-wish save from-these punishments

‘And every man who should wish to save himself from these punishments...’

That both *se* and full DPs can occur after the non-finite verb is shown in (31) and (32), respectively. Compare this with (33), where an object DP climbs to precede the finite verb and (34), where *se* does the same. That lexical DPs and the clitic *se* can occupy the same position suggests that they are the same element, subject to the same kind of movement operations.

Following Rivero (1986), I take this as an indication that *se* is, in fact, a DP at this stage rather than a head. *Se* at this stage is in complementary distribution with DPs, which is not the case in the modern language.

Other authors have claimed that OldS clitics are maximal projections (v. Rivero 1986, 1991, 1997; Barbosa 1993, 1996; Fontana 1993, 1997; Halpern & Fontana 1994; Halpern 1995).

²⁶ Anonymous, *Libro de los fueros de Castiella*, para. 318; c. 1284

²⁷ Anonymous, *Fuero viejo de Castilla*, para. 9; 1356

²⁸ Anonymous, *Fuero de Cuenca*, Book 4, sect. 1; 1284-1295

²⁹ Anonymous, *Fuero viejo de Castilla*, para. 1; 1356

Fontana (1993) claims that clitics, including direct and indirect object clitics, are second-position clitics and thus XPs in OldS but are reanalyzed as heads in ModS.³⁰ He analyzes interpolation as taking XP-clitics that adjoin either to the right or left of the first XP position dominated by IP via scrambling or as being substituted into Spec,T. Nevertheless, unlike Latin, OldS interpolation is rather rare and more constrained, occurring primarily in embedded clauses and limited to specific types of interveners. This leads to the conclusion that, while *se* was a DP in both Latin and OldS, it was undergoing reanalysis among some speakers, hence the rarity of interpolation. Thus, OldS was not firmly at stage (a) of the cycle but rather moving towards stage (b), where reanalysis as a head takes place.

4.2.2.2 Auxiliary selection

More evidence that *se* is a DP argument in OldS comes from auxiliary selection. In some Romance languages such as Italian and French, passives and unaccusatives in the compound past tense select the BE auxiliary, Italian *essere*, as in (35) and (36), respectively.

(35) La lettera è scritta.	(36) Giovanni è venuto.
the letter is written	Giovanni is come
‘The letter is written.’	‘Giovanni came.’

Thus, both have intransitive syntax and, consequently, a derived subject (McGinnis 2004). With transitive verbs, the compound past is formed with the HAVE auxiliary, *avere*, as in (37).

(37) Giovanni ha accusato Luigi.
Giovanni has accused Luigi
‘Giovanni accused Luigi.’

³⁰ According to Pancheva (2005) and Bošković (2015), second-position clitics do not allow doubling cross-linguistically. OldS clitics in second-position cliticize to any sentence-initial word including complementizers. This may be further evidence that in OldS, *se* is not as closely connected with the verb as in ModS and, thus it is a DP rather than a head.

The same pattern applies if a verb is used reflexively with the full DP *se stesso*, but if the reflexive clitic *si* is used instead, the BE auxiliary is selected. This is illustrated by the data in (38) and (39) below.

(38) Giovanni ha accusato se stesso	(39) Giovanni si è accusato.
Giovanni has accused self very	Giovanni Refl _{se} is accused
‘Giovanni accused himself.’	‘Giovanni accused himself.’

Since other intransitives select the BE auxiliary, the data in (39) serve as evidence that Italian *si*-constructions have intransitive syntax; i.e., a single argument which merges in object position and moves to subject position.

OldS, unlike ModS, also had auxiliary selection but the patterns with respect to compound tense formation do not match what was shown above for Italian. In his study of auxiliary selection in OldS, Aranovich (2003) notes that, unlike the situation in Italian, most reflexive verbs select the HAVE auxiliary.³¹ Consider the following data.

(40) Minaya Alvar Fáñez essora es llegado. ³²	
Minaya Alvar Fáñez then is arrived	
‘Minaya Alvar Fáñez then arrived.’	
(41) como ninguno de los athenienos no se a vestido de negro por mi. ³³	
because none of the Athenians not Refl _{se} has dressed of black for me	
‘...because none of the Athenians has dressed in black for me.’	

³¹ The exception, according to Aranovich (2003), is “quasi-reflexives,” using Manacorda de Rosetti’s (1961) study. These apparently include inherent reflexives and anticausative reflexives.

³² Anonymous, *Cid*, line 2449; 1207

³³ Juan Fernández de Heredia, *Traducción de Vidas paralelas de Plutarco*, III, fol. 176v; 1379-1384

(42) mas es necesario que quando el se ha echado en tierra...³⁴

but is necessary that when he Refl_{se} has thrown on ground

‘But it is necessary that when he has cast himself to the ground...’

The data in (40) are a control example to show that unaccusatives selected the BE auxiliary in OldS. This property is lost in Early Modern Spanish, with the latest examples occurring in the 17th century. According to Benzing (1931), cited in Aranovich (2003:4), the verb *llegar* continues to select BE until the 16th century. Examples (41) and (42) contain compound reflexives, and both select the HAVE auxiliary, a pattern not expected given what takes place in Italian as discussed above. I take this as evidence that reflexive *se*-constructions in OldS were transitive structures, with *se* being the internal argument. In (42), it is clear that *se* is the internal and not the external argument, since *el* is the agent DP which would merge in Spec,*v* to receive its theta-role.

From the discussion above we may conclude that in OldS *se* was at the end of stage (a), moving toward stage (b). As a DP internal argument it could undergo more constrained types of phrasal movement as demonstrated by the interpolation data.³⁵ However, it could not be coordinated or modified as it could be in Latin, which suggests that in OldS *se* was reanalyzed by some speakers as a head; i.e., stage (b).³⁶

³⁴ Ferrer Sayol, *Libro de Pallado BNM 10211*, para. 115; 1380-1385

³⁵ Martins (2002, 2003) argues that direct, indirect, and reflexive (*se*) clitics were in fact heads (Xs) in Old Spanish (and Old Portuguese), while the oblique pronouns *i* and *en* were weak pronouns and thus XPs. In order to accept her reasoning we must reconcile it with the auxiliary selection data, which she does not address. Nevertheless, if Martins is correct, the only part of my analysis that would require modification is placing OldS firmly in stage (b) rather than "moving toward" it.

³⁶ I did find 15 tokens of *se* possibly modified by *mismo* or *mesmo* without the differential object marker. However its occurrence is so rare as to be negligible, and perhaps attributable to orthographic error.

4.2.3 Stage (b) - Middle Spanish (1400-1600)

4.2.3.1 Loss of interpolation

At stage (b) of the cycle, the DP pronoun is reanalyzed as a D head. Since it is now a head, interpolation of *se* should disappear, on the analysis of interpolation as a type of XP movement. The historical record is consistent with this prediction and interpolation is indeed lost during this period. The data in (43) and (44) below show that interpolation does still occur.

(43) no os deberíais matar ni perder por ninguna cosa que os aviniese, cuanto más not you should kill nor lose for no thing that you happens how-much more por hecho de mujeres que se ligeramente gana y pierde.³⁷
for deed of women that Pass_{se} easily wins and loses

‘You should not be killed nor lost on account of anything that may befall you, how much more on account of the action of women, which is easily won and lost.’

(44) que se non ficiese deservicio de Dios e daño de la tierra.³⁸
that Pass_{se} not do disservice of God and damage of the earth
‘...that a disservice not be made to God and damage to the earth...’

In order to quantify this loss, I conducted a CORDE search for interpolation patterns with negation, which is always the most frequent type.³⁹ The results of this search are given in Table 3 below.

³⁷ Garcí Rodríguez de Montalvo, *Amadís de Gaula*, Book II; 1475-1500

³⁸ Jerónimo Zurita, *Anales de la corona de Aragón, Primera Parte*, para. 232; 1562

³⁹ CORDE search conducted Oct. 12, 2015 at 4:00 p.m. See also Cheshire (1905) and Ramsden (1963) for more quantitative data on interpolation.

Century		Pattern 1: SE + Neg + V	Pattern 2: Neg + SE + V
Old Spanish	1200-1300	25.2	74.8
	1300-1400	11.8	81.2
Middle Spanish	1400-1500	6.2	93.8
	1500-1600	.5	99.5
Modern Spanish	1600-1700	.1	99.9
	1700-1800	0	0

Table 3. Loss of interpolation by century.

As Table 1 shows, interpolation (Pattern 1) is already a minority pattern in the OldS period, declining from 25.2% to 11.8% from the thirteenth to the fourteenth century. By the MidS period, it is quite rare and in the modern period it is almost nonexistent, though sporadic examples do occur in the sixteenth century (Keniston 1937).

4.2.3.2 Auxiliary selection

In MidS, auxiliary selection is still active, with unaccusatives like *llegar* selecting BE, as in (45) below.

(45) a tal punto el miserable ombre es llegado.⁴⁰

to such point the miserable man is arrived

‘The miserable man has arrived to such a point.’

At this stage, reflexives continue to pattern with transitive verbs in selecting HAVE, as in (46) and (47) below.

(46) ca el mismo se ha cortado la lengua.⁴¹

because he himself Refl_{se} has cut the tongue

‘Because he himself has cut out his own tongue.’

⁴⁰ Enrique de Villena, *Traducción y glosas de la Eneida Libros I-III*, para. 24; 1427-1428

⁴¹ Antón de Zorita, *Árbol de Batallas* (de Honoré Bouvet), para. 166; c. 1440-1460

(47) hasta agora no se ha guardado la merced que vuestra Alteza hizo.⁴²

until now not Pass_{se} has guarded the mercy that your highness made

‘Up until now the mercy that your Highness commanded has not been protected.’

Thus, the analysis of *se* as the internal argument in OldS proposed in Section 4.2.2.2 above holds in MidS as well.

4.2.4 Stage (c) - Early Modern and Contemporary Modern Spanish (1600-present)

In ModS, *se* has been reanalyzed from a D head to a Voice/v head via the Late Merge Principle (LMP). Rather than merging as D and moving up to Voice, it is now the spell out of the Voice head. It is clear that *se* is now a head since interpolation is no longer possible; see (52b) below. Auxiliary selection is lost in the seventeenth century (Aranovich 2003), so it can no longer be used as a diagnostic. This development suggests that ModS is in stage (c) of the ROC.

4.2.4.1 Evidence for ModS *se* as agreement

ModS *se* has been independently argued by others to be an affix (Cuervo 2003, 2014; Folli & Harley 2007). In what follows I present some of the evidence that has been put forth to support this and I present additional corroborating data. First, according to Fábregas & Scalise (2012), inflectional morphology has two properties: 1) it does not change the grammatical category of the base, and 2) it does not produce new words, like derivational morphology, but rather different forms of a single word, as with the subject agreement suffix in (48a). The same properties apply to *se* as is shown in (48b) with Imp_{se}.

⁴² Anonymous, *Cortes de Madrid*, para. 19; 1551

(48) a. Juan com-e mucho. b. Juan se-lava mucho.
John eat-3S a-lot John Refl_{se}-washes a-lot
'John eats a lot.' 'Juan washes himself a lot.'

In (48a), subject agreement is affixed to the verbal stem but it does not change the category nor does it create a new word. In (48b), the *se* morpheme is prefixed to the verb, which is still a verb.

Additionally, inflectional morphology follows a strict ordering pattern. This also applies to *se*. Consider the following data.

(49) a. Tú habl -a -ba -s	(50) a. Se -pre -dice el futuro.
you speak-TV-Asp-2S	Pass _{se} -pre-says the future
‘You were speaking.’	‘The future is predicted.’
b. *Tú habl-ba-s-a	b. *Pre-se-dice el futuro.

In (49a), the imperfective aspect marker *-ba-* must come after the thematic vowel and precede the subject agreement marker; any other order is unacceptable (49b). The same pattern applies to *se*. In (50a), *Pass_{se}* must precede the preverb *pre-* and not the other way round (50b). When *se* occurs with other clitics, ordering restrictions still hold whereby *Imp_{se}*, for example, must precede an object clitic.⁴³

Also, no lexical material may intervene between an affix and the base. This is shown in (51a), where *pre-* cannot have negation come between it and the verb (51b).

⁴³ We might expect rigid ordering to be a diachronic development since, at earlier stages, *se* was a DP. I have not found any evidence of flexible clitic or affix ordering at any stage. Pescarini (in press) notes that with respect to spurious (dative) *se* and accusative clitics, the more archaic order within Romance is accusative then dative, but in Ibero-Romance, the order has always been dative then accusative since the earliest documents. Ordóñez (2002, 2012) observes that some non-standard varieties of Spanish do allow first- and second-person clitics to precede *se*.

(51) a. Juan no pre-dice el futuro. b. *Juan pre-no-dice el futuro.
 Juan not pre-dicts the future Juan pre-not-dicts the future

This extends to *se* as shown in (52).

(52) a. Juan no se lava. b. *Juan se no lava.
 Juan not Refl_{se} washes Juan Refl_{se} not washes

Se also behaves like inflection in that it can interact morphophonologically with other inflection.

Consider the following data taken from Halle & Harris (2005).

	<u>Normative</u>	<u>Alternative</u>
(53) a.	Váyan-se.	b. Váyan-se-n.
	go-AntiC _{se}	go-AntiC _{se} -n
	‘Go!/Leave!’	‘Go!/Leave!’
(54) a.	Sírvan-se.	b. Sírvan-se-n.
	serve-Refl _{se}	serve-Refl _{se} -n
	‘Serve yourselves!’	‘Serve yourselves!’

The normative forms in (53a) and (54a) have *se* coming after the plural agreement affix *-n*.

However, in different registers, speakers place an additional affix after the *se*, as shown in (53b) and (54b). Halle & Harris (2005) note that this behavior occurs most frequently with *se*, *me*, and *le*, and less frequently with direct object clitics. They also emphasize that it is not a performance error but rather an alternative form that occurs in many different varieties of Spanish.

Franco (1993) points out that *se* behaves like an inflectional affix in subject-verb contexts. When forming a question, the verb moves to a position preceding its subject, but the affix moves with the verb rather than being left behind. This is shown with the subject agreement affix *-a* in (55).

(55) a. Juan lav-a su coche. b. ¿Con qué frecuencia lav-a Juan su coche?
 Juan wash-es his car with what frequency wash-es Juan his car
 ‘Juan washes his car.’ ‘How often does Juan wash his car?’

The same type of movement takes place with *se* as shown in (56).

(56) a. Juan nunca se ducha. b. ¿Con qué frecuencia se ducha Juan?
 Juan never Refl_{se} showers with what frequency Refl_{se} showers Juan
 ‘Juan never showers.’ ‘How often does Juan shower?’

Thus, just as inflectional affixes have a tight relationship with the verbal base, so does *se*.

Additional evidence comes from patterns of repetition in coordinated VPs. This diagnostic has been used to show that subject clitics in French pattern like agreement affixes (Kayne 1975, Culbertson 2010, *inter alios*). Since agreement is obligatory it must be repeated on each verb in two conjoined VPs. In (57) below, for example, the subject agreement affix cannot be omitted in the second conjunct.

(57) a. Juan com-e y habl-a.
 ‘John eats and talks.’
 b. *Juan com-e y habl-Ø

According to Culbertson (2010), this is true of subject clitics in Modern Colloquial French, which suggests they are full subject agreement prefixes. Now consider how ModS *se* behaves in this environment:

(58) a. No se puede y no se debe fumar.
 ‘One cannot and should not smoke.’
 b. *No se puede y no debe fumar.

Example (58) is unacceptable; i.e., *se* cannot be omitted.⁴⁴ Thus, it patterns like an agreement affix. Given the preceding data and discussion we may conclude that ModS *se* has a distribution and behavior similar to that of agreement, which is to be expected outcome of the cycle.

4.2.4.2 Paradigmatization

Another step we might expect to see in this grammaticalization process is "paradigmatization." This is the stage where a single form is extended to the entire paradigm for all persons and numbers. As discussed in Hopper & Traugott (2003) and Ottosson (2004, 2008), this happened with the Old Norse reflexive pronoun *sik* during its grammaticalization as a valency marker as discussed above in Section 4.1. The *-sk* affix was already extended to second person singular and plural while the *-mk* affix derived from the reflexive first-person singular pronoun was extended to the first-person plural. In some modern Scandinavian languages like Danish, Icelandic, and Swedish, the descendant of *-sk* is invariable for all persons and numbers, the logical outcome of paradigmatization.

Concerning the extension of *se* to all persons in Spanish, this process has not taken place yet in the "standard" variety. However, Benito Moreno's (2015) investigation into syncretism in the reflexive paradigm in Spanish and Catalan provides some interesting data in this respect.

Consider the following data, taken from Benito Moreno (2015):

(59) ...yo quería reírse.

I wanted-1S laugh-Inher_{se}

'I wanted to laugh.'

⁴⁴ Compare (58) with (43) above, where in Middle Spanish the *se* can be omitted.

(60) Pues, se comprabas el tinte, lo ponías en una cacerola.⁴⁵

well Refl_{se} bought-2S the dye it put-2S in a saucepan

‘Well, you had to buy the dye, put it in a saucepan...’

(61) ...éramos pequeños, pero se escondíamos como las ratas.

were-1P small but Refl_{se} hid-1P like the rats

‘We were small but we hid ourselves like rats.’

(62) Se laváis to los días a desgusto.⁴⁶

Refl_{se} wash-2P all the days to unpleasure

‘Every day you guys take a shower while complaining.’

These data exemplify the use of *se* with the first-person (59) and second-person singular (60); and the first (61) and second-person plural (62). The social and geographic variation of this phenomenon is complicated and Benito Moreno deals with it in detail. For our purposes, this type of extension serves as further evidence of the grammaticalization process. As noted by Benito Moreno and others, paradigmatization has taken place more extensively in other varieties of Romance such as Surselvan, Brazilian Portuguese, varieties of Italian, etc. Whether this process will continue in Spanish is difficult to predict at this point, but it would not be surprising given the cross-linguistic synchronic and diachronic patterns. In sum, paradigmatization relates to feature loss because the less features *se* has, the more different types of verbs and verbal forms, it can combine with.

⁴⁵ I am not sure how to classify this use of *se* so I label it generically as Refl_{se}, though it may be closer to an aspectual *se*, dative of interest, or benefactive.

⁴⁶ The grammatical gloss is mine; the translation is Benito Moreno's.

4.2.5 *Summary*

In the above section I demonstrated that Latin *sē*, the ancestor of ModS *se*, was a pronominal DP as supported by the fact that it can be modified, coordinated, and separated from the verb via XP movement. That *sē* was already undergoing grammaticalization is evident from the phonological attrition it suffered from Old Latin to Classical Latin to Proto-Romance. In OldS, *se* retains its DP status since it can still be separated from the verb (interpolation) and it can occupy the same position as stressed pronouns and full nominals. Additional evidence for *se* as a DP argument in OldS comes from auxiliary selection. Compound *se*-constructions select the HAVE auxiliary rather than the BE auxiliary, showing that they have transitive rather than intransitive syntax. In MidS, interpolation is lost but auxiliary selection shows that *se* merged as an argument and moved as a head. In ModS, *se* displays many properties shared with inflectional morphemes. It is now a Voice head, a valency marker. All this evidence points to a change in categorial status. These properties all correlate with the diachronic doubling data, which is the topic of the next section.

