UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE COMUNICAÇÃO E EXPRESSÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

Giuseppe Freitas da Cunha Varaschin

A SIMPLER SYNTAX OF ANAPHORA

Florianópolis
2021

Giuseppe Freitas da Cunha Varaschin

A SIMPLER SYNTAX OF ANAPHORA

Tese submetida ao Programa de Pós-graduação em Linguística da Universidade Federal de Santa Catarina para obtenção do título de doutor em Linguística.

Orientador: Prof. Heronides M. M. Moura, Dr.

Coorientador: Prof. Peter W. Culicover, Ph.D.

Florianópolis
2021

Giuseppe Freitas da Cunha Varaschin

A Simpler Syntax of Anaphora

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Heronides M. M. Moura, Dr.
Universidade Federal de Santa Catarina (UFSC)

Prof. Marcus Vinicius da Silva Lunguinho, Dr.
Universidade de Brasília (UnB)

Prof. Philip Miller, Ph.D.
Université de Paris

Profa. Roberta Pires de Oliveira, Dra.
Universidade Federal de Santa Catarina (UFSC)

Prof. Robert D. Levine, Ph.D.
The Ohio State University (OSU)

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de Doutor em Linguística.

_____
Coordenação do Programa de Pós-graduação

_____
Prof. Heronides M. M. Moura, Dr.
Orientador

_____
Prof. Peter W. Culicover, Ph.D.
Coorientador

Florianópolis, 2021.

To Rafa

ACKNOWLEDGEMENTS

Even though this is not the first acknowledgment section I have found myself writing, the situation I am in is surprisingly unique. For the past four and a half years, I have benefited from unprecedented amounts of academic generosity, intelectual insight and emotional support. In spite of all the hardships, these experiences were responsible for keeping me reasonably sane during the preparation of this thesis, partly carried through during a global pandemic.

I would like to thank my advisor, Heronides Moura, for bearing with my eccentricities throughout all of these years and for trusting me with the freedom to push my (often bizarre) ideas as boldly as I can. I also owe a tremendous debt to my co-advisor, Peter Culicover, whose general perspective on language has had a profound influence on me. Peter has generously offered critical feedback on nearly every word I wrote over the past four years. Through our friendship and collaboration, he has provided me not only with thrilling intellectual insights that never cease to surprise me, but also with a model of how to conduct good work in academia.

It is impossible to overestimate the impact that my period as a Visiting Scholar at The Ohio State University had on my life. I learned more about syntax and semantics during the months I spent in Columbus than at any other time in my life. Most of the insights that appear throughout this thesis were gestated somewhere in between Oxley Hall and Goodale Park. I can specifically link some paragraphs here to specific bus stops along the route out of the 8 line that runs through Neil Avenue. I am especially indebted to the folks that participated on the Synners discussion group, where many of the proposals in this thesis first saw the light of day. The following people were particularly helpful to me during our discussions: Aswhini Deo, Daniel Puthawala, Jordan Needle, Martha Johnson, Noah Diewald and Symon Stevens-Guille.

I am very grateful to Bob Levine for all of the support and advice throughout the last couple of years. Bob taught me how to think about language in an entirely novel way – his course on syntax was among the best I've ever took on any topic and hints of what I learned with him appear in all of my work since. Thanks also to Roberta Pires de Oliveira, who was a very helpful interlocutor throughout a good part of my doctorate. I am also thankful to the other members of my thesis committee – Marcus Lunguinho and Philip Miller. Both of them have made important contributions to this work that have made it substantially better than it was when I first thought it was "ready".

Conversations with several other researchers have played fundamental role in shaping my thoughts about some of the topics explored here and linguistics in general. I would specifi-

"Seek simplicity and distrust it."

Alfred North Whitehead

"There's a part of you always standing by,

Mapping out the sky, finishing a hat,

Starting on a hat, finishing a hat...

Look, I made a hat

Where there never was a hat. "

Stephen Sondheim

RESUMO

Esta tese estuda o fenômeno da anáfora intrasentencial de NPs sob a perspectiva da Sintaxe Mais Simples (CULICOVER; JACKENDOFF, 2005). Como uma alternativa à abordagem universalista e configuracional defendida pela Gramática Generativa Mainstream, argumento que parte das restrições interpretativas que pesam sobre formas pronominais individuais devem ser capturadas por construções específicas a cada língua, as quais estipulam correspondências bem-formadas entre arranjos particulares de funções gramaticais e estruturas de variáveis ligadas na semântica, incorporando insights básicos da teoria da reflexividade-e-cadeias de Reinhart e Reuland (1993) em uma abordagem não-derivacional. Outras restrições interpretativas vêm de fatores não-sintáticos relacionados à perspectiva discursiva (Sells 1987) e a princípios de inferência pragmática explorados na teoria neo-Griceana de Horn (1985). Uma vez que a abordagem construcional reduz princípios gerais da gramática a propriedades do léxico, ela permite toda a gama de diversidade encontrada em sistemas anafóricos nas línguas ao redor mundo. As fontes de universais são basicamente restritas à semântica (noções como reflexividade e ligação de variáveis), ao vocabulário de funções gramaticais (noções como predicado sintático e GF-comando), aos procedimentos de aprendizagem (e.g. preempção estatística) e a princípios pragmáticos gerais. Embora esta tese se detenha principalmente na microvariação interna às línguas germânicas e românicas, a abordagem desenvolvida aqui explica naturalmente o fato de que diferentes línguas podem expressar a categoria semântica universal de reflexividade por meios sintáticos radicalmente diferentes. A ideia de que construções são violáveis (com a violação possivelmente fazendo emergir significados marcados a partir de implicaturas de Modo) é também é uma adição útil à abordagem construcional, uma vez que ajuda a explicar o comportamento heterogêneo de formas pronominais específicas dentro de uma língua individual. Explorando esta última conjectura, proponho que os reflexivos de longa distância do inglês violam condições gramaticais, mas podem se tornar aceitáveis em virtude do fato de que uma interpretação logofórica é acionada como uma implicatura (MENUZZI 1999, 2004). Argumento que essa implicatura de logoforicidade apenas surge em contextos onde a construção padrão associada aos reflexivos coargumentais no inglês não tem como ser satisfeita de modo pragmaticamente feliz. Também proponho reduzir os efeitos sintáticos da Condição B da Teoria da Ligação a uma propriedade construcional que deve ser abertamente adquirida pelos aprendizes de cada língua e está sujeita à variação. Ao fazer isso, minha teoria explica por que muitas línguas simplesmente carecem de um princípio que estipula anti-localidade sintática para pronomes pessoais do tipo visto em inglês. Como mostro, isso é o que encontramos em certos dialetos do português brasileiro falados no sudeste do País. O tipo de efeito de anti-localidade que vemos para os pronomes pessoais nessas variedades é melhor compreendido como consequência de um princípio pragmático que associa formas não-marcadas a significados não-marcados (HORN, 1985; RETT, 2020). O traço comum que atravessa a maioria dos argumentos nesta tese é o objetivo de tornar o entendimento teórico dos pronomes (em inglês e outras línguas) o mais simples possível.

Palavras-chave: anáfora, pronomes, reflexivos, Simpler Syntax, Teoria da Ligação

RESUMO EXPANDIDO

Introdução

Esta tese estuda o fenômeno da anáfora intrasentencial de NPs sob a perspectiva da Sintaxe Mais Simples (CULICOVER; JACKENDOFF, 2005). Como uma alternativa à abordagem universalista e configuracional defendida pela Gramática Generativa Mainstream, argumento que parte das restrições interpretativas que pesam sobre formas pronominais devem ser capturadas por construções específicas a cada língua, as quais estipulam correspondências bem-formadas entre arranjos particulares de funções gramaticais e estruturas de variáveis ligadas na semântica, incorporando insights básicos da teoria da reflexividade-e-cadeias de Reinhart and Reuland (1993) em uma abordagem não-derivacional. Outras restrições interpretativas vêm de fatores não-sintáticos relacionados à perspectiva discursiva (SELLS, 1987) e a princípios de inferência pragmática explorados na teoria neo-Griceana de Horn (1984).

Objetivos

Os objetivos específicos desta tese são: (i) propor um quadro teórico e formal dentro do qual seja possível formular hipóteses específicas a respeito de restrições interpretativas a pronomes; (ii) propor minhas próprias hipóteses a respeito de quais seriam as restrições interpretativas específicas às formas pronominais no Inglês, Português Brasileiro, Holandês e algumas outras línguas. Esse exercício tem como objetivo de fundo discernir o que há de específico a cada língua e o que há de universal em fenômenos anafóricos.

Metodologia

A metodologia empregada nesta tese seguiu, em linhas gerais o método introspectivo característico da linguística gerativa desde os seus primórdios (CHOMSKY, 1957). Este método consiste em elaborar hipóteses formuladas em uma linguagem formal a respeito da aceitabilidade de um certo conjunto de estruturas linguísticas e testá-las diante da intuição de falantes nativos de cada língua. A maior parte dos julgamentos reportados ao longo do trabalho não foram feitos por mim, mas recolhidos da longa literatura linguística sobre o fenômeno da anáfora sentencial. Evitei reportar julgamentos feitos apenas por autores isolados e procurei me ater a dados cuja

validade não é objeto de debate ou contestações. Nas (poucas) ocasiões em que precisei utilizar dados ausentes na literatura prévia, fiz coletas informais de julgamentos entre falantes nativos. Em algumas ocasiões, recorri também a dados de corpora e a resultados de experimentos conduzidos por outros autores. As hipóteses elaboradas ao longo deste trabalho partiram de uma pesquisa bibliográfica extensiva em busca de fenômenos de diversas línguas (com especial ênfase no Inglês, Holandês e Português Brasileiro) tidos como problemáticos para as teorias sobre anáfora vigentes.

Resultados e Discussão

Argumento ao longo da tese que as restrições interpretativas sobre formas pronominais são ora epifenômenos de fatores extra-gramaticais, ora construções específicas que estabelecem correspondências entre arranjos de configurações gramaticais e estruturas de variável ligada na semântica. De acordo com essa conjectura, as fontes de universais anafóricos passam a ser basicamente restritas à semântica (noções como reflexividade e ligação de variáveis), ao vocabulário de funções gramaticais (noções como predicado sintático e GF-comando), aos procedimentos de aprendizagem (e.g. preempção estatística) e a princípios pragmáticos gerais (capturados pela teoria neo-Griceana de Horn (1984)). Ao reduzir os efeitos da Condição A da Teoria da Ligação clássica a um conjunto de construções, minha teoria prevê que pode haver um grau de variação razoável na forma como as línguas marcam sintaticamente a reflexividade. A ideia de que construções são violáveis (com violações possivelmente fazendo emergir significados marcados a partir de implicaturas de Modo) é também é uma adição útil à abordagem construcional, uma vez que ajuda a explicar o comportamento heterogêneo de formas pronominais específicas dentro de uma língua individual. Explorando esta última conjectura, proponho que os reflexivos de longa distância do inglês violam condições gramaticais, mas podem se tornar aceitáveis em virtude do fato de que uma interpretação logofórica é acionada como uma implicatura (MENUZZI, 2004). Argumento que essa implicatura de logoforicidade surge apenas em contextos onde a construção padrão associada aos reflexivos coargumentais no inglês não tem como ser satisfeita de modo pragmaticamente feliz. Também proponho reduzir os efeitos sintáticos da Condição B da Teoria da Ligação a uma propriedade construcional que deve ser abertamente adquirida pelos aprendizes de cada língua e está sujeita à variação. Ao fazer isso, minha teoria explica por que muitas línguas simplesmente carecem de um princípio que es-

tipula anti-localidade sintática para pronomes pessoais do tipo visto em inglês. Como mostro, isso é o que encontramos em certos dialetos do português brasileiro falados no sudeste do país. O tipo de efeito de anti-localidade que vemos para os pronomes pessoais nessas variedades é melhor compreendido como consequência de um princípio pragmático que associa formas não-marcadas a significados não-marcados (HORN, 1984; RETT, 2020).

Considerações Finais

Uma vez que a abordagem construcional proposta pela Sintaxe Mais Simples reduz princípios gerais da gramática a propriedades do léxico (em um sentido expandido), a diversidade encontrada nos sistemas anafóricos das línguas ao redor mundo passa a poder ser capturada de forma simples e direta. Embora esta tese se detenha principalmente sobre a microvariação interna às línguas germânicas e românicas, a abordagem desenvolvida aqui explica naturalmente o fato de que diferentes línguas podem expressar a categoria semântica universal de reflexividade por meios sintáticos radicalmente diferentes. Igualmente, ao reduzir a essência da Condição B da Teoria da Ligação a uma propriedade construcional, minha hipótese prevê que a anti-localidade sintática associada a pronomes pessoais deve ser explicitamente aprendida. Isso fornece uma nova maneira de interpretar o fato de que as crianças não parecem ter um conhecimento robusto da Condição B até os sete anos de idade (ELBOURNE, 2005; HAMANN, 2011; BAAUW, 2018). Também passa a não ser surpreendente que muitas línguas simplesmente careçam de um princípio de anti-localidade para pronomes do tipo visto em inglês. Essa hipótese foi fundamental para fornecer uma explicação para os padrões peculiares de anti-localidade que encontramos para os pronomes pessoais na variedade do português brasileiro falando em São Paulo e Minas Gerais. Apesar de seus fundamentos relativamente novos, minha proposta também tem um sabor tradicional, compartilhado pelas teorias de Pollard and Sag (1992) e Reinhart and Reuland (1993). Assim como Jespersen (1933), proponho que o domínio relevante no qual a gramática regula a interpretação dos pronomes é o predicado sintático - ou seja, mesmo o domínio sobre o qual a valência sintática e funções gramaticais como Sujeito e Objeto podem ser definidas. Esta reformulação captura a localidade de reflexivização e a anti-localidade de pronominalização (em línguas onde a última existe) de uma maneira mais direta do que abordagens baseadas em condições de licenciamento configuracionais (CHOMSKY, 1981; HORNSTEIN, 2001; CHARNAVEL, 2019). O desafio de tornar essa abordagem compatível com reflexivos discursivos não-

locais e com a tendência ubíqua de usar formas reflexivas especiais para sinalizar ligação local é transferido para a pragmática. O traço comum que atravessa a maioria dos argumentos nesta tese é o objetivo de tornar o entendimento teórico dos pronomes (em inglês e outras línguas) o mais simples possível. Uma vez que muitos artifícios teóricos e foram postulados com base em fatos sobre anáfora, uma teoria mais simples sobre fenômenos anafóricos pode contribuir para uma teoria gramatical geral mais simples — um objetivo bem-vindo para muitas abordagens gramaticais contemporâneas.

Palavras-chave: anáfora, pronomes, reflexivos, Simpler Syntax, Teoria da Ligação

ABSTRACT

This thesis studies the phenomenon of intra-sentential NP anaphora from the perspective of Simpler Syntax (CULICOVER; JACKENDOFF, 2005). As an alternative to the universalist and phrase-structure approach championed by Mainstream Generative Grammar, I argue that part of the interpretive constraints on individual pronoun forms should be captured by language-specific constructions specifying well-formed correspondences between particular arrangements of grammatical functions and bound variable structures in semantics, incorporating the core insights of Reinhart and Reuland's (1993) reflexivity-and-chains theory into a non-derivational setting. Other interpretive constraints come from non-syntactic factors related to discourse perspective (Sells 1987) and principles of pragmatic inference drawn from Horn's (1985) neo-Gricean theory. Since the constructional approach reduces general principles of grammar to properties of the lexicon, it allows for the full-range of diversity found in anaphoric systems across the world's languages. The sources of universals are mostly confined to semantics (notions like reflexivity and variable binding), the vocabulary of grammatical functions (notions like syntactic predicate and GF-command), the learning procedures (e.g. statistical preemption) and domain-general pragmatic principles. Though this thesis focuses mostly on the micro-variation among Germanic and Romance languages, the framework developed here naturally accounts for the fact that different languages may express the universal semantic category of reflexivity by radically different syntactic means. The idea that constructions are violable (with violation possibly giving rise to marked meanings as Manner implicatures) is also a useful addition to the constructional stance, since it helps to account for the diverse behavior of specific pronoun forms within an individual language as well. Exploring this latter conjecture, I propose English long-distance reflexives violate grammatical conditions, but may become acceptable in virtue of the fact that a logophoric interpretation emerges as an implicature (MENUZZI 1999, 2004). I argue that the logophoricity implicature only emerges in contexts where the core construction associated with English coargument reflexives has no way of being felicitously fulfilled. I also propose to reduce the syntactic effects of Condition B of the Binding Theory to a constructional property, which has to be overtly acquired by language learners and may vary from language to language. By doing so, my theory explains why many languages simply lack a syntactic anti-locality principle for pronominals of the kind seen in English. As I show, this is what we find in certain southeastern dialects of Brazilian Portuguese. The kind of anti-locality effect we see for pronominals in these varieties is best understood as a consequence of a pragmatic principle that associates unmarked forms with unmarked meanings (HORN, 1985; RETT, 2020). The common thread that runs through most of the arguments in this thesis is the goal of making the theoretical account of pronouns (in English and other languages) as simple as possible.

Keywords: anaphora, pronouns, reflexives, Simpler Syntax, Binding Theory

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AVM	Attribute-Value Matrix
BNC	British National Corpus
BP	Brazilian Portuguese
CB	Condition B of the Classical Binding Theory
CBT	Classical Binding Theory
CC	Chain Condition
CR	Conjunction Reduction
CRP	Constraint on Reflexive Predicates
D-REFL	Discursive Reflexive Structure
ERC	English Reflexive Construction
FAC	Form-Accessibility Correlation
GF	Grammatical Function (or Gramatical Function Tier)
GFO	Grammatical Function Ttier Object
IET	Identity Erasure Transformation
LF	Logical Form
LFG	Lexical-Functional Grammar
LSAC	Longman Spoken American Corpus
MGG	Mainstream Generative Grammar
ME	Meaningful Expression
MTF	Model-Theoretic Formalism
N-PASS	Non-Peculiar Passive Structure
N-REFL	Non-Discursive Reflexive Structure

PHON	Phonological Structure
POS	Posessivization
P-PASS	Peculiar Passive Structure
PR	Pronominalization Rule
PRED	Syntactic Predicate
PRON	Pronominal
PTF	Proof-Theoretic Formalism
QNP	Quantified Noun Phrase
R	Referential Independence
RCA	Reflexivity Condition A
RCB	Reflexivity Condition B
RCR	Relative Clause Reduction
RCT	Reflexivity and Chains Theory
RR	Reflexivization Rule
RTO	Raising to Object
RTS	Raising to Subject
SE	Simplex Expression Anaphor
SEM	Semantic Structure
SSH	Simpler Syntax Hypothesis
SGF	Set of Grammatical Functions
SiSx	Simpler Syntax
SYN	Syntactic Structure
TRH	Two Reflexives Hypothesis
UG	Universal Grammar

CONTENTS

INTRODUCTION

During his Harvard Lectures, renowned composer and conductor Leonard Bernstein exercised a short-lived career as a popularizer of linguistics. In the midst of this endeavor, he stumbled upon the importance of pronouns for human language:

> *John was glad that Harry persuaded John to take up golf*. That is clearly an impermissible sentence. You would never say it, although perhaps you could not cite a rule that forbids it. Yet such a rule does exist in the human mind, a transformational rule called [. . .] pronominalization, which is much easier to understand than to pronounce: it simply means the substitution of the correct and relevant pronoun for the repeated name. And so the second John is rewritten as the pronoun "him", and a grammatical sentence is born. Imagine I having to go around saying things like: *John promised that John would do John's homework the minute John finished John's dinner*. Nobody talks that way and it is perfectly self-evident why. (BERNSTEIN, 1976, pg. 75-76)

Bernstein's quote serves to inspire the curiosity of a general audience towards an interesting topic in linguistics, demonstrating a willingness to be puzzled by a fact most people take for granted in their day-to-day lives. However, his very last remark inappropriately points in the opposite direction. Far from being "self-evident", the phenomena he takes note of, in fact, raises many tricky questions – many of which are still unanswered, much like the conflict between tonal and atonal music Bernstein talks about during most of his Harvard Lectures.

What does it mean to "substitute" a linguistic form for another? Is this an inferential process speakers actually have to mentally perform while uttering or comprehending sentences? Or is the notion of "substitution" to be understood metaphorically? If that is the case, what is it a metaphor *of*? What counts as the "correct and relevant pronoun" for any given name? For instance, why couldn't the second occurrence of *John* in *John was glad that Harry persuaded John to take up golf* be replaced by the pronoun *himself*? Since the present Thesis is concerned with pronouns and anaphora, these questions need to be addressed.