4.3 Doubling with *se* diachronically

4.3.1 *Doubling in Old and Middle Spanish*

As was discussed in Chapter 3, on van Gelderen's (2011) account of the OAC, doubling only becomes available when the clitic is no longer an argument. Thus, it is not expected at stage (a); i.e., Latin and Old Spanish. At stage (b), an additional coreferential XP is not allowed because the clitic merges as the internal argument and then moves as a D head. As the clitic loses features and becomes reanalyzed as a functional head doubling becomes possible. In Chapter 3, I proposed *contra* van Gelderen that object clitic doubling is possible at stage (b), albeit via a different derivation. Since the ROC is a sub-type of OAC, one might expect the

same rules to apply. Before addressing this issue however, let us consider the patterns of reflexive doubling that are found in the data.

In OldS, I found only three instances of reflexive doubling in my searches, as in (63) and (65) below. The more frequent pattern, however, is for *a sí mismo* to occur without *se*, as in (64) and (66).

(63) el se alaba a si mismo.⁴⁷

he Refl_{se} praises DOM self same

‘He praises himself.’

(64) e desí que lavasse a sí mismo...⁴⁸

and thus that he-wash DOM self same

‘and thus that he wash himself...’

(65) el se offrecio a si mismo.⁴⁹

he Refl_{se} offered DOM self same

‘He offered himself.’

(66) Demostenes saco su guchiello y ferio a si mismo.⁵⁰

Demosthenes took his knife and injured DOM self same

‘Demosthenes took his knife and injured himself.’

The above data are particularly interesting since both the doubled and non-doubled versions occur in the same authors, which suggests a change in progress at this time, or perhaps two competing grammars.

Examples of doubling start to increase in frequency during the MidS period, as in (67).

⁴⁷ Alfonso X, *General Estoria, Segunda Parte*, para. 53; c. 1275

⁴⁸ Alfonso X, *General Estoria, Primera Parte*, para. 17; c. 1275

⁴⁹ Juan Fernández de Heredia, *Traducción de Vidas paralelas de Plutarco, III*, para. 35; 1379-1384

⁵⁰ Juan Fernández de Heredia, *Traducción de Vidas paralelas de Plutarco, III*, para. 42; 1379-1384

(67) matóse a sí mismo.⁵¹

he-killed-Refl_{se} DOM self same

‘He killed himself.’

Some putative examples of doubling are actually clitic-left dislocation, as in (68) below.

(68) pues a sí mismo se condena quien al que yerra perdona.⁵²

since DOM self very Refl_{se} condemns who DOM-the that errs pardons

‘For he condemns himself, he who pardons the one that errs.’

The data in (69) shows that for some speakers *se* was still in complementary distribution with *a sí mismo* during this period.

(69) el que a sí mismo aborrece, él se juzga a mal.⁵³

he who DOM self very abhors he Refl_{se} judges to evil

‘He who abhors himself, he judges himself to be evil.’

In order to determine how frequent doubling with *a sí mismo* was, a CORDE search was conducted for clitic doubling for the OldS and MidS periods, the results of which are given in Table 4.⁵⁴

Period / Century		Tokens	Percentage
Old Spanish	1200-1300	53/3	5.6%
	1300-1400	47/3	6.3%
Middle Spanish	1400-1500	186/21	12%
	1500-1600	541/253	46.8%

Table 4. Clitic doubling with *se* by century.

⁵¹ Pedro de Corral, *Crónica del rey don Rodrigo, postrimero rey de los godos (Crónica sarracina)*, para. 30; c. 1430

⁵² Diego de San Pedro, *Cárcel de Amor*, para. 6; 1482-1492

⁵³ Fray Diego de Valencia, *Sobre la predestinación y sobre la Trinidad y la Encarnación*, para. 22 (1486-1487)

⁵⁴ CORDE search conducted on Nov. 19, 2015 at 3:40 p.m. I searched for *a sí mismo* with and without the written accent, since orthography was not completely standardized yet. In the tokens column, the number on the left is the total occurrences of *se*; the number on the right is the number doubled.

In the OldS period, doubling is rare. In the fifteenth century, however, doubling increases twofold and in the sixteenth century almost half of all instances of *a sí mismo* are doubled by *se*. The loss of interpolation, auxiliary selection of reflexives with HAVE, and dramatic increase of doubling suggest that MidS was at stage (b) in the cycle, with incipient progression toward stage (c), where doubling should be allowed.

4.3.2 Doubling in Modern Spanish

In ModS, doubling of *se* is common. Since *se* is a Voice head and no longer an argument, the object position is open for other DPs to merge. The data in Table 2 above show that doubling starts to increase dramatically in the sixteenth century, the last century of "Middle Spanish." However, by the seventeenth century, doubling with *a sí mismo* increases to 75.4%, suggesting that for most speakers, *se* is now a head.⁵⁵ Examples from the Early Modern period of doubling are given below in (70) and (71).

(70) el que se aborreciere a sí mismo...este tal segura tiene la vida.⁵⁶

he who Refl_{se} abhors DOM self very this so secure has the life

'He who abhors himself...this one hold his life secure.'

(71) él a sí mismo se admire.⁵⁷

he DOM self very Refl_{se} admires

'He himself admires.'

For some Early Modern speakers, however, *se* is not yet obligatory in reflexive constructions, as (72) shows.

⁵⁵ Calculated from the same CORDE search discussed above.

⁵⁶ San Juan Bautista de la Concepción, *Pláticas a los religiosos*, para. 3; 1603-1607

⁵⁷ Sor Juana Inés de la Cruz, *Poesía (Lírica personal)*, para. 353; 1666-1695

(72) El que atiende a sí mismo por Dios, hace el todo.⁵⁸

he who attends DOM self very for God does the everything

‘He who attends to himself for the sake of God, accomplishes everything.’

The present-day state-of-affairs is that Refl_{se} can be optionally doubled by *a sí (mismo)*, or a prepositional phrase with a pronominal complement, in which case a bound reading is induced (Babcock 1970, Torrego 1995), as is shown in (73).

(73) Juan_i se_i lava (*pro_i* / a sí_i / a sí mismo_i).

John Refl_{se} washes him DOM self DOM self very

‘John washes himself.’

Crucially, however, Refl_{se} cannot be omitted in the presence of *a sí mismo* as a complement (74).

(74) Juan *(se) lava a sí mismo.

Thus, since in ModS *se* is obligatory, grammaticalization of *se* is complete.

4.3.3 Analysis

In Chapter 3, the focus was on third-person DO clitics. I showed that accusative clitic doubling (ACD) in “standard” ModS was restricted to pronominal objects while in Rioplatense Spanish it is unrestricted since lexical objects can also be doubled. ModS is at stage (b) of the OAC while Rioplatense is at stage (c), yet both exhibit ACD. In order to reconcile these facts, I proposed that stage (b) doubling in ModS is derivationally distinct from that which is seen in Rioplatense. I adopted an analysis based on Harianov (2014) and Kramer (2014) whereby in ModS ACD the object merges as a DP argument and moves to Spec, v where it undergoes merger with v . Both copies of the object are pronounced but the higher copy is realized as a

⁵⁸ Miguel de Molinos, *Guía espiritual*, para. 369; 1675-1675

clitic due to m-merger. In Rioplatense, on the other hand, ACD is true agreement. The clitic is a *v*-head, leaving the complement position open for lexical objects or *pro*.

Given that I have proposed two different analyses of doubling for object clitics, how should we approach the derivation of reflexive doubling? Is it like stage (b) or stage (c) ACD? While ModS third-person object clitics are at stage (b), there is evidence that suggests ModS *se* is at stage (c) rather and thus it is already a Voice or *v* head. Consequently, ModS reflexive doubling would be agreement. Let us consider some of additional doubling evidence as compared with third-person direct object clitics.

In Section 4.2.4.1 above I claimed that ModS *se* is more like agreement inflection since it patterns like inflection in various respects. If it is indeed a type of agreement inflection, then ModS *se* is at stage (c) of the ROC. Doubling data further corroborates this claim. Torrego (1995:237) observes that nonreferential quantifiers are not acceptable when serving as antecedents of *lo(s)/la(s)*, as in (75) below.

(75) A pocos políticos (**los*) admiramos.

DOM few politicians them we-admire

‘We admire few politicians.’

This same constraint, however, does not hold for *se*.

(76) Pocos políticos se critican a sí mismos.

few politicians Refl_{se} criticize DOM self very

‘Few politicians criticize themselves.’

Thus, *se* does not pattern like the third-person DO clitics in doubling configurations. *Se* is an inflectional morpheme rather than a determiner clitic like *lo(s)/la(s)*. In conclusion, reanalysis of *se* as a Voice/*v* head is complete in ModS.

Since *se* is at stage (c) of the ROC, then the agreement analysis of stage (c) doubling applies, similar to ACD in Rioplatense. The structure of ModS reflexive doubling is represented in (77) below.

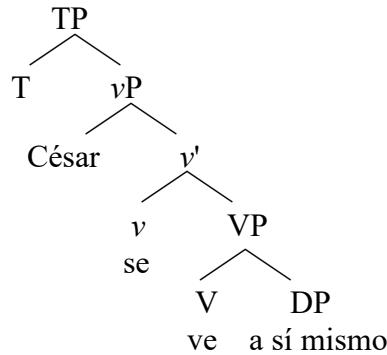
(77) Stage (c) ROC - Modern Spanish

a. César se ve a sí mismo.

César Refl_{se} DOM self same

‘César sees himself.’

b.



What about MidS? It was shown in Section 4.3.1 above that there are sporadic occurrences of reflexive doubling starting in the 15th century and dramatically increasing in the 16th century. Furthermore, based on interpolation and auxiliary selection patterns I concluded that MidS was at stage (b). Given the conclusions I reached about ACD in Chapter 3, the movement and copy analysis based on Harizanov (2014) Kramer (2014) would apply to this type of doubling. If we extend this to reflexives, which these previous authors do not, then the reflexive merges as verbal complement and moves to Spec, v. M-merger takes place and the higher copy is spelled out as a reflexive clitic. Nevertheless, there may be reasons that this analysis does not work for reflexives specifically, especially given the known differences between reflexives and object pronouns with respect to binding and coreferentiality. Since this issue will take us too far afield for the purposes of this dissertation, I leave it aside for now.

4.3.4 Doubling with other types of *se*

It should be noted that doubling does not take place with anticausative *se* (Anti_{se}), hence the ungrammaticality of (78).

(78) *El bosque se quemó a sí mismo.

the forest Anti_{se} burned DOM self very

Intended: ‘The forest burned.’

As previous authors have noted (Mendikoetxea 1999, Schäfer 2008, Koontz-Garboden 2009, MacDonald 2017), Anti_{se} can occur with a causal *by*-phrase, as in (79).

(79) El bosque se quemó por sí solo anoche.⁵⁹

the forest Anti_{se} burned by self alone last-night

‘The forest burned by itself last night.’

However, this does not count as doubling since the PP is an adjunct rather than a constituent merged in the complement position originally occupied by *se* prior to reanalysis. Nevertheless, we should not expect doubling to occur with Anti_{se} since it is an intransitive structure, having a single DP argument.

As for passive and impersonal *se*, the situation is more complicated. Several authors (Mendikoetxea 2008, MacDonald 2017) have argued that these types of *se* consist of a single DP as the internal argument, *se* as the head of Voice, and a *pro* in the external argument position. A possible analysis is that *se* is doubling *pro*, but the sake of simplicity, I leave this issue unaddressed for now. See Chapter 5, Section 5.3, for an analysis of the development of reflexive constructions as the result of cyclic interaction between the ROC and the Subject Agreement Cycle, wherein *pro* plays a significant role.⁶⁰

⁵⁹ Adapted from MacDonald (2017).

⁶⁰ Chapter 5, Section 5.3, is based on Maddox & MacDonald (in contract).

Additionally, MacDonald & Huidobro (2010) note that Asp_{se} also cannot be doubled, as in (80) below.

(80) Juan se comió la paella (*a sí mismo).⁶¹

Juan Asp_{se} ate the paella DOM self very

‘Juan ate (up) the paella.’

In this case, within the analysis I am proposing, we would not expect doubling because there is already a direct object, *la paella*, occupying complement position and *a sí mismo* has no potential merging sites.

4.4 Chapter summary

The Reflexive Object Cycle is a grammaticalization process that takes a reflexive object pronoun and turns into a valency-marking morpheme, a Voice or v -head. In this chapter I presented evidence that in Latin and earlier stages of Spanish, the reflexive pronoun was a full DP. It was later reanalyzed as a D-head and then a Voice head. This reanalysis is supported by diagnostics of interpolation, modification, coordination, and doubling. The possibility of interpolating, modifying, and coordinating the reflexive are lost going from Latin to Spanish, while patterns of doubling only start to emerge later. This is to be expected if the reflexive changed from a DP to a head.

⁶¹ Adapted from MacDonald & Huidobro (2010).

Chapter 5: Agreement Cycles and the Licensing of Null Arguments

Introduction

The goal of this chapter is to explore how null subjects and null objects can develop in a language as a result of the Subject Agreement Cycle (SAC) and the Object Agreement Cycle (OAC), respectively. The primary focus is Spanish but I also discuss data from throughout Romance including French, Brazilian and European Portuguese, etc. Building on van Gelderen (2011), the common thread connecting the SAC and the OAC together in this chapter is that once the relevant pronoun is reanalyzed as uninterpretable features on T for subjects or *v* for objects (stage c), the argument position where it originally merges at an early stage in the cycle is open for *pro* to merge. At stages (a) and (b) *pro* is not able to merge because that position has already been occupied by either a subject or object pronoun/clitic, depending on the cycle. Thus, only a language that has arrived at stage (c) of the cycle will allow referential null arguments. I also discuss how clitic-left dislocation, accusative clitic doubling, and null objects only become available at specific stages of the OAC and I illustrate how interaction between the SAC and the Reflexive Object Cycle (ROC) relate to the development of reflexive clitic constructions.

The format of this chapter is as follows. In Section 5.1 I review the primary theoretical work on the licensing of null arguments that I am adopting. In Section 5.2 I show how null objects develop in a language as a result of the OAC. I argue that the synchronic and diachronic distribution of object clitic left dislocation (CLDT) and accusative clitic doubling (ACD), which may or may not involve null objects, is a direct result of what stage of the OAC a language is in. More specifically, I show that CLDT is available at every stage of the OAC while ACD is only possible at stages (b) and (c). Crucially, these constructions appear identical on the surface but

their derivations vary depending on the stage. I extend Holmberg's (2005, 2010) and Holmberg et al's (2009) D-feature and topic-identification analysis of null subjects to null objects. My main claim is that the D-feature in *v* that licenses null objects is there due to the reanalysis of clitics that is a part of the OAC. Section 5.3 focuses on null subjects and reflexive constructions.¹ I propose that null subjects are a result of the SAC. Building on Bahtchevanova & van Gelderen (2016), I show that the SAC and the ROC have interacted historically to give rise to passive *se* constructions. The SAC yields subject *pro* and the ROC changes *se* into a Voice or *v* head, both elements which are needed to form passive *se* per MacDonald's (2017) analysis. A prediction falls out of this whereby a language can only develop passive *se* if it has null subjects and has reanalyzed the reflexive pronoun as a Voice/*v* head. This prediction holds in Spanish, French, and other languages. I summarize the findings of this chapter in Section 5.4

5.1 Background: Licensing null arguments

Since this chapter revolves around the relationship between null arguments and grammaticalization, I review here the main theories from which I draw many of my own assumptions regarding how null arguments are licensed. First, I review Holmberg (2005, 2010) and Holmberg et al (2009), according to whom a D-feature in T is the primary trigger for allowing null subjects. I then present Roberts (2010ab) who gives a deletion analysis for null subjects that involves incorporation into T and chain reduction.

5.1.1 Holmberg (2005, 2010), Holmberg et al (2009): *D-in-T*

In subsequent sections my analysis relies on the idea that null objects are present in clitic-left dislocation (CLDT) and accusative clitic doubling (ACD). I take as a theoretical point of departure Holmberg's (2005, 2010) investigations of null subjects which rest on the presence of

¹ Section 5.3 is based on Maddox & MacDonald (in contract).

a D-feature in T. Languages that allow null subjects have this feature, while non-null subject languages lack it. The idea that T is somehow pronominal in nature has been argued previously, but Holmberg brings it up to date with assumptions made in the Minimalist Program.

Holmberg (2005) takes the null subject to be a radically deficient sort of pronoun per Cardinaletti & Starke (1999), a φP , that has an unvalued D-feature. Since there is a valued D-feature in T, when this φP merges it is valued as definite and thus receives a definite, referential interpretation. Null subject languages have D-in-T and so null subjects are always definite. These types of languages resort to other mechanisms such as impersonal-reflexives to derive a generic subject. Presumably the reflexive morpheme overrides the D-feature. The φP 's D-feature then remains unvalued and is interpreted as generic.

Holmberg et al (2009) modify Holmberg's (2005) analysis by stipulating that the D-feature in T is unvalued. They also adopt the notion of topic-licensing following Frascarelli (2007). This is necessary since it has been shown that unless a null subject has a coreferential topic that has been introduced in previous discourse, it will still be unacceptable (Samek-Lodovici 1996). Thus, every clause with a null subject has an Aboutness topic in the left periphery whose referential index it shares. With this adjustment, the D-feature in T is valued by the topic; i.e., it copies its referential index. The φP merges and then incorporates into T following, in part, Roberts' (2010b) analysis to be discussed below. Holmberg further states that Spec,T is consequently not projected and that the EPP is satisfied by the topic.

Given his analysis, Holmberg (2005, 2010) proposes a typology of null subject languages. Consistent null subject languages (cNSLs) have null referential subjects as a consequence of D-in-T. This is the usual null subject case like Italian and Spanish. Other languages, like Brazilian Portuguese and Modern Hebrew, are partial null subject languages

(pNSLs) which lack D-in-T. They can have null subjects but they will only receive a generic interpretation unless bound by a DP in a higher clause. The referential null subject is in Spec,T where it checks EPP. Generic null subjects remain in Spec,v; the EPP is checked by “some other category” (Holmberg 2010:102). Languages without null subjects are non-null subject languages (nNSLs) like English and Swedish. These languages are similar to pNSLs since they lack D-in-T. However, they also have a PF-conditioned EPP according to which the subject in Spec,T must also be phonetically realized. Holmberg et al (2009) reformulate this as the presence of a P-feature in T. Finally, there are radical or discourse null subject languages (dNSLs) such as Chinese and Japanese. These languages are distinct from other types in that while they do have null subjects, they are not related to φ -features because these languages lack φ -features completely. Thus, the D-in-T analysis does not apply to this dNSLs.²

5.1.2 *Roberts (2010ab): incorporation and deletion*

Roberts investigates the phenomena of null subjects and object clitics. While these are related in that they both involve merging a deficient pronominal-type argument, they are derived in different ways. He analyzes subject *pro* as phonologically null due to a process of deletion that takes place either in the syntax or at PF. Romance object clitics, however, undergo incorporation into *v*. Let us start with his analysis of null subjects. Following Holmberg (2005, 2010), Roberts adopts the notion of D-in-T being relevant to the licensing of null subjects but adds a novel component whereby the null subjects are not pronounced because they are deleted. Roberts proposes that subject *pro* is a simultaneously minimal/maximal element, $D^{\min/\max}$. T has unvalued φ -features, a D-feature, and an EPP feature. *Pro* moves to Spec,T where it values T’s φ - and D-features, and satisfies EPP. In cNSLs, T has an EPP feature which *pro* lacks and so

² See Holmberg & Phimsawat (2015) for an analysis of generic pronouns in discourse null subject languages such as Thai.

pro's features are a subset of T's, rendering *pro* a defective goal per Roberts' (2010b:70) definition of defectivity.³ He further proposes a generalization whereby defective goals are either deleted and/or not pronounced independently of their probe. Hence, *pro* deletes. In cNLs there is a D-in-T by virtue of it having fully specified morphological features per Müller's (2005) account. Languages with impoverished subject agreement morphology lack D-in-T.

Turning now to object clitics, Roberts analyses them as $\varphi^{\min/\max}$; i.e., a deficient pronominal bearing φ -features. Cliticization in Romance is a matter of Agree and incorporation. Object clitics are probed by *v*. The clitic values *v*'s φ -features and then incorporates into *v*. Roberts distinguishes between this type of Agree and that which takes place when EPP is involved. In the latter case, the object will check features of *v* and move to Spec,*v*. Since clitics are deficient and *v* lacks EPP in languages like ModS, there is no movement to Spec,*v* but rather incorporation, which is why object clitics are phonetically realized affixed to their verbal host. This makes the prediction that languages which have object clitics of the Romance kind will lack EPP in *v*. Thus, Roberts (2010b:68) notes that SOV languages lack object clitics since SOV order is derived by EPP-triggered movement (Kayne 1994). In sum, the object clitic merges as an argument and then moves as a head to incorporate into *v*. This accounts for simple patterns of Romance cliticization where there is no doubling.

How does Roberts account for clitic doubling? There are two types of clitic doubling that are relevant here: subject clitic doubling (SCD) and accusative clitic doubling (ACD). The first

³ Roberts (2010b:76) - "The postulation of the D-feature on T in null-subject languages means that *pro* counts as a defective goal in such languages. Its features, phi and D, are properly included in T's. But T also has an EPP-feature, which *pro* can satisfy, as we have seen."

occurs in some varieties of non-standard Italian. Roberts discusses Northern Italian and Tuscan varieties. An example of SCD in Montesover Trentino is provided from Poletto (2000):

(1) El popo *(el) magna el pom.

the child he eats the apple

‘The child eats the apple.’

In (1) the subject *el popo* is obligatorily doubled by the clitic *el*. In languages that allow SCD, T has EPP, D-, and φ -features. In SCD constructions, the overt DP’s D-feature is realized on T via Agree. The D- and φ -features of the DP incorporate into T and the remnant DP moves to Spec,T for EPP. The clitic is the realization of D- and φ -features on T. Since these languages have D-in-T, they also license subject *pro*.

As for ACD, Roberts considers the possibility of extending the same analysis he gave for SCD, only in this case the clitic would be agreement features on *v*. He dispenses with this, however, since *v* lacks D in languages that exhibit ACD, as evidenced by the absence of referential null objects. Additionally, he claims such an analysis would not account for Kayne’s Generalization and the specificity effects that arise with ACD. Instead he adopts and modifies Uriagereka’s (1995) “big DP” analysis. First, he postulates that the clitic phrase is a φ P, consistent with what was seen for subject clitics, rather than a DP *contra* Uriagereka. Additionally, whereas Uriagereka places a *pro* as complement to the D-head in his big DP, Roberts argues the doubled constituent is there and not in Spec,D. Roberts’ DP also has an *n*P layer. After N raises to *n*, *n*P raises to Spec, φ , which is triggered by an EPP-feature on a part with EPP in T.

Roberts argues against the presence of a D-feature in *v*, the most notable evidence being the lack of null referential objects in languages with ACD like Rioplatense.⁴ Thus, while SCD involves EPP, D-, and φ -features, ACD appears to be more complicated and does not relate to a D-feature at all. This raises the question that if both types of doubling are actually instantiations of agreement, why are they derived differently? Below I pursue an analysis in which SCD and ACD are considered to be the same process at work in distinct domains. SCD occurs within TP while ACD takes place in *vP*. The observable differences between these types of doubling is simply a matter of the status of the clitic and which features are present on the relevant functional head. Where the clitic is a DP/D, doubling will be restricted. If the clitic has been fully reanalyzed as φ -features on T or *v*, doubling will be unrestricted. In both types of doubling, where D is present in T or *v*, a null subject or object is licensed.

5.2 Null objects as a result of the Object Agreement Cycle

Modern Spanish exhibits both clitic-left dislocation (CLDT), as in (2), and accusative clitic doubling (ACD), as in (3) below:⁵

(2) Los libros _i los compré ayer _i	(3) Juan la _i abrazó (a ella) _i / % (a María) _i
the books them I-bought yesterday	Juan her he-hugged DOM she DOM María
‘I bought the books yesterday.’	‘Juan hugged her / María.’

Historically, CLDT appears prior to the development of ACD. This seems odd since both constructions involve a clitic that corefers with a lexical DP, albeit in different positions. Furthermore, standard Modern Spanish (ModS) disallows null referential objects (NROs) while

⁴ Actually there is evidence of null referential objects in Rioplatense per Schwenter (2006), as I discuss in Section 5.2.1 below.

⁵ In previous literature, clitic left dislocation has typically been abbreviated as CLLD and clitic doubling as CLD. Since CLLD and CLD are only separated by a single <L>, which can be confusing for the reader, I have adopted CLDT and ACD instead.

non-standard varieties like Rioplatense allow NROs (Schwenter 2006). While these patterns may seem unrelated, in this section I argue that they are related since each one is tied to the reanalysis of object clitics, i.e., the Object Agreement Cycle. Additionally, I extend Holmberg's (2005) D-in-T analysis of null referential subjects to NROs. My main claim here is that a D-feature is present on *v* due to the Object Agreement Cycle (OAC), by which object clitics become object agreement morphemes.

I propose that at stage (a) of the OAC, CLDT is actually a dislocated topic with a coreferential full DP resumptive pronoun. At stage (b), CLDT involves movement of the object clitic as a DP to Spec,*v* followed by m-merger, similar to the analysis of ACD from Chapter 3. At stage (c) the clitic in CLDT is a *v*-head with a D-feature, and either *pro* or a lexical object can be in complement position. In Chapter 3, I proposed that m-merger is reanalysis. Stage (b) is where m-merger takes place. The DP clitic is reanalyzed via m-merger as a complex D-*v* head. This is how the D-feature ends up on *v*. Once a language has D-in-*v*, NROs are licensed. At first, D-in-*v* will be realized as a clitic. Thus, what I will refer to as “simplex cliticization” as in (4) below actually involves an NRO at stage (c).

(4) Juan lo_i leyó pro_i.

Juan it read it

‘Juan read it.’

Example (4) is also possible at stage (b) but it does not involve *pro*. Instead, the clitic merges as argument and moves to *v*. There is no m-merger because in this case because only pronominal objects are deficient enough to undergo m-merger, which is how ACD is derived. In simplex cliticization structures like (4) eventually the overt realization of the D-feature (the clitic) will be lost and NROs are licensed without it, as in Rioplatense, which is stage (c) of the OAC. This

accounts for the historic pattern of CLDT appearing before ACD. A prediction falls out of my analysis: a language with clitic-less NROs will have developed unrestricted ACD first. This prediction holds throughout Romance. As a consequence, I propose a typology of null object languages similar to Holmberg's (2005, 2010) typology of null subject languages.

5.2.1 *Diachronic/synchronic distribution: CLDT, ACD, and null objects*

5.2.1.1 Clitic left dislocation (CLDT)

CLDT is a topicalization operation whereby a constituent in the left periphery has a coreferential clitic immediately adjacent to the verb, as in (5) below:⁶

(5) Las flores_i las_i compré ayer.
the flowers them I-bought yesterday
'The flowers, I bought them yesterday.'

Olarrea (2012) identifies the following properties of CLDT: 1) the dislocate can be any type of XP, including non-argumental ones such as PPs; 2) the co-referring element must be a clitic or "an empty pronominal licensed by agreement;" 3) CLDT can be embedded without restriction; 4) the dislocate and the clitic must have identical Case and subcategorization; 4) multiple constituents may be dislocated; 5) CLDT is insensitive to weak islands such as wh-islands; 6) the dislocate may be separated from the rest of the clause by a comma or intonational break, but the break is much weaker than that of hanging left-dislocated topics.

Formal analyses of CLDT often make explicit the contrast with hanging topic left dislocation (HTLD). Since the former is more "connected" grammatically to the entire clause (case and subcategorization), some have analyzed it as movement and HTLD as base-generation. Most agree that the dislocate in HTLD is base-generated, but the analyses vary with respect to

⁶ Modern Spanish data and discussion adapted from Olarrea (2012), unless otherwise indicated.

CLDT. The problem is that certain diagnostics show that in some ways CLDT patterns with wh-movement while others show the contrary. For example, CLDT is sensitive to strong syntactic islands which motivates a movement analysis as in Cinque (1977), Dobrovie-Sorin (1990), Kayne (1994), *inter alios*. On the other hand, CLDT does not trigger subject-verb inversion, and so a base-generation approach might make more sense (Anagnostopoulou 1997, Zagona 2002). See Olarrea (2012) for more data and discussion.

Let us now consider the distribution of CLDT, starting with OldS. CLDT occurs in the earliest documents, including *El poema del Cid*, though it is apparently rather rare and increases in later centuries (Company Company 2003). The following are examples from OldS taken from Riho (1988).

(6) [La tierra del Rey Alfonso]_i esta noch la_i podemos quitar.⁷

the land of-the king Alfonso this night it we-can leave

‘Tonight we can leave King Alfonso’s land.’

(7) [vuestras mannas]_i bien las_i sabemos.⁸

your abilities well them we-know

‘We know your abilities well.’

Recall that at this point clitic pronouns are still full DPs in OldS per the OAC, thus any analysis of OldS CLDT that adopts the stages of the cycle must take this into account.

As was shown above, ModS has CLDT, as do all varieties of Spanish, as far as I am aware. Interestingly, however, authors such as Suñer (2006) and Estigarribia (2017) have observed variation in Rioplatense Spanish. Consider the following data from Suñer (2006):

⁷ Anonymous, *El Cid*, l. 423; c. 1207

⁸ Anonymous, *Razones d’Amor*, l. 175; c. 1205

(8) [A mi mejor amiga]_i, la_i vi [a esa loca linda]_i el jueves.

DOM my best friend her I-saw DOM that crazy beautiful the Thursday

‘I saw my best friend, that crazy beautiful girl, on Thursday.’

(9) [A Menem]_i, nadie lo_i votará [a ese estafador sinvergüenza]_i.

DOM Menem no-one him will-vote DOM that swindler shameless

‘Menem, no one will vote for that shameless swindler.’

Examples (8) and (9) are similar to typical CLDT since a topic on the left has a resumptive clitic with which it matches in case, number, and gender. However, these data differ from standard CLDT since there is a nominal object after the verb. Suñer (2006) labels the postverbal object in these cases the “epithet.” CLDT with epithets pose a challenge for analyses where it is assumed the topic merges first as complement and then moves to the left periphery, since there is now an additional object to account for. Suñer also notes that CLDT with epithets is allowed with non-human animates as well, but not with inanimates, as her data below show:

(10) A nuestro gato, mi hija no lo quiere más a ese bribón.

DOM our cat mi daughter not it loves more DOM that naughty

‘Our cat, mi daughter does not love that naughty thing anymore.’

(11) *Esa motocicleta, ¿a qué demente se le ocurrió comprarla a ese

that motorcycle DOM what deranged it to-him occurred to-buy-it DOM that

máquina infernal?

machine infernal

‘That motorcycle, to what deranged individual did it occur to buy that infernal machine?’

Before moving on to ACD, it should be noted that I have not found any examples of CLDT with epithets in my OldS corpus. In fact, it would be unexpected to get these configurations in OldS based on the stages of the OAC and the analysis that I give below in Section 5.2.2.

5.2.1.2 Accusative clitic doubling (ACD)

Patterns of ACD variation are discussed in Chapter 3, with data taken from Suñer (1988), van Gelderen (2011), and Ormazabal & Romero (2013). I refer the reader there for the details, but the essential patterns can be summarized as follows. In “standard” ModS, ACD is obligatory with pronominal direct objects and is unacceptable otherwise. Thus, lexical direct objects of any type cannot be doubled as ACD, as in (12) below.

(12) Juan lo_i vio a él_i / *a Miguel_i / *al carro_i.

Juan him saw DOM he DOM Miguel DOM-the car

‘Juan saw him / Miguel / the car.’

This is what I refer to as “restricted” ACD. Other varieties, on the other hand, display “unrestricted” ACD since they can double proper names, inanimates, and quantifiers. These varieties include Rioplatense, Limeño, Malinche, and Basque Spanish. I argued in Chapter 3, following van Gelderen (2011), that the reason these varieties have unrestricted ACD is because the object clitic is at stage (c) of the OAC, while in ModS object clitics are at stage (b).

According to Gabriel & Rinke’s (2010) quantitative study, “true” doubling of pronominal objects starts in the 15th century. Gabriel & Rinke’s corpus shows that ACD of pronominal direct objects was not the majority pattern until the 16th century. Prior to that, pronominal objects occur more frequently without a coreferential clitic. Naturally, distinguishing between CLDT and ACD is more difficult to do diachronically given the nature of written texts. The primary diagnostic is whether the coreferential lexical DP appears post-verbally (ACD) or

preverbally (CLDT). Below are some early examples of ACD. Example (13) is interesting since it illustrates that ACD was possible when enclisis was still allowed.

(13) e matáronlo a él e a uno de los que yvan con él.⁹

and they-killed-him DOM he and DOM one of those that went with he

‘And they killed him and one of those that went with him.’

(14) y después lo prendieron a él, como diremos...¹⁰

and afterwards him they-captured DOM he as we-will-tell

‘And afterwards they captured him, as we will tell...’

Nevertheless, ACD was by far the exception to the typical pattern of non-doubled pronominal objects, as in (15) below.

(15) y ella muy bien veía a él.¹¹

and she very well saw DOM he

‘And she saw him very well.’

Gabriel & Rinke (2010) observe that ACD was quite rare in older texts and when resumptive pronouns occur they tend to corefer with a dislocated object. Thus, most apparent examples of ACD in early OldS are actually clitic right dislocation (CLRD). This position is also maintained by Eberenz (2000). How do these authors distinguish between ACD and CLRD given that the surface strings could be identical? CLRD is well-known to be phonologically distinct from ACD in that it involves a prosodic break from the rest of the utterance. This can be indicated via a comma or by means of a caesura in poetic texts such as the *Poema de mio Cid*.¹²

⁹ Anonymous, *Crónica de Juan II de Castilla*, para. 201; 1406-1411

¹⁰ Pedro Cieza de León, *Las guerras civiles peruanas*, para. 577; c. 1553-1584

¹¹ Anonymous, *Libro del conde Partinoplés*, para. 91; c. 1500

¹² See Fontana (1993) and Eberenz (2000) for more discussion of this point.

5.2.1.3 Null objects

Null direct objects are common in some languages of the world like Pashto. However, while ModS is a conventional null subject language, the extent to which null objects are licensed is much more limited. Consider the following ModS data from Schwenter (2006):

(16) Fui a la tienda a comprar café_i pero no *pro_i* tenían.

I-went to the store to buy coffee but not it they-had

‘I went to the store to buy coffee but they did not have it.’

(17) Fui a la tienda a comprar el periódico_i pero no *lo_i* / **pro_i* tenían.

I-went to the store to buy the newspaper but not it it they-had

‘I went to the store to buy the newspaper but they did not have it.’

In (16) the direct object of both clauses is *café*, but in the second clause it is null. The object in (17) is *el periódico* and it cannot be null in the second clause but must be resumed by an object clitic. The difference between (16) and (17) is the type of direct object involved. In (16) the object is non-specific or non-referential, while in (17) is specific and referential. Thus, ModS only allows null objects of the former type and not the latter. Typical null object languages like Pashto regularly license null referential objects (NROs), and so ModS is not a null object language in the conventional sense. An additional environment where null objects are allowed in ModS is given in (18) below, which is a conversation between two interlocutors.

(18) a. ¿Compraste pan_i / el libro_j ?

‘Did you buy bread / the book?’

b. Sí, compré *pro_i* / **pro_j*.

‘Yes, I bought it.’

The null object in (18b) is the same type as what was seen in (16) above; i.e, non-specific/non-referential.¹³

While ModS does not allow NROs, there is dialectal variation in this regard. According to Schwenter (2006:28), NROs (he uses the term “specific” rather than “referential”) are found in the Spanish spoken in the following locations in South America: northwest Argentina, Bolivia, Ecuador (Quito), Peru, Uruguay (on the Brazilian border), Paraguay, and “River Plate Spanish,” also known as “Rioplatense,” which is the term I have been using throughout.¹⁴ An example of an NRO in Rioplatense is in Schwenter’s data in (19) below.

(19) a. Queremos el postre_i.

we-want the dessert

‘We want the dessert.’

b. Ya traigo *proi*.

now I-bring it

‘I’m bringing it now.’

Importantly, Schwenter, following Masullo (2003), notes that these NROs are only allowed when “the referent can be anaphorically recovered from the immediate context of utterance.” This suggests a connection with information structure and the licensing of null objects to which I return in my analysis. As for OldS, I have found no examples of either NROs or non-referential null objects.

5.2.1.4 Summary

The patterns discussed in this section are summarized in Table 5 below.

¹³ For simplicity sake I am not going to dwell on the differences between specificity and referentiality. Henceforth I use the terms ‘referential’ and ‘non-referential.’

¹⁴ See Cyrino & Matos (2016) for an analysis of null objects as VP ellipsis.

	Old Spanish	“Standard” Modern Spanish	Rioplatense Spanish
Critic-left dislocation	✓	✓ (epithets disallowed)	✓ (epithets allowed)
Accusative clitic doubling	✗	✓ (restricted)	✓ (unrestricted)
Null referential objects	✗	✗	✓

Table 5. Diachronic and synchronic variation in Spanish

5.2.2 Analysis: $OAC \rightarrow D\text{-}in\text{-}v$

The distributional patterns seen above for CLDT, ACD, and NROs are tied to the stages of the OAC. Recall that the OAC has three stages per van Gelderen (2011). At stage (a) the pronoun is a full DP that merges as complement and can undergo further movement afterwards. On van Gelderen’s original analysis, at stage (b), the pronoun/clitic merges as a DP complement and then moves as a D-head to v . However, in Chapter 3 I observed that ACD is possible at both stage (b) and stage (c) and, consequently I proposed a different analysis for stage (b) ACD based on Harizanov (2014) and Kramer (2014). On this analysis, in ACD the DP object merges as verbal complement and moves to Spec, v . At this point, m-merger creates a complex head between the DP object and v . Two copies of the object are now present and both are pronounced, only the higher copy is realized as a D-clitic. This differs from van Gelderen since the whole DP moves, rather than just the head. Nevertheless, the end result is the same: association of a D-clitic with v . For stage (c), I remain consistent with van Gelderen’s (2011) original analysis. This final stage is when the clitic is reanalyzed as the features on v , thus it is the spell-out of v . At this point, a lexical object or *pro* can merge as complement and the cycle is renewed. Thus, at stage (b) ACD involves object movement and m-merger while stage (c) ACD is true agreement between v (spelled-out as the clitic) and a verbal complement.

For the languages under consideration here, OldS represents stage (a), ModS represents stage (b), and Rioplatense is at stage (c). All three varieties allow CLDT, but only ModS and Rioplatense have ACD, while only Rioplatense displays NROs. Nevertheless, since each variety is at a different stage in the OAC, the derivation of CLDT and ACD will not be the same, though it yields identical surface strings. While both CLDT and ACD involve clitics, the clitic is not the same category across stages of the OAC, and so the derivations are distinct. Let us consider how.