The understanding of pronominalization as a kind of substitution resonates with the way most of us were taught to think of pronouns – pronouns are supposed to be words that *stand for* other words, in particular, nouns and noun phrases. Though somewhat misleading, this particular way of talking hints at an intuition which I believe is fundamentally correct: the idea that pronouns are words that lack a specific lexical content beyond certain grammatical features like number, gender and person. As a result of their lack of content, pronouns typically inherit their semantic value from some other expression in the discourse (i.e. their antecedent).

A standard way to formalize this intuition is by having pronouns semantically correspond to *individual variables*, in the sense used in first-order predicate-calculus (MONTAGUE, 1974; COOPER, 1979; BACH; PARTEE, 1980; REINHART, 1983; HEIM; KRATZER, 1998; BÜRING, 2005, i.a.). Gender, person and possibly number specifications can be understood as presuppositions or subtypings that determine the kinds of entities pronoun variables can range over. A pronoun like *she*, on this approach, corresponds to a variable whose value must necessarily be picked up from the set of female singular entities who are distinct from both the speaker and the hearer.

Even though all pronouns correspond to variables, not all pronouns are alike in every respect. Suppose Joanne and Robert are quietly sitting in the bar of a nightclub. All of the sudden, Joanne spots her husband, Larry, performing what she perceives to be a ridiculous dance on the dance floor. If Joanne decides to call Robert's attention to this fact, she can point to Larry and utter something like (1a), but not (1b).

(1)   a.   Look at him!

         b.   *Look at himself!

The contrast in (1) shows that, unlike personal pronouns, reflexive pronouns cannot be used as deictics – i.e. they cannot pick out their referents from the situation of utterance. Their meaning is inherently dependent on the existence of a pre-established referent in the discourse. Following Chomsky (1981) and Pollard and Sag (1994), I will use the term PRONOUN for the general category of NPs that function semantically like individual variables. I reserve the more specific term ANAPHOR for pronouns that are referentially defective (such as English reflexives) and the more specific term PRONOMINAL for pronouns that are referentially autonomous (e.g. personal pronouns such as him *him* and *her*).

I will assume that the opposition between anaphors and pronominals is universally established on the basis of this semantic property of referential (in)dependence. This is not to say, however, that all languages have pronouns belonging to both of these classes. It is logically possible that there are languages whose pronouns are all anaphors and languages whose pronouns are all pronominals. It is also possible that there are languages where the distinction between pronoun types is dependent on the linguistic (or extra-linguistic) context.

Anaphors and pronominals also differ other respects, which may vary from language to language. In English, anaphors must have antecedents that are sufficiently LOCAL (i.e. that

occur within the same clause as the anaphor) and PROMINENT (e.g. not embedded within some other NP). When an antecedent is either not local or not prominent, an anaphor is unacceptable and a pronominal must be used, as the following examples illustrate:[1]

(2)   *Laurie* told me that Curly loves {*herself* / *her*}.

(3)   An old boyfriend of *Laurie* still loves {*herself* / *her*}.

However, when the locality and prominence requirements are met, pronominals cannot be used and anaphors are the *only* way to express semantic identity:

(4)   *Laurie* loves {*herself* / *her*}.

This is not a property we typically find in formal languages. The well-formedness of a logical formula containing a variable generally does not depend on whether there is a covalued individual constant or variable elsewhere in the expression. This sensitivity to the global structure of expressions appears to be a unique property of how variables are encoded in natural languages (MCCAWLEY, 1981; REINHART, 2006).

With the plausible exception of *wh*-words, English pronouns are solid candidates for the most hotly debated lexical items among the vast inventory of forms found in the world's languages. This situation is somewhat surprising, since, before a recent profusion of debate in the latter half of the twentieth century, traditional grammars were not particularly excited about the topic. For instance, as Reuland (2014, p. 1) points out, all that Jespersen had to say about reflexive anaphors in his nearly 400 page grammar of English can be summarized in the following:

> When the subject and object are identical, we use for the latter a so-called reflexive pronoun, formed by means of *self* e.g. *I defend myself.* [...] The reflexive pronouns are also used after prepositions: *He looked at himself in the glass.* (JESPERSEN, 1933, p. 111-112)

Given the amount of ink spent over the last fifty years, the above quote can sound almost naive. But it is not obvious why a particular subset of pronouns in a language should be a

---

[1] In order to avoid the intrusion of particular theoretical assumptions in the presentation of the data, I adopt the notation in Postal (1970), signaling semantic identity between expressions by means of italics (see also Safir (2004) and Reuland (2011)). When an example occurs within the main text in italics, coreferential expressions are signaled in boldface. These conventions are only dropped when a particular theory involving indexing is being discussed – in which case numerical indices are used. As usual when it comes to studies of anaphora, judgments refer to string-interpretation pairs (and not merely to strings themselves). Most of the judgments in this Thesis were gathered from the fifty-year literature on pronouns in Generative Grammar, but others (which were not previously mentioned in the literature) were informally collected by me from native speakers. Unless otherwise indicated, reported judgments were consistently agreed upon by all of my informants.

major concern for linguists. Far from suggesting a linear progress towards a consensus, the sheer number of diverse proposals in recent decades is a testament to how much there is to be discovered even in apparently banal territories of grammar.

Mysteries turn up in a wealth of puzzling facts. Along with well-behaved contrasts like the ones in (2)-(4), which clearly conform to Jespersen's intuition (and to virtually all accounts since then), one can readily find data in which reflexive anaphors do not appear to signal semantic identity between clausemates (e.g. subjects and objects), such as (5), based on naturally occurring examples from the iWeb Corpus. In all these contexts, unlike what we see in (2)-(4), reflexive anaphors and pronominals can be freely interchanged keeping the meanings of the sentences constant:

(5)   a.   A little part of *me* still hates {*myself / me*} for letting that happen.

        b.   *Pears* believes some customers would be similar to {*herself / her*}.

        c.   The picture of {*himself / him*} that *he* gave his girlfriend will remind her with what she originally fell in love.

        d.   *She* found herself experimenting with materials and soon had a vision to create a clothing line that would not only give her an emotional outlet, but would empower other young women like {*herself / her*}.

        e.   This post is an almost accurate description of {myself / me}.

In (5a), the antecedent is embedded within the subject, just like in (3). In (5b)-(5d), pronouns and their antecedents are not clausemates, as also occurs in (2). Example (5e) is akin to (1b) in that there is no antecedent in the sentence at all. Nonetheless, a reflexive anaphor is acceptable in all of these structures. No widely accepted account of cases such as (5) exists – the truth about them is still pretty much up for grabs.

Nevertheless, far-reaching theoretical moves are frequently drawn on the basis of shaky foundations. Again, with the exception of the syntacticians's favorites *who*s and *what*s, it is hard to think of a class of expressions which has motivated so many conclusions about the nature of language as English pronouns. A few such conclusions deserve to be mentioned.

In the earliest study on the topic in Modern Generative Grammar, Lees and Klima (1963) claimed facts about English pronouns (anaphors and reflexives) vouched decisive theoretical tenets of the theory put forth by Chomsky (1957): rule ordering and generalized transformations. Ross (1967) and Chomsky (1973) used interpretive properties of pronouns of different

types to motivate the hypothesis that rules of grammar are subject to general conditions (e.g. locality) – a move which, despite the challenges posed by data like (5b)-(5d), paved the way for the Principles and Parameters framework. On similar grounds, Langacker (1969) and Reinhart (1976) were among the first to suggest that constraints on the interpretation of each type of pronoun were sensitive to abstract phrase-structure relations such as (c-)command.

Later, in the Government and Binding era, the interpretive properties of pronouns were crucial for motivating an entire module of Universal Grammar – the Binding Theory – in which complex syntactic notions like that of *government* were claimed to play a crucial role (CHOM-SKY, 1980, 1981). In this framework, anaphors were likened to traces of NP movement. This identification, in its turn, suggested two equally radical and mutually incompatible outcomes: the reduction of movement rules to rules of construal – as exemplified by Koster's (1978) work, as well as by non-transformational approaches (BRESNAN; KAPLAN, 1982a; POLLARD; SAG, 1994) – and the reduction of construal rules to movement – as proposed by Lebeaux (1983), Hornstein (2001) and Kayne (2005). Both of these alternatives are still alive today.

At the dawn of the minimalist program, Chomsky (1993) invoked facts about reflexive anaphors in A′ positions to support the Copy Theory of Movement and the elimination of S-Structure from the architecture of grammar. Starting with Chomsky and Lasnik (1993), attempts were made to purge narrow syntax of some of the theory-internal resources required by the earlier Binding Theory (e.g. syntactic indices, binding domains) and to offer more natural accounts in terms of principles of economy and the *bare output conditions* that operate at the interface with semantics (CHOMSKY, 1995). At around the same time, Belletti and Rizzi (1988), Larson (1990), Pesetsky (1995) and others explored data concerning English anaphors to justify more abstract phrase structure configurations in which thematic roles could be transparently encoded, as predicted by the Uniformity of Theta Assignment Hypothesis (BAKER, 1997).

The work of Reinhart and Reuland (1991, 1992, 1993) and Pollard and Sag (1992, 1994) popularized a very different kind of approach. Rather than stating interpretive rules for pronouns solely in terms the vocabulary of phrase structure (using notions like dominance, syntactic categories and c-command), these proposals argued that a proper theory of anaphora should also be sensitive to predicate-argument structure: e.g. notions pertaining to lexical semantics, syntactic valence and grammatical functions (e.g. subject, object, oblique).

In this Thesis, I will address some of the reasons why people feel there is so much to be said about pronouns of different sorts and try to evaluate how a simpler (hopefully *the* simplest)

alternative fares with respect to these issues. The theory proposed in this Thesis is mostly indebted to the predicate-based theory championed by Reinhart and Reuland (1991, 1992, 1993), so special attention will be devoted to discussing their approach.

This work is organized in seven chapters. Chapter 1 introduces some basic concepts that will be assumed throughout most of this Thesis. It also gives a brief historical overview of two classic theories of anaphora: an early transformational approach, represented mainly by Lees and Klima (1963), and the Classical Binding Theory made famous by Chomsky (1981, 1986b). I will argue that each of these instantiate a specific strategy in handling putative counterexamples to constraints on pronouns. The former favors enriching structures and derivations in order to make them compatible with a simple unified principle, while the latter favors enriching the principle in order to handle recalcitrant data. I will show that both of these tactics prove to be insufficient to handle the diversity of the phenomena at hand.

Chapter 2 summarizes an alternative approach to the problems of anaphora: Reinhart and Reuland's (1991, 1992, 1993, 1995) reflexivity-and-chains theory (henceforth RCT). I show that the RCT is fundamentally correct in stating interpretive constraints on pronouns in terms of argument structure (rather than phrase structure) as well as in eliminating the Binding Theory as an autonomous module of grammar. These conclusions warrant a more natural and typologically adequate theory of anaphora than the Classical Binding Theory was.

However, as I show in Chapter 3, the RCT winds up running into many of the same difficulties that haunted previous approaches to anaphors and pronominals. Furthermore, some of its innovations also introduce further problems. For instance, RCT's approach to anaphors winds up entailing that English reflexives can only have long-distance antecedents when they are *exempt* from grammatical constraints (where exemption is defined in purely syntactic terms). I will argue that this leads to empirically incorrect results, which motivate a different kind of theory. In addition to offering a systematic critique of the RCT, Chapter 3 also points towards the solutions to the problems of anaphora I spell out in subsequent chapters.

The theory sketched in Chapters 2-3 appeals, for the most part, to an outdated theoretical vocabulary drawn from Reinhart and Reuland's (1993, 1995) early work from the 1990s.[2] Chapter 4 presents and motivates a different grammatical framework which is suitable for formalizing hypotheses about anaphora that overcome some of the difficulties that upset the previous

---

[2] A translation of the RCT into minimalism was accomplished by Reuland (2001, 2011, 2017). Since the empirical content of original formulation of RCT is basically preserved in this updated version, I will mostly ignore the technical innovations introduced therein and recast the original insight of the RCT in my own terms instead.

approaches: a version of the theory of Simpler Syntax (SiSx) (CULICOVER; JACKENDOFF, 2005; CULICOVER, 2009, 2021; JACKENDOFF; AUDRING, 2020) with an enriched component for representing grammatical functions (BRESNAN; KAPLAN, 1982a). I interpret SiSx as a variant of construction grammar (GOLDBERG, 1995, 2013; MÜLLER, 2020).

The general aim of the constructional outlook is to minimize the role of language-specific innate knowledge (i.e. UG) in the explanation of facts about anaphora. This can be done by reanalyzing the principles that Chapters 2-3 attribute to UG in terms of extra-linguistic factors (e.g. pragmatics or economy considerations) or, alternatively, in terms of lexically-stored constructions. Condition A of the RCT, for instance can be reinterpreted as a construction. This means that, though the notion of a reflexive semantic predicate may be a universal property of the human conceptual system (CULICOVER, 2021), the way particular languages choose to convey these predicates is, to a large extent, a matter of arbitrary form-meaning pairings.

In Chapter 5, I propose a specific constructional account of English reflexives along these lines. Inspired by Menuzzi's (1999, 2004) work, I argue that one of the ways of purging the theory of anaphora from the unwelcome consequences of the previous approaches is by adopting a very simple general condition on reflexives and allowing it to be *violated* in certain contexts. The generality of the condition accounts for the *unity* of reflexives and its violability accounts for the fact that reflexives may behave differently in different contexts. This condition will be essentially a constructional formulation the following:

(6)   CONDITION A (simplified):

If R is a reflexive, then R is bound by a higher argument of a syntactic predicate P where R serves as a syntactic argument, thereby making P reflexive.

The notion of domain appealed to in (6) (the syntactic predicate) is the same notion over which valence is defined: it includes a head and its subcategorized arguments. When a syntactic predicate is such that one of its arguments is bound to another, we can say that it corresponds to a REFLEXIVE PREDICATE. Within the framework of SiSx, (6) can, therefore, be understood as a construction, because it establishes a language-specific correspondence between a particular form (i.e. morphosyntactic reflexivity) and meaning (a reflexive predicate).

In this approach, non-coargument reflexives like those in (5) above are neither treated as *exempt* from (6), nor licensed by a separate constraint. Rather, they are viewed as as tolerable violations of (6). What distinguishes these tolerable violations from unacceptable violations

like (2)-(4) is the fact that they occur in contexts where (6) *could not have been fulfilled* – i.e. either when reflexives are *not arguments* of any predicate, or when they are arguments of *anti-reflexive* predicates (predicates which cannot be interpreted as reflexive).

In Chapter 6, I present a theory of pragmatics that explains why violations of (6) are rendered acceptable precisely in these circumstances (HORN, 1984; GRICE, 1989; RETT, 2020). What I argue there is that violations of the grammatical constraint on reflexives are interpreted as contributing to markedness. A Manner implicature is responsible for linking markedness to atypical interpretations. In the case of reflexives, atypical interpretation conveyed by marked occurrences convey notions like perspective, empathy and point of view (SELLS, 1987, CHIERCHIA, 1989; STIRLING, 1993; OSHIMA, 2007; CHARNAVEL, 2019). The fact that these Manner implicatures may be conventionalized does not change the fact that the interpretations triggered by violations of (6) are pragmatically motivated (MORGAN, 1978).

Lastly, in Chapter 7, I propose my own account of the nature of constraints on pronominals across different languages, with a special emphasis on Brazilian Portuguese. I argue that local disjointness effects involving pronouns, typically captured in terms of a single innate syntactic constraint (Condition B of the Binding Theory) do not, in fact correspond to a unified phenomenon. As an alternative, I explore the hypothesis that the work previously assigned to Condition B should be distributed into two separate factors: (i) language-specific binding constructions that enforce disjointness; (ii) a pragmatically-grounded constraint on the morphosyntactic encoding of reflexive predicates in semantics.

In addition to explaining certain facts about the interpretation of pronominals in Brazilian Portuguese, the second of these factors sheds light on why languages tend to develop special devices to convey to convey reflexivity. In line with functionalist literature (FALTZ, 1985; FARMER; HARNISH, 1987; LEVINSON, 1991; HUANG, 2000; HASPELMATH, 2008), I argue that this cross-linguistically ubiquitous urge to mark semantic reflexivity is a byproduct of a principle that associates unmarked forms with unmarked meanings (HORN, 1984). The general idea is that, since pronominals are unmarked in contrast to reflexives, they typically cannot be used to convey reflexivity, because reflexive events are, generally speaking, less prototypical than non-reflexive ones. As we will see, this also provides a natural explanation for the behavior of what Reinhart and Reuland (1993) call "inherently reflexive predicates" (e.g. **shave**$'$ and **wash**$'$) and for the historical development of reflexives in English (KÖNIG; SIEMUND, 2000a; KÖNIG; GAST, 2002; CULICOVER, 2021).

Part I

PREVIOUS THEORIES

1 CONCEPTUAL AND HISTORICAL BACKGROUND

This first chapter presents the conceptual and historical background for the theory of anaphora proposed in this Thesis. Though the discussion here is largely informal, it serves as an initiation to the topics and hypotheses explored in subsequent chapters. I will introduce some basic terminology, as well as survey the landmarks in the history of the study of pronominals and reflexive anaphors in Mainstream Generative Grammar (MGG).

My focus will be on UNIVERSAL and UNIFIED accounts – i.e. those which include constraints that apply to *all* reflexives and pronominals in *all* languages, without exempting particular forms (within a language or cross-linguistically) from structural requirements. The main goal is to understand what went wrong with previous efforts of this sort.

My narrative mostly replicates MGG's emphasis on English as a source of data for theories of anaphora, though facts about other languages will also be adduced. However, insofar as most of the theories discussed here confer upon their principles an allegation of universality, presenting counterexamples in English suffices as a prima facie demonstration of their flaws.

In Section 1.1, I clarify some of the basic concepts that will be employed throughout this Thesis. In Section 1.2, I begin the historical overview of previous theories, starting with the transformational approach epitomized by Lees and Klima (1963) – i.e. the earliest theory of anaphora in the MGG tradition. Section 1.3 covers the Classical Binding Theory (CBT) of Chomsky (1981, 1986b) – i.e. the most influential theory of anaphora in MGG. Each of these strands exemplifies a different strategy in facing counterexamples to their postulates: the former favors enriching structures and derivations, while the latter favors enriching the principles that account for the interpretation of pronouns. Both of these strategies are ultimately insufficient. Section 1.4 elaborates on the problems affecting the CBT, which is the more influential of these two approaches. This paves the way for the proposals in Chapters 5–7.

1.1 BASIC CONCEPTS

Before we begin, it is useful to define a few basic concepts which will be reoccurring in the story that follows: concepts like (A-)BINDING, COREFERENCE, ANAPHOR and REFLEXIVE. Some of these were not clearly formulated (or even recognized) throughout most of the history of inquiry into anaphora. Nonetheless, even when discussing older theories, I will try to employ them in a relatively univocal fashion. This means that I will sometimes have to resort to minor conceptual anachronisms for the sake of ensuring commensurability with my own fa-

vored conceptual framework – which roughly follows the one in Reinhart (1983). My particular theoretical implementation of these concepts is discussed in Chapters 5–7.

I use the term ANAPHORA as a general label for the phenomenon of coconstrual between linguistic expressions. The term COCONSTRUAL is borrowed from Safir (2004) to simultaneously cover two different kinds of interpretive dependencies: (INTENDED) COREFERENCE and BOUND-VARIABLE relations (cf. POSTAL, 1970; LASNIK, 1976; EVANS, 1980; REINHART, 1983, 2006; BÜRING, 2005, i.a.). Both are customarily signaled, since Chomsky (1965), by means of coindexing – a practice I mostly abandon in favor of italics.

Lumping coreference and binding together under the rubric of coconstrual allows me to speak neutrally about semantic identity *in general*. This is useful for describing examples where both of these interpretive relations are instantiated, as well as theories in which the distinction between them is not recognized. However, since this ambiguity can also lead to confusion, it is important to have a proper understanding of the difference between these notions. A paradigm illustration is given in (1) (REULAND, 2017, pg. 15):

(1) a. *The soldier* has a gun. Will *he* attack?

   b. **No soldier* has a gun. Will *he* attack?

   c. *No soldier* thinks *he* will attack.