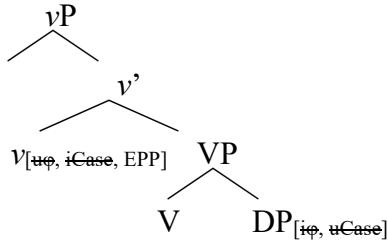
Following Harizanov's (2014:1072) analysis of CLDT in Bulgarian, I assume CLDT in ModS also involves m-merger. In both cases the DP object merges as verbal complement and moves to Spec, v . One might then assume that for CLDT the DP object moves up to a topic position in the left periphery, which I assume is CP.¹⁵ However, this is not possible because m-merger creates a complex D- v head, after which point the resulting head (spelled out as a clitic) is not accessible for further syntactic movement. How then does the coreferential argument end up in CP? Harizanov solves this by arguing that when the A' -probe searches for a goal, the only one available is the lower copy of the DP object in argument position since it has no access to the copy in Spec, v which is now a complex head. This original lower copy then moves to CP. Why is the original lower copy in argument position then not pronounced at linearization? Only the c-commanding copy in the left periphery is pronounced due to how chain reduction works following Nunes (2004). The steps in the derivation of ModS CLDT are summarized in (20) below:

¹⁵ Harizanov (2014:1073) labels the topic projection FP.

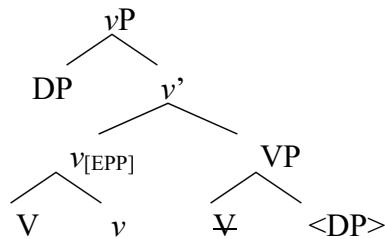
(20) [Las flores]_i las_i compré ayer.

‘The flowers, I bought them yesterday.’

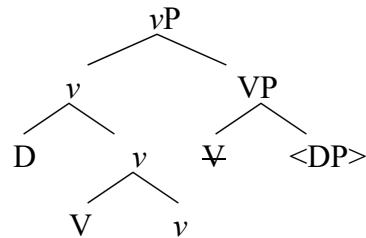
Step 1 → Object DP merges and Agrees with *v*.



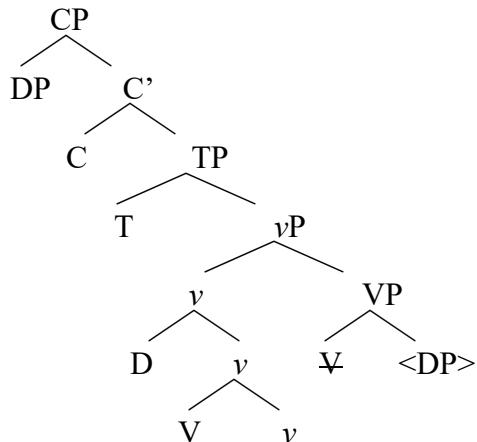
Step 2 → V moves to *v*; object DP moves to Spec,*v*.



Step 3 → M-merger between object DP and *v* yield D-*v* complex head.



Step 4 → Lower original copy of object DP moves to Spec,C.

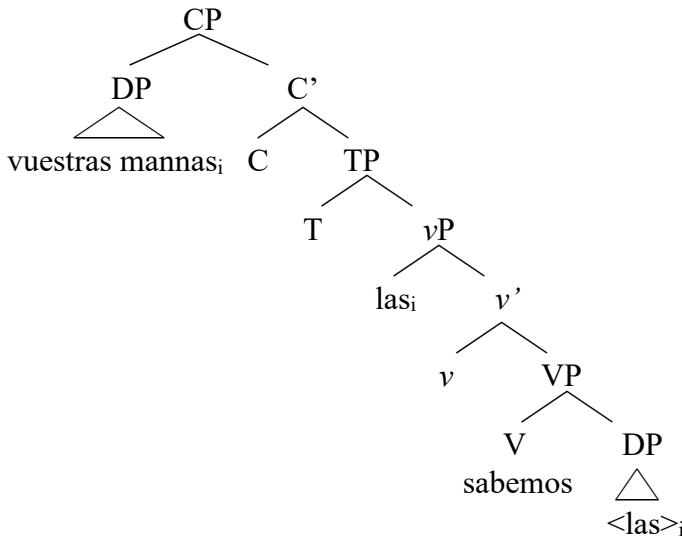


Thus, in CLDT there are three copies of the object involved. In ACD there are only two. In both cases there is m-merger with v , resulting in the presence of the clitic.

OldS also exhibits CLDT. Nevertheless, in OldS, the clitic is a full DP. In fact, it is more accurate to refer to it as an object pronoun since it was shown in Chapter 3 that it patterns as such. Given the status of the OldS object pronoun being at stage (a) in the OAC, I propose that what looks superficially like CLDT in OldS is actually a base-generated topic with a resumptive pronoun. In putative CLDT structures, the clitic merges as complement where it checks Case and receives its theta-role. It also values the φ -features on v . Object movement in OldS is triggered by an optional EPP-feature on v (see Chapter 3). This feature is also at work in OldS CLDT; i.e., the object clitic moves to Spec, v . As for the dislocated topic, I assume it is base-generated in the left-periphery, since the clitic is the argument. The structure of the CLDT example in (7) above is represented in (21) below:

(21) vuestras mannas_i bien las_i sabemos.¹⁶

‘We know your abilities well.’



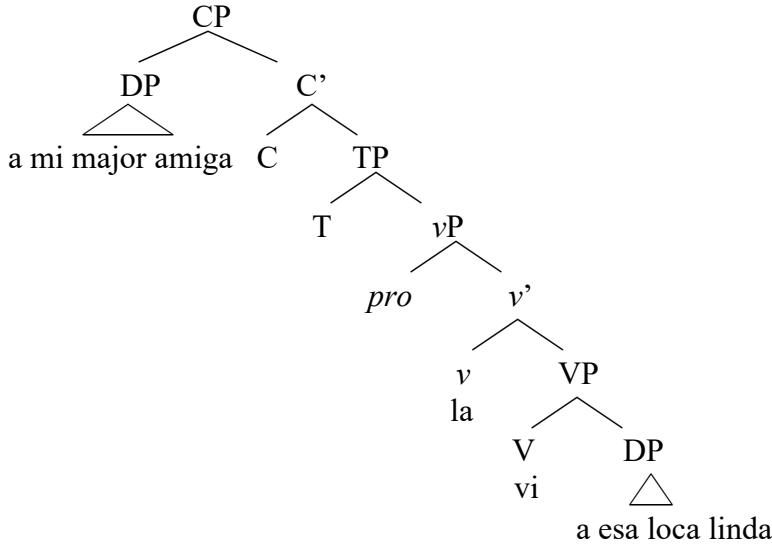
¹⁶ Anonymous, *Razones d'Amor*, l. 175; c. 1205

Thus, OldS CLDT is derivationally distinct from ModS CLDT because in the former the object “clitic” is just resumptive full DP pronoun. Other than diagnostics of interpolation, movement, etc., as was shown in Chapter 3, is there any additional evidence to back up this analysis of OldS CLDT? Recall that OldS object pronouns were inherited from Latin. In Latin object pronouns were also full pronouns and not clitics, so Latin did not have CLDT. Nevertheless, according Bortolussi (2017), Latin did have topicalization with resumptive pronouns. Thus, the same strategy was used in both Latin and OldS since both languages were at stage (a) of the OAC.

CLDT in Rioplatense allows epithets as in (8) above. This poses a potential problem since we now have three coreferential constituents to account for: the topic, the clitic, and the epithet. However, considering the status of the clitic in Rioplatense, it makes sense that epithets are allowed. Rioplatense is at stage (c) of the OAC, where the clitic is the spell-out of *v*. At this stage, it does not merge as the complement and so that position is open for a lexical object or *pro*, which is what happens in unrestricted ACD. This is essentially the same for CLDT. The topic is base-generated on the left, the clitic is the realization of *v*, and the epithet merges as the verbal complement. Thus, whereas in ModS CLDT there is movement of the original copy of the object to Spec,C, in Rioplatense CLDT, there is no object movement. The structure of Rioplatense (7) above is represented in (22) below:

(22) A mi mejor amiga_i, la_i vi a esa loca linda_i el jueves.

‘I saw my best friend, that crazy beautiful girl, on Thursday.’



For more details on the derivations of ACD, the reader is referred to Chapter 3, Section 3.3.4, since the discussion here is consistent with my proposals there. In sum, ModS exhibits restricted ACD and it is at stage (b) of the OAC. Per the Kramer-type analysis, the DP object merges as complement and then moves to Spec,*v*. After m-merger with *v* the higher copy is pronounced as a D-clitic. M-merger is the synchronic equivalent of diachronic reanalysis. Stage (b) ACD serves as Primary Linguistic Data to language learners, who reanalyze the complex head formed by the D-clitic and *v* as object agreement (stage c). Thus, ACD in Rioplatense is different. There is no object movement and the clitic is the realization of the *v*-head and the doubled object is in argument position; i.e., object agreement. OldS represents stage (a) since it lacks ACD. The object pronoun can undergo movement to Spec,*v* and possibly higher, but it is not targeted for m-merger.

Moving on to NROs, as was observed in the preceding section, they are allowed in Rioplatense and they are disallowed in OldS and ModS. I propose that this is to be expected

given how the OAC works; i.e., null objects are a by-product of the cycle. Furthermore, this is consistent with Holmberg's (2005, 2010) analysis, which was summarized in Section 5.1.1 above. While his analysis is concerned primarily with null subjects, in what follows I extend it to null objects as well.

Let us consider how a Holmberg-type analysis could work for null objects. As far as I am aware, Anders Holmberg has not explored the possible extension of his analysis of null subjects to null objects. Roberts' (2010b:78) does consider the role of D for null objects, albeit briefly. He observes that languages with rich null object agreement morphology would have a D-feature in *v* and so they allow null objects; e.g., Pashto. Roberts does not go into further detail, but it does seem desirable for a theory of null arguments that both subjects and objects would be licensed in a similar manner.

On a Holmberg-type analysis, there are three crucial ingredients required for a null argument: 1) a D-feature on a functional head (T for subjects), 2) incorporation of a φP , and 3) a base-generated topic. Thus, in Rioplatense all three of these elements must be present. Let us consider the derivation of the NRO in (23) below, from Schwenter (2006:28).

(23) a. Tengo un calmante; para dormir.
I-have a sedative to sleep
'I have a sedative in order to sleep.'

b. No tomes *pro_i*. Te va a hacer mal.
not take it you it-goes to make ill
'Don't take it. It will make you ill.'

Since we are concerned with null objects here rather than null subjects, the relevant functional head is *v*. The first assumption we need to make is that there is an unvalued D-feature on *v*, just

as there is an unvalued D-feature in T per Holmberg's (2005) analysis of null subjects. A φP merges as verbal complement and v probes the φP to have its unvalued φ -features valued. The φP has its Case feature valued by v . The φP is a defective probe since v 's features a superset of the φP 's; i.e. v has the unvalued D-feature which the φP lacks. There might also be an EPP-feature on v but I leave that aside for now. Since the φP is defective it incorporates into v , chain reduction applies and the φP is unpronounced. What about v 's D-feature? Notice that the data in (23) above and the majority of the data in Schwenter's (2006) study are in conversational exchange or question/answer format. In Schwenter's corpus NROs are allowed primarily when there is an immediate discourse referent available; i.e., a topic. This gives us the third ingredient for licensing a null argument. In (23b), the topic is *un calmante*, which was introduced into the discourse by the interlocutor in (23a). I propose, following Frascarelli (2007) and Holmberg et al (2009) that in (23b) the topic is null and is base-generated in the vP left periphery as a low topic (L-TopP), where it values the D-feature on v . As in Holmberg et al's (2009) analysis of null subjects, the topic checks the EPP-feature which, for null objects, is on v . The derivation of the NRO in (23b) above is represented in (24a), with the feature makeup of the relevant elements indicated in (24b).

(24) a. $[_{\text{TP}} \text{T} [_{\text{NegP}} \text{no} [_{\text{L-TopP}} \text{un calmante}; [_{\text{vP}} \text{v} [_{\text{VP}} \text{tomes} \phi \text{P}_i]]]]$

b. $\begin{array}{lll} \text{v} & \rightarrow & [\text{D: } _, \phi: _, \text{Case:Acc}] \\ \phi \text{P} & \rightarrow & [\phi: 2\text{S}, \text{Case: } _] \\ \text{un calmante} & \rightarrow & [\text{D: i}] \end{array}$

The essential take-away from the analysis of null objects I have proposed here is that they are licensed by a D-feature on v , just as null subjects are licensed by a D-feature in T . How do CLDT, ACD, and NROs then tie into the stages of the OAC? Each construction becomes possible at a different stage because of the categorial status of the clitic. In OldS and ModS

CLDT, there is no D-feature in v . The clitic is either a full DP or a DP that moves as to Spec, v and undergoes m-merger with v (DP/D- v). The only D-feature involved is on the clitic itself. In ModS ACD the clitic undergoes m-merger with v , which feeds reanalysis of it as a v -head. Once the clitic is fully reanalyzed as v , NROs become possible, as in Rioplatense, because v now has a D-feature. As is expected in grammaticalization cycles, the agreement morphology (the clitic) will eventually disappear.¹⁷ Now the overt realization of the D-feature is no longer expressed, but there is still a D-feature on v .

(25) Stage (a): clitic = DP; only CLDT allowed

Stage (b): clitic = DP/D- v ; CLDT and ACD allowed

Stage (c): clitic = v ; CLDT + epithets, ACD, and NROs allowed

CLDT without epithets are still available in Rioplatense as well. In this case, CLDT is a null object construction and the derivation is similar to the one I have given for NROs, only the topic is overt and in high topic position (CP) and the clitic is pronounced. There is a null object as the verbal complement by virtue of the D-feature on v . In ModS CLDT there is a “null object,” but on the analysis I have given it is the lowest copy of the original DP object and not a φP .

5.2.3 *Consequences*

5.2.3.1 Predictions and cross-linguistic patterns

My analysis accounts for the distribution of CLDT, ACD, and NROs in Romance.

CLDT appears before ACD because OldS CLDT is actually a topic with a resumptive DP pronoun, which is generally available in probably all languages with DP pronouns, like English: *That guy, I hate him*. ModS CLDT looks like OldS “CLDT,” but they are derivationally distinct. A prediction regarding NROs and ACD falls out of this which is not necessarily directly related

¹⁷ See van Gelderen (2011:42).

to the historical distribution of CLDT. If a language allows NROs it will have developed ACD first. NROs arise after ACD because it is in ACD where m-merger of D and *v* takes place which is the synchronic parallel of diachronic reanalysis. Since NROs are licensed by D-in-*v*, there must be an operation whereby D is associated with *v*. On my analysis it is m-merger that formally results in there being a D-feature on *v*.

Now let us see how these predictions fare across different Romance languages. First, consider how the different varieties of Spanish map onto the stages of the OAC as in (26) below.

(26) OldS → Stage (a): clitic = DP; only CLDT allowed
 ModS → Stage (b): clitic = DP/D-*v*; CLDT and ACD allowed
 Rioplatense → Stage (c): clitic = *v*; CLDT + epithets, ACD, and NROs allowed

Within Spanish, different synchronic and diachronic varieties represent different stages of the OAC. The constructions that occur in these varieties also correlate with stages of the OAC. OldS is the most conservative while Rioplatense is the most advanced with respect to CLDT, ACD, and NROs.

Now let us look outside of Spanish. One might expect there to be other languages that pattern like OldS such that the object clitic is a full DP and CLDT occurs but ACD and NROs do not. Incidentally, Italian and French both have CLDT but lack ACD and NROs.¹⁸ Given my prediction, French and Italian should also be at an earlier stage of the OAC. What evidence is there for this? In Section 3.2.2.2, following Culbertson (2010), I showed that obligatory repetition of a clitic in VP conjuncts is evidence that the clitic is becoming object agreement. OldS (27a) allowed the second clitic to be omitted but in ModS (27b) it is obligatory for most

¹⁸ Italian and French, like ModS, do allow generic or indefinite null objects.

speakers, just like subject agreement is obligatory. Thus, ModS is further along in the OAC than OldS.

(27) a. lo_i mató y Ø_i despedaçó...¹⁹ (Old Spanish)
b. lo_i mató y *(lo_i) despedazó. (Modern Spanish)
him it-killed and him it-tore-apart
'It killed him and tore him apart...'

Now compare the following data for French (from Kayne 1975:95) and Italian (from Luraghi 1997):

(28) Paul l_i'a insulté e Ø_i mis à la porte. (French)
Paul him.Aux insulted and him threw to the door
'Paul insulted him and threw him out the door.'

(29) L_i'ho baciato e Ø_i abbracciato. (Italian)
him-I-have kissed and him hugged
'I kissed him and hugged him.'

As in OldS, the object clitic in the second conjunct can be omitted and so French and Italian are less advanced in the OAC.

Portuguese is interesting since the Brazilian (BP) and European (EP) varieties do not pattern alike here. EP has CLDT but lacks ACD (Barrie 2000), so it should pattern like Italian and French with respect to the less grammaticalized status of its object clitics. Consider the following data from Luis & Kaiser (2016):

¹⁹ Pedro Mejía, *Silva de varia lección*, para. 264; c. 1540-1550

(30) Apenas a minha mãe me ajudou e (me) incentivou.

only the my mother me helped and me encouraged

‘Only my mother helped me and encouraged me.’

(31) Se me não engano...

if me not mistake

‘If I am not mistaken.’

Compare (30) with Italian (29). In both cases, the second object clitic can be omitted. The datum in (31) is particularly revealing since it is an example of interpolation, which following van Gelderen (2011) and others I have adopted as evidence for the phrasal status of the object clitic in OldS. Since interpolation is absent in French and Italian, EP is actually more conservative than they are.

BP is on the other end of the spectrum in that it is even more advanced than Rioplatense Spanish. Evidence for this is that while CLDT is possible in BP, the clitic tends to be replaced by a full pronoun (p.c. Janayna Carvalho), as in (32) below:

(32) A minha amiga_i, eu a_i / ela_i vi na quinta.

the my friend I her/her saw on-the farm

‘My friend, I saw her on the farm.’

In (32), the the stressed pronoun *ela* is preferred to the clitic *a*. BP also exhibits ACD (Machado & Rocha 2016) and allows NROs as in (33) below from Schwenter (2006):

(33) O João comprou um livro novo. Ontem ele trouxe Ø à aula.

the Juan bought a book new yesterday he brought it to class

‘Juan bought a new book. Yesterday he brought it to class.’

On my account, NROs are expected in BP since there is now D-in-v as a result of the OAC.²⁰ In BP, NROs can occur without a clitic because clitics (object agreement morphology) are being lost. This is consistent with the diachronic process known as “deflection,” whereby overt morphology disappears with time; e.g., loss of subject agreement morphology in French.

5.2.3.2 Towards a typology of null object languages

Recall the typology of null subject languages discussed above in Section 5.1.2, according to which there are four types: consistent null subject languages (cNSLs), partial null subject languages (pNSLs), non-null subject languages (nNSLs), and radical/discourse null subject languages (dNSLs). These classifications have been fairly useful in the realm of subject pro-drop, and given the discussion above it begs the question, why not have a similar typology for null objects? Since at least Rizzi (1986) it has been acknowledged that this property could be used typologically. Indeed, Roberts (2010b:78) tacitly makes this suggestion when he notes that the “arbitrary” null objects studied by Rizzi and others may be similar to null subjects in pNSLs. However, since the null objects that are allowed in Italian do not appear to be connected with verbal agreement, Roberts does not take the parallel further, other than pointing out that languages that do have referential null objects are likely to be OV and to have overt object agreement on the verb; e.g., Pashto.

Nevertheless, if we consider what distinguishes the different types of NSLs, the same patterns can apply to null objects as well. First, imagine what a consistent null object language would look like. A cNSL has a D-feature in T and has “rich” subject agreement on the verb.

²⁰ A couple languages to also consider here are Latin, since it is the ancestor of Romance, and Romanian. In Latin, object pronouns were full DPs so it lacked CLDT and ACD. However, NROs are allowed in limited contexts (Luraghi 1997). Romanian allows CLDT and ACD but not NROs (Avram & Coene 2009). Latin then would be even more conservative than EP while Romanian patterns like ModS. It should also be noted that EP and Old Italian display limited null objects; v. Raposo (1986) and Luraghi (2004). This is still consistent with my prediction since one would expect older varieties to be closer to Latin.

Null subjects are identified by a topic per the analyses I have adopted (Frascarelli 2007, Holmberg et al 2009). If we extend this to objects then we expect a consistent null object language (cNOL) to have a D-feature also, but in *v* rather than *T*, and rich overt object agreement. Pashto then would qualify as a cNOL.

Next consider pNSLs like Finnish and Icelandic. The most salient characteristic of these is that they disallow null subjects except when bound by a higher argument or when generic/indefinite. These languages also tend to have impoverished subject agreement paradigms. Subject agreement in Brazilian Portuguese, for example, is greatly reduced when compared with European Portuguese (Holmberg et al 2009, *inter alios*), which is why EP is a cNSL while BP is a pNSL. What would this look like for objects? I have been arguing throughout this chapter that object clitics in Spanish are becoming object agreement morphemes, but they not fully grammaticalized yet. In other words, Spanish does not have a full object agreement paradigm since Type 1 clitics (*lo/la/los/las*) still behave, at least in some respects, like DP/D rather than *v*-heads. One might hypothesize then that Spanish is a pNOL.

There are at least a couple of points that support this conjecture. First, there is abundant dialectal variation in Spanish with regards to ACD and object clitic morphology. This resembles what is seen in pNSLs such as the non-standard varieties of French and northern Italian dialects. Roberts (2010ac) discusses the extensive variation with respect to verbal subject agreement and subject clitic paradigms. These languages vary with respect to whether they display a full set of distinctions in person and number for verbal inflection and/or subject clitics. More importantly, they vary with respect to the type of subject clitic doubling (SCD) they allow. Building on data from Poletto (2000), Roberts (2010a:110) observes that in Central Veneto, for example, SCD is

obligatory with pronominal objects (35a), but optional otherwise (35b) and unacceptable with negatively quantified subjects (35c).

(35) a. Ti *(te) magni sempre.

you you eat always

‘You are always eating.’

b. Nane (el) magna.

John he eats

‘John eats.’

c. Nisun (*el) magna.

nobody he eats

‘Nobody eats.’

In Montesover-Trentino, SCD is obligatory with definite subjects (36a) but not allowed with nonreferential quantified subjects (36b):

(36) a. El popo *(el) magna el pom.

the child he eats the apple

‘The child eats the apple.’

b. Nissun (*el) me capis.

nobody he me understands

‘Nobody understands me.’

In Milanese, SCD occurs with quantified and non-quantified subjects but not when the subject has undergone wh-movement.²¹ And in Friulian, SCD occurs “in all environments.” These patterns are similar to those that were discussed for ACD in Chapter 3 above. Recall, for

²¹ See Roberts (2010a:112) for these data.

example, that ACD is obligatory in ModS for pronominal objects, but disallowed otherwise; i.e., restricted doubling. In Rioplatense, on the other hand, ACD is unrestricted, much like Friulian. Thus, different dialects of “Italian” display distinct patterns of SCD just as different dialects of “Spanish” display distinct patterns of ACD. This is because subject clitics are undergoing grammaticalization throughout Italian like object clitics are throughout Spanish. This is typical of a partial null subject/object language. Since the agreement paradigm is not completely grammaticalized, in some varieties doubling is “true” doubling and in others it is agreement.

The second piece of supporting evidence is the types of null objects that are allowed in Spanish. Crucially, they are generic and indefinite. The only way to get a referential/definite null object in ModS is with a co-occurring object clitic. This is like pNSLs where the subject clitic is obligatory with null referential subjects. It should be pointed out that since standard Italian also has generic/indefinite null objects it could qualify as a pNOL as well. Nevertheless, ACD is notably absent from standard Italian, and so it has not reached pNOL status as of yet.

Finally, what are we to make of Latin? In this language referential null objects occur with some frequency but there is no overt object morphology and no object clitics since object pronouns are full DPs. Furthermore, consistent with Roberts’ observation, Latin was OV. Perhaps object agreement morphology that was present at an earlier Proto-Indo-European stage disappeared via deflection, but the D-feature remained and was able to license null objects, similar to what I have proposed for BP. Alternatively, Latin could be a sort of discourse or radical null object language, like Japanese is a radical null subject language. See Luraghi (2004) for more discussion on the reconstruction of null objects in Proto-Indo-European. Naturally, this is a first step in developing a typology of null object languages, and further research is needed.

5.3 Cyclic interaction: null subjects and passive *se*²²

Having explored how null referential objects have developed in Rioplatense Spanish as a result of the Object Agreement Cycle, let us now turn to another type of unpronounced argument, null subjects. In this section, I examine the development of reflexive clitic constructions in the history of Spanish and French and how it relates to the presence of absence of subject *pro*. As was mentioned in Chapter 4, ModS displays various “types” of *se*; e.g., reflexive (37), anticausative (38), passive (39).

(37) Juan se lava.	(38) Se quemó el bosque.	(39) Se vendieron los pisos.
John Refl _{se} washes	AntiC _{se} burned the forest	Pass _{se} sold the flats
‘John washes himself.’	‘The forest burned down.’	‘The flats were sold.’

In what follows I argue that one of these types, Pass_{se}, is the result of interaction between two “linguistic cycles” in the sense of Bahtchevanova & van Gelderen (2016). The two main claims made are as follows: i) *pro*, which MacDonald (2017) argues to occur in Spec, Voice in Pass_{se}, can merge in Spec, Voice due to the Subject Agreement Cycle (SAC); and ii) *se* is a Voice head due to the Reflexive Object Cycle (ROC). The SAC was discussed briefly in Chapter 1 and I revisit it here in both French and Spanish. The different *se* structures are derived cross-linguistically by the presence/absence of *pro* and *se*’s status as Voice head rather than a DP argument. A resulting prediction is that the types of reflexive constructions that a language has depends on whether it has subject *pro* and whether it has grammaticalized the reflexive pronoun/clitic as verbal inflection. I demonstrate that this prediction holds for Latin, Spanish, French, and some non-Romance languages like German.

²² This section is based on Maddox & MacDonald (in contract).

5.3.1 Background: Bahtchevanova & van Gelderen (2016)

Modern Colloquial French is at stage (c) of the Subject Agreement Cycle; i.e., subject clitics serve as agreement that license *pro*. Bahtchevanova & van Gelderen (2016) show that French object clitics (participants in the OAC) interact with subject clitics (participants in the SAC) because they intervene between agreement affixes (subject clitics) and the verb.²³

(40) subject clitic + object clitic + verb → Je le mange. (I eat it.)

The two cycles are targeting structurally adjacent elements, which leads to cyclic interaction. For subject clitics to be reanalyzed as verbal morphology, they need to have no elements intervening between them and the verb. Nevertheless, object clitics do intervene.

Bahtchevanova & van Gelderen show that this can affect the OAC in three ways: 1) reanalysis of preverbal object clitics to being postverbal; i.e., replacement of the clitic with a full postverbal DP or PP; 2) object clitics can be reanalyzed as agreement, which is the expected outcome of the OAC; or 3) the object clitic may be lost completely; i.e., intransitivization (see references in Bahtchevanova & van Gelderen 2016:126). These changes could lead to different results: acceleration of the OAC to stage (b) or the skipping of stages (b) and jumping to stage (c).

Some phonological changes point to the reanalysis of object clitics as agreement in French. Bahtchevanova and van Gelderen (2016) observe, for example, that object clitics can form portmanteau morphemes with subject clitics. Additionally, doubling of objects like that in (41) below do occur:

(41) Celui-là, je l'ai pas vu.

that-there I it-have not seen

'I haven't seen that one.'

²³ Bahtchevanova & van Gelderen (2016) also consider the adverbial clitics *y* and *en* as part of the OAC, but for simplicity sake we focus here on the direct object clitics.

Nevertheless, given the prosodic pause, (41) is likely an example of clitic left-dislocation rather than agreement. The authors argue instead that preverbal object clitics are simply being replaced by postverbal constituents. This is based on data like (42) and (43) below, taken from the Orléans Corpus:

(42) [La langue de Bretagne]; je ne la_i comprends pas.

the language of Brittany I not it understand not

‘I don’t understand the language of Brittany.’

(43) Le gouvernement aura compris ça.

the government will-have understood that

‘The government will have understood it.’

In (42), the dislocated object is doubled by the object pronoun *la*. This is the expected pattern if the language were at stage (b). In (43) the object is replaced by the demonstrative pronoun instead of doubling. While in this corpus the first pattern is still more frequent than the second, Bahtchevanova & van Gelderen (2016) argue that since the second pattern is still quite common, some speakers are skipping stage (b) in the OAC. They conclude that this is due to interference from the subject clitics which are on their own cycle. The crucial point to take away here for what is to follow is that different types of interactions between cycles can occur when they target structurally adjacent elements.

5.3.2 *The SAC revisited*

As discussed in Chapter 1, subject agreement affixes are often derived from subject pronouns via a process of grammaticalization (Givón 1976, Fuß 2005, van Gelderen 2011).²⁴ This process has been documented in other language groups like Basque (Tauli 1958), Bantu

²⁴ Fuß (2005:3) gives an extensive list of typologically distinct languages that have developed agreement affixes from full pronouns.

(Givón 1976), Pama-Nyungan (Hale 1973), Iroquoian (Mithun 1991), Nahuatl (Haugen 2004), etc. The SAC as outlined in van Gelderen's (2011) study is comprised of three stages. At Stage (a) the pronoun is a DP merged in Spec, v and moved to Spec, T to contribute interpretable φ -features. The DP pronoun can be freely separated from the verb via shifting or scrambling. During stage (b) the pronoun/clitic is reanalyzed as a D-head and feature loss begins.²⁵ No additional coreferential nominal is allowed during this stage; i.e., no doubling. Stage (c) occurs when the clitic is reanalyzed as uninterpretable φ -features on T, triggering another element to merge. An additional coreferential constituent can merge; i.e., doubling is allowed.

5.3.2.1 SAC in French and Spanish

In Chapter 1, I presented subject clitics in Modern Colloquial French (MCF) as an example of the SAC at stage (c), such that clitics are now subject agreement. This is based off of previous work by Lambrecht (1981) and van Gelderen (2011). For the French details the reader is referred to Chapter 1, Section 1.3.1. What follows is a summary of that discussion. The SAC is best understood with Latin as a starting point. French inherited subject agreement from Latin, but the endings have undergone morphological attrition since. While Latin had distinct endings for all six persons, currently in “standard” Modern French (ModF), only the first- and second-person plural have distinct phonetically realized endings. Additionally, Latin and Old French (OldF) subject pronouns could be interpolated and elided. In ModF interpolation is no longer possible and in MCF all subject clitics must be repeated in VP conjuncts. Finally, subject clitics can undergo doubling in MCF, suggesting that MCF is at stage (c) of the SAC.

²⁵ In Chapter 3 I modified stage (b) of the Object Agreement Cycle in order to account for clitic doubling at both stage (b) and (c). I leave the question as to whether this same modification should be made to the SAC for future work.

Another development in French related to the SAC is the loss of null subjects.²⁶ While OldF was a null subject language, ModF is a non-null subject language. Adams (1987:42) identifies the loss of subject *pro* to have occurred in Late Middle French; i.e., late 1400s. During the medieval period, null subjects occurred primarily in matrix clauses and in V2 contexts; in embedded clauses with a coreferential argument in the matrix clause, the subject tends to be overt. The following data are examples of the typical environments for subject *pro* in OldF:

(44) si firent *pro* grant joie la nuit.²⁷

thus made they great joy the night

‘Thus they made great joy that night.’

(45) Lors s’acorderent *pro*_i que il_i diroient que il_i l’avoient baillié par le commandement Nichodemus.²⁸

command Nicodemus

‘Then they agreed that they would say that they had held him by the command of Nicodemus.’

Adams’ (1987) analysis is that in OldF *pro* is licensed by Infl, but only when “canonically” governed; i.e., government to the right in a head-first language. For this to happen in OldF, the verb must move up so that it governs Spec,T. Thus, loss of rich inflection is one factor in the loss of null subjects; the other is loss of V2 which meant the verb did not always move up to govern Spec,T.²⁹

²⁶ Data in this section adapted from Adams (1987), Roberts (2014)

²⁷ Robert de Clari, *La Conquête de Constantinople*; 13th century

²⁸ Chrétien de Troyes, *Le Roman du Graal*, 25-6; 1135-1190

²⁹ Roberts (2014) takes a similar approach to Adams (1987), but claims there are different ways to lose *pro* in null subject languages like French and Brazilian Portuguese.

As in French, Spanish subject affixes were inherited from Latin. The details of the SAC in Spanish were presented in Chapter 2, Section 2.3.2.1. Here I summarize that discussion. Unlike ModF, ModS maintains six distinct endings for subject agreement which it also inherited from Latin. In OldS, subject pronouns were full DPs since they could be separated from the verb and be coordinated. This continues to be the case in ModS. Additionally, in both OldS and ModS, the subject pronoun can be omitted in VP conjuncts. Finally, OldS and ModS subject pronouns cannot undergo the type of doubling exhibited by subject clitics in MCF. Thus, Spanish has remained at the end of stage (c). It has overt subject agreement on the verb and subject pronouns are topics. Spanish, unlike French, has also remained *pro*-drop since the earliest texts (MacKenzie & van der Wurff 2012), though some varieties of Caribbean Spanish are starting to show new developments in this regard (Martínez-Sanz 2011, Camacho 2016), particularly those that are losing distinct subject agreement endings (Duarte & Figueiredo Silva 2016).

5.3.2.2 Null subjects as a result of the SAC

Based on the discussion above, I propose that the ability to merge *pro* in Spec, Voice is due to the grammaticalization of subject pronouns as subject agreement morphemes; i.e., the SAC.³⁰ Since the pronoun is now the spell-out of the Voice head, *pro* is able to merge in Spec, Voice. Thus, null subjects are the result of the SAC. This has been claimed previously by others. Klausenburger (2000), for example, states that the pro-drop status of OldF is a result of the low-level of grammaticalization of subject pronouns at that stage. Van Gelderen (2011:44) makes a similar claim that “pro-drop is connected to the stages of the cycle,” but she adds that it is not the sole factor in determining where a language is at in the SAC.

³⁰ Similar claims have been made by Klausenburger (2000), Fuß (2005), and van Gelderen (2011).

Previous analyses of null subjects are easily adapted to the proposal above since they typically connect the subject agreement morphology to the licensing of *pro*. Rizzi (1982), for example, posits two conditions for null subjects to occur: licensing and identification. The licensing condition requires that *pro* be licensed by a parametrically defined head such as Infl, V, P, etc. According to the identification condition, feature identification will take place via binding of *pro* by that head. Roberts' (2010b) deletion analysis has the subject being deleted since it shares a subset of features with the licensing head T. Note that, based on my proposal, these features are only on T due to the SAC. Holmberg (2005) also relates the licensing of *pro* to a T head. On his analysis, a language will have consistent null subjects if that language has a D-feature in T. Given that I have adopted Holmberg's analysis of the licensing of null arguments, we can reframe stage (c) of the SAC as the point at which the D-feature originally contributed by the subject pronoun is reanalyzed on T.

5.3.3 *Cyclic interaction at work*

In Section 5.3.1 above, I reviewed Bahtchevanova & van Gelderen's (2016) study which showed that the SAC and OAC were interacting in French with various results affecting the distribution of object clitics. Both subject and object clitics occur preverbally and are immediately adjacent to each other, and both are undergoing grammaticalization. The SAC and the OAC interact because object clitics intervene between the subject clitic and the verb, which prevents subject clitics from being reanalyzed as verbal morphology. In this section I examine another type of cyclic interaction between the SAC and the Reflexive Object Cycle discussed in Chapter 4. I argue that the different types of reflexive constructions arise as a result of what stage in the SAC and ROC a language is in.

5.3.3.1 Spanish

In order to develop Pass_{se} , two elements are required: 1) subject *pro* and 2) *se* as a Voice head, not occupying $\text{Spec}, \text{Voice}$ where *pro* merges. The former is a result of the SAC; the latter is due to the ROC. Thus, the structure of Pass_{se} is as shown below:

$$(46) \quad \text{Pass}_{\text{se}} \rightarrow [\text{Voice}_P \text{ } pro \text{ } \text{Voice}_{\text{se}} \text{ } [\text{VP} \text{ } \text{DP} \text{ }] \text{ }]$$

In a sense, the two cycles feed each other. Since the two cycles target *se* and *pro*, two elements that are structurally adjacent, this is similar to the interaction studied by Bahtchevanova & van Gelderen (2016). OldS was a null subject language (it had subject *pro*) and *se* was reanalyzed as a head in Middle Spanish, as I argued in Chapter 4. Thus, it is not surprising to find Pass_{se} during this period. Spanish is the clear case where cyclic interaction yielded Pass_{se} .

5.3.3.2 French³¹

A prediction falls out of the aforementioned proposal regarding cyclic interaction: if a language has reached the stage of the ROC where the reflexive is reanalyzed as a Voice head and it has reached the stage of the SAC where subject clitics are agreement that license *pro* in $\text{Spec}, \text{Voice}$, Pass_{se} may develop. To see whether this prediction holds or not, consider the situation in French.

First, there appears to be little consensus in the literature as to when Pass_{se} developed in French. Naturally, dating such a development with any precision is rendered more difficult by external factors. Nevertheless, Nyrop (1899:204) claims the development of Pass_{se} in French took place during the 13th century, with its expansion during the Renaissance. Thomasett & Ueltschi (2007), on the other hand, claim that Pass_{se} already existed in the oldest French texts, which date back to the eighth and ninth centuries. Other studies such as Jensen (1990) and

³¹ Some data in this section adapted from Nyrop (1899), Cennamo (1993), and Wolfsgruber (2016).

Wolfsgruber (2016) suggest Pass_{se} hardly occurs in OldF. Commonly cited examples of Pass_{se} in OF include the following:

(47) or se cante.³² (48) Car amors ne se puēt celer.³³

now Pass_{se} sings for love not Pass_{se} can hide

‘Now it is sung.’ ‘For love cannot be hidden.’

(49) par qui l’evēsque se fera.³⁴

by whom the-bishop Pass_{se} will-make

‘By whom the bishop will be appointed.’

(50) ne faison rien qui soit ou Dieu ne se nomme.³⁵

not we-do nothing that be where God not Pass_{se} names

‘Let us not do anything that may be where God is not named.’

(51) La chose se deliveroit par le dit Thomas.³⁶

the thing Pass_{se} delivered by the called Thomas

‘The thing will be delivered by the one called Thomas.’

The earliest of the data above, example (48), dates from the 12th century, during a time when French was a null subject language.³⁷ In other words, at this time, OldF had subject *pro*, which is the first element needed to form Pass_{se} , based on the prediction. The second ingredient needed is for *se* to be a Voice head. That this was the case in OldF can be inferred by the lack of interpolation. Recall that this pattern occurs frequently in Latin and early OldS. In French, however, some authors have claimed interpolation never occurs (Ramsden 1963, Batllori et al

³² Anonymous, *Aucassin et Nicolete*; 12/13th century. This is the heading of a musical notation; it occurs several times in the manuscript; instructions to the reader.