The mini-discourse in (1a) illustrates a typical instance of coreference. The pronoun is a referring expression on its own, which happens to share its referent with the definite NP *the soldier* (its antecedent). Another way of thinking about this involves saying that *the soldier* invokes an individual as a discourse referent, which is subsequently retrieved by the pronoun. This option is not available in (1b), because the antecedent of the pronoun is not the kind of expression that typically invokes a discourse referent. The pronoun cannot *co*refer with *no soldier* because the latter, as a quantified NP, does not *refer* to an entity in the first place.[1]

If coreference was the only possible mode of coconstrual, a pronoun could never be anteceded by a quantified NP. However, the acceptability of (1c) shows that this is not so. Pronouns can be anteceded by quantified NPs as long as they are *bound* by them – which is exactly what

---

[1] In some contexts, a quantified NP (e.g. *every soldier*) might make a set of entities contextually salient (e.g. the set of soldiers). The latter can then serve as an antecedent for a plural pronoun:

(i)   *Every soldier* has a gun. I think *they* are about to attack.

   Note, however, that the predicates *has a gun* and *are about to attack* are not predicated of the same discourse referent: the former is predicated of each element in the set of soldiers individually and the latter is predicated of this set as an aggregate entity. So it is still not the case that *every soldier* and *they* corefer in (i).

happens in (1c). In that example, the pronoun *he* does not refer to a linguistic-external entity – rather, it behaves like *bound variable* in first-order logic, depending on *no soldier* for its interpretation. As Reuland (2011, pg. 29) puts it: "Binding does not interpret an expression by assigning it a value directly [as in coreference]. It computes its interpretation by linking it to another expression [i.e. *no soldier*, in (1c)] and determining the interpretation of the latter."

The relevant logical structure of binding in (1c) can be represented as (2), which follows the λ-notation introduced by Sag (1976, pg. 103).

(2)   No soldier ($\lambda x.\ x$ thinks (x will attack))

The facts in (1) suggest that variable binding is subject to specific structural conditions which do not bear on coreference: e.g. it can only be established intra-sententially.[2] The contrast in (3) indicates a further difference in the syntactic manifestation of these two interpretive strategies (thereby reinforcing the idea that they are indeed distinct):

(3)   a.   [$_{NP}$The captain who trained [$_{NP}$ *the soldier*]] thought *he* would attack.

         b.   *[$_{NP}$The captain who trained [$_{NP}$ *every soldier*]] thought *he* would attack.

In a context where the antecedent NP is embedded within another NP, coreference is possible (cf. (3a)) but binding is generally not (cf. (3b)). This shows that, even intra-sententially, binding is governed by constraints that do not apply to coreference.

In all of these examples, quantified antecedents are used to evince the presence of bound readings for pronouns. However, whether a pronoun is interpreted as a bound-variable is actually *independent* of the semantics of its antecedent. Even fully referential NPs can bind pronouns – a point first observed by Keenan (1971) (see also Sag (1976), Williams (1977) and Reinhart (1983)). Consider the example in (4):

(4)   a.   *Maria* thinks *she* is a good singer.

         b.   Maria ($\lambda x.\ x$ thinks (x is a good singer))   (binding)

         c.   Maria ($\lambda x.\ x$ thinks (y is a good singer)) & $y = Maria$ (coreference)

The sentence in (4a) is formally ambiguous between the bound reading in (4b) and the corefer-

---

[2]   In fact, as I will discuss in Chapter 2, Reinhart's (1983) original insight is that only binding is subject to structural conditions (cf. GRODZINSKY; REINHART, 1993; REULAND, 2011; CULICOVER; VARASCHIN; WINKLER, in press). According to her, coreference is governed solely by domain-general interpretive heuristics (e.g. Gricean maxims, processing economy). This means that non-coreference effects – such as the ones attributed to CBT's Condition C (LASNIK, 1976, 1989; CHOMSKY, 1981) – are not encoded in the grammar.

ence reading in (4c). Two different properties are ascribed to Maria in each of these: in (4b), it is the property of considering oneself to be a good singer, whereas in (4b) it is the property of considering Maria to be a good singer.

Ultimately, the two readings above are truth-conditionally indistinguishable. But there do exist contexts in which they are not equivalent, such as VP-ellipsis:

(5) a. *Maria* thinks *she* is a good singer, and Elsa does too.

　　 b. Elsa ($\lambda x.\ x$ thinks (x is a good singer))　　　　　　　(binding: sloppy reading)

　　 c. Elsa ($\lambda x.\ x$ thinks (y is a good singer)) & $y = Maria$　(coreference: strict reading)

The sentence in (5a) is ambiguous (ROSS, 1967). The ambiguity lies within the second conjunct of (5a), which can be interpreted as claiming either that Elsa thinks Elsa is a good singer (the SLOPPY reading) or that Elsa thinks Maria is a good singer (the STRICT reading). In the former, the elided predicate is one where the pronoun is construed as a bound-variable (cf. (5b)), while in the latter it is construed as coreferential with the subject of the first conjunct (cf. (5c)).

Sentences involving the focus-sensitive operator *only* are yet another context where bound and coreferential readings for pronouns with referential antecedents are not equivalent:

(6) a. Only *Maria* loves *her* children.

　　 b. Only Maria ($\lambda x.\ x$ loves x's children)　　　　　　　　　(binding)

　　 c. Only Maria ($\lambda x.\ x$ loves y's children) & $y = Maria$　　　(coreference)

As in (5), the two possible meanings of (6a) hinge on whether the property denoted by *loves her children* is one in which the variable corresponding to *her* is bound or merely coreferential with the subject. The bound reading in (6b) entails that people other than Maria do not love their own children. The coreferential reading in (6c) on the other hand, entails that people other than Maria do not love *Maria's* children (so they might well love their own children).

Up to now, I have been speaking of binding as a relationship between arguments of predicates – i.e. between pronouns and their antecedents. It is this notion which is sensitive to the structural properties of sentences, giving rise to contrasts like the one in (3). However, this constitutes a slight deviation from the *logical* concept of binding, which is not a relation between arguments, but between variables and logical operators (e.g. $\forall$, $\exists$, and λ).

The notion of binding which is relevant in the statement of grammars for natural languages is distinct from the purely logical notion, but, nonetheless, it can be easily derived from it, as

in (7). Reinhart (2006, pg. 171) calls this derived notion A(RGUMENT)-BINDING. It correctly captures the difference between binding and coreference in all of the relevant examples above (cf. (4)-(6)). (Unless the distinction between logical binding and A-binding is directly pertinent to the issue at hand, I will keep referring to the latter by the name of *binding*.)

(7)   A(RGUMENT)-BINDING

β A-binds α iff β is the sister of a λ-predicate whose operator logically binds α.[3]

It is important to note that the definition in (7) does not apply at the level of narrow syntax, but at the level of SEMANTIC STRUCTURE (or, as Reinhart (2006) calls it, LOGICAL SYNTAX). This level is understood as a formal representation which supports inference and can be subject to a model-theoretic interpretation (REULAND, 2011, pg. 34). I assume that semantic structure is distinct from NARROW SYNTAX, which primarily serves to encode dominance relations between syntactic categories (e.g., S, NP, VP). In this picture, part of what grammars have to do is to specify a set of correspondence rules which relate surface structures (i.e. the superficial parse of strings into constituents) to this level of semantic representation.

According to the definition in (7), the only elements that can be bound are those which correspond directly to VARIABLES in semantic structure. I assume, following Bouchard (1984), that what makes a natural language expression correspond to a variable is the lack of specific lexical content beyond features like PERSON, NUMBER and GENDER (known in the literature as ϕ-features). Among NPs, this includes solely the category of PRONOUNS, broadly construed to include reciprocals like *each other*, *one another*, etc. Names, descriptions and quantified NPs form the separate category of R(EFERENTIAL)-EXPRESSIONS. If we accept, following Reinhart (1983), that the only kind of grammatically relevant coconstrual relation is A-binding, pronouns will be the only kinds of NPs whose interpretations are subject to grammatical constraints.

There are different ways to subdivide pronouns in light of their morphosyntactic properties: singular *vs.* plural, clitics *vs.* non-clitic, oblique *vs.* accusative, etc. With respect to their

---

3   Given this definition, we can now see that there is actually a *third* way in which two NPs can be coconstrued with each other apart from (A-)binding and coreference. As Heim (1998) notes, variables can also be *co-bound* by the same antecedent. She gives the following example:

(i)   a.   *Every wife* thinks that only *she* respects *her* husband.
         b.   Every wife (λx. x thinks (only x (λy. y respects y's husband)))   (binding)
         c.   Every wife (λx. x thinks (only x (λy. y respects x's husband)))   (co-binding)

The sentence in (ia) is ambiguous in much the same sense as (6a). On the reading represented in (ib) (where the variable corresponding to *she* binds the possessive) the sentence entails that every wife thinks that other wives do not respect their husbands, while reading (ic) (where *she* and *her* are cobound by *every wife*) entails that every wife thinks that other wives do not respect her husband.

anaphoric behavior, however, the most important distinction internal to the class of pronouns is the one between ANAPHORS – e.g. *myself*, *herself*, *themselves* – and PRONOMINALS – e.g. *his*, *her*, *them* – as it appears in Chomsky (1981) and Pollard and Sag (1994).

A theory of anaphora is essentially a set of empirically motivated constraints concerning the linguistic environments under which each of these pronoun types can be bound. As Burzio (1991) notes, this presupposes that the typology of pronouns can be established on the basis of features that are independent of those which are prescribed by the grammatical constraints on anaphora themselves. Otherwise, these constraints will be interpreted as *definitions* of the pronoun types ANAPHOR and PRONOMINAL. The problem with this move is that it robs the theory of anaphora of its empirical content, making its statements true *by stipulation*.[4]

In order to avoid this conceptual problem, I propose to ground the taxonomy of pronouns in terms of the independent property of REFERENTIAL INDEPENDENCE. This is, in fact, the standard view in most of the literature (CHOMSKY, 1986b; SAFIR, 2004; REULAND, 2011; SPORTICHE, 2013). An ANAPHOR, in this framework, is a pronoun that is unable to support independent reference. In virtue of this semantic property, anaphors are always dependent on previously established discourse referents and, therefore, cannot be used as deictics (SAFIR, 2003). A PRONOMINAL, in turn, is a kind of pronoun which *can*, but need not, be dependent on pre-established referents. Since pronominals are not referentially deficient like anaphors, they can be used deictically to pick out an entity which is not given in the prior discourse.

The deictic scenario in (8) clearly distinguishes English reflexives from ordinary personal pronouns in this respect. The former are genuine anaphors and the latter are pronominals, according to the classic definitions I assume here:

(8)   [Context: Joanne and Robert are sitting in a bar. Joanne spots her husband, Larry, in the dance floor. She decides to call Robert's attention to Larry.]

    a.   Look at him!   [pointing gesture]

    b. *Look at himself!   [pointing gesture]

---

[4] The absence of a characterization of NP types which is independent of their structural distribution *vis à vis* the anaphora conditions undermines some proposals in the literature. Everaert (1991), for example, suggests that NPs are pronominals or anaphors depending on whether they have a local or a long-distance antecedent (i.e. whether they comply to Conditions A or B of the CBT). This approach strikes me as circular. As Burzio (1991, 84) puts it, "if an anaphor is *defined* as an element that occurs locally bound, then of course it will always be true that it is". This circularity is often avoided in the literature on English anaphora because there is a tacit agreement on an independent morphological criterion for distinguishing anaphors from pronominals: namely, the presence/absence of the *-self* morpheme. However, this clearly does not do as a general basis for establishing a taxonomy of pronouns across different languages.

Kiss (2009) claims that cases like (9) – first discussed by Ross (1970) – are counterexamples to the claim that anaphors (such as English reflexives) are referentially dependent:

(9)   a.   This paper was written by Susan and myself.

       b.   This is a picture of myself.

I do not think (9) poses a real threat to the anaphor status of English reflexives. Even though the reflexives therein occur without being coconstrued with other expressions in the prior linguistic context, they are still inherently dependent on a pre-established discourse referent: the speaker. The only difference between this dependency and the one in (8b) is that the referent in (9) is one that does not need to be independently invoked by an overt NP. In other words, the anaphoric requirement of 1SG reflexives is trivially satisfied because the speaker always counts as given in any discourse context (SELLS, 1987; STIRLING, 1993; SCHLENKER, 2005a).[5] If the reflexives in (9) are switched to 3rd person, the result is just as bad as (8b):

(10)   a.   *This paper was written by Susan and himself.

       b.   *This is a picture of himself.

The contrast between (9) and (10) shows that English reflexives are *anaphoric* in the relevant sense: i.e. they are *always* (i.e. in all morphological realizations) referentially dependent.

Being an anaphor, however, is not the same thing as being a REFLEXIVE. In order to count as a reflexive, a pronoun must have a REFLEXIVIZING FUNCTION: i.e. the persistent capacity to express coconstrual between thematic arguments of a predicate (FALTZ, 1985). The contrast in (11) shows how the pronoun *her* differs from *herself* with respect to this property:

(11)   *Laurie* admires {*herself* / **her*}.

In the case of *himself*, the capacity to express local coconstrual is also PERSISTENT – i.e. it is not confined to a lexically restricted class of predicates. This is what makes *himself* a genuine reflexive, in contrast to an expression like the Dutch anaphor *zich* (REULAND, 2011, pg. 199). The latter, though deictically deficient, can only express semantic identity between thematic coarguments in a subset of the cases where English reflexives can, cf. the contrast between (12) and (13) (where *zich* is glossed as SE, standing for SIMPLEX EXPRESSION ANAPHOR).

---

[5]   This applies to all 1SG forms, including personal pronouns like *I*. Since the speaker is always present, there is no way to *know* whether these forms are referentially dependent. Therefore, in order to determine their place in the typology of pronouns, we always need to look at their 3rd person counterparts (as I did for reflexives).

(12) a. *Bobby* shaves *himself.*

      b. *Bobby* hates *himself.*

(13) a. *Max* scheert *zich.*
          Max shaves SE

      b. **Max* haat *zich.*
          Max hates SE.

The notional characterization of reflexives as expressions that signal a specific kind of interpretation for their predicates comes from traditional grammar, as well as from early transformational work on reflexivity (JESPERSEN, 1933; LEES; KLIMA, 1963; POSTAL, 1971). The segregation between the reflexivizing function property (which defines the category of reflexives) and the property of being referentially dependent (which defines the category of anaphors) provides semantic grounds for classifying pronouns according to the typology in Table 1:

	Reflexive Anaphor	Non-reflexive Anaphor	Pronominal
Referential Independence	-	-	+
Reflexivizing Function	+	-	-

Table 1 – The typology of pronouns (adapted from Reinhart and Reuland (1993, pg. 659))

English reflexives fall into the category of reflexive anaphors because they lack referential independence and have a reflexivizing function. The Dutch pronoun *zich* is a non-reflexive anaphor because, even though it is also referentially dependent, its reflexivizing function is not persistent. English personal pronouns (e.g. *she*, *her*, *they*) have neither of these properties and are, therefore, neither reflexive nor anaphoric. This places them in the category of pronominals.[6]

---

[6] The typology in Table 1 also leaves open the theoretical possibility of there being REFLEXIVE PRONOMINALS. Anagnostopoulou and Everaert (1999) argue that this possibility is realized by the Greek expression *o eaftos tu*. Unlike English reflexives, these forms can appear as subjects in contexts like (i) (where CL stands for clitic):

(i)   *O eaftos tu*   tu   aresi *tu Petru*
      the self   his-NOM CL.DAT likes the Petros-DAT
      '*Himself* pleases *Petros* '   (ANAGNOSTOPOULOU; EVERAERT, 1999, pg. 108)

However, the fact that a particular kind of pronoun is used as a subject does not prove that it is not an anaphor – at least not according to the independent characterization of anaphors in terms of referential dependence adopted here. It only proves that the pronoun in question does not abide by constraints which restrict anaphors to non-subject positions. In fact, Chiou (2010, pg. 140) uses examples like (ii) to argue that *o eaftos tu* is in fact an anaphor in our sense – i.e. it cannot bear autonomous reference:

(ii) **O eaftos tu   ine pu   mas kani mathima.*
      the self   his-NOM is   that us   does lesson
      'It is himself who teaches us.'

If Chiou (2010) is right, all that (i) shows is that universal principles which restrict anaphors to non-subject positions have to be abandoned.

One of the main advantages of the typology in Table 1 is that it is based on two semantic properties which are independent of any specific theory regulating the conditions under which NPs can be bound. This allows us to say that the property of being an anaphor, for instance, carries no implication about whether anaphoric NPs must have antecedents that obey particular structural requirements (e.g. locality). In this framework, the correlation between NP types and structural requirements is entirely an empirical matter – i.e. it is part of what theories of anaphora should be *about*, rather than part of what they should *presuppose*.

The next section argues that concrete attempts to state the structural requirements governing English pronouns in terms of single unified principles have failed. To an extent, this is due to a lack of clarity concerning some of the basic notions we examined throughout this section (e.g. the distinction between binding and coreference).

1.2 A TRANSFORMATIONAL THEORY OF ANAPHORA

Lees and Klima (1963) launched the first systematic investigation of anaphora in the tradition of Mainstream Generative Grammar (MGG). Their proposal inaugurated an influential paradigm which was further elaborated and refined by Postal (1966a, 1966b), Ross (1967, 1969) and Lakoff (1968). In this section, I present a version of Lees and Klima's (1963) theory incorporating some of these subsequent modifications. In spite of its anachronism, this line of work is a useful starting point for my discussion because it presents, in its virtues and flaws, a particularly clear statement of a unified approach to reflexives and pronominals.

A virtue of Lees and Klima's theory is the simplicity of its rule system. However, as we will see, this virtue is counterbalanced by the complexity of the underlying structures and derivations Lees and Klima (1963) have to posit in order to make the full range of phenomena compatible with their simple rules. This exemplifies a trade-off which is familiar to linguistic theorizing: when the purpose is to account for a heterogeneous set of data, the simpler the rules, the more complex the underlying structures typically have to be (MILIORINI, 2018).

Lees and Klima's (1963) motivation was to employ the (then recently developed) machinery of transformational analysis to shed light into regularities that were not captured by what they took to be purely semantically-based theories of anaphora. Their main hypothesis was that the distribution and interpretation of pronouns were subject to *syntactic rules* whose formulation required some of the tools put forth by Chomsky (1955, 1957).

The rules in question were assumed to be TRANSFORMATIONS – i.e., rules that map

strings with a given constituent structure (phrase-markers) into new strings with derived constituent structure (CHOMSKY, 1957, pg. 44). The applicability of a transformation to a phrase marker is conditioned by the satisfaction of the structural analysis each transformation specifies. Since structural analyses are formulated in terms of a syntactic vocabulary (i.e. string sequences consisting of Vs, NPs, PPs, etc.), transformations are a natural way to express that aspects of interpretation are syntactically determined. The proposal to use transformations to model anaphoric relations can be traced back to Chomsky (1955, chap. 9).

Two transformations were proposed to account for the interpretation and distribution of pronouns in English. The following are slightly revised versions of the rules Lees and Klima (1963, pg. 23) proposed, incorporating insights by Postal (1966b) and Ross (1967):

(14)   REFLEXIVIZATION RULE (RR)

$$\underbrace{X - NP - Y - NP' - Z}_{\text{Structural Analysis}} \Rightarrow \underbrace{X - NP - Y - NP' + \textit{Refl} - Z}_{\text{Structural Change}}$$

Conditions: (i) NP=NP$'$; (ii) NP and NP$'$ are within the same simplex sentence.

(15)   PRONOMINALIZATION RULE (PR)

$$\underbrace{X - NP - Y - NP' - Z}_{\text{Structural Analysis}} \Rightarrow \underbrace{X - NP - Y - NP' + \textit{Pron} - Z}_{\text{Structural Change}}$$

Condition: (i) NP=NP$'$

These rules apply the order given above. The symbols *Refl* and *Pron* are to be understood as abstract syntactic features which are mapped to actual phonological forms representing *herself*, *her*, *himself*, etc. by means of specific morphophonemic rules. This allows RR and PR to be generalized to other languages. In fact, Postal (1966a,1966b) hypothesizes that RR and PP are universal and that the only language-specific part of the theory of anaphora are the constraints which govern how NPs marked with *Refl* and *Pron* get realized in terms of phonological forms. Since Lees and Klima propose a *single* rule to account for *each* type of pronoun in English (with plausible applications to other languages), their approach counts as a unified one.

The purpose of these rules is to convert the second of two identical NPs in an underlying structure into a reflexive or a pronominal. RR applies to NPs which cohabit the same simplex sentence (a clause containing a single main verb), while PR applies to NPs in all domains in an unrestricted fashion. This system thus treats reflexivization as a special instance of pronominalization: basically, as form of pronominalization confined to a local domain.[7]

---

[7]   This point is not clear in Lees and Klima's own exposition, but it is made explicit by Postal (1966b, pg. 202),

This astonishingly simple system accounts for important distributional facts concerning reflexives and pronominals in English. The derivation of (16a) is given in (16c) (where *MR* stands for the morphophonemic rule mapping NPs bearing *Refl* to their overt reflexive forms).