³³ Béroul, *Tristan*; 12th century Norman

³⁴ Gautier de Coincy, *Miracles de Nostre-Dame*; early 13th century

³⁵ Anonymous, *La Farce de Maître Pathelin*; 1457

³⁶ *Chartes du Fores*, cited in Cennamo (1993:79) from Stefanini (1962:640).

³⁷ See Wolfsgruber (2017:264ff) for a discussion of the null subject status of Medieval French.

1995), while at least one study, Kaiser (2002), has shown that it does occur rarely, but with subject pronouns. Nevertheless, no study that I am aware of has found interpolation in OldF with the reflexive pronoun. This suggests that in OldF, *se* was indeed already a head. Thus, French actually conforms with the prediction above. It had null subjects and *se* was a head rather than a full DP. In this case, the prediction holds.

5.3.3.3 Other languages

Now consider whether this prediction holds when other languages are examined. In Latin, for example, the reflexive pronoun was a full DP and not a head, as was shown above in Section 4.2.1. Latin was also a null subject language. Thus, given the relevant prediction it should not have Pass_{se} because it has only one of the required ingredients. Latin conforms to this prediction since it does lack Pass_{se} . The structure for Latin Refl_{se} is given in (52) below:

$$(52) \quad [\text{VoiceP} \ pro \ \text{Voice} [\text{VP} \ s\bar{e}]]$$

Since Latin *sē* is a full DP it occupies the internal argument position while $\text{Spec}, \text{Voice}$ is filled with *pro*.

Given that Latin is the ancestor of Romance, it might not be the best test case for the prediction. How does the prediction fare among less related languages? German has an expletive *pro* but it lacks a referential subject *pro* (Roberts & Holmberg 2010). Additionally, the reflexive pronoun *sich* has been argued to be a full DP based on its free word order status, the presence of abstract case, and the appearance of the HAVE auxiliary (Schäfer 2008, Alexiadou et al 2015). Thus, German has neither of the required elements necessary to develop Pass_{se} . While German does have reflexive (53a) and anticausative *se* (53b), it indeed lacks Pass_{se} , as is shown by the unacceptability of (53c).

(53) a. Johann wäscht sich.

John washes self

b. Die Tür öffnete sich.

the door opened self

c. *Die Wohnungen verkauften sich.

the apartments sold self

Intended: 'The apartments were sold.'

A final case to consider is Swedish and Norwegian. In these languages there are two reflexive elements: 1) an independent DP *seg* or *sig*, and 2) an affixal form *-s*, derived historically from the reflexive pronoun (Geniušienė 1987). The full DP forms occur in reflexive and anticausative constructions while the affix *-s* only occurs with passives. This makes sense given the prediction. The independent DP *seg/sig* cannot appear in passives because there is no empty position for *pro* to merge. In contrast, the affix *-s* does not occupy an argument position, so *pro* can merge there. However, Swedish and Norwegian are non-null subject languages (Platzack 1987, Kinn 2016), so how can they have Pass_{se}? The problem is resolved in a similar manner to French. The ancestor of Swedish and Norwegian, Old Norse, was a null subject language and thus the passive reflexive construction probably developed at this point and was inherited as a remnant in the modern descendants. In fact, as was discussed in Section 4.1, a passive affix *-sk* did exist in Old Norse and therefore the prediction holds. Below are couple more examples of this form, taken from Barnes (2008).

(54) hann fyrirdœmisk af illum mœnum.

he condemned-SK by wicked men

'He is condemned by wicked men.'

(55) á hans dógum byggðisk Ísland.

in his days settled-SK Iceland

‘In his days Iceland was settled.’

5.4 Chapter summary

In this chapter I have shown how null subjects and null objects relate to the stages of the SAC and the OAC. In sum, null arguments are allowed in a language only if that language has reached stage (c) of the relevant cycle, such that the pronoun is reanalyzed as φ -features on T (for subjects) or v (for objects). I extended a D-feature and topic-identification type of analysis based on Holmberg (2005, 2010) and Holmberg et al (2009) to the licensing of null objects. Thus, null objects are licensed by a D-feature in v . I argued that this D-feature is only present on v in some varieties of Spanish because the clitic’s D-feature has been reanalyzed as a feature of v , which is due to how the OAC works. I showed how clitic left-dislocation and accusative clitic doubling are tied to the stages of the OAC, which accounts for their cross-linguistic distribution. As for null subjects, I illustrated how the SAC and the ROC have interacted in the history of Spanish to give rise to passive *se*. In order to develop passive *se*, two elements are need: null subjects and a reflexive Voice head. These elements are present due to the SAC and the ROC, respectively.

Chapter 6: Conclusion

In this dissertation I investigated two linguistic cycles in Romance per van Gelderen's (2011) framework: the Subject Agreement Cycle (SAC) and the Object Agreement Cycle (OAC). These cycles turn pronouns into agreement morphology on the verb. Both cycles are comprised of three stages. At stage (a) the pronoun is a full DP. At stage (b) the pronoun is reanalyzed as a D-head and at stage (c) it is reanalyzed as a T-head in the SAC or a *v*-head in the OAC. I extended the SAC to account for the grammaticalization of impersonal pronouns. I showed that Modern French *on* is at stage (c) of the Impersonal Subject Cycle (ISC). Old Spanish (OldS) *omne* was on this cycle but it disappeared only to be replaced by Modern Spanish (ModS) *uno*. I proposed that the reason for this disappearance was that impersonal subject pronouns will only be reanalyzed if personal subject pronouns are being reanalyzed first via the SAC. In Modern Colloquial French (MCF), subject pronouns/clitics are at stage (c) of the SAC and impersonal *on* is at stage (c) of the ISC. In Spanish, subject pronouns are at stage (a) of the SAC and generic *uno* is at stage (a) of the ISC.

I also built upon van Gelderen (2011) by examining the OAC in Spanish in more detail. I showed that in OldS, object clitics were full DPs and thus OldS was at stage (a) of the OAC. Based on diagnostics of coordination, modification, and movement, I determined that ModS object clitics are more deficient than OldS object clitics. Patterns of clitic doubling are evidence that standard ModS is at stage (b) of the OAC while Rioplatense Spanish is at stage (c). In ModS, only pronominal objects are doubled by the clitic while in Rioplatense, lexical DP objects can be doubled. I adopted an analysis of accusative clitic doubling (ACD) based on Harizanov (2014) and Kramer (2014) whereby the object merges and moves to Spec,*v* (object shift) as a DP.

M-merger creates a complex D-*v* head and this head is spelled out as a clitic. In ACD, both the high clitic copy and the low original copy are pronounced. The reason that ACD starts with pronominal objects is that prounoun clitics are deficient enough to undergo m-merger. Once the object clitic is reanalyzed as *v* per the OAC, there is no m-merger and so any lexical object can merge. I also showed how object movement feeds the OAC. Object movement results in the object pronoun being in Spec,*v*, where it is associated with the *v*-head. Since ACD depends upon object shift to Spec,*v*, we expect languages that have developed ACD to have had object movement at an earlier period. This is the case for Romance.

I showed that the reflexive clitic *se* in ModS has been subjected to a type of OAC which I labelled the “Reflexive Object Cycle” (ROC). The ROC is a grammaticalization cycle that takes a reflexive object pronoun and turns it into a valency-marking morpheme, a Voice or *v*-head. I presented evidence that in Latin and OldS, the reflexive pronoun was a full DP. It was later reanalyzed as a D-head and then a Voice head. This reanalysis is supported by diagnostics of interpolation, modification, coordination, and doubling. The possibility of interpolating, modifying, and coordinating the reflexive are lost going from Latin to Spanish, while patterns of doubling only start to emerge later. This is to be expected if the reflexive changed from a DP to a head.

I demonstrated that null subjects and null objects relate to the stages of the SAC and the OAC. In sum, null arguments are allowed in a language only if that language has reached stage (c) of the relevant cycle, such that the pronoun is reanalyzed as features on T (for subjects) or *v* (for objects). I extended a D-feature and topic-identification type of analysis based on Holmberg (2005, 2010) and Holmberg et al (2009) to the licensing of null objects. Thus, null objects are licensed by a D-feature in *v*. I argued that this D-feature is only present on *v* in some varieties of

Spanish because the clitic's D-feature has been reanalyzed as a feature of v , which is due to how the OAC works. I showed how clitic left-dislocation (CLDT) and ACD are tied to the stages of the OAC, which accounts for their cross-linguistic distribution. Languages like French and Italian lack ACD but do have object CLDT. This is because they are at stage (a) of the OAC while ModS is at stage (b). ModS ACD is restricted to pronominal objects while Rioplatense ACD is unrestricted. Further evidence that Rioplatense is at stage (c) is that in CLDT, an epithet can merge in argument position. This is because the object clitic is v and so the complement position is open for the epithet to merge. I proposed a typology of null object languages based on Holmberg's (2005, 2010) typology of null subject languages. As for null subjects, I illustrated how the SAC and the ROC have interacted in the history of Spanish to give rise to passive *se* (Pass_{se}). In order to develop passive *se*, two elements are needed: null subjects and a reflexive Voice head. These elements are present due to the SAC and the ROC, respectively. This also accounts for the presence or absence of passive reflexive constructions crosslinguistically. Pass_{se} develops when a language has subject *pro* and *se* as a Voice head.

Though this research has yielded interesting findings as summarized above, there is still more work to do. One issue to pursue is the prediction made regarding the ISC such that impersonal pronouns only undergo grammaticalization if personal pronouns do so first. While this prediction holds in French and Spanish, more languages need to be investigated in this regard. For the SAC, patterns of doubling should be looked at in more detail. Since I observed that doubling is actually possible at both stage (b) and stage (c) of the OAC, one might expect the same for the SAC. Some evidence supporting this would be if pronominal subjects double before lexical subjects. I would also like to consider if the movement and m-merge analysis I adopted for ACD could apply to reflexive doubling as well, and what ramifications this might

have for the ROC. Finally, since I claimed that the D-feature that licenses null objects only ends up on v due to CLDT, it should be investigated whether the same could be said for the D-feature in T that licenses null subjects.

Null arguments were an important part of this dissertation as well and I have, for the most part, avoided discussion of the Null Subject Parameter and the vast body of literature that has been conducted in that area. Future research should aim at determining how an understanding of linguistic cycles might inform work in parametric theory. The general framework I have adopted here is compatible with the conceptualization of parameters as a feature value on a functional head, as per the Borer-Chomsky Conjecture. However, if null arguments are just the result of grammaticalization cycles, which are natural linguistic processes, how does that relate to the Null Subject Parameter? For the time being, I leave this question and the other issues mentioned above for future work.

References

Achar, Michel. 2016. *Impersonals and Other Agent Defocusing Constructions in French*. Amsterdam: John Benjamins.

Adams, J.N. 2013. *Social variation and the Latin language*. Cambridge: UP.

Adams, Marianne. 1987. *Old French, Null Subjects, and Verb Second Phenomena*. Phd Dissertation: UCLA.

Alexiadou, A. & E. Anagnostopoulou. 1997. Toward a Uniform Account of Scrambling and Clitic Doubling. In Werner Abraham & Elly van Gelderen (eds.), *German: Syntactic Problems-Problematic Syntax*, 143-161. Tübingen: Max Niemeyer Verlag.

Alexiadou, Artemis, Elena Anagnostopoulou & Florian Schäfer. 2015. *External arguments in transitivity alternations: A layering approach*. Oxford: Oxford University Press.

Alkire, Ti & Carol G. Rosen. 2010. *Romance languages: a historical introduction*. Cambridge: UP.

Anagnostopoulou, E. 1997. Clitic left dislocation and contrastive left dislocation. In Elena Anagnostopoulou, Henk C. van Riemsdijk & Franz Zwarts (eds.), *Materials on Left-Dislocation*, 151-192. Amsterdam: John Benjamins.

Anderson, Stephen. 2005. *Aspects of the Theory of Clitics*. Oxford: UP.

Aoun, Joseph. 1981. The Formal Nature of Anaphoric Relations. Phd dissertation, MIT.

----- 1999. Clitic-Doubled Arguments. In Kyle Johnson & Ian Roberts (eds.), *Beyond Principles and Parameters: Essays in Memory of Osvaldo Jaeggli*, 13-42. Dordrecht: Kluwer.

Aoun, Joseph & D. Sportiche. 1981. The Domain of Weak Cross-Over Restrictions. *MIT Working Papers in Linguistics* 3:43-52.

Aranovich, Raúl. 2003. The semantics of auxiliary selection in Old Spanish. *Studies in Language* 27:1-37.

Augerot, James E. & Florin D. Popescu. 1971. *Modern Romanian*. University of Washington Press: Seattle.

Avram, Larisa & Martine Coene. 2009. Null objects and accusative clitics in Romanian.

Published online at <https://www.researchgate.net/publication/260060174>

Babcock, Sandra Scharff. 1970. *The Syntax of Spanish Reflexive Verbs: The Parameters of the Middle Voice*. The Hague: Mouton.

Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: UP.

----- 2008. *The Syntax of Agreement and Concord*. Cambridge: UP.

Badia Margarit, A.M. 1947. Los complementos pronominalo-adverbiales derivados de IBI e INDE en la Península Ibérica. *Revista de Filología Española, Anejo 38*. Madrid: S. Aguirre.

Bahtchevanova, Mariana & Elly van Gelderen. 2016. The interaction between the French subject and object cycles. In E. van Gelderen (ed.), *Cyclical Change Continued*. Amsterdam: John Benjamins.

Barbosa, Pilar. 1993. Clitic placement in Old Romance and European Portuguese. *Papers from the Twenty-Ninth Regional Meeting of the Chicago Linguistics Society*. Chicago: UP.

----- 1996. Clitic Placement in European Portuguese and the position of subjects. In Aaron L. Halpern & Arnold M. Zwicky (eds.), *Approaching Second: Second Position Clitics and Related Phenomena*, 1-40. Stanford: CSLI Publications.

Barnes, Michael. 2008. *A New Introduction to Old Norse*. Exeter: Short Run Press Limited.

Barrie, Michael. 2000. The Syntax of Clitic Placement in European Portuguese. *Revue québécoise de linguistique* 28:91-109.

Batllori, M., M.-Ll. Hernanz, C. Picallo & F. Roca. 2005. *Grammaticalization and Parametric Variation*. Oxford: UP.

Batllori, M., Carlos Sánchez & Avellina Suñer. 1995. The Incidence of Interpolation on the Word Order of Romance Languages. *Catalan Working Papers in Linguistics* 4:185-209.

Belletti, Adriana. 1999. Italian/Romance clitics: Structure and derivation. In Riemsdijk, Hen van (ed.), *Clitics in the Languages of Europe*. New York: Mouton de Gruyter.

----- 2001. Inversion as focalization. In Aafke C.J. Hulk & Jean-Yves Pollock (eds.), *Subject Inversion in Romance and the Theory of Universal Grammar*, 60-90. New York: Oxford UP.

Benincà, Paola. 2014. “Subject” clitics and particles in Provençal. *Probus* 26:183-215.

Benito Moreno, Carlota de. 2015. Pero se escondíamos como las ratas: syncretism in the reflexive paradigm in Spanish and Catalan. *Isogloss* 1:95-127.

Bleam, Tonia. 1999. Leísta Spanish and the Syntax of Clitic Doubling. PhD dissertation, University of Delaware.

Bloom, Lois. 1970. *Language Development*. Cambridge: MIT Press.

----- 1973. *One Word at a Time*. The Hague: Mouton.

Borer, Hagit. 1984. *Parametric Syntax*. Dordrecht: Foris.

Bortolussi, Bernard. 2017. Topicalizations, left dislocations and the left periphery. *Catalan Journal of Linguistics* 16:101-123.

Bortolussi, Bernard & L. Sznajder. 2014. Topicalization versus Left-Dislocation in Biblical Latin. *Journal of Latin Linguistics* 13:163-195.

Bošković, Željko & Deiko Takahashi. 1998. Scrambling and Last Resort. *Linguistic Inquiry* 29:347-366.

Brown, Charles Barrett. 1931. The Disappearance of the Indefinite Hombre from Spanish. *Language* 7:265-277.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The Evolution of Grammar*. Chicago: UP.

Cabredo Hofherr, Patricia. 2004. Impersonal Pronouns in French, German, and Somali. Talk at Syntax of the World's Languages 1, Leipzig.

Campos, Héctor. 1986. Inflection Elements in Romance. PhD dissertation. University of California, Los Angeles.

Cardinaletti, Anna & Michal Starke. 1999. The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns. In Henk van Riemsdijk (ed.), *Clitics in the Languages of Europe*. Berlin: Mouton de Gruyter.

Cennamo, Michela. 1993. *The Reanalysis of Reflexives: A Diachronic Perspective*. Naples: Liguori Editore.

----- 1999. Late Latin Pleonastic Reflexives and the Unaccusative Hypothesis. *Transactions of the Philological Society* 97:103-150.

Cennamo, Michela, Thórhallur Eyþórsson & Jóhanna Barðal. 2015. Semantic and (morpho)syntactic constraints on anticausativization: Evidence from Latin and Old Norse-Icelandic. *Linguistics* 53:677-729.

Chenery, W.H. 1905. *Object pronouns in dependent clauses: A study in Old Spanish word order*. Publications of the Modern Language Association of America.

Chomsky, Noam. 1986. *Barriers*. Cambridge: MIT Press.

----- 1991. Some notes on economy of derivation and representation. In R. Freidin (ed.), *Principles and Parameters in Generative Grammar*, 417-54. Cambridge, Mass.: MIT Press.

----- 1995. *The Minimalist Program*. Cambridge: MIT Press.

----- 2000. Minimalist Inquires: The Framework. In R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89-155. Cambridge, Mass.: MIT Press.

----- 2001. Derivation by Phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1-52. Cambridge, Mass.: MIT Press.

Cinque, G. 1977. The Movement Nature of Left-Dislocation. *Linguistic Inquiry* 8:397-411.

Company Company, C. 2003. Transitivity and grammaticalization of object: The struggle of direct and indirect object in Spanish. In G. Fiorentino (ed.), *Romance Objects: Transitivity in Romance Languages*, 217-260. New York: Mouton de Gruyter.

Company Company, Concepción & Julia Pozas Loyo. 2009. Los indefinidos compuestos y los pronombres genérico-impersonales *omne* y *uno*. In C. Company Company (ed.), *Sintaxis histórica de la lengua española. Segunda parte: La frase nominal*, 1073-1222. México, D.F. : Universidad Nacional Autónoma de México, [Institututo de Investigaciones Filológicas].

Corbett, Greville. 2006. *Agreement*. Cambridge: UP.

Creissels, Denis. To appear. Impersonal pronouns and coreference: the case of French on. In Manninen, S., K. Hietam, E. Keiser & V. Vihman (eds.), *Passives and Impersonals in European Languages*.

Cuervo, María Cristina. 2003. Datives at Large. PhD Dissertation, MIT.

----- 2014. Alternating unaccusatives and distribution of roots. *Lingua* 141:48-70.

Culbertson, Jennifer. 2010. Convergent Evidence for Categorial Change in French: From Subject Clitic to Agreement Marker. *Language* 86:85-132.