(16) a. *Julie* protects *herself*.

b. **Julie* protects *her*.

c. Julie protects Julie $\overset{RR}{\Longrightarrow}$ Julie protects Julie+*Refl* $\overset{MR}{\Longrightarrow}$ Julie protects herself

The example in (16b) cannot be similarly derived. In order to generate (16b), PR would need to apply to the underlying kernel *Julie protects Julie* before RR inserts *Refl* into the structure. But this cannot happen, because RR is ordered to apply before PR.

This illustrates a general point: whenever two NPs cohabit the same simplex sentence at the point of RR, the second one will be turned into a reflexive before PR gets a chance to turn it into a pronominal. Ordering (14) before (15) thus makes reflexivization bleed pronominalization in local environments. This entails that reflexives and pronominals must be in complementary distribution: where a reflexive can occur, a pronominal cannot occur (and *vice versa*).

The example in (17) confirms the other side of this observation. A reflexive cannot appear in (17a) because, in the underlying kernel of this structure, the two instances of *Laurie* occupy different clauses at the point where RR is supposed to apply. Since RR cannot apply, PR is free to convert the second occurrence of *Laurie* into a pronominal. This is shown in the simplified derivation in (17b), which skips he morphophonemic rules applying after RR/PR. (*N* represents an abstract noun which gets replaced by a sentence by an embedding transformation.)

(17) a. *Laurie* thought that Curly hated {*herself* / **her*}.

b. {Laurie thought *N*; Curly hated Laurie} $\Rightarrow$ Laurie thought that Curly hated Laurie $\overset{PR}{\Longrightarrow}$ Laurie thought that Curly hated her

The fact that the structural change part in both RR and PR specifies a linear order where the antecedent NP has to precede the pronoun explains why sequences like (18) are blocked:

(18) a. **Herself* protects *Julie*.

b. **She* thought that Curly hated *Laurie*.

---

who claims that "reflexivization can be taken as that subtype of pronominalization relevant to identical NP within the same simple sentence structure" (see Ross (1967) for a similar proposal). This idea, which had a lasting impact on the field, goes back to Chomsky's (1955, pg. 657) early transformational treatment.

Lees and Klima (1963) also makes accurate predictions for cases like (19). These sentences are produced by interactions between (14)-(15) and other transformations, such as the Identity Erasure Transformation (IET) (ROSENBAUM, 1967), later known as Equi-NP Deletion, and Conjunction Reduction (CR) (CHOMSKY, 1957). In order to yield the right results, the transformations have to be applied in precisely the order given in (20).

(19)   a.   *Julie* tried to protect {*herself* / **her*}.

       b.   Billy told *Julie* to protect {*herself* / **her*}.

       c.   *Billy* told Julie to protect {**himself* / *him*}.

       d.   *Laurie* prefers Curly and {*herself* / **her*} to Jud.

(20)   a.   {Julie tried to *N*, Julie protect Julie} $\Rightarrow$ Julie tried to Julie protect Julie
            $\overset{RR}{\Rightarrow}$ Julie tried to Julie protect herself $\overset{IET}{\Rightarrow}$ Julie tried to protect herself

       b.   {Billy told Julie to *N*, Julie protect Julie}
            $\Rightarrow$ Billy told Julie to Julie protect Julie $\overset{RR}{\Rightarrow}$ Billy told Julie to Julie protect herself
            $\overset{IET}{\Rightarrow}$ Billy told Julie to protect herself

       c.   {Billy told Julie to *N*, Julie protect Billy} $\Rightarrow$ Billy told Julie to Julie protect Billy
            $\overset{PR}{\Rightarrow}$ Billy told Julie to Julie protect him $\overset{IET}{\Rightarrow}$ Billy told Julie to protect him

       d.   {Laurie prefers Laurie to Jud $\overset{RR}{\Rightarrow}$ Laurie prefers herself to Jud;
            Laurie prefers Curly to Jud} $\overset{CR}{\Rightarrow}$ Laurie prefers Curly and herself to Jud

As the derivations above make evident, full lexical NPs (e.g. *Julie*, *Laurie* and *Billy*) were the sources for pronouns like *her* and *himself* in the the underlying kernel sentences. The transformational outlook preserved, therefore, the traditional grammar view according to which pronouns are words that "replace" nouns.

This position became unpopular after Jackendoff (1969), Dougherty (1969) and Chomsky (1971) proposed a full-fledged interpretive system for anaphora. They argued that reflexives and pronominals should no longer be *derived* from fully-specified NPs, but base-generated as real lexical items. Rules like (14)-(15) were supplanted by rules of *construal*, which regulate the *interpretation* of NPs, as schematized in (21) (adapted from Jackendoff (1969, pg. 45)):

(21)   [NP^2 + *Refl/Pron*] is interpreted as +/- coconstrued with NP^1 in the context X.

The interpretive stance has been close to a consensus since these early formulations – with Lidz and Idsardi (1998), Hornstein (2001) and Kayne (2005) as virtually the only outliers.

There were some important empirical reasons for this shift towards interpretive theories of anaphora. One, which is implicit in Jackendoff's (1969) formulation of (21), was that rules of construal can specify *disjoint reference* just as easily as coreference. This seemed necessary to account for the meaning of obviative expressions like *else* (as in *Jud loves Laurie but Laurie loves someone else*) (CULICOVER; JACKENDOFF, 1995) and for the effects of what later came to be called Conditions B and C of the CBT (LASNIK, 1976, 1981; CHOMSKY, 1981).

While transformational systems based solely on RR and PR can handle simple disjointness effects like (18) above, they face a much harder task when plural antecedents are involved (POSTAL, 1966a; GRINDER; POSTAL, 1971). It is not clear what the underlying sources of *us* and *we* have to be in order to predict the following deviant cases:

(22)   a. ??*I* scratched *us*.

       b. ??*We* scratched *me*.

Moreover, all proposals which derive pronouns from full NPs run into what came to be known as the Bach-Peters Paradox: they have no way of generating (23a) without positing an infinitely recursive underlying structure like (23b) (BACH, 1970, pg. 121):

(23)   a.   The man who deserves it will get the prize he wants.

       b.   The man who deserves [the prize [which the man [who ...]] wants] will get the
            prize [the man [who deserves the prize [which ...]]]] wants.

Transformational theories also have no natural way of representing the distinction between coreference and binding discussed in Section 1.1. As a consequence, they are forced to claim that sentences like (24) – where the pronoun is interpreted as a variable bound by *every actress* – are synonymous with (25) (JACKENDOFF, 1968, pg. 433). This is clearly false.[8]

(24)   a.   *Every actress* admires *herself*.

       b.   *Every actress* thinks that *she* should win the Oscar.

---

[8]   Quantifiers were something of a blind spot to transformational accounts of other coconstrual phenomena. Partee's (2015) examples in (i) and (ii) illustrate similar mismatches between derived transforms and their underlying source structures in Equi-NP Deletion and Conjunction Reduction. It is unclear how recent attempts to recast coconstrual phenomena in transformational terms (HORNSTEIN, 2001) circumvent these problems.

(i)   a.   Every candidate wanted to win
       b.   ≠ Every candidate wanted every candidate to win.

(ii)   a.   No number is even and odd.
       b.   ≠ No number is even and no number is odd.

(25)   a.   Every actress admires every actress.

            b.   Every actress thinks that every actress should win the Oscar.

All of these difficulties arise in a system like Lees and Klima's (1963), where coconstrual between two surface forms is defined as identity on an underlying level of structure. There is no reasonable solution to these problems other than switching to interpretive theories.

However, not all of the obstacles Lees and Klima (1963) face are directly related to their outmoded transformational mindset. Some of them stem from the very *simplicity* of their rule system. Simple rules like RR and PR require artificial assumptions to handle data like (26)-(28):

(26)   a.   *The men* found a smokescreen around *them*.

            b.   *The men* threw a smokescreen around *themselves*.

(27)   a.   *John* saw many pictures of *him*.

            b.   *John* saw many pictures of *himself*.

(28)   a.   **Mary*'s father supported *herself*.

            b.   *Mary*'s father supported *her*.

Assuming that the sentences in (26)-(27) are simplex, only reflexives are expected to occur. The fact that *both* reflexives and pronominals are equally acceptable is puzzling. As we saw above, in line with most of the literature up to Huang (1983), Lees and Klima (1963) take complementary distribution between the two kinds of pronouns to be an inviolable norm. In the particular version of their theory I am discussing here, the complementarity assumption derives from RR being ordered to apply before PR.

The contrast in (28) presents an even more obvious difficulty, because it is exactly the *opposite* of what RR and PR apparently predict. Note that the structural analyses in both of these rules – as was the case for transformations in general in early MGG (CHOMSKY, 1957) – do not refer to fine-grained aspects of the constituent structure of the strings to which they apply. All that is required for RR to apply, for example, is that two NPs cohabit the same simplex sentence domain – a condition which (at first sight) is met in (28a). No other structural relations (e.g. c-command) were imposed in the structural analysis, because the notion of domain relevant for anaphora was, basically, the simplest one imaginable: the (simplex) sentence.

In order to circumvent these problems while keeping their rules intact, Lees and Klima (1963) had to attribute richer (and largely unmotivated) covert structures to the strings (26)-(28).

To account for the non-complementarity within locative PPs illustrated in (26), Lees and Klima (1963) proposed that (26a) and (26b) result from two distinct underlying derivations:

(29) a. {The men found a smokescreen; A smokescreen was around the men}

⇒ The men found a smokescreen which is around the men

$\overset{PR}{\Longrightarrow}$ The men found a smokescreen which is around them

⇒ The men found a smokescreen around them

b. The men threw a smokescreen around the men.

$\overset{RR}{\Longrightarrow}$ The men threw a smokescreen around themselves

According to Lees and Klima, the surface similarity between (26a) and (26b) turns out to be misleading, because only in (26b) – understood as (29b) – the nominal which is a target for RR is actually part of the same simplex sentence as its antecedent. The guiding intuition seems to be that each kernel sentence should encode a separate predication. Since the PP in (29a) is not a part of the predicate headed by *found* (i.e. it is an ADJUNCT), it should be generated as an independent kernel. This means that it must be embedded in the main clause as a separate sentence at some point. If this is the point in which RR and PR apply, only PR will be able to go through. The PP in (29b), on the other hand, is an ARGUMENT of *throw*. This allows it to be generated directly as part of the main sentence kernel, making the repeated NP within it a target for RR, thereby blocking PR. A similar analysis of these data is also offered by Partee (1965).

One of the problems the proposal assumed in (29) faces is that it predicts, contrary to facts, that reflexives that have subjects as antecedents cannot appear within adjunct PPs and pronominals that have subjects as antecedents cannot appear within argument PPs:

(30) a. *The men* found a smokescreen around *themselves*.
      b. *The men* threw a smokescreen around *them*.

Though cases like (30a) have been reported as marginal by some speakers (CHOMSKY, 1981; FALTZ, 1985; HESTVIK, 1991) the judgments are not, by any means, as clear as Lees and Klima imply (cf. REINHART; REULAND 1993, pg. 686; SAFIR, 2004, pg. 100). Sentences like (30b) seem to be universally accepted. The only way to accommodate (30) within Lees and Klima's system would be to posit some kind of structural ambiguity, whereby locative PPs can count simultaneously as arguments and adjuncts for the same predicates (WILKINS, 1988).

For the possessives in (28), the proliferation of underlying structures is even more telling.

In order to explain how reflexivization is unacceptable (and pronominalization is acceptable) in an apparently "local" context like (28), Lees and Klima (1963) propose to treat possessives in English as transformational in origin, deriving from full sentences: e.g. *Mary's father* is derived from *the father that Mary has* (which is itself a derived phrase-marker). In the latter, *Mary* is embedded within a relative clause. This entails that there is a point in the derivation of (28) in which *Mary* is not within the same sentence as the direct object of the verb *support*. If this point is the point where RR and PR apply, then only the conditions for PR would be met, because the two nominals which refer to Mary would not be in the same sentence:

(31)   {The father supported Mary; Mary has a father}
       ⇒ The father that Mary has supported Mary
       $\overset{PR}{\Longrightarrow}$ The father that Mary has supported her ⇒ Mary's father supported her

For cases like ***John saw many pictures of him*** in (27a), Lees and Klima (1963, pg. 23) suggest the derivation in (32). As in (29a), each kernel encodes a separate predication:

(32)   {John saw many pictures; The pictures are of John}
       ⇒ John saw many pictures which are of John
       $\overset{PR}{\Longrightarrow}$ John saw many pictures which are of him ⇒ John saw many pictures of him

The problem with this derivation, however, is that it accounts solely for pronominals and fails to explain how reflexives appear in (27b) (***John saw many pictures of himself***). As Lees and Klima (1963, pg. 24) admit, this is a particularly tricky issue because the pair in (27), unlike the one in (26), consists of strings which are identical, *minus* the pronouns.

The only alternative for Lees and Klima (1963) is to claim a kind of structural ambiguity. This requires enriching the base component to yield, along with the kernels in (32), more complex kernels like the one in (33), where the two occurrences of *John* occupy the same simplex sentence. It is not clear whether this move is actually compatible with the overall framework – especially given the assumption that each kernel sentence should encode a separate predicate.[9]

---

[9]   This perceived tension is one of the motivations for the TRH, which is inaugurated by Partee's (1965) proposal to treat apparent violations of RR as generated from underlying structures containing a pronominal and an emphatic form homophonous with reflexives, as in ***John saw many pictures of him himself***. According to this theory, the pronominal to which the emphatic attaches can be subsequently deleted, giving rise to ***John saw many pictures of himself***. This was also the solution adopted by Helke (1970). A different early implementation of the TRH is found in Postal (1971, pg. 188), who does not take RR violations to be derived from emphatics, but posits a *late reflexivization* rule. Unlike the RR, *late reflexivization* applies *after* pronominalization and whatever rule reduces the relative clause to a simple PP in (29a) and (32).

(33)   John saw many pictures of John $\overset{RR}{\Longrightarrow}$ John saw many pictures of himself

The problems that pronouns embedded within NPs pose for the transformational theory go even deeper than examples like (32)-(33) suggest. More complex cases reveal the emergence of ordering paradoxes. The examples below were discussed in Chomsky (1993):

(34)   a.   Curly knows that this picture of *herself*, *Laurie* will like.

        b.   *Curly* knows that this picture of *himself*, Laurie will like.

In order to explain how Lees and Klima's system could yield (34a), one would have to stipulate a kernel *Laurie will like this picture of Laurie* to which RR could apply, as in (33). A topicalization transformation (which "moves" the *picture* NP to the front of the embedded clause) could only apply after RR. However, precisely the opposite ordering would have to be stipulated to account for (34b), where the reflexive within the *picture* NP takes the matrix subject as its antecedent. If RR applied before topicalization, the antecedent of the reflexive could only ever be the embedded subject Mary, as in (34a).[10]

Another ordering paradox emerges for pronouns within *picture* NPs take possessives as their antecedents (JACKENDOFF, 1969, pg. 39). Consider the following contrast:

(35)   a.   Curly saw *Laurie*'s pictures of *herself*.

        b.   *Curly saw *Laurie*'s pictures of *her*.

If PR precedes Posessivization (POS) and Relative Clause Reduction (RCR) – as is required to explain the pronominals in (31)-(32) – (35b) should be acceptable. This is the case regardless of whether POS applies before or after RCR. The following derivation illustrates:

(36)   {{Curly saw pictures; Laurie took pictures}

        $\Rightarrow$ Curly saw pictures that Laurie took, The pictures are of Laurie}

        $\Rightarrow$ Curly saw pictures that Laurie took that are of Laurie

        $\overset{PR}{\Longrightarrow}$ Curly saw pictures that Laurie took that are of her

        $\overset{RCR}{\Longrightarrow}$ Curly saw pictures that Laurie took of her

        $\overset{POS}{\Longrightarrow}$ Curly saw Laurie's pictures of her

To avoid this overgeneration and account for the availability of the reflexive in (35a), we need

---

[10]   Moreover, it is unclear to which string RR could apply in (34b), since it is implausible that *Curly knows that this picture of Curly* is construed as a part of simplex sentence anywhere in the derivation.

RR to apply after POS and RCR. This would give us the following partial sequence:

(37) Curly saw pictures that Laurie took that are of Laurie

$\overset{RCR}{\Longrightarrow}$ Curly saw pictures that Laurie took of Laurie

$\overset{POS}{\Longrightarrow}$ Curly saw Laurie's pictures of Laurie $\overset{RR}{\Longrightarrow}$ Curly saw Laurie's pictures of herself

The problem is that this particular derivational history incompatible with the basic ordering between RR and PR, since PR has to precede both POS and RCR (cf. (31)-(32)). Furthermore, the application of RR after POS also runs into conflict with the ordering required to explain why a reflexive is unacceptable in **Mary's father supported **herself*** (cf. (31)).

The reason why Lees and Klima's theory gives rise to these ordering paradoxes is because it does not anticipate the existence of non-sentential domains for anaphora. In particular, it does not recognize that the *picture* NP in (35) is in some sense similar to a simplex sentence, insofar as it requires a reflexive and rules out a pronominal.

There were attempts to save the transformational theory by claiming that *picture* nouns (e.g. *picture*, *description*, *painting*) are derived from underlying verbs. Ross (1967, pg. 155-157) attributes this suggestion to Warshawsky. On this account, the underlying form of ***Laurie's pictures of herself*** would, in fact, be a full sentence (something like *Laurie depicted Laurie*). If RR and PR apply before this sentence is converted into an NP, a reflexive is correctly predicted to be obligatory (since two occurrences of *Laurie* occupy the same sentence).

The problem, as Jackendoff (1969, pg. 39) notes, is that this approach does not extend to cases like (34). In addition, it is also stipulative, as it requires the postulation of underlying verbs for which there is no independent evidence. Though one could argue that *picture*, *description* and *painting* are derived from corresponding verbs, this is far less plausible for nouns like *poem*, *novel* and *biography*, which, nonetheless, display a similar behavior with respect to anaphora.

Another ordering paradox emerges in connection to Raising to Subject (RTS):

(38) a. *Laurie* seems to admire *herself*.

   b. *Laurie* seems to *herself* to have solved the problem.

Example (38a) suggests that RR applies early on, before *Laurie* gets raised to the subject position of *seems*, as the derivation below illustrates (Δ stands for an empty node):

(39) {Δ seems *N*; Laurie admires Laurie $\overset{RR}{\Longrightarrow}$ Laurie admires herself}

   ⇒ Δ seems Laurie admire herself $\overset{RTS}{\Longrightarrow}$ Laurie seems to admire herself

Example (38b), on the other hand, suggests RR applies *after* Laurie has undergone RTS:

(40)   {Δ seems to Laurie *N*; Laurie to have solved the problem}

  ⇒ Δ seems to Laurie to Laurie have solved the problem

  $\overset{RTS}{\Longrightarrow}$ Laurie seems to Laurie to have solved the problem

  $\overset{RR}{\Longrightarrow}$ Laurie seems to herself to have solved the problem

The only way to avoid this ordering paradox – wherein RR has to apply both before (cf. (39)) and after (cf. (40)) RTS – is to view RR as a CYCLIC RULE, as Jackendoff (1969) proposed within an interpretive framework. A cyclic rule is one that applies repeatedly in pre-defined stages in the derivation (e.g. every time an S node is encountered).

Raising to Object (RTO) is tricky for a different reason. While the distribution of pronouns in (41a) implies that the position adjacent to *expect* is part of the same simplex S as *Laurie*, the one in (41b) implies that this position is part of the same simplex S as the object of *defend*:

(41)   a.   *Laurie* expects {*herself* / **her*} to win.

  b.   Laurie expects *Curly* to defend {*himself* / **him*}.

These observations could suggest that all NPs in (41) populate the same simplex S. However, this is patently wrong, since it fails to account for the following distributional pattern:

(42)   *Laurie* expects Curly to defend {**herself* / *her*}.

What we need to cover the facts above is for the RR domain of the embedded object to include the position adjacent to *expect* and exclude the matrix subject and for the RR domain of the matrix subject to include the position adjacent to *expect* and exclude the embedded object. In other words, the position adjacent to *expect* hast to be a part of two otherwise disjoint simplex Ss: the main clause (as in (41a)) and the embedded clause (as in (41b)). The only way to capture this, within Lees and Klima's theory, is by positing a Raising to Object (RTO) transformation that "moves" the embedded subject to the object position in the main clause.

But this move is technically unstatable within the simple transformational theory. The difficulty stems from the fact that RTO would be a STRING VACUOUS RULE (i.e. a rule that keeps the input string intact in the output, changing only its underlying structure), whereas traditional transformations, as defined by Chomsky (1955, 1957), are designed precisely to

yield overt manipulations of strings (CHOMSKY, 1973, pg. 253- 254).[11]

Lastly, note that the existence of cataphoric and deictic uses of pronominals is also problematic in Lees and Klima's simple transformational theory of anaphora. Their PR makes it obligatory for there to be an NP antecedent which precedes every pronominal (LANGACKER, 1969; ROSS, 1969). This is contradicted by ordinary examples like (43):

(43)   a.   If *he* gets rich, *Enoch* will buy a house.

        b.   Julie likes him!   [pointing at Billy]

These were the kinds of examples that led Chomsky (1955, pg. 649) to suggest that at least some pronominals must be introduced as real lexical items. There is, however, a way to avoid this conclusion. We can preserve the view that all instances of pronouns are derived transformationally (in keeping with Lees and Klima's simple theory) if we assume that the pronouns in (43) are, in fact, repetitions of preceding lexical NPs in some underlying level.

For (43a), we could claim that the subordinate *if*-clause actually starts out to the right of the main clause. If PR applies at this point, as in (44a), the pronoun is licensed. The case in (43b) is more tricky. Nevertheless, we could posit a conjunction structure containing a demonstrative clause (e.g. *That is Billy*) in the stage where PR applies. This clause would then get subsequently deleted, leaving the impression that *him* occurs without a linguistic antecedent, as the sample derivation in (44b) illustrates:[12]

(44)   a.   Enoch will buy a house if Enoch gets rich

            $\overset{PR}{\Longrightarrow}$ Enoch will buy a house if he gets rich

            $\Rightarrow$ If he gets rich, Enoch will buy a house

        b.   That is Billy and Julie likes Billy $\overset{PR}{\Longrightarrow}$ That is Billy and Julie likes him

            $\Rightarrow$ Julie likes him

Needless to say, even though some of these solutions technically work, they are ex-

---

[11] In order to circumvent this technical conundrum, the obstinate transformationalist could posit a phonologically null complementizer to serve as a marker for the boundary of simplex sentences. RTO could then be stated as follows (POSTAL, 1974; POSTAL; PULLUM, 1988):

(i)   RAISING TO OBJECT (RTO) : X-V-C-NP-Y $\Rightarrow$ X-V-NP-C-Y

    After decades of heated debate, a strategy equivalent to (i) was eventually incorporated into MGG (LASNIK; SAITO, 1991; CHOMSKY, 1995, 2013).

[12] The derivation for (43b) might look outlandish. However, a similar proposal is suggested in Kayne (2005), who holds that, in cases like (43b), *him* is either anteced by an unpronounced demonstrative or by an overt antecedent in a previous sentence, which is integrated with (43b) via some kind of covert conjunction.

tremely artificial and theoretically dubious. They essentially rely on transformational mechanisms which were shown to be excessively powerful (CHOMSKY, 1970; EMONDS, 1970; JACKENDOFF, 1972; PETERS; RITCHIE, 1973) – e.g. the rule in (44b), which deletes the first conjunct of a coordination. Furthermore, even if these language-specific and non-structure preserving transformations are accepted, the proposed ordering of these rules with respect to the RR and PR is also entirely *ad hoc*.

Anachronisms aside, Lees and Klima's (1963) proposals are useful for my exposition because they flaunt a dilemma which affects all unified approaches to anaphora since then. In light of potential counterexamples to simple unified principles such as RR and PR, theorists can take one of the two following strategies:

(i)   ENRICH STRUCTURE: claim that the apparent counterexamples are not *really* counterexamples, but cases where there is a mismatch between overt and covert structure (where the latter in fact *confirms* the favored principle); or

(ii)   ENRICH PRINCIPLE: revise the principle and make it more complex (which either burdens learning, if the principle is learned, or UG, if the principle is innate).

Lees and Klima (1963) produced the perfect example of a simple unified theory because they opted for (i) in every case. This allowed them to keep their simple rule system intact in the face of diverse data – but only at the cost of having to postulate complex derivational mechanisms. Lees and Klima never considered the introduction of an additional structural relation (e.g. c-command) in order to refine their notion of domain to handle (26)-(28). Rather, they stuck to their crude definition of domains as simplex sentences and proceeded to proliferate complex covert structures for cases where their simple notion of domain apparently fails. This overall strategy, which is a corollary of syntactocentrism (CULICOVER; JACKENDOFF, 2005), is still common to diverse treatments of anaphora within MGG.

1.3   THE CLASSICAL BINDING THEORY

Even though the appeal to abstract syntactic structure to avoid counterexamples remained fashionable, it was generally agreed that the complex derivational mechanisms required by Lees and Klima (1963) were excessive and unmotivated. In accordance with the ENRICH PRINCIPLE strategy, a more nuanced view of the rules responsible for licensing reflexives and pronominals was called for. Instances of this tactic were Jackendoff's (1969) Thematic Hierarchy Condition,

Langacker's (1969) primacy relations, Chomsky's (1973) Conditions framework and Reinhart's (1976) c-command requirement. All of these were attempts to impose *further* constraints on anaphora beyond what is suggested by RR and PR, while still maintaining a unified approach.

This line of development crystalized in works such as Chomsky (1981, 1986b), Reinhart (1983) and Huang (1983). These proposals have enough in common so as to merit a single label: they are all versions of what I call the CLASSICAL BINDING THEORY (CBT).

Before getting into details, it is important to keep in mind that the CBT is conceptually a very different theory from Lees and Klima's. First, it is fully interpretive: pronouns are taken to be real lexical items, and not spell-outs of transformations – this alone avoids many of the problems we saw in Section 1.2. Second, its postulates are explicitly interpreted as principles of UG, whereas RR and PR were for the most part treated as language-specific rules (but see Postal (1966b)). Third, the CBT fits into a framework which allows recursive phrase-structure rules, as first suggested by Chomsky (1965). Fourth, coconstrual is signaled by means of syntactic coindexing, and not by identity of phrase-markers in covert structures. Lastly, as its name suggests, the CBT is primarily concerned with BINDING (construed as narrow syntax relation) and only derivatively concerned with the distribution of pronouns. Though, of course, insofar as binding is syntactically defined, distributional effects are expected.

I will call the narrow syntax notion of binding SYN-BINDING, in order to distinguish it from Reinhart's (2006) semantic notion of A(RGUMENT)-BINDING given in (7) above.[13] The concept of syn-binding is defined as follows (CHOMSKY, 1981, pg. 184):

(45) α is SYN-BOUND by β iff α and β are coindexed and β c-commands α;

     a. β C-COMMANDS α iff neither α nor β dominate the other and the first branching node that dominates β also dominates α.

---

[13] The concept of syn-binding denotes a relation between syntactic constituents whereas A-binding applies specifically to a level of *semantic structure*, where the notion of c-command plays no independent role. To be sure, the two concepts overlap to an extent. Given the way λ-predicates are formed compositionally, if β A-binds α in semantics, β is likely to syn-bind α in narrow syntax as well. In fact, if one abides to a strict version of type-driven compositionality, semantic A-binding entails syn-binding – see Büring (2005, pg. 91). The converse, however, is not true: syn-binding does not entail A-binding – at least in a system where coindexing can be interpreted as coreference (FIENGO; MAY, 1994). A syn-binding representation like (ia), for instance, can correspond either to the binding semantic structure in (ib) or to the coreference structure in (ic):

(i)   a.   Curly$_5$ thinks that he$_5$ is clever.
     b.   Curly ($\lambda x.$ x thinks (x is clever))
     c.   Curly ($\lambda x.$ x thinks (y is clever)) & $y =$Curly

Though syn-bound in (ia), the pronoun under the reading in (ic) is *not* semantically bound: it acts as a free variable whose value is contextually supplied. Some consequences of this are discussed in Chapter 2.

The CBT is essentially a theory about how each type of NP must be, can be or shouldn't be syn-bound. The typology of pronouns in the CBT is simpler than the one in Table 1 above. The CBT does not grant a privileged status to the notion of reflexive – rather, it treats reflexives (along with reciprocals like *each other*) as archetypes of the category of anaphors as a whole. This implies that all referentially dependent NPs behave alike with respect to narrow syntax. As in Table 1, personal pronouns (*him*, *her*, etc.) are classified as pronominals.

The principles of the CBT relevant for anaphors and pronominals – Binding Conditions A and B – have subtly different incarnations, but they all agree on the following:

(46)   CLASSICAL BINDING THEORY (CBT) (CHOMSKY, 1986b, pg. 166)

      a.   CONDITION A: An anaphor is syn-bound in a local domain.

      b.   CONDITION B: A pronominal is free (i.e. not syn-bound) in a local domain.

The concept of local domain invoked in (46) varied among versions of the CBT. Its most prominent rendition is Chomsky's (1981, pg. 211) notion of GOVERNING CATEGORY:

(47) γ is a GOVERNING CATEGORY for α if and only if γ is the minimal category containing α, a governor of α, and a SUBJECT accessible to α.

Chomsky (1981, pg. 209) introduces the concept of SUBJECT as a technical term intended to cover both subjects in the usual sense and AGR (the bundle of agreement features in verbal inflections).[14] The definition of accessibility is given in terms of the *i*-within-*i* condition in (48) (CHOMSKY, 1981, pg. 212):

(48) $[_{\psi} \dots \delta \dots]$ is ill-formed if ψ and δ bear the same index.[15]

(49) σ is ACCESSIBLE to α if and only if σ is in the c-command domain of α and assignment to α of the index of σ would not violate (48).

Governing categories and simplex sentences make equivalent predictions for the more

---

[14]   This is a way of unifying the Specified Subject and the Tensed Sentence Conditions of Chomsky (1973).

[15]   It is unclear whether the *i*-within-*i* condition is independently motivated (as Chomsky's formulation here suggests) or whether it should simply be seen as a stipulation inherent to the definition of accessibility. Independent motivation appears to come from structures like *[his$_1$ father]$_1$ and *[a picture of itself$_1$]$_1$ which (48) correctly rules out. But there are counterexamples to *i*-within-*i* as a general constraint (CULICOVER, 1997, pg. 71):

(i)   a.   One finds [many books about themselves$_1$]$_1$ in Borges's literary output.
        b.   [That woman listening to her$_1$ own voice on the radio]$_1$ is Barbra Streisand.

    Fanselow (1986) gives numerous similar counterexamples in German.

basic cases of reflexivization and pronominalization discussed in Section 1.2. Some of these examples are repeated below:

(50)   a. $\langle$Julie$_1$ protects {herself$_1$ / *her$_1$}$\rangle$

       b.   Julie$_2$ tried $\langle$PRO$_2$ to protect {herself$_2$ / *her$_2$}$\rangle$

       c.   Billy told Julie$_8$ $\langle$PRO$_8$ to protect {herself$_8$ / *her$_8$}$\rangle$

       d.   Billy told Julie$_6$ $\langle$PRO$_6$ to protect {*himself$_6$ / him$_6$}$\rangle$

       e.   Laurie$_3$ thought that $\langle$Curly hated {*herself$_3$ / her$_3$}$\rangle$

In each of the sentences in (50), the minimal category including the pronoun, its governor (i.e. the verb that selects the pronoun) and an accessible SUBJECT is the simplex sentence, as indicated by the angled brackets. Within this domain, Condition A requires anaphors to be syn-bound (as they are in (50a)-(50c)) and Condition B requires pronominals to be free (as they are in (50d)-(50e)).

However, if we examine some of the phenomena that posed problems for the transformational theory in Section 1.2, a notion of domain like (47), which is more intricate and internally structured than simplex sentences, starts to look more appealing. Governing categories, precisely in virtue of their complexity, permit a significant reduction of structure in the explanation of some of these data. Consider first the issue of binding by possessives within NPs:

(51)   a.   Curly saw $\langle$Laurie$_3$'s pictures of herself$_3$$\rangle$

       b.   *Curly saw $\langle$Laurie$_1$'s pictures of her$_1$$\rangle$
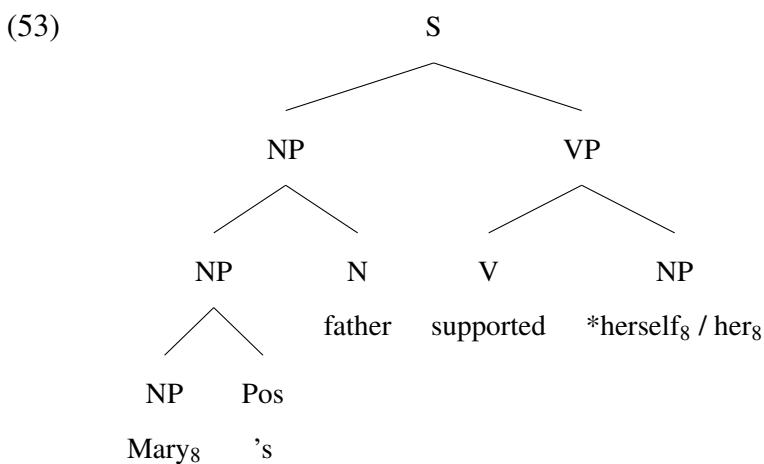
In order to explain (51) within the simple Lees and Klima inspired transformational system of Section 1.2, it was necessary to posit underlying structures where *picture* NPs are represented as full sentences. This is not needed in the CBT. If we assume that possessives are subjects, the *picture* NPs meet all of the requirements for being the governing categories for the pronouns in (51): they are the minimal categories containing the pronouns themselves, a governor for the pronouns (the preposition *of*) and a SUBJECT (*Laurie's*) accessible to the pronouns. Since Condition A says anaphors need to be syn-bound within their governing category, (51a) is acceptable. Pronominals, in their turn, are prohibited from being syn-bound within their governing category by Condition B, so (51b) is ruled out.

The richer conception of locality in terms of governing categories also avoids the need to posit a formally dubious RAISING TO OBJECT transformation to explain data like (52):

(52)   a. $\langle$Laurie$_1$ expects {herself$_1$ / *her$_1$}$\rangle$ to win

        b.   Laurie expects $\langle$Curly$_5$ to defend {himself$_5$ / *him$_5$}$\rangle$

        c.   Laurie$_2$ expects $\langle$Curly to defend {*herself$_2$ / her$_2$}$\rangle$

The anaphor and the pronominal in (52a) are governed by *expect*. This makes the matrix sentence, which includes *Laurie* as an accessible SUBJECT, the relevant governing category for *herself*/*her*. Since the syn-binding relationship is established internally to this category, the anaphor has to be used obligatorily. In (52b)–(52c), however, the embedded pronouns are governed by the verb *defend*. The minimal category containing these pronouns, their governor and the accessible subject *Curly* is the embedded clause. In (52b), the pronouns are syn-bound inside this clause, so an anaphor is obligatory. In (52c), on the other hand, the syn-binding relation is *not local*, so only the pronominal is acceptable.

Aside from a domain condition – which empirically improves on Lees and Klima's (1963) "simplex sentence" proviso – (46a) introduces an additional requirement for anaphors like reflexives: they must be syn-bound, and, therefore, C-COMMANDED by their antecedents (cf. (45)). The c-command relation, originally formulated by Reinhart (1976) with antecedents in Klima (1964), is taken to be independently necessary to account for other phenomena of language, such as unbounded dependencies, government and scope. Even though this requirement adds further complexity to the rule system, it allows a reduction of structure in the explanation of the pattern with possessives in (28), whose syntactic representation is sketched out below:

(53)
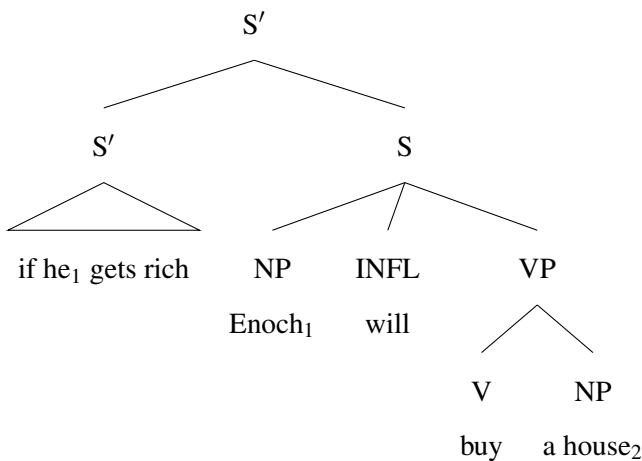
There is no need to resort to a complex derivation here. A reflexive is ruled out in (53) because its putative antecedent *Mary* does not c-command it. A pronominal is fine because pronominals are only subject to a negative requirement: they must not be locally syn-bound, as
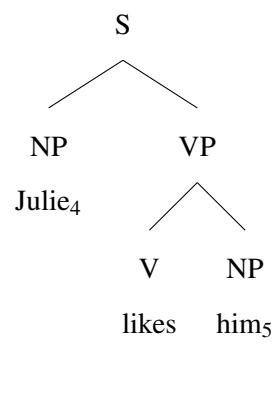
per Condition B. Since the pronominal in (53) is not syn-bound at all, the structure is fine.

The same applies to cataphoric and deictic uses of pronominals (cf. (43)). Unlike Lees and Klima's (1963) PR, Condition B does not require pronominals to be preceded by their antecedents. A pronominal is at liberty to not have an antecedent at all or to precede its antecedent, insofar as it is not c-commanded by the latter. This allows the underlying structures of (43) to be much simpler than the derivations in (44), as (54) exemplifies:

(54)   a.

```
                    S′
          ┌─────────┴─────────┐
         S′                    S
      ┌───┴───┐          ┌─────┼─────┐
  if he₁ gets rich      NP   INFL    VP
                      Enoch₁  will  ┌──┴──┐
                                    V     NP
                                   buy  a house₂
```

b.

```
             S
          ┌──┴──┐
         NP     VP
       Julie₄ ┌──┴──┐
              V     NP
            likes  him₅
```

Note, though, that this simplification of structures is done at the expense of enriching the rule system. Since the derivational history of a sentence is no longer sufficient to determine the reference of pronouns contained within it, the CBT needs to appeal to independent principles of construal to determine interpretation (see Hornstein (2001, pg. 12-13) for a similar point).

The CBT is embedded as a module within the general architecture of the Revised Extended Standard Theory (CHOMSKY, 1975, 1980, 1981), which is illustrated in Figure 1.

```
        D-Structure
            │
            │
        S-Structure
          ┌─┴─┐
   Phonetic Form   Logical Form
```
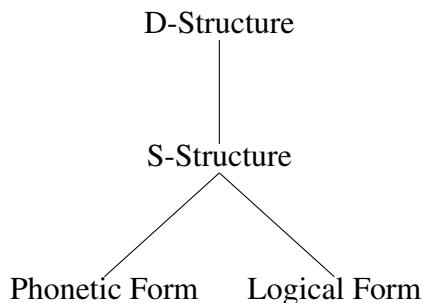
Figure 1 – The Revised Extended Standard Theory

A single transformation Move α is responsible for mapping D-Structure into S-Structure and S-Structure into Logical Form. Move α displaces a constituent from one structural position to another, leaving behind a coindexed trace, which serves to keep a record of the derivational history of a sentence at each of its derivational steps (CHOMSKY, 1975, 1977).

Given that traces are NPs like any other, they can enter into syn-binding relations that are visible for the Binding Conditions. This eliminates the ordering paradox that emerges when Lees and Klima's transformational theory is applied to Raising to Subject constructions. We no longer need to say that the reflexive is licensed *before* raising in (55a) and and *after* raising in (55b). Since raising is an instance of Move α, we can say that that reflexives are licensed *after* raising in both of these cases. In (55a), the reflexive is locally syn-bound by the trace t_1; in (55b), it is locally syn-bound by *Laurie*.

(55) a. [Laurie$_1$ seems [$\langle t_1$ to admire herself$_1\rangle$]]

     b. [$\langle$Laurie$_1$ seems to herself$_1\rangle$ [t_1 to have solved the problem]]

Though the CBT was empirically and conceptually superior to previous theories, it still struggled to provide a satisfactory account for many phenomena. The absence of complementary distribution between reflexives and pronominals within *picture* NPs illustrated in (27), repeated below as (56), is a case in point.:

(56) a. *John* saw many pictures of *him*.

     b. *John* saw many pictures of *himself*.

The crucial problem here stems from the fact that, like Lees and Klima, most versions of the CBT take complementarity between reflexives and pronominals to be the norm. The domain in which reflexives must be syn-bound and the domain in which pronominals must be free is defined to be one and the same: the governing category, as defined in (47).

Assuming (47), the CBT predicts that the reflexives and pronominals in (56) should meet their respective binding-theoretic requirements in the main clause, which is the minimal category containing their governor (the case-marking preposition *of*) and an accessible SUBJECT (*John*). This is apparently falsified by (56a): the pronominal therein is *not* free in the main clause as Condition B seemingly requires it to be. The dilemma (56) raises can be phrased as a question: how can a *single* notion of domain for anaphors and pronominals be made compatible with the existence of syntactic contexts in which they appear to occur in free variation?