Cyrino, Sonia & Gabriela Matos. 2016. Null Objects and VP Ellipsis in European and Brazilian Portuguese. In Wetzels, W. Leo, Sérgio Menuzzi & João Costa (eds.), *The Handbook of Portuguese Linguistics*. Wiley- Blackwell

D'Alessandro, Roberta. 2007. *Impersonal Si Constructions*. Berlin: Mouton de Gruyter.

----- 2014. Death and contact-induced rebirth of impersonal pronouns: A case study. *Probus* 26:249-274.

Danckaert, Lieven. 2017. *The Development of Latin Clause Structure*. Oxford: UP.

Demonte, V. 1995. Dative alternation in Spanish. *Probus* 7:5-30.

Diesing, Molly. 1992. Bare Plurals and the Derivation of Logical Representations. *Linguistic Inquiry* 23:353-380.

----- 1996. Semantic Variables and Object Shift. In Höskuldur Thráinsson, Samuel Epstein & Steve Peter (eds.), *Studies in Comparative Germanic Syntax*, 66-84. Dordrecht: Kluwer.

Diessel, Holger. 2004. *The Acquisition of Complex Sentences*. Cambridge: UP.

Devine, A.M. & Laurence D. Stephens. 2006. *Latin Word Order: Structure Meaning and Information*. Oxford:UP.

Dindelegan, Gabriela Pană. *The Syntax of Old Romanian*. Oxford: UP.

Dobrovie-Sorin, Carmen. 1990. Clitic doubling, wh-movement, and quantification in Romanian. *Linguistic Inquiry* 21:351-398.

----- 1998. Impersonal Se Constructions in Romance and the Passivization of Unergatives. *Linguistic Inquiry* 29:399-437.

Donaldson, Bryan. 2016. Preverbal subjects, information structure, and object clitic position in Old Occitan. *Journal of Linguistics* 52:37-69.

Donohue, Mark. 2003. Morphological Templates, Headedness, and Applicatives in Barupu. *Oceanic Linguistics* 42:111-143.

Duarte, Inês & Maria Cristina Figueiredo Silva. 2016. The Null Subject Parameter and the Structure of the Sentence in European and Brazilian Portuguese. In W. Leo Wetzels, João Costa & Sergio Menuzzi (eds.), *The Handbook of Portuguese Linguistics*, 234-253. Hoboken, NJ: John Wiley & Sons.

Eberenz, R. 2000. *El español en el otoño de la edad media*. Madrid: Gredos.

Egerland, Verner. 2003. Impersonal pronouns in Scandinavian and Romance. *Working Papers in Scandinavian Syntax* 71: 75-102.

Estigarribia, Bruno. 2017. The Semantics of Spanish Clitic Left Dislocations with Epithets. *Probus* 29: 287-321.

Faarland, Jan Terje. 2004. *The Syntax of Old Norse*. Oxford: UP.

Fábregas, Antonio & Sergio Scalise. 2012. *Morphology: From Data to Theories*. Edinburgh:UP.

Faltz, Leonard. 1977. Reflexivization: A Study in Universal Syntax. PhD dissertation, University of California-Berkeley.

----- 2008. A Reflexive Cycle? Paper presented at Linguistic Cycles Workshop, Tempe, Arizona.

Fischer, Susann. 2003. Rethinking the Tobler-Mussafia Law: Data from Old Catalan.
Diachronica 20:259-288.

Fischer, Susann, Mario Navarro & Jorge Vega Vilanova. 2016. The Clitic Doubling Cycle: A Diachronic Reconstruction. Talk given at 18th Diachronic Generative Syntax (DiGS), June 29-July 1.

Flores-Ferrán, Nydia. 2009. Are you referring to me? The variable use of UNO and YO in oral discourse. *Journal of Pragmatics* 41: 1810-1824.

Folli, Raffaella & Heidi Harley. 2007. Causation, obligation and argument structure: On the nature of little v. *Linguistic Inquiry* 38:197-238.

Fontana, J. M. 1993. *Phrase Structure and the Syntax of Clitics in the History of Spanish*. Diss. University of Pennsylvania.

Foulet, Lucien. 1936. L'estension de la forme oblique du pronom personnel en ancien français.
Romania 62:27-91.

Franco, Jon Andoni. 1993. *Object Agreement in Spanish*. Phd Dissertation: University of Southern California.

Franzén, Torsten. 1939. *Étude sur la syntaxe de pronoms personnels sujets en ancien français*. Uppsala: Almqvist & Wiksell.

Frascarelli, Mara. 2007. Subjects, Topics, and the Interpretation of Referential Pro: An Interface Approach to the Linking of (Null) Pronouns. *Natural Language and Linguistic Theory* 25:691-734.

Frascarelli, Mara & R. Hinterhölzl. 2007. Types of topics in German and Italian. In S. Winkler & K. Schwabe (eds.), *On Information Structure, Meaning, and Form*, 87-116. Amsterdam: John Benjamins.

Fuß, Eric. 2005. *The Rise of Agreement*. Amsterdam: John Benjamins.

Gabriel, Christoph & Esther Rinke. 2010. Information Packaging and the Rise of Clitic Doubling in the History of Spanish. In Gisella Ferraresi & Rosemarie Lühr (eds.), *Diachronic Studies on Information Structure: Language Acquisition and Change (Language, Context, and Cognition 10)*, 55-77. Berlin: Mouton de Gruyter.

Gallego, Ángel. 2013. Object Shift in Romance. *Natural Language and Linguistic Theory* 31:409-451.

Gelderen, Elly van. 2004a. *Grammaticalization as Economy*. Amsterdam: John Benjamins.

----- 2004b. Economy, Innovation, and Prescriptivism: From Spec to Head and Head to Head. *Journal of Comparative Germanic Linguistics* 7:59-98.

----- 2007. The Definiteness Cycle in Germanic. *Journal of Germanic Linguistics* 19:275-305.

----- 2008. Where Did Late Merge Go? Grammaticalization as Feature Economy. *Studia Linguistica* 62:287-300.

----- 2011. *The Linguistic Cycle: Language Change and the Language Faculty*. Oxford: UP.

Geniušienė, Emma. 1987. *The Typology of Reflexives*. Berlin: Mouton de Gruyter.

Gerdts, D.B & M.Q. Hinkson. 2004. The Grammaticalization of Halkomelem FACE into a Dative Applicative Suffix. *International Journal of American Linguistics* 70:227-250.

Giacalone Ramat, Anna & Andrea Sansò. 2007. The indefinite usage of UOMO ('man') in early Italo-Romance: Grammaticalization and Areality. *Archivio Glottologico Italiano* 92:65-111. Dowloaded from Academia.edu.

https://www.academia.edu/1033376/The_indefinite_usage_of_uomo_man_in_early_Italo-Romance_Grammaticalization_and_areality

Givón, Taly. 1976. Topic, Pronoun, and Grammatical Agreement. In Charles Li (ed.), *Subject and Topic*, 151-188. New York: Academic Press.

Gupton, Timothy & Sarah Lowman. 2013. An F Projection in Cibeno Dominican Spanish. In Jennifer Cabrelli Amaro, Gillian Lord, Ana de Prada Pérez & Jessi Elana Aaron (eds.), *Selected Proceedings of the 16th Hispanic Linguistics Symposium*, 338-348. Somerville, MA: Cascadilla Proceedings Project.

Haider, Hubert. 1984. The Case of German. In Jindrich Toman (ed.), *Studies in German Grammar*, 65-101. Dordrecht: Foris.

----- 2006. Mittelfeld Phenomena (Scrambling in Germanic). In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell Companion to Syntax, Volume I*, 204-274. Oxford: Blackwell.

Haider, Hubert & Inger Rosengren. 1998. Scrambling. *Sprache und Pragmatik* 49:1-104.

Halle, Morris & James Harris. 2005. Unexpected Plural Inflections in Spanish: Reduplication and Metathesis. *Linguistic Inquiry* 36:195-222.^[1]

Halpern, Aaron. 1995. *On the Placement and Morphology of Clitics*. Stanford: CSLI Publications.

Halpern, Aaron & J.M. Fontana. 1994. X⁰ Clitics and Xmax Clitics. *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL XII)*. Stanford: CSLI Publications.

Harizanov, Boris. 2014. Clitic doubling at the syntax-morphophonology interface: A-movement and morphological merger in Bulgarian. *Natural Language and Linguistic Theory* 32:1033-1088.

Harris, Martin. 1978. *The Evolution of French Syntax: A Comparative Approach*. Working Papers in Functional Grammar 1. University of Amsterdam.

Haspelmath, Martin. 1990. The Grammaticization of Passive Morphology. *Studies in Language* 14:25-72.

----- 1997. *Indefinite Pronouns*. Oxford: Clarendon Press.

Hastings, Robert. 1994. L'espressione del soggetto indefinite in un dialetto EA. *L'Italia dialettale* LVII (new series XXXIV).

Heim, Irene. 1982. The Semantics of Definite and Indefinite Noun Phrases. PhD dissertation, University of Massachusetts-Amherst.

Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: UP.

Heine, Bernd, Tom Güldemann, Christa Kilian-Hatz, Donald. A Lessau, Heinz Roberg. Mathias Schladt, and Thomas Stolz. 1993. *Conceptual Shift: a Lexicon of Grammaticalization Processes in African Languages*. University of Cologne: Institut für Afrikanistik. (Afrikanistische Arbeitspapiere 34/35)

Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: UP.

Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske Verlag.

Herman, József. 1954. Recherche sur l'ordre des mots dans les plus anciens textes français en prose. *Acta Linguistica Academiae Scientiarum Hungaricae* 4: 69-94.

Hill, Jane. 1987. Spanish as a Pronominal Argument Language: The Spanish Interlanguage of Mexicano Speakers. *Coyote Papers* 6:68-90.

Hock, Hans Henrich. 1991. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.

Holmberg, Anders. 1986. Word Order and Syntactic Features in Scandinavian Languages. PhD dissertation, University of Stockholm.

----- 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36:533-564.

----- 2010. The Null Generic Subject Pronoun in Finnish: A Case of Incorporation. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric Variation: Null Subjects in Minimalist Theory*, 200-230. Cambridge: UP.

Holmberg, Anders & Christer Platzack. 1995. *The Role of Inflection in Scandinavian Syntax*. Oxford: UP.

Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three Partial Null Subject Languages: Brazilian Portuguese, Finnish, and Marathi. *Studia Linguistica* 63:59-97.

Hoop, Helen de. 1992. Case Configuration and Noun Phrase Interpretation. PhD dissertation, Groningen University.

Hoop, Helen de & Wim Kosmeijer. 1995. Case and Scrambling: D-Structure vs. S-Structure. In Hubert Haider, Susan Olsen & Sten Vikner (eds.), *Studies in Comparative Germanic Syntax*, 139-159. Dordrecht: Kluwer.

Hopper, Paul & Elizabeth Traugott. 2003. *Grammaticalization*. Cambridge: UP.

Jaeggli, Osvaldo. 1982. *Topics in Romance Syntax*. Dordrecht: Foris.

----- 1986. Three Issues in the Theory of Clitics: Case, Doubled NPs, and Extraction. In Hagit Borer (ed.), *The Syntax of Pronominal Clitics*, 15-42. New York: Academic Press.

Jespersen, Otto. 1917. *Negation in English and Other Languages*. Copenhagen: A.F. Høst.

Kaiser, Georg. 1992. *Die klitischen Personalpronomina im Französischen und Portugiesischen: Eine synchronische und diachronische Analyse*. Vervuert Verlag: Frankfurt am Main.

Kalulli, Dalina & Liliane Tasmowski. 2008. *Ctic Doubling in the Balkan Languages*. Amsterdam: John Benjamins.

Kandybowicz, Jason. 2007. On Fusion and Multiple Copy Spell-Out. In Norbert Corver & Jairo Nunes (eds.), *The Copy Theory of Movement on the PF Side*, 119-150. Amsterdam: John Benjamins.

Kärde, S. 1943. Quelque Manières D'Exprimer L'Idée d'un Sujet Indéterminé ou Général en Espagnol. Thesis. Uppsala.

Kayne, Richard S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, Mass.: MIT Press.

----- 1989. Facets of Romance Past Participle Agreement. In P. Benincà (ed.), *Dialect Variation and the Theory of Grammar*, 85-104. Dordrecht: Foris.

----- 1991. Romance Clitics, Verb Movement, and PRO. *Linguistic Inquiry* 22:647-686.

----- 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.

Kemenade, Ans van. 2000. Jespersen's Cycle Revisited. In Susan Pintzuk (ed.), *Diachronic Syntax*, 51-74. Oxford: UP.

Kempchinsky, Paula. 2004. Romance SE as an Aspectual Element. In J. Auger et al (eds.), *Contemporary Approaches to Romance Linguistics*, 239-256. Amsterdam: John Benjamins.

----- 2006. Teasing apart the middle. In Itziar Laka & Beatriz Fernández (eds.) *Andolin gogoan/Homenaje a Andolin Eguzkitza*. University of the Basque Country Press, 532-547.

Keniston, Hayward. 1937. *The Syntax of Castilian Prose*. Chicago: UP.

King, J.E. & C. Cookson. 1888. *The Principles of Sound and Inflection as Illustrated in the Greek and Latin Languages*. Oxford: Clarendon Press.

Kinn, Kari. 2016. Referential vs. Non-referential Null Subjects in Middle Norwegian. *Nordic Journal of Linguistics* 39:277-310.

Kiss, Katalin É. 1994. Sentence Structure and Word Order. In Ferenc Kiefer & Katalin É. Kiss (eds.), *The Structure of Hungarian Syntax*, 1-90. New York: Academic Press.

Klausenburger, Jurgen. 2000. *Grammaticalization: Studies in Latin and Romance Morphosyntax*. Amsterdam: John Benjamins.

Koontz-Garboden, Andrew. 2009. Anticausitivazation. *Natural Language & Linguistic Theory* 27:77-138.

Kramer, Ruth. 2014. Clitic doubling or object agreement: the view from Amharic. *Natural Language and Linguistic Theory* 32:593-634.

Kroch, Anthony. 1989. Reflexes of Grammar in Patterns of Language Change. *Journal of Language Variation and Change* 1:199-244. Kuczaj, S. 1976. -Ing, -s, and -ed: A Study of the Acquisition of Verb Inflections. Dissertation, University of Minnesota.

Lambrecht, Knud. 1981. *Topic, Antitopic, and Verg Agreement in Non-Standard French*. Amsterdam: John Benjamins.

Ledgeway, Adam. 2012. *From Latin to Romance: Morphosyntactic Typology and Change*. Oxford:UP.

Legate, J.A. 2014. *Voice and v: Lessons from Acehnese*. Cambridge: MIT Press.

Lehmann, Christian. 1985. Grammaticalization: synchronic variation and diachronic change. *Lingua e Stile* 20:303–318.

----- 1995. *Thoughts on Grammaticalization*. Munich: Lincom.

Lightfoot, David. 1979. *Principles of Diachronic Syntax*. Cambridge: UP.

Lindsay, Wallace Martin. 1894. *The Latin Language: An Historical Account of Latin Sounds, Stems, and Flexions*. Oxford: Clarendon Press.

Longobardi, Giuseppe. 1994. Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form. *Linguistic Inquiry* 25:609-665.

Luraghi, Silvia. 1997. Omission of the Direct Object in Classical Latin. *Indogermanische Forschungen* 102:239-257.

----- 2004. Null Objects in Latin and Greek and the Relevance of Linguistic Typology for Language Reconstruction. In Karlene Jones-Bley, Martin E. Huld, Angela Della Volpe & Miriam Robbins Dexter (eds.), *Proceedings of the Fifteenth Annual UCLA Indo-European Conference, Journal of Indo-European Monograph Series, No. 49*, 234-256. Washington D.C.: Institute for the Study of Man.

MacDonald, Jonathan E. 2017. An Implicit Projected Argument in Spanish Impersonal- and Passive-Se Constructions. *Syntax*. doi:10.1111/synt.12146

MacDonald, Jonathan E. & Susana Huidobro. The Lack of Spanish Non-Argumental Clitic Doubling. In Claudia Borgonovo, Manuel Español-Echevarría & Philippe Prévost (eds.), *Selected Proceedings of the 12th Hispanic Linguistics Symposium*, 50-62. Somerville, MA: Cascadilla Proceedings Project.

MacDonald, Jonathan E. & Matthew Maddox. 2018. Passive Se in Romanian and

Spanish: A Subject Cycle. *Journal of Linguistics* 54:389-427.

Machado Ramos, Ricardo & Jânia Martins Ramos. 2016. Clitic Doubling and Pure Agreement Person Features. *Revista de Estudos da Linguagem* 24:378-416.

MacKenzie, Ian & Wim van der Wurff. 2012. Relic Syntax in Middle English and Medieval Spanish: Parameter Interaction in Language Change. *Language* 88:846-876.

Maddox 2017. Spanish and French HOMO-derived Impersonal Pronouns: Stalled Grammaticalization. 23rd meeting of the International Conference on Historical Linguistics (ICHL), University of Texas at San Antonio.

Maddox, Matthew L. 2018. Licensing Conditions on Null Generic Subjects in Spanish. In Lori Repetti & Francisco Ordóñez (eds.), *Romance Languages and Linguistic Theory*, 185-199. Amsterdam: John Benjamins.

Maddox, Matthew L. & Jonathan E. MacDonald. In contract. Reflexive Constructions in German, Spanish, and French as a Product of Cyclic Interaction. In Sam Wolfe & Christine Meklenborg Salvesen (eds.), *Continuity and Variation in Germanic and Romance*. Oxford: UP.

Marantz, Alec. 1993. A Late Note on Late Insertion. In Y.S. Kim et al (eds.), *Explorations in Generative Grammar: A Festschrift for Dong-Whee Yang*, 396-413. Seoul: Hankuk Publishing.

----- 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of your Own Lexicon. In A. Dimitriadis et al. (eds.), *University of Pennsylvania Working Papers in Linguistics*

Martins, Ana Maria. 2002. Tipologia e Mudança Linguísticas: Os Pronomes Pessoais do Português e do Espanhol. *Santa Barbara Portuguese Studies* 6:340-386.

----- 2003. Deficient pronouns and linguistic change in Portuguese and Spanish. In Josep Quer, Jan Schrotten, Mauro Scorretti, Petra Sleeman, Els Verheugd (eds.), *Romance Language & Linguistic Theory 2001*, 213-230. Amsterdam: John Benjamins.

Masullo, Pascual José. 2003. Clitic-less definite object drop in River Plate Spanish. Paper presented at 33rd LSRL, Bloomington, IN.

Mathieu, E. 2009. On the Germanic Properties of Old French. In Paola Crisma & Giuseppe Longobardi (eds.), *Historical Syntax and Linguistic Theory*. Oxford: UP.

Matushansky, Ora. 2006. Head Movement in Linguistic Theory. *Linguistic Inquiry* 37:69-109.

Mayer, Elisabeth. 2003. Clitic Doubling in Limeño. Ms. Australian National University.

McGinnis, Martha. 2004. Lethal Ambiguity. *Linguistic Inquiry* 35:47-95.

Mendikoetxea, Amaya. 1999. Construcciones inacusativas y pasivas. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, vol. 2, ch. 25. Madrid: Espasa-Calpe.

----- 2008. Clitic Impersonal Constructions in Romance: Syntactic Features and Semantic Interpretation. *Transactions of the Philological Society* 106: 290-336.

Menschling, Guido. 2012. Parameters in Old Romance Word Order: A Comparative Minimalist Analysis. In Charlotte Galves, Sonio Cyrino, Ruth Lopes, Filomena Sandalo & Juanito Avelar (eds.), *Parameter Theory and Linguistic Change*, 21-42. Oxford: UP.