Chomsky's (1986b, pg. 167) answer to this is an instance of the strategy which proliferates covert structure to keep the rule system intact. Assuming a variant of (47), he argues that the NPs in (56) contain a null PRO subject. When this local PRO subject is not coindexed to the pronoun it c-commands, as in (57a), a pronominal is required. When PRO is coindexed with

the pronoun, as in (57b), an anaphor is mandatory.

(57)   a. $John_8$ saw many [$_{NP}$ PRO_2 pictures of him_8]

          b. $John_8$ saw many [$_{NP}$ PRO_8 pictures of $himself_8$]

There are fair reasons to be suspicious about this analysis. First, it presupposes that NPs have PRO subjects – a position which was shown, on quite independent grounds, to be untenable (WILLIAMS, 1985). Second, even if PRO subjects are tolerated, it is unclear how they are supposed to be interpreted in structures like (57). As Lebeaux (1985, pg. 347) notes, there is no detectable meaning difference between pairs like (57a)-(57b). There is no suggestion, for instance, that the pictures in (57a) (where PRO is not coindexed with *him*) belong to some unspecified person, whereas in (57b) they belong to John. PRO and its index appear to be semantically inert – they only function as *ad hoc* devices to save the CBT.

The only alternative compatible with the spirit of the CBT is to devise two distinct notions of domain: one for reflexives and one for pronominals. The domain in which a pronominal must be free has to be, in some sense, "smaller" than the domain in which a reflexive must be bound.

This was more or less the gist of Huang's (1983) modification of Chomsky's (1981) system. Huang proposes to relativize the notion of governing category according to the type of pronoun contained within each category. He states that only governing categories for *anaphors* have to contain an accessible SUBJECT in the sense of (49).[16] Huang also redefines the concept of SUBJECT so as to include nominal heads (HUANG, 1983, pg. 557-558):

(58) γ is a GOVERNING CATEGORY for α if and only if γ is the minimal category containing α, a governor of α, and a SUBJECT that, if α is an anaphor, is accessible to α.

(59)   The SUBJECT of a maximal phrase ψ is the subject of ψ (in the usual sense) or the nominal head of ψ.

According to (58), the minimal governing category for the pronominal in (56a) is the postverbal NP *pictures of him*, since it contains the pronominal, its governor (i.e. the preposition *of*) and a SUBJECT – namely, the nominal head *pictures*, *as per* (59). Since the pronominal is *not syn-bound* in this domain, (56a) is fine. The governing category for the anaphor in (56b), on the

---

[16] For Huang, the accessibility requirement is viewed not as a notion solely invented for the purpose of defining a governing category, but as an expression of the fact that anaphors need to have antecedents. Since pronominals do not need antecedents, it does not make sense to require them to be free in a domain coinhabited by a potential antecedent (an accessible SUBJECT), as Chomsky does in his definition of governing category.

other hand, is predicted to be the whole S, because only S contains the anaphor, its governor and an *accessible* SUBJECT.[17] Given that the anaphor in (56b) *is* syn-bound in S (its local domain), (56b) also supports Huang's (1983) definitions.

Cases of non-complementarity across finite clause boundaries, like (60), fall under the same explanation (HUANG, 1983, pg. 554):

(60)   a.   *They* expected that books about *them* would be on sale.

        b.   *They* expected that books about *themselves* would be on sale.

The governing category for pronominal in (60a) is the NP, for the same reason as in (56a). On the other hand, the governing category for reflexive in (60b) is the matrix sentence, since only the matrix sentence contains an accessible SUBJECT. The nominal head *books* is not accessible to *themselves* because coindexing both would be a violation of the *i*-within-*i* condition: [$_{NP}$ books$_3$ about themselves$_3$]$_3$.

Huang's partitioning of binding domains thus accommodates into the CBT the lack of complementarity between pronominals and reflexives within NPs without having to resort to unmotivated covert structures. To be sure, this is done at the expense of introducing further intricacies into the system, some of which have a dubious flavor (e.g. the stipulation that nominal heads are SUBJECTs in (59)). Nonetheless, Huang's innovations were not unusual at the time and his account works reasonably well for the relatively simple cases mentioned so far. However, in the next section, I will discuss some problematic phenomena which jeopardize not only Huang's specific suggestions, but the CBT as a whole.

1.4   PROBLEMS WITH THE CLASSICAL BINDING THEORY

The CBT earned important payoffs from wagering on a complex rule system stated in terms of constructs like c-command, governing categories, SUBJECT and *i*-within-*i*. Given that these notions refer to very specific syntactic configurations, the Binding Conditions came to be used as probes into the fine-grained structure of sentences in diverse languages. Moreover, since each Binding Condition is, itself, intricately intertwined within a web of independent concepts

---

[17]   The nominal head, though a SUBJECT in the technical sense defined in (59), must not count as an *accessible* SUBJECT. For this to work, Huang (1983, pg. 558) has to stipulate that the index of a nominal head percolates from its maximal projection. In that case, the N head of NP is never accessible to an anaphor α contained within that NP, since coindexing α with N in this configuration would instantiate the *i*-within-*i* schema in (48); i.e. it would always yield something like [$_{NP}$ N$_i$ α_i]$_i$.

and assumptions, potential counter-examples can be handled simply by making relatively minor tweaks to specific definitions, as we saw in connection to Huang's (1983) proposal. These factors explain part of the popularity the CBT enjoyed throughout the 1980s.

However, little can be accomplished by minor tweaks if there is something fundamentally flawed about the basic assumptions of the CBT. I will argue that this is indeed the case. Most of the trouble stems from CBT's joint commitment to unification and universality: i.e. the idea that Conditions A and B apply indiscriminately to all pronouns in all constructions of every human language. We will see in this section that, both within a single language as well as cross-linguistically, there are instances of anaphors that need not be locally syn-bound and of pronominals that need not be locally free. Another related set of problems comes from the CBT's exclusive emphasis on phrase-structure, disregarding the role of semantics and discourse relations in determining admissible anaphoric patterns.

1.4.1 Problems with Condition A

The precise empirical consequences of Condition A vary among different incarnations of the CBT, depending on details about how local domains, syn-binding and c-command are defined (see Chomsky (1980), Higginbotham (1983), Lebeaux (1983) and Koster (1984) for versions of the CBT which differ from the one presented here). There are, however, a shared set of ideas which constitute the essence of Condition A, regardless of implementation. This common core consists of the following hypotheses (POLLARD; SAG, 1992, pg. 263):

(61)   a.   All anaphors behave like reflexives with respect to narrow syntax.

      b.   Anaphors are never coconstrued with NPs outside the domain defined by the nearest specified subject (in the sense of Chomsky (1973)).

      c.   Anaphors must be c-commanded by their antecedents.

      d.   Anaphors never have sentence-external antecedents.

None of these propositions stand in light of a wider survey of data. Assumption (61a) is contradicted by pronominal systems of languages other than English – even closely related ones, such as Dutch. The pronoun system in Dutch consists of pronominals (e.g. *hem*), reflexive anaphors (e.g. *zichzelf*, *hemzelf*) and simplex anaphors (e.g. *zich*, glossed as SE). The latter pattern with reflexives in some contexts (cf. (62a)) and with pronominals in others (cf. (62b)).

(62)   a.   *Max* wast   {*zichzelf / zich / *hem*}
            Max washes {REFL   / SE   / PRON}
            '*Max* washes *himself*.'

       b.   *Max* haat   {*zichzelf / *zich / *hem*}
            Max hates {REFL   / SE   / PRON}
            '*Max* shaves *himself*.'   (REULAND; REINHART, 1995, pg. 242)

A similar tripartite system is found Icelandic (MALING, 1984; ANDERSON, 1986; HYAMS; SIGURJÓNSDÓTTIR, 1990; MANZINI, 1992), where we see regular personal pronouns that function like pronominals (e.g. *hann*), complex anaphors which function like reflexives (*sjálfan sig*) and an ambivalent third element, the simplex anaphor *sig* (also glossed as SE). The simplex anaphor patterns with the reflexive in contexts like (63) and with pronouns in contexts like (64) (REULAND; SIGURJÓNSDÓTTIR, 1997, pg. 324):

(63)   *Jón* rakaði   {*sjálfan sig / sig / *hann*}
       John shaved {REFL   / SE / PRON}
       '*John* shaved *himself*.'

(64)   *Jón* segir   [að María elski   {**sjálfan sig / sig / hann*}].
       John say.PRES.IND that Mary love.PRES.SBJV {REFL   / SE / PRON}
       '*John* says that Mary loves *him*.'

Note that *zichzelf* and *zich* in Dutch and *sjálfan sig* and *sig* in Icelandic are all genuine anaphors. None of these forms can be used deictically to pick out a sentence external antecedent, unlike the personal pronouns *hem* and *hann*.

This ambivalent nature of *zich* and *sig* reveals that Chomsky's (1981) English-centered typology of NPs is mistaken. In particular, the assumption that all anaphors are subject to the same syntactic constraints as reflexives must be dropped: *zich* and *sig* are anaphors, but they don't always behave like reflexives. A distinction is needed between *reflexive* anaphors (e.g. *zichzelf*, *sjálfan sig*, *himself*) and *non*-reflexive anaphors (e.g. *zich* and *sig*) – as was already anticipated in Table 1 of Section 1.1. In other words, the predicate-based notion of a reflexive has syntactic significance beyond that which the CBT attributes to the notion of anaphor.

The failure to recognize the concept of reflexive as grammatically relevant leads the CBT to conflate, within Condition A, two requirements that are best kept separate: the requirement to be in a certain structural relation to an antecedent (e.g. c-command) and the requirement that this relation should be established locally (e.g. within a governing category).

The Icelandic anaphor *sig* is a typical representative of a LONG-DISTANCE ANAPHOR. This term refers to a cross-linguistically well-attested class of pronouns whose only obligation

is to have antecedent meeting certain structural conditions, with no strict syntactic locality requirement (see Koster and Reuland (1991) for an overview). The following examples from Chinese (POLLARD; XUE, 1998, pg. 289) and European Portuguese (MENUZZI; LOBO, 2016, pg. 345) succinctly corroborate the existence of this class of anaphors in other languages:

(65)   Zhangsan$_1$ zhidao [Li xiaojie$_2$ xihuan ziji$_{1/2/*3}$].
Zhangsan   know   Li   miss   like   SE
'Zhangsan$_1$ knows that Miss Li$_2$ likes herself$_2$/him$_{1/*3}$.'

(66)   Maria$_1$ soube do   João [que alguém$_2$ tinha falado   mal de si$_{1/2/*3}$].
Maria   knew   from.the João that   somebody had   spoken bad of SE
'Mary$_1$ heard from João that somebody$_2$ spoke badly of her$_{1/2/*3}$.'

The crucial structural licensing condition for *ziji* and *si* seems be the presence of a c-commanding antecedent (BRANCO; MARRAFA, 1999). Note that *ziji* and *si* can be locally bound in the domain of their specified subjects (indicated by index 2) as well as be non-locally bound by the matrix subject (indicated by index 1). Both of these forms qualify as anaphors because they cannot refer deictically to a third element (indicated by index 3).

All of these cases directly contradict another fundamental tenet of the CBT in addition to (61a): the locality assumption expressed in (61b). In fact, in the case of Icelandic, the distance between *sig* and its antecedent can be arbitrarily long – as long as none of the intervening clauses is in the indicative mood (IND), as in (67) illustrates (ANDERSON, 1986, pg. 66-67):

(67)   a.   *Jón*   segir [að María viti   [að Haraldur vilji   [að Billi
        John says that Mary knows.SBJV that Harold   wants.SBJV that Bill
        meiði   *sig*]]].
        hurts.SBJV SE
        '*John* says that Mary knows Harold wants Bill to hurt *him*.'

   b.   **Jón*   segir [að María viti   [að Haraldur vill   [að Billi meiði
        John says that Mary knows.SBJV that Harold   wants.IND that Bill   hurts.SBJV
        sig]]].
        SE

There are languages in which pronouns that meet the notional characterization of anaphors deviate even more radically from the syntactic distribution of English reflexives. The Japanese anaphor *zibun* – which is also referentially deficient (SAFIR, 2003) and is used in the same contexts as *himself* (cf. (68a)) – can not only find a binder outside of its subject domain (cf. (68b)), but can also c-command its local antecedent (cf. (68c)), in violation of (61c).

(68) a. *John*-wa *zibun*-o   hometa.
          John-TOP REFL-ACC praised
          '*John* praised *himself*.'   (KURODA, 1973, pg. 385)

      b. *Takasi*-wa Bill-ni [Yosiko-ga   *zibun*-o   nikundeiru koto]-o   hanasita.
          Takasi-TOP Bill-DAT [Yosiko-NOM REFL-ACC be-hating   COMP]-ACC told
          '*Takasi* told Bill that Yosiko hated *him*.'   (SELLS, 1987, pg. 453)

      c. *Zibun*-o   *Hanako*-ga   utagatte   iru.
          REFL-ACC Hanako-NOM suspicious be
          '*Hanako* doubts *herself*.'   (BÜRING, 2005, pg. 14)

Though the challenges typologically diverse languages pose to the CBT are severe, counterexamples to each of the assumptions in (61) can be found even in English. Regarding (61a), it has often been remarked that the adjunction of *own* to possessive pronominals creates a hybrid complex possessive that has many of the characteristics of an anaphor (LEES; KLIMA, 1963; SAXON, 1991; SAFIR, 1996). The impossibility of deictic use is illustrated in (69):

(69) Use {*his own / his} pencil!   [pointing gesture]

As Higginbotham (1985) shows, *own* adjunction can neutralize the disjointness restriction associated with a bare pronominal in a way that mimics the effect of using a reflexive:

(70) a. ??*John* is *his* enemy. (cf. **John* hates *him*.)

      b. *John* is *his own* enemy. (cf. *John* hates *himself*.)

In contexts like (71), *own*-adjunction triggers a locality requirement that also parallels the effect of using a reflexive anaphor: in (71), neither *their own* nor *themselves* can be bound across the intervening subject *people* (the judgment for (71a) reflects an unstressed use of *own*):

(71) a. **The realtors* are encouraging people to buy *their own* houses.

      b. **The realtors* are encouraging people to defend *themselves*.

However, the parallel between reflexives and and *own* possessives breaks down in contexts like (72) (ZRIBI-HERTZ, 1995, pg. 364). In these cases, *his own* behaves like the pronominal *his* because it is capable of finding non c-commanding (cf. (72a)) and sentence-external antecedents (cf. (72b)). The presence of *own* seems to have no effect for the binding possibilities.

(72) a. *Benjamin*'s sanity was compromised by {*his own* / *his*} desire for revenge.

      b. *Bobby* was sad. {*His own* / *His*} friends decided to leave the party.

Like Dutch *zich* or Norwegian *sig*, *own*-possessives are pronouns which qualify as an anaphors by independent criteria (cf. (69)), but only behave like a garden-variety English reflexives in *some* contexts, thereby contradicting the assumption in (61a).

Most important for my purposes, however, are assumptions (61b)-(61d) (POLLARD; SAG, 1992, pg. 263). The relevant English counterexamples to these were known at least since Partee (1965), Jackendoff (1969) and Cantrall (1969). Consider the following cases – two of which were already mentioned in (9) above:[18]
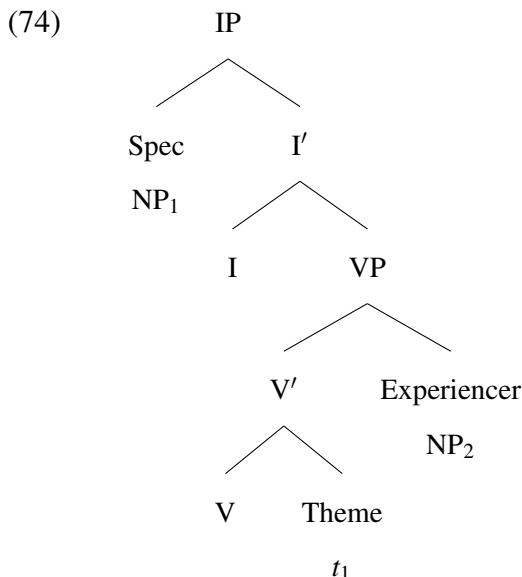
(73) a. *Albert* was never hostile to laymen who couldn't understand what physicists like *himself* were trying to prove.

   b. *John* believes that Mary would never marry a man less wealthy than *himself*.

   c. *Brad* warned Janet that she shouldn't trust anyone but *himself*.

   d. *Max* boasted that the queen invited Lucie and *himself* for a drink.

   e. That description of *himself* annoyed *DeGaulle* more than you know.

   f. Unflattering descriptions of *himself* have been banned by *LBJ*.

   g. The picture of *himself* in Newsweek shattered the peace of mind that *John* had spent the last six months trying to restore.

   h. *John*'s favorite topic of conversation is *himself*.

   i. *Joyce*'s defenses are well inside *herself*.

   j. This is a picture of myself.

   k. Physicists like myself were never too happy with the parity principle.

   l. This paper was written by Ann and myself.

Like the Chinese, Portuguese and Icelandic examples in (64)-(67), (73a)-(73d) show, *contra* (61b), that anaphors *can* be covalued to NPs across specified subjects. The c-command requirement (61c) is violated in (73e)-(73l). The assumption that anaphors never refer to individuals established extra-sententially in the discourse (61d) runs into trouble with the examples in (73j)-(73l). Note also that all of the anaphors in (73) could be replaced by pronominals without loss in acceptability. Insofar as the data in (62)-(73) contradict all of the fundamental

---

[18] The fact that some of these examples have been discussed since the 1960s shows that, more commonly than most would like to admit, proponents of the CBT had to sweep known data under the rug to make a unified theory seem plausible – a point made by Zribi-Hertz (1989, pg. 703) and Reuland (2011, pg. 44) alike. Examples (73a) and (73j)-(73l) come from Ross (1970, pg. 228-233), (73b)-(73c) from Safir (1992, pg. 3), (73d) from Reinhart and Reuland (1993, pg. 670), (73e) from Postal (1971, pg. 188), (73f) from Jackendoff (1969, pg. 40), (73g) from Pollard and Sag (1992, pg. 278), (73h) from Partee (1965, pg. 44), and (73i) from Zribi-Hertz (1989, pg. 718).

assumptions in (61), they present serious problems for the CBT framework in its entirety.

Pesetsky (1987, 1995) and Belletti and Rizzi (1988) attempt to save the CBT from some of these counterexamples by appealing to the ENRICH STRUCTURE strategy mentioned above. The idea is to set up an underlying representation for (73) where Condition A is fulfilled, despite appearances suggesting the contrary. Having cases like (73e) in mind, they posit the structure in (74), which is designed to reflect the fact that Experiencers can bind into Themes.

(74)

There is, in fact a grain of insight in these proposals which reveals an important property of the examples in (73). It seems to be true that entities that stand in an Experiencer-like relation to portions of the discourse are, in a sense, favored antecedents for reflexives which contradict the conditions in (61). This is clear in Pollard and Sag's (1992, pg. 277) contrast in (75). Note that *John* is an Experiencer in (75a), but not in (75b).

(75)   a.   The picture of *himself* in Newsweek bothered *John*.

   b.   *The picture of *himself* in Newsweek bothered *John*'s mother.

Since (74) places Experiencers higher than Themes (in D-Structure), *John* only c-commands the reflexive in (75a). In this way, the contrast in (75) falls out of the c-command requirement in Condition A, which is assumed to apply before the movement of the Theme to [Spec, IP].

However, it is clear that this cannot be a general solution to the problems posed by (73). Apparent violations of (61) do not gravitate around a class of verbs with specific thematic properties: examples in (73a), (73c) and (73f)-(73l) don't involve Experiencers (or psych-verbs) at all. Even if we assume an underlying structure like (74) for a broader range of examples, the

antecedent would still not c-command the reflexive in (73g)-(73i). Moreover, the problem of long-distance reflexives (cf. (73a)-(73d)) would remain untouched, so we would still need to rethink the assumption that reflexives must have local antecedents (i.e. (61b)).

What has generally been assumed is that the reflexives in (73) and (75) are conditioned by the *perspective* assumed in the discourse – i.e. they must refer to persons whose points of views are represented in the context in which they are used (KUNO; KABURAKI, 1977; POLLARD; SAG, 1992; OSHIMA, 2007; CHARNAVEL, 2019). This is why they are called, somewhat loosely, LOGOPHORIC REFLEXIVES (REINHART; REULAND, 1991, 1992, 1993), in analogy to the logophoric pronouns of African languages (CLEMENTS, 1975; HAGÈGE, 1974). Since the notion of perspective which licenses these reflexives is a global property of the discourse and not a narrow syntactic notion, I will call them DISCURSIVE REFLEXIVES. However, I also use the term LOGOPHORIC REFLEXIVES whenever the purpose is to highlight the perspectival aspect of the meaning of these expressions (i.e. their LOGOPHORIC READINGS).

Insofar as sentences with Experiencers are not the *only* ones which license discursive reflexives, proposing ingenious covert structures designed to make the odd binding properties of these expressions compatible with the CBT addresses, at best, only a subset of a larger problem. What really explains the contrast in (75) is the fact that Experiencers are good candidates for being bearers of a point of view wherever they appear.[19] It is simply much easier to think of John as a perspective-bearer for the discourse context in (75a) than in (75b). It is unlikely that a sentence like (75b) can be read as an expression of John's subjective viewpoint in the absence of any kind of specific contextual support indicating otherwise.

However, if a sentence structurally analogous to (75b) is embedded in a broader context where John's perspective is more plausibly being assumed, the result is not as bad:

(76)   John was crying. What else could he do now? These revelations might ruin his great

         career. And the ugly picture of *himself* in Newsweek would no doubt bother *his* mother.

---

[19] The correlation between point of view and particular θ-roles seems to be encoded as part of the meaning of certain verbs (MINKOFF, 1994; HINTERWIMMER, 2019). Propositional attitude verbs, for instance, carry the implication that their Agent/Experiencer argument is an entity whose point of view is represented in the context of their embedded clauses. Assuming that the reflexives in (i)-(ii) are logophoric, we predict the following contrasts (see Pollard and Sag (1992, pg. 272) and Kuno (1987, pg. 123) for similar examples):

(i)   a.   ?It was believed by *John* that there would soon be a picture of *himself* in the post-office wall.
        b.   *It was believed of *John* that there would soon be a picture of *himself* in the post-office wall.

(ii)   a.   ?Mary heard from *John* that physicists like *himself* were a godsend.
        b.   *?Mary said about *John* that physicists like *himself* were a godsend.

Likewise, if (75b) is switched to first person, as in (77), the result is also moderately acceptable, even in the absence of a specific context. The reason for this is that speakers count as plausible perspective-bearers in *every* context (SELLS, 1987; STIRLING, 1993).

(77)   The picture of *myself* in Newsweek bothered *my* mother.

In fact, any environment in which a third-person reflexive can have a non-local antecedent is also one where a first-person reflexive can naturally occur without an overt antecedent at all – a point first noted by Ross (1970) (see also Reinhart and Reuland (1992)). As the examples in (73) suggest, this works not only for picture NPs (cf. (73g) and (73j)), but also for comparatives (cf. (73a) and (73k)) and coordinated NPs (cf. (73d) and (73l)).

Another advantage of a discursive treatment of the reflexives in (73) over structural proposals based on (74) is the potential to cover paradigms like (78)-(79). These cases are analogous to (75) except for the fact that the antecedent is not an Experiencer in the sense defined by θ-theory. (In (79), the antecedent is also not in the same sentence as the anaphor.)

(78)   a.   Anonymous posts about *herself* on the internet hurt *Lucy*'s feelings.

         b.   *Anonymous posts about *itself* on the internet hurt *the camera*'s sales.

                                                                    (CHARNAVEL; ZLOGAR, 2015, pg. 5)

(79)   a.   *John* was going to get even with Mary. That picture of *himself* in the paper would really annoy her, as would the other stunts he had planned.

         b.   *Mary was quite taken aback by the publicity *John* was receiving. That picture of *himself* in the paper had really annoyed her, and there was not much she could do about it.   (POLLARD; SAG, 1992, pg. 274)

What explains the shifting judgments here is the fact that the antecedent's point-of-view is expressed in (78a)/(79a), but not in (78b)/(79b). In (78a), the antecedent is locally prominent in the sense of Hinterwimmer (2019) – i.e. it is a person whose feelings are explicitly being talked about in the sentence, in contrast to (78b), where it is an inanimate object which lacks a mental perspective altogether. In (79a), John's status as a point of view bearer is established by more global properties of the narrative (e.g. topicality) which are absent from (79b). The notion of a syntactically represented Experiencer is, therefore, too narrow to offer a generalization for all of the cases where a reflexive appears to acceptably violate Condition A.

The fact that the reflexives in (73)-(79) are interpreted relative to a point-of-view bearer

can be independently shown by comparing them to another class of perspective-sensitive expressions: possessive NPs such as the ones highlighted in (80) (KUNO; KABURAKI, 1977; SELLS, 1987; SPEAS; TENNY, 2003; CHARNAVEL, 2019).

(80) a. Carrie kissed **her dear husband**.

b. Carrie invited Julie to visit **her beloved cottage by the sea**.

Adjectives like *dear* and *beloved* are interpreted relative to a perspectival evaluation which can only be made by the individual who actually experiences the emotions they denote: in (80), only Carrie has direct access to the knowledge that her husband and her house are dear/beloved to her. Using these terms within third person possessive like the ones highlighted in (80) thus requires speakers to empathize with the emotional viewpoint of someone other than themselves (i.e. Carrie in both (80a) and (80b)). Possessives like *his/her dear* or *his/her beloved* can, therefore, be used as a test to identify whether a particular referent counts as a perspective bearer. The particular type of perspective bearer that can serve as an antecedent for these expressions is what Charnavel (2019, pg. 165) defines as an EMPATHY LOCUS:

> I define an empathy locus as an event participant that the speaker identifies with from a sensory or emotional perspective. [. . .] Under empathic conditions, the speaker puts herself in the empathy locus's shoes to report his first-personal perception: instead of reporting beliefs or speech, as in the case of attitude reports, (s)he reports an experience by expressing what the empathy locus could say if he had to reflect on and formulate his feelings at the first-person.

Whether a discursive reflexive can refer to a particular entity seems to hinge on whether this entity can be construed as an empathy locus in Charnavel's sense. We can see this by noting that there is significant overlap between the contexts in which a discursive reflexive can be anteceded by a particular NP and the contexts in which *dear* or *beloved* can be interpreted relative to this NP (CHARNAVEL; ZLOGAR, 2015, pg. 10):

(81) a. The picture of *his* dear father in Newsweek bothered *John*.

b. ??The picture of *his* dear father in Newsweek bothered *John*'s mother.

(82) a. Anonymous posts about *his* beloved daughter on the internet hurt *John*'s feelings.

b. ??Anonymous posts about *its* beloved lens on the internet hurt *the camera*'s sales.

The parallels between (75)/(78) and (81)/(82) suggest that discursive reflexives need to refer to empathy loci. This confirms their logophoric nature. However, referring to an empathy locus is not sufficient to license a discursive reflexive – it is merely a necessary condition. Even

if we assume that empathy is kept constant in the contrasts in (83)-(85), only the Condition A violations in the a-cases are actually acceptable (REINHART; REULAND, 1991, pg. 289).

(83) a. *Joyce*'s defenses are well inside *herself*.

  b. **Joyce*'s defenses failed *herself*.

(84) a. *John* believes that Mary would never marry a man less wealthy than *himself*.

  b. **John* believes that Mary would never marry *himself*.

(85) a. *Max* boasted that the queen invited Lucie and *himself* for a drink.

  b. **Max* boasted that the queen invited *himself* for a drink.

The possessive+*beloved* diagnostic shows that antecedents in (83b)/(84b)/(85b) do indeed count as possible empathy loci:

(86) a. *Joyce*'s defenses failed *her* beloved family.

  b. *John* believes that Mary would never marry *his* beloved brother.

  c. *Max* boasted that the queen invited *his* beloved wife for a drink.

The problem with (83b)/(84b)/(85b) is, therefore, not the absence of a potential logophoric antecedent in the discourse. The contrasts in (83)-(85) suggest that English reflexives can only function logophorically in specific linguistic contexts. For instance, whatever triggers the logophoric reading in (85a) has to be ultimately related to the presence of the conjunction – as witnessed by the unacceptability of (85b). As Reuland (2017, pg. 21) notes, this observation undermines proposals which regard the duality between CBT-abiding and CBT-violating reflexives as a mere instance of homonymy (e.g. Baker (1995) and Hornstein (2001)).

The particular environments which allow acceptable discursive reflexives are strikingly parallel to those in which the complementarity between pronominals and reflexives breaks down. As I noted above, all of the reflexives in (73) can replaced by pronominals without loss in acceptability. Some examples are given in (87) (cf. (73e), (73a), (73c) and (73l)).

(87) a. That description of {*himself* / *him*} annoyed *DeGaulle* more than you know.

  b. *Albert* was never hostile to laymen who couldn't understand what physicists like {*himself* / *him*} were trying to prove.

  c. *Brad* warned Janet that she shouldn't trust anyone but {*himself* / *him*}.

  d. This paper was written by Ann and {*myself* / *me*}.

The overlap between non-complementary and discursive contexts raises the question of whether cases of coconstrual which fall under the standard CBT, such as (27) (repeated below as (88)) should be thought of as instances of discursive reflexives as well. If this is so, the difference between the example below and the typical discursive occurrences in (87) becomes inessential: like any discursive occurrence, the reflexive in (73) refers to a perspectival empathy locus. However, in that case, the empathy locus just happens to be invoked by an NP that locally c-commands the reflexive.

(88)   *Bobby* saw many pictures of {*himself* / *him*}.

Recall that, for CBT and Lees and Klima (1963) alike, reflexives in non-complementary environments like (88) are essentially the same kinds of beasts those in standard complementary contexts like (11), repeated below as (89). Both are subject to the same restrictions.

(89)   *Laurie* admires {*herself* / **her*}.

This is exactly what makes these theories *unified*: according to them, the grammar does not include more than one rule for generating/interpreting reflexives. The same principle that is responsible for (89) should be responsible for (88) as well. It is this assumption which, upon confrontation with data like (88), led Lees and Klima (1963) and Huang (1983) to the ENRICH STRUCTURE and ENRICH PRINCIPLE strategies characterized in Section 1.2.

As we have seen, both of these strategies are insufficient to handle the discursive reflexives gathered in (87).[20] If, on the other hand, (88) is seen simply as a discursive reflexive (on a par with (87)), the *ad hoc* tweaks Lees and Klima (1963) and Huang (1983) had to make in order to account for (88) can be avoided. The condition governing reflexivization and the structures of sentences containing reflexives can be simplified accordingly. The former can be restricted to local coargument contexts like (89) and *all cases* of non-complementarity can be delegated to a separate condition – one that is sensitive to the discourse factors which influence discursive reflexives in general. This is the correct insight grounding the TRH.

Since some sort of extra-grammatical discursive account is independently required for (87), extending it to mildly problematic cases like (88) is essentially costless. Aside from

---

[20]   In fact, as we saw above, Lees and Klima's (1963) unalloyed application of the ENRICH STRUCTURE strategy failed to offer a consistent account for (88) as well. However, even if we accept their tentative solution (which consists in arbitrarily enriching the base component of their grammar to yield kernels like *John saw many pictures of John*), there is no way they could account for typical discursive reflexives like the ones in (87).

simplifying the overall theory, this alternative also has empirical virtues. I will show in Chapter 2 that reflexives in structures like (88) are in fact much more similar to the discursive reflexives in (87) than to the coargument reflexives in (89) with respect to crucial properties.[21]

1.4.2 Problems with Condition B

In all versions of the CBT, Conditions A and B exhibit a certain symmetry with respect to each other insofar as the latter denies a property attributed by the former. In practice, this parallelism entails a bias towards the view that anaphors and pronominals are in complementary distribution. What is appealing about complementarity is that it accords the possibility of deriving the distribution of one type of pronoun from the distribution of the other.

As we have seen, this does quite not work in Huang's (1983) version of the CBT because the notions of domain for anaphors and pronominals are not the same. But the basic complementarity picture is partially retained in his system because the contexts where anaphor and pronominal domains overlap are signaled out as exceptional: they are restricted to environments where pronouns are governed within the domain of a subjectless NP.

There are many empirical problems with the complementarity assumption which is built into the CBT. One important set of problems comes from the instances of locally free anaphors we saw in Section 1.4.1. Now we will turn to some of cases which question the complementarity from the Condition B side of the story: namely, instances of locally syn-bound pronominals.

The first class of counterexamples to Condition B involves focus constructions like (90a) and peculiar discourse contexts like (90b)-(90c) (adapted from Safir (2004)):

---

[21] Seeing the reflexive in (88) as an instance of discursive anaphora involves saying that it is not licensed by strictly syntactic conditions – in particular, that it does not abide to locality (cf. (61b) above). Some authors claim that this conclusion is too extreme, at least for *picture* NP reflexives (though perhaps not for discursive reflexives in general) (SAFIR, 1992; DALRYMPLE, 1993). The fact that *himself* cannot refer to *Bill* in (i), skipping the intervening subject *Tom*, seems to suggest *some* form of purely syntactic locality constraint.

(i)   *?*Bill* remembered that Tom saw a picture of *himself* in the post office.

However, syntactic locality cannot be the issue here because, as we have already seen, antecedents for discursive reflexives can be as far away as in a different sentence (cf. (79a)), and, in the case of first person discursive reflexives, they can be lacking altogether (cf (73j)). Moreover, if the intervening subject is inanimate (cf. (iia)), quantified (cf. (iib)) or expletive (cf. (iic)), the intervention effect illustrated in (i) largely disappears. The examples below are from Pollard and Sag (1992, pg. 272) (see Büring (2005, pg. 226-227) similar data).

(ii)   a.   ?*Bill* remembered that the Times had printed a picture of *himself* in the Sunday edition.
         b.   *Bill* thought that nothing could make a picture of *himself* in the Times acceptable to Sandy.
         c.   *Bill* suspected that there would soon be a picture of *himself* on the post office wall.

The intervention effect is too gradient and contextually defeasible to be imputable to grammar. For this reason, it seems more plausible to view it as a byproduct of processing or other contextual factors, along the lines proposed in Pollard and Sag (1992, pg. 273).

(90) a. Everybody hates Donald. At the end of the day, only *Donald* voted for *him*.

b. I know what Biden and Warren have in common. Warren voted for Biden and *Biden* voted for *him* too.

c. If everyone voted for Biden, then certainly *Biden* must have voted for *him*.

What these examples have in common is that the relevant coconstrual relation that holds between the pronominal and its antecedent is not A-binding, but coreference. The property that (90a) says only Donald possesses is the property of voting for Donald ('$\lambda x.\ x$ voted for Donald') and not the property of voting for oneself ('$\lambda x.\ x$ voted for x'), which would be the one obtaining under a bound-variable reading of *him*. Since (90a) is interpreted under the coreference reading, it is compatible with a scenario where other candidates cast votes on themselves (e.g. where Biden voted for Biden). Similarly, it is the property '$\lambda x.\ x$ voted for Biden' which is claimed to be shared by Warren and Biden in (90b) and by everyone in (90c).

Since the CBT is concerned with the relation of syn-binding (which does not distinguish between coreference and A-binding) Condition B incorrectly rules out local coreference in all of the cases above. Reinhart's (1983) solution to this problem, as we will see in Chapter 2, involves restricting the Binding Conditions so that they apply only to A-binding configurations. On this picture, Condition B would have nothing to say about coreference cases like (90). However, this idea is not easy to implement within the CBT because the Binding Conditions are stated as constraints on coindexing, which is used ambiguously to signal both A-binding and coreference.

The second class of counterexamples to Condition B emerges in connection to locative PPs in English, whose problematic status was pointed out above in connection to Lees and Klima's (1963) examples in (26). These examples are repeated below along with further data adapted from Chomsky (1981, pg. 290) and Hestvik (1991, pg. 464).

(91) a. *The men* threw a smokescreen around *them*.

b. *John* drew the book towards *him*.

c. *Rose* pulled the blanked over *her*.

(92) a. *The men* found a smokescreen around *them*.

b. *Curly* saw a snake behind *him*.

c. *Billy* found a dollar bill in front of *him*.

Contrary to Lees and Klima's ((1963) suggestion, it does not make any difference whether the PP from which the pronominal is syn-bound is an argument (cf. (91)) or an adjunct (cf.

(92)). Moreover, the syn-binding relationships depicted in (91)-(92) count as *local* according to the definitions of governing category in both Chomsky (1981) and Huang (1983):

(93) γ is a GOVERNING CATEGORY for α if and only if γ is the minimal category containing α, a governor of α, and a SUBJECT accessible to α, where the SUBJECT of γ is the subject of γ (in the usual sense) or the AGR features on γ.

(94) γ is a GOVERNING CATEGORY for α if and only if γ is the minimal category containing α, a governor of α, and a SUBJECT that, if α is an anaphor, is accessible to α, where the SUBJECT of γ is the subject of γ (in the usual sense) or the nominal head of γ.

According to Chomsky's definition in (93), the governing category in (91)-(92) is the whole S, because S is the minimal category containing the pronominals, their governors (the locative Ps) and an accessible SUBJECT. Huang's definition in (94) yields the same result because P heads also do not count as SUBJECTs in his system. Given that the pronominals in (91)-(92) are syn-bound within S, all of these cases violate Condition B.[22]

Moreover, unlike the cases in (90), the form of coconstrual implied by these examples is perfectly compatible A-binding, as evidenced by the possibility of quantified antecedents and the licensing of sloppy readings in ellipsis contexts:

(95)   a.   *Every actress* pulled a blanket over *her* in order to get dressed.

     b.   *Every man* who saw a snake behind *him* on the plane is advised to press charges against the airline company.

(96)   a.   *Bobby* pushed the book away from *him* and so did *Larry*.

        ... Larry ($\lambda x.\ x$ pushed the book away from x)   (sloppy reading)

     b.   *Billy* found a dollar bill in front of *him* and so did *Enoch*.

        ... Enoch ($\lambda x.\ x$ found a dollar bill beside x)   (sloppy reading)
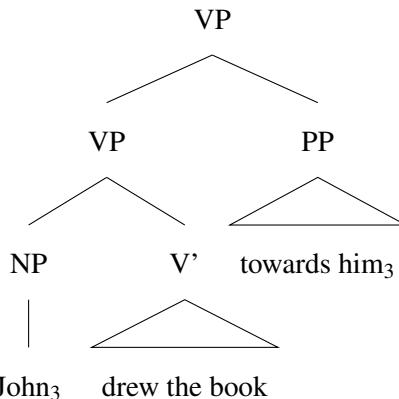
These examples demonstrate that simply making Condition B deal only with A-binding is not sufficient to explain the PP facts. The actual content of the Condition B needs to be changed.

It has been occasionally suggested in the literature that this conclusion may be avoided by meddling with the underlying syntactic structures ascribed to (91)-(92). This is an instance of ENRICH STRUCTURE strategy mentioned in Section 1.2. One idea is to exploit the c-command

---
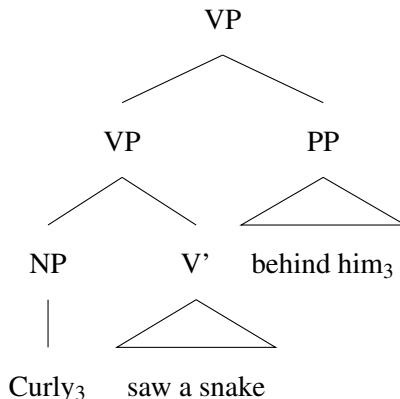
[22] Non-locative PPs behave in accordance to Condition B (MENUZZI, 1999, pg. 125): (e.g. ***Bobby** *relies on* **him**, ***Bobby** *talks to* **him** *all the time*, ***Bobby** *is proud of* **him**). This is also a problem for the CBT, insofar as the locative vs. non-locative distinction can hardly be cashed out in purely structural terms.

condition, requiring locative PPs to be adjoined in a position higher than the position occupied by the subject NP (cf. (JACKENDOFF, 1990a, pg. 453)):

(97)   a.

```
                    VP
                   /  \
                 VP     PP
                /  \    / \
              NP    V'  towards him₃
              |    / \
           John₃  drew the book
```

b.

```
                    VP
                   /  \
                 VP     PP
                /  \    / \
              NP    V'  behind him₃
              |    / \
          Curly₃ saw a snake
```

Due to the different height of attachment of locative PPs in (97) it turns out that the pronominals are not c-commanded (and, hence, not syn-bound) by their antecedents. This avoids a Condition B violation. Note, however, that the structures in (97) embody an unusual assumption: namely, that locative PPs are Chomsky-adjoined to their host VPs regardless of whether they themselves are arguments (cf. (97-a)) or adjuncts (cf. (97-b)).[23]

This generalized Chomsky-adjunction structure resuscitates the Raising to Subject ordering paradox that emerged in connection to Lees and Klima's transformational theory in Section 1.2. Note that (97) presupposes the VP-INTERNAL SUBJECT HYPOTHESIS of Koopman and Sportiche (1991). If external arguments have to move to [Spec, IP] to check nominative case and satisfy the EPP, the well-formedness of (97) implies that Conditions B can be checked *before* such movement takes place. This is clearly contradicted by (98), which was among the main motivations for the view that Binding Conditions apply at LF:

(98)   *Bobby* seems to *him* to be a genius.