Meyer-Lübke, Wilhelm. 1900. *Grammaire des langues romanes*. Parise: A. Franck.

Mithun, Marianne. 1991. The Development of Bound Pronominal Paradigms. In Winfred Lehmann & Helen-Jo Jakusz Hewitt (eds.), *Language Typology*, 1988, 85-104. Amsterdam: John Benjamins.

----- 2003. Pronouns and Agreement: The Information Status of Pronominal Affixes. *Transactions of the Philological Society* 101:235-278.

Monge, Félix. 1954. *Las frases pronominales de sentido impersonal en español*. In Archivo de Filología Aragonesa VII: 7-102.

Müller, Gereon. 1995. *A-Bar Syntax: A Study in Movement Types*. Berlin: Mouton de Gruyter.

----- 2005. Pro-drop and Impoverishment. In P. Brandt & E. Fuss (eds.), *Form, Structure and Grammar: A Festschrift Presented to Günther Grewendorf on the Occasion of His 60th Birthday*, 93-115. Tübingen: Narr.

Müller, Gereon & Wolfgran Sternefeld. 1994. Scrambling as A'-Movement. In Norbert Corver & Henk C. van Riemsdijk (eds.), *Studies on Scrambling: Movement and Non-Movement Approaches to Free Word-Order Phenomena*, 331-385. Berlin: Mouton de Gruyter.

Muller, Henri F. 1924. The Passive Voice in Vulgar Latin. *Romanic Review* 15:68-93.

Mussafia, A. 1886. Una particolarità sintattica della lingua italiana dei primi secoli. In G.I. Ascoli et al (eds.), *Mescellanea di filologia e linguistica in memoria de Napoleone Caix e Ugo Angelo Canello*. Florence: Le Monni.

Muysken, P. 1982. Parametrizing the notion "head." *Journal of Linguistic Research* 2:57-75.

Navarro, Mario, Susann Fischer & Jorge Vega Vilanova. 2017. Reconstruyendo un Ciclo: Doblado de Clíticos y Gramaticalización en las Lenguas Romanes. In Silvia Gumiell-Molina, Manuel Leonett & Isabel Pérez-Jiménez (eds.), *Investigaciones actuales en Lingüística, Vol. III: Sintaxis*, 111-126. Publicaciones de la Universidad de Alcalá.

Nicolae, Alexandru. 2019. *Word Order and Parameter Change in Romanian: A Comparative Romance Perspective*. Oxford: UP.

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge: MIT Press.

Nyrop, Kr. 1899. *Grammaire Historique de la Langue Française*. Copenhagen: E. Bojesen.

Odenkirchen, Carl J. 1949. A Preliminary *Chrestomathy of Old Catalan*. Chapel Hill, N.C.

Olarrea, Antxon. 2012. Word Order and Information Structure. In José Ignacio Hualde, Antxon Olarrea & Erin O'Rourke (eds.), *The Handbook of Hispanic Linguistics*, 603-628. Oxford: Blackwell.

Ordóñez, Francisco. 1997. Word Order and Clausal Structure of Spanish and Other Romance Languages. PhD dissertation, City University of New York.

----- 1998. Post-verbal Asymmetries in Spanish. *Natural Language and Linguistic Theory* 16:313-346.

----- 2000. *The Clausal Structure of Spanish: A Comparative Study*. New York: Garland.

----- 2012. Clitics in Spanish. In José Ignacio Hualde, Antxon Olarrea & Erin O'Rourke (eds.), *The Handbook of Hispanic Linguistics*, 423-451. Chichester, West Sussex: Wiley-Blackwell.

Ordóñez, Francisco & Esthela Treviño. 1999. Left Dislocated Subjects and the Pro-Drop Parameter: A Case Study of Spanish. *Lingua* 107:39-68.

Ormazabal, Javier & Juan Romero. 2013. Object Clitics, Agreement, and Dialectal Variation. *Probus* 25: 301-344.

Otero, Carlos P. 1986. Arbitrary Subjects in Finite Clauses. In I. Bordelois et al (eds.), *Generative Studies in Spanish Syntax*. Dordrecht:Foris.

Ottosson, Kjartan. 2004. *The Old Nordic Middle Voice*. Paper presented at Center for Advanced Studies, Oslo, Norway.

----- 2008. The Old Norse Middle Voice in the pre-literary period: Questions of grammaticalisation and cliticization. In Josephson, Folke & Ingmar Söhrman (eds.), *Diachronic and typological perspectives on verbs*, 185-219. Amsterdam: John Benjamins.

Paden, William D. 1998. *An Introduction to Old Occitan*. Modern Language Association of America: New York.

Parodi, Claudia. 1995. Participle Agreement and Object Shift in Old Spanish. In Héctor Campos & Paula Kempchinsky (eds.), *Evolution and Revolution in Linguistic Theory*, 276-301. Washington, D.C.: Georgetown UP.

Penny, Ralph. 2002. *A History of the Spanish Language*. Cambridge UP: Massachusetts.

Perini, Mário A. 2002. Modern Portuguese: A Reference Grammar. Yale UP: New Haven.

Pescarini, Diego. In press. Clitic Clusters. In Martin Everaert & Henk van Riemsdijk (eds.), *The Companion to Syntax, Second Edition*. Oxford: Blackwell.

Peterson, David A. 2007. *Applicative Constructions*. Oxford: UP.

Pintzuk, Susan. 1991. Phrase Structures in Competition: Variation and Change in Old English Word Order. PhD dissertation, University of Pennsylvania.

Platzack, Christer. 1987. The Scandinavian Languages and the Null Subject Parameter. *Natural Language & Linguistic Theory* 5:377-401.

Poletto, Cecilia. 2000. *The Higher Functional Field: Evidence from Northern Italian Dialects*. Oxford: UP.

----- 2014. *Word Order in Old Italian*. Oxford: UP.

Poletto, Cecilia & Jean Yves Pollock. 2004. On wh-clitics, wh-doubling and apparent wh in situ in French and some north eastern Italian dialects. *Probus* 16:241-273.

Pollock, Jean Yves. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20:365-424.

Poole, G. 2007. Interpolation and the Left Periphery in Old Spanish. In M. Hussein, M. Kolokante & C. Wright (eds.), *Newcastle Working Papers in Linguistics* 13:188-216.

Pozas Loyo, Julia. 2016. Rise and Fall Impersonal Pronouns in Spanish. Online: <http://archive.sfl.cnrs.fr/sites/sfl/IMG/pdf/impersMay2010PozasLoyoOmneUno.pdf>
Downloaded: Nov. 14, 2016.

----- 2009. “Los indefinidos compuestos y los pronombres genérico-impersonales *omne* y *uno*” en *Sintaxis Histórica de la Lengua Española. Segunda parte: La frase nominal*, Concepción Company (dir.), México: FCE, UNAM, pp. 1072-1219. En coautoría con Concepción Company.

Ramsden, H. 1963. *Weak Pronoun Position in the Early Romance Languages*. Manchester: UP.

Raposo, E. 1986. On the Null Object in European Portuguese. In O. Jaeggli & C. Silva Corvalán (eds.), *Studies in Romance Linguistics*, vol. 24, 373-390. Dordrecht: Foris.

Raposo, Eduardo & Juan Uriagereka. 1996. Indefinite Se. *Natural Language and Linguistic Theory* 14:749-810.

Revol, Thierry. 2005. *Introduction à l'ancien français*. Paris: Colin.

Riho, Timo. 1988. *La redundancia pronominal en el iberromance medieval*. Tübingen: Niemeyer.

Rivas, Alberto M. 1977. A Theory of Clitics. PhD dissertation, MIT.

Rivero, M. 1986. Parameters in the Typology of Clitics in Romance and Old Spanish. *Language* 62:774-807.

----- 1991. Clitic and NP Climbing in Old Spanish. In H. Campos & F. Martínez-Gil (eds.), *Current Studies in Spanish Linguistics*, 241-282. Georgetown: UP.

----- 1997. On two Locations for Complement Clitic Pronouns: Serbo-Croatian, Bulgarian and Old Spanish. In A. van Kemenade & N. Vincent (eds.), *Parameters of Morphosyntactic Change*, 170-206. Cambridge: UP.

----- 2000. On Impersonal Reflexives in Romance and Slavic and Semantic Variation. In Camps et al (eds.), *Romance Syntax, Semantics, and L2 Acquisition*. Amsterdam: John Benjamins.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

----- 1997. The Fine Structure of the Left Periphery. In L. Haegeman (ed.) *Elements of Grammar*, Kluwer, Dordrecht, 281-337.

----- 2006. On the Form of Chains: Criterial Positions and ECP Effects. In L. Cheng & N. Corver (eds.), *Wh-Movement: Moving On* (pp. 97-133). Cambridge: MIT Press.

Roberts, Ian. 2004. *Some Consequences of a Deletion Analysis of Null Subjects*; paper presented at the LAGB Annual Meeting, University of Surrey, Roehampton.

----- 2007. *Diachronic Syntax*. Oxford: UP.

----- 2010a. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. Cambridge: MIT Press.

----- 2010b. A deletion analysis of null subjects. In Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric Variation: Null Subjects in Minimalist Theory*, 58-87. Cambridge: UP.

----- 2010c. Varieties of French and the Null Subject Parameter. In Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric Variation: Null Subjects in Minimalist Theory*, 303-327. Cambridge: UP.

Roberts, I. & A. Holmberg. 2010. Introduction: Parameters in Minimalist Theory. In Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric Variation: Null Subjects in Minimalist Theory*, 1-57. Cambridge: UP.

Roberts, Ian & Anna Roussou. 1999. A Formal Approach to "Grammaticalization." *Linguistics* 37:1011-1041.

----- 2003. *Syntactic Change*. Cambridge: UP.

Salvesen, Christine Meklenborg & Hans Petter Helland. 2013. *Challenging Clitics*. Amsterdam: John Benjamins.

Samek-Lodovici, V. 1996. *Constraints on Subjects: An Optimality Theoretic Analysis*. Phd dissertation, Rutgers University.

Sánchez Lancis, Carlos. 2001. The Evolutions of the Old Spanish Adverbs *ende* and *ý*: A Case of Grammaticalization. *Catalan Working Papers in Linguistics* 9: 101-118.

Sánchez, Liliana. 1999. Null Objects and D-features in Contact Spanish. In J.-Marc Authier, Barbara E. Bullock & Lisa A. Reed (eds.), *Formal Perspectives on Romance Linguistics*, 227-242. Amsterdam: John Benjamins.

Sansò, Andrea. 2006. 'Agent defocusing' revisited: Passive and impersonal constructions in some European languages. In W. Abraham, L. Leisioe (eds.), *Passivization and Typology: Form and Function*. Amsterdam: John Benjamins.

----- 2011. Grammaticalization paths and prototype effects: A history of the agentive reflexive passive in Italian. *Folia Linguistica Historica* 32: 1-34.

----- To appear. Where Do Antipassives Come From? A Study in Diachronic Typology. *Diachronica* 34.

Santorini, Beatrice. 1989. The Generalization of the Verb-Second Constraint in the History of Yiddish. PhD dissertation, MIT.

Sato, Yosuke & Nobu Goto. 2014. Scrambling. In Andrew Carnie, Dan Siddiqi & Yosuke Sato (eds.), *The Routledge Handbook of Syntax*, 264-282. New York: Routledge.

Schäfer, Florian. 2008. *The Syntax of (Anti-)causatives*. Amsterdam: John Benjamins.

Schwenter, Scott A. Null Objects Across South America. In Timothy L. Face & Carol A. Klee (eds.), *Selected Proceedings of the 8th Hispanic Linguistics Symposium*, 23-36. Somerville, MA: Cascadilla Proceedings Project.

Sibille, Jean. 2012. Les clitiques sujets dans le parler occitan de Chiomonte et des Ramats (Italie). *Revue de linguistique romane*, Société de linguistique romane, pp. 401-435. <hal-00952429>

Siewierska, Anna. 2008. Verb Person Marking. In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Linguistic Structures Online*, chapter 102. Munich: Max Planck Digital Library. Available online at <https://wals.info>

Sigurðsson, Halldór Ármann. 1990. V1 declaratives and verb raising in Icelandic. In J. Maling & A. Zaenen (eds.), *Syntax and Semantics 24: The Syntax of Modern Icelandic*. Orlando: Academic Press, 41-69.

Sigurðsson, Halldór Ármann & Verner Egerland. 2009. Impersonal null-subjects in Icelandic and Elsewhere. *Studia Linguistica* 63: 158-185.

Simpson, A. & Zoe Wu. 2001. IP-Raising, Tone Sandhi, and the Creation of Sentence-Final Particles. *Journal of East Asian Linguistics* 11:67-99.

Sitaridou, Ioanna. 2011. Word Order and Information Structure in Old Spanish. *Catalan Journal of Linguistics* 10:159-184.

----- 2012. A Comparative Study of Word Order in Old Romance. *Folia Linguistica* 46:553-604.

----- 2016. Word Order in Old Spanish: V2 or non-V2? Presentation given at the 18th Diachronic Generative Syntax (DiGS) conference, Ghent University, June 29.

Spencer, Andrew & Ana R. Luís. 2012. *Clitics: An Introduction*. Cambridge: UP.

Sportiche, Dominique. 1996. Clitic constructions. In J. Rooryck & L. Zaring (eds.), *Phrase Structure and the Lexicon*, 213-276. Dordrecht: Kluwer.

----- 1999. Subject Clitics in French and Romance Complex Inversion and Clitic Doubling. In K. Johnson & I. Roberts (eds.), *Beyond Principles and Parameters*, 189-221. Dordrecht: Kluwer.

Strozer, Judith Reina. 1976. Clitics in Spanish. PhD dissertation, UCLA.

Suñer, Margarita. 1988. The Role of Agreement in Clitic-Doubled Constructions. *Natural Language and Linguistic Theory* 6:391-434.

----- 2000. Object-shift: Comparing a Romance Language to Germanic. *Probus* 12:261-289.

----- 2006. Left-dislocations with and without epithets. *Probus* 18:127-158.

Tauli, Valter. 1956. The Origin of Affixes. *Finnisch-Ugrische Forschungen* 32:170-225.

Taylor, Ann. 1990. Clitics and Configurationality in Ancient Greek. PhD dissertation, University of Pennsylvania.

Thráinsson, Höskuldur. 2001. Object Shift and Scrambling. In Mark Baltin & Chris Collins (eds.), *The Handbook of Contemporary Syntactic Theory*, 148-202. Oxford: Blackwell.

Tobler, A. 1875. Besprechung von J. Le Coultre, 'De l'ordre des mots dans Chrétien de Troye.' *Vermische Beiträge zur französischen Grammatik* 5:395-414.

Tomić, Olga. 2007. Variation in Clitic-Climbing in South Slavic. *Syntax and Semantics* 36:443-468.

----- 2008. Towards Grammaticalization of Clitic Doubling: Clitic Doubling in Macedonian and Neighbouring Languages. In Dalina Kalulli & Liliane Tasmowski (eds.), *Clitic Doubling in the Balkan Languages*, 65-87. Amsterdam: John Benjamins.

Torrego, E. 1995. From argumental to non-argumental pronouns: Spanish doubled reflexives. *Probus* 7:221-241.

----- 1998. *The Dependencies of Objects*. Cambridge: MIT Press.

Traugott, Elizabeth & Bernd Heine. 1991. *Grammaticalization*. Amsterdam: John Benjamins.

Travis, Lisa. 1984. Parameters and Effects of Word Order Variation. PhD dissertation, Massachuts Institute of Technology.

Tse, Keith Ka-Kei. 2016. Grammaticalization and 'Lateral' Grammaticalization: New Perspectives on Linguistic Interfaces and Functional Categories. MA Dissertation, University of York.

Uriagereka, Juan. 1995. Aspects of the Syntax of Clitic Placement in Western Romance. *Linguistic Inquiry* 26:79-124.

Vikner, Sten. 2006. Object Shift. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell Companion to Syntax, vol. III*, 392-436. Oxford: Blackwell.

Vilanova, Jorge Vega, Mario Navarro & Susann Fischer. 2016. The Clitic Doubling Cycle: A Diachronic Reconstruction. Paper presented at DiGS, Ghent, Belgium.

Vilanova, Jorge Vega, Susann Fischer & Mario Navarro. 2018. The Clitic Doubling Cycle: A

Diachronic Reconstruction. In Gabriela-Pană Dindelegan, Adina Dragomirescu, Irina Nicula & Alexandru Nicolae (eds.), *Comparative and Diachronic Perspectives on Romance Syntax*, 117-134. Newcastle upon Tyne: Cambridge Scholars.

Wackernagel, J. 1892. Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1:333-435.

Wanner, Dieter. 1987. *The Development of Romance Clitic Pronouns: From Latin to Old Romance*. Berlin: Mouton de Gruyter.

----- 1991. The Tobler-Mussafia Law in Old Spanish. In Héctor Campos & Fernando Martínez-Gil (eds.), *Current Studies in Spanish Linguistics*. Washington, D.C.: Georgetown UP.

Welton-Lair, Lisa K. 1999. The evolution of the French indefinite pronoun *on*: A corpus-based study in grammaticalization.

Wheeler, Max, Alan Yates & Nicolau Dols. 1999. *Catalan: A Comprehensive Grammar*. London: Routledge.

Wolfsgruber, Anne. 2016. Why is French Different? Evidence for Diverging Paths in the Grammaticalization of *Se* in Medieval French and Spanish. Paper presented at Workshop on Romance SE/SI. University of Wisconsin-Madison.

----- 2017. Medieval Romance *Se*: Its Grammaticalization and Syntactic Status. PhD dissertation, Salzburg University.

Wu, Zoe. 2004. *Grammaticalization and Language Change in Chinese*. London: Routledge Curzon.

Zagona, Karen. 2002. *The Syntax of Spanish*. Cambridge: UP.

Zaring, Laurie. 1998. Object Shift in Old French. In Armin Schwiegler, Bernard Tranell &

Myriam Uribe-Etxebarria (eds.), *Romance Linguistics: Theoretical Perspectives*, 319-332. Amsterdam: John Benjamins.

Zdrojewski, Pablo & Liliana Sánchez. 2014. Variation in accusative clitic doubling across three Spanish dialects. *Lingua* 151:162-176.

Zimmerman, Michael. 2014. *Expletive and Referential Subject Pronouns in French*. Berlin: De Gruyter.

Zwicky, Arnold M. 1977. *On Clitics*. Bloomington, IN: Indiana University Linguistics Club.

----- 1985. Clitics and Particles. *Language* 61:283-305.

Zwicky, Arnold & Geoffrey Pullum. 1983. Cliticization vs. Inflection: English n't. *Language* 59:502-513.

Databases:

The Packard Humanities Institute. Database. *PHI Latin Texts* (online).
 <<http://latin.packhum.org/index>

Corpus Informatizado do Português Medieval. <http://cipm.fcsh.unl.pt/>

Galves, Charlotte, and Pablo Faria. 2010. *Tycho Brahe Parsed Corpus of Historical Portuguese*. URL: <http://www.tycho.iel.unicamp.br/~tycho/corpus/en/index.html>.

Torruella, Joan, Manuel Pérez Saldanya & Josep Martínez. *Corpus Informatitzat del Català Antic*. <http://www.cica.cat/>

Scrivner, Olga, Michael Paul McGuire, Sandra Kübler, Barbara Vance & E.D. Blodgett. *Old Occitan Digital Collection*. <http://cl.indiana.edu/~mpmcguir/introduction.php>

REAL ACADEMIA ESPAÑOLA: Banco de datos (CORDE) [en línea]. *Corpus diacrónico del español*. <<http://www.rae.es>>