A more viable alternative is to claim that the PPs (91)-(92) are part of a larger small clause (SC) structure whose subject is the internal argument. This encodes the intuition that there is a predication relation holding between the PP and the accusative object of the verb.

(99)   a.   John₂ drew [SC the book₁ towards him₂].

       b.   Curly₂ saw [SC a snake₁ behind him₂].

---

[23] The restricted use of Chomsky-adjunction structures to represent adjuncts is, in itself, far from a consensus – arguments against it are found in Larson (1988, 1990) and in Culicover and Jackendoff (2005, chap. 4). The idea of generalizing Chomsky-adjunction to argument PPs is arguably even more problematic.

The SCs in (99) count as a governing category according to (93) and (94) because they contain a pronominal, its governor (P) and a SUBJECT (*the book* and *a snake*). Since the pronominals are free within these SCs, no Condition B violation ensues. The SC account also makes a surprising prediction: when the pronominal is syn-bound by the object of the verb, the result should be bad. This is in fact correct (REINHART; REULAND, 1993, pg. 687):

(100)   *$Bobby_2$ rolled [$_{SC}$ the $carpet_1$ over it_1].

Though *prima facie* appealing, this alternative is rejected by Chomsky (1981, pg. 291), Koster (1987, pg. 334), Hestvik (1991, pg. 472) and Reinhart and Reuland (1993, pg. 688). One problem with it is that the prediction illustrated by (100) is not always borne out:

(101) $Bobby_2$ examined the $carpet_1$ underneath it_1.

If *the carpet* and the PP form an SC governing category in (101), the example would be inadequately ruled out as violation of Condition B.

The SC analysis also seems inadequate to handle (102). The PPs in these examples seem to be semantically inert in that they can be omitted without much difference to the interpretation. Note that the prepositional objects therein can neither be replaced by non-pronominal NPs nor be questioned, as in (103) (JACKENDOFF, 1990b, pg. 60; WILKINS, 1988, pg. 202):

(102)   a.   *Bobby* likes to take work home with {*him* / *Harry}.

          b.   *This melody* has a haunting character to {*it* / *the song}.

          c.   *The list* includes my name on {*it* / *the page}.

(103)   a.   *Who does Bobby like to take work home with?

          b.   *What does this melody have a haunting character to?

          c.   *What does the list include my name on?

Since the PPs in (102) are semantically inert, they arguably do not establish a predication relation with the accusative objects to their left (CHOMSKY, 1981, pg. 290).

Furthermore, in virtue of the complementarity assumption deeply ingrained in the CBT, any attempt to enrich the structures ascribed to (91)-(92) in order to make the pronominals therein come out as locally free immediately causes a problem for Condition A, because anaphors are acceptable in all of these environments ((105) is often reported as slightly marked):

(104)   a.   *The men* threw a smokescreen around *themselves*.

   b.   *John* drew the book towards *himself*.

   c.   *Rose* pulled the blanked over *herself*.

(105)   a.   *The men* found a smokescreen around *themselves*.

   b.   *Curly* saw a snake behind *himself*.

   c.   *Billy* found a dollar bill in front of *himself*.

This problem affects the SC proposal (which makes the pronouns in (104)-(105) not local with respect to the subjects) as well as the Chomsky-adjunction analysis in (97) (which makes the pronouns in (104)-(105) not syn-bound by the subjects).

If we look beyond English, even to other Germanic languages, we find counterexamples to Condition B that simply cannot be avoided by enriching underlying structures or by redefining binding domains. Frisian, for instance, has a binary system of pronouns consisting of reflexive anaphors like *himsels* and pronominals like *him*. Like English, these forms are in complementary distribution in contexts like (106), where *himsels* must be locally syn-bound and *him* must be used as a deictic (REULAND; REINHART, 1995, pg. 258):

(106)   Willem$_1$ bewûnderet {himsels$_{1/*2}$ / him$_{*1/2}$}.
   Willem admires   {REFL   / PRON}
   'Willem admires {himself$_{1/*2}$/ him$_{*1/2}$}.'

Also like English, *himsels* and *him* can both be coconstrued with the subject of S from inside a PP contained within S, as (107) illustrates (REULAND, 2011, pg. 269):

(107)   *Klaas* treau   de karre foar   {*him*   / *himsels*} út.
   Klaas pushed the cart   before {PRON / REFL}   out
   '*Klaas* pushed out the car in front of {*him* / *himself*}.'

However, unlike what Condition B would lead us to expect on the basis of English, *him* can also be locally syn-bound in transitive structures involving grooming predicates (e.g. *wash*, *bathe*, *shave*) and raising-to-object perception verbs (e.g. *feel*, *see*, *sense*):

(108)   a.   Jack$_1$ wasket him$_{1/2}$.
      Jack   washed PRON
      'Jack$_1$ washed him(self)$_{1/2}$.'

   b.   Jack$_1$ fielde [him$_{1/2}$ fuortglieden].
      Jack   felt   PRON   slip-away
      'Jack$_1$ felt him(self)$_{1/2}$ slip away.'

An even clearer violation comes from Traditional Jambi Malay, which contains an ambivalent pronoun form *dio*. This pronoun is preferably associated with a long-distance and deictic interpretation, which indicates that it should be classified as a pronominal. However, in contradiction to what Condition B predicts, *dio* can also be locally syn-bound in simple transitive structures, as (109) illustrates (COLE; HERMON; YANTI, 2015, pg. 148):

(109) a. Yanti$_1$ neŋoʔ dio$_{1/2}$ di tipi.
           Yanti saw PRON on TV
           'Yanti$_1$ saw her(self)$_{1/2}$ on TV.'

       b. Yanti$_1$ pikir [Arna$_2$ cinto dio$_{1/2/3}$].
           Yanti thinks Arna loves PRON
           'Yanti$_1$ thinks Arna$_2$ loves her(self)$_{1/2/3}$.'

Another interesting class of counterexamples to Condition B comes from 1st and 2nd person clitics in Romance (BURZIO, 1991; SAFIR, 2004). As the French and Brazilian Portuguese examples in (110) and (111) show, these forms (glossed as CL) can be used without an overt antecedent, with a long-distance antecedent and with a local antecedent:

(110) a. Jean me   parle.
           Jean CL.1SG speak.
           'Jean speaks to me.'

       b. *Je* pense que Jean *me*   parle tous les jours.
           I   think that Jean CL.1SG speak all   the days
           '*I* think Jean speaks to *me* every day.'

       c. *Je me*   parle.
           I   CL.1SG speak.
           '*I* speak to *myself*.'

(111) a. O David te   defendeu.
           the David CL.2SG defended
           'David defended you.'

       b. *Tu* disse que o David *te*   defendeu.
           you said that the David CL.2SG defended
           '*You* said that David defended *you*.'

       c. *Tu te*   defendeu.
           You CL.2SG defended
           '*You* defended *yourself*.'

Unless 1st and 2nd person Romance clitics are declared, by stipulation, to be ambiguous between a reflexive and a pronominal uses (in which case, as Burzio (1991) notes, the CBT

would become unfalsifiable), the examples in (110c) and (111c) are unexpected from the point of view of Condition B. Like the examples from Frisian and Traditional Jambi Malay mentioned earlier, French *me* and Brazilian Portuguese *te* are, to all appearances, pronominals that are capable of being syn-bound within their governing categories. In the following chapters, we will examine other instances of locally bound pronominals in other languages, including Brazilian Portuguese and Middle English.

1.5 CONCLUSION

In this chapter, I presented an overview of two prominent unified approaches to anaphora within Mainstream Generative Grammar (MGG): the transformational theory of Lees and Klima (1963) and the Classical Binding Theory (CBT) of Chomsky (1981, 1986b). Each of these is representative of a particular strategy in handling potential objections to unified principles: Lees and Klima's theory opts for enriching structure and the CBT almost always opts for making the principles responsible for the interpretation of reflexives more complex.

Neither of these two strategies offers a general way to save a unified theory of anaphora from data involving discursive reflexives and locally syn-bound pronominals: one cannot "get around" the counterexamples by positing covert structures or by enriching the rule system.

Examples like (73) suggest that (some) anaphors are not governed by Condition A of the CBT at all. Anaphors in these contexts are licensed according to whether they refer to an empathy locus in Charnavel's (2019) sense, rather than by solely syntactic conditions. This undermines the syntactocentric outlook of Lees and Klima (1963) and the CBT. The existence of non-reflexive anaphors like Dutch *zich* (cf. (62)), and Icelandic *sig* (cf. (63)–(64)) shows, moreover, that the notion of a *reflexive* (i.e. an NP that can persistently function as a marker of coconstrual between coarguments) is syntactically relevant.

Examples like (90a)–(90c) show that local syn-binding of pronominals in English is possible as long as it does not correspond to an A-binding in in semantics. However, even local A-binding is allowed when pronominals are placed within PPs, as the examples in (91)–(92) show. In other languages, such as Traditional Jambi Malay (cf. (109)) and Frisian (cf. (108)), locally A-bound pronominals can appear even as arguments of simple transitive clauses.

The considerations above call for a revision of the core assumptions of the CBT. This is exactly what Reinhart and Reuland (1993) set out to do in their reflexivity-and-chains theory (RCT), which will be examined in the next chapter.

2 THE REFLEXIVITY-AND-CHAINS THEORY OF ANAPHORA

In Chapter 1, I argued that unified theories of anaphora within MGG suffered from several major shortcomings. One is their inadequate typology of pronouns which fails to recognize the existence of non-reflexive anaphors (like *zich* in Dutch). Another is their inability to explain the behavior of discursive reflexives in languages like English. Yet another is their failure to account for the behavior of locally syn-bound pronominals in various languages.

An additional issue, which was only barely mentioned in Chapter 1, but which is discussed in detail below, concerns MGG's ambivalence concerning the semantic interpretation of pronouns. The CBT, for example, uses the same device – syntactic coindexing – to represent the two modes of coconstrual defined in Chapter 1: variable binding and coreference. Though this practice is, to a large extent, unexamined, it rests on the underlying assumption that grammars regulate both of these interpretive dependencies by means of the same principles (i.e. the Binding Conditions). I argue below that we have solid empirical and theoretical reasons for abandoning this view, which leads to an unnecessary proliferation of covert structure.

No amount of conceptual contortion is likely to help MGG address these issues within the frameworks of Chapter 1. In light of all of these problems, it seems necessary to rebuild the theory of anaphora on entirely different grounds – as many recognized after Chomsky's (1986b) last statement of the CBT (see Pica (1987), Wilkins (1988), Rizzi (1990), Jackendoff (1992), Pollard and Sag (1992), Dalrymple (1993), Cole and Sung (1994), Burzio (1998), i.a.).

The reflexivity-and-chains theory (RCT) of Reinhart and Reuland (1991, 1992, 1993, 1995) is an attempt to do just that. In this chapter, I argue that the RCT was, for the most part, successful in this endeavor. First, it is rests upon a much more solid understanding of how grammar specifies the interpretation of pronouns: following Reinhart (1983), the RCT abandons CBT's ambiguous use of indices and states its grammatical principles so that they determine exclusively A-binding (coreference is analyzed in pragmatic terms). Second, the RCT is more parsimonious: it eliminates the Binding Theory as a module of grammar, reducing its effects to properties of independent subsystems. Third, the RCT is more empirically adequate: e.g. it recognizes the influence of semantic factors on binding phenomena as well as existence of non-anaphoric reflexives, giving rise to interesting typological predictions.

However, though certainly an improvement over the earlier theories, the RCT is responsible for perpetuating old mistakes as well as creating new problems of its own. These will be

largely ignored throughout this chapter. My critical appraisal of the RCT and similar approaches is concentrated in Chapter 3. Specific problems that arise when applying RCT's account of locally bound pronominals to Brazilian Portuguese are discussed in depth in Chapter 7.

The RCT, in its original formulation, is a theory of the early 1990s. As such, it does not hesitate to attribute a rich system of principles to Universal Grammar (UG). Since most of the discussion in this chapter is couched in these terms, some of it might look outdated from a contemporary viewpoint. But one of the virtues of the early RCT, as opposed to some of its more modern renditions (REULAND, 2011), is that it has explicit definitions and a clear deductive structure. Since this design is convenient for expository purposes, I will frame my discussion around the original RCT, pointing out some conceptual problems along the way.

What I will ultimately suggest in Chapters 5-7 is that the allegedly universal principles invoked by the RCT can be rephrased in constructional terms within the theory of Simpler Syntax (CULICOVER; JACKENDOFF, 2005). The leading insights of the RCT will be preserved in their essentials, but their conceptual foundations will be drastically modified. RCT's more surface-oriented outlook and its explicit recognition of an autonomous level of semantic representation will be particularly convenient in this connection, since they open the possibility of simplifying the structures that a theory of anaphora requires from syntax. The insight that constraints on anaphors and pronominals are sensitive to properties of predicate-argument structure will also be instrumental to my approach.

In Section 2.1, I take an excursus into two conceptions of how variable binding can be encoded by grammars. The theory I will argue for – Reinhart's (1983) theory – says that variable binding is the only kind of anaphoric dependency which is syntactically encoded. But this, by itself, does not say which kinds of pronouns are capable of encoding variable binding under what syntactic conditions. In Section 2.2 I outline a purported answer to this question: Reflexivity Theory and Chain Theory, which conjointly constitute the RCT.

2.1   VARIABLE BINDING IN GRAMMAR

In this section, I will survey two conceptions of the relationship between variable binding and grammar: the theory which underlies the CBT and Reinhart's (1983) theory, which is presupposed by the RCT. The CBT presumes that coreference and bound variable interpretations are grammatically constrained in terms of the same principles. Reinhart's (1983) theory, in turn, proposes that the only kind of interpretation grammar specifies for pronouns are bound variable

readings; coreference, in the sense defined in Chapter 1, is taken to be a pragmatic process. I will argue that Reinhart' theory, is, in fact, the best one. Investigating this involves asking how different modes of coconstrual (i.e. semantic identity) are represented in grammar.

Answers to this *semantic* issue are largely independent of the statement of particular *syntactic* constraints on pronouns of different types. The latter, which address the *domains* and *configurations* in which anaphors and pronominals can or must be bound, are the subject of the RCT *per se*, as we will see in Section 2.2.

2.1.1 A surface-based theory of variable binding

Interpretive theories of anaphora purport to offer an account of the principles regulating the coconstrual possibilities of NPs. The CBT does this in a rather roundabout way, by framing its basic statements as constraints on the distribution of syntactic objects called *indices*.[1] In the theory outlined in Chomsky (1981), each NP is generated with a freely assigned index, which may be identical to the index of another NP, yielding coindexation, as in (1). Some of the resulting coindexed structures are grammatical, such as (1a) and (1b), and some are filtered out by the Binding Conditions, such as (1c)-(1d).

(1)   a. $Curly_3$ admires $himself_3$.

        b. $Curly_1$ said that Laurie kissed him_1.

        c.   *$Curly_7$ said that Laurie kissed $himself_7$.

        d.   *$Curly_2$ admires him_2.

Coindexing is assumed to correspond, in a systematic way, to some *semantic* relation between pronouns and their antecedents. It is only in virtue of this presumed correspondence that the CBT makes concrete predictions about the coconstrual possibilities of NPs. The mapping procedure between the syntactic indexing and semantic interpretation is often left implicit, but the standard view is that coindexation is ambiguous between a coreference and a bound-variable interpretation. The latter is mandatory for (2), where the antecedent is a quantified NP.

---

[1]   The exposition in Chapter 1 obscures this fact somewhat, because, following usual practice, Conditions A and B are stated there in terms of the concept of syn(tactic)-binding. However, since the latter is defined as "coindexation under c-command", the following formulation is equivalent to the one presented earlier:

(i)   a.   CONDITION A: An anaphor is coindexed to a c-commanding NP in a local domain.

        b.   CONDITION B: A pronominal is not coindexed to a c-commanding NP in a local domain.

The formulation in (i) makes it clear that the CBT is, in fact, a theory regulating the occurrence of indices.

(2)   Every cowboy$_8$ thinks that people admire him$_8$.

Every cowboy ($\lambda x.\ x$ thinks that people admire x)

This ambiguity does not present a serious threat for compositional interpretation because, in this system, the semantics of the antecedent can fully determine whether coindexing will be interpreted as coreference or A-binding. If a pronoun is coindexed to a referential antecedent, coindexation is interpreted as coreference; if the antecedent is a quantified NP, a bound-variable reading follows (see Büring (2005, chap. 4) for a proposal along these lines).

Note that this view is *prima facie* falsified by data in Section 1.1 which suggest that bound-variable readings are *not* restricted to quantification environments. To take one example, the fact that (1b) has a sloppy reading in ellipsis contexts like (3) shows that the pronoun therein, despite the referential nature of its antecedent, can also be interpreted as a bound-variable:

(3)   *Curly* said that Laurie kissed *him*, and so did Jud.

Curly ($\lambda x.\ x$ said (Laurie kissed x)) & Jud ($\lambda x.\ x$ said (Laurie kissed x))   (sloppy)
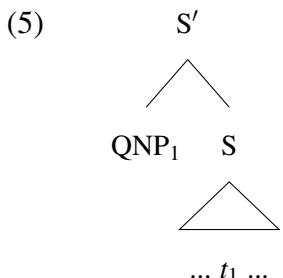
However, there are several ways these kinds of facts can be accommodated within CBT's ambiguous indexing system. The standard solution involves saying that referential NPs are optionally interpreted as quantifiers. This is made possible by the fact that any expression whose semantic type is e can be type-lifted into an expression of type $\langle\langle e,t\rangle,t\rangle$ (i.e., a generalized quantifier, in the sense of Montague (1974)). This particular type-shifting principle can be stated as follows (where ε is a meta-variable over e-typed meanings) (PARTEE, 1986, pg. 362):

(4) $\varepsilon \rightarrow \lambda P_{\langle e,t\rangle}.P(\varepsilon)$

So, for example, the antecedent of the pronoun in (3), instead of denoting the individual named Curly, can be type-lifted to denote the set of properties that Curly has (i.e. $\lambda P.P(\text{curly})$). This quantifier meaning is what triggers the bound-variable interpretation required to generate the sloppy reading. What the first conjunct of (3) says, under this reading, is that the property expressed by the predicate (i.e. '$\lambda x.\ x$ said (Laurie kissed x)') belongs to the set of properties that Curly has – i.e. to the denotation of the quantifier.

Semantically-savvy advocates of the CBT also argue that, in order to be properly interpreted by standard structure-driven compositionality, quantified NPs – regardless whether they are "primitive" (e.g. *every cowboy*) or generated via type-lifting – undergo QUANTIFIER RAISING at the level of Logical Form (LF) (HEIM; KRATZER, 1998, chap. 7)

Quantifier Raising is a type of covert A′-movement which adjoins NPs that are interpreted as quantifiers (QNPs) to higher S nodes at LF (MAY, 1985). The S node which the QNP c-commands is interpreted as its *logical scope*. The resulting structure is always the following:

(5)

$$S'$$

QNP$_1$ S

... t_1 ...

By making coreference/bound-variable readings dependent upon the structural position of antecedents at LF, the CBT can provide a systematic mapping from narrow syntax to an unambiguous semantic representation *in spite of* its inherently ambiguous use of indices. This trick does a reasonably good job at modeling the facts examined thus far, but it requires ancillary assumptions to explain the ungrammaticality of (6) and the absence of a sloppy reading (which signals a bound-variable interpretation) in (7):

(6)   a.   *The directors who know *her* adore *every actress*.

        b.   *The directors who know *every actress* adore *her*.

(7)   The directors who know *Dolly* adore *her* and the directors who know Irene do too.

        (... the directors who know Dolly adore Dolly/*Irene)   (no sloppy reading)

These facts are puzzling because coreference (and, hence, coindexing) is allowed in a structurally indistinguishable examples like (8).

(8)   a.   The directors who know *her* adore *Dolly*.

        b.   The directors who know *Dolly* adore *her*.

The contrast between (8) and (6)-(7) demonstrates that the syntactic environments which allow coreference are not identical to those in which bound-variable anaphora comes to light. Condition B of the CBT is unable to make this distinction precisely because it is stated as a constraint on *indexing*, which is neutral between coreference and variable binding. As a result the CBT licenses coindexing in all of the structures above.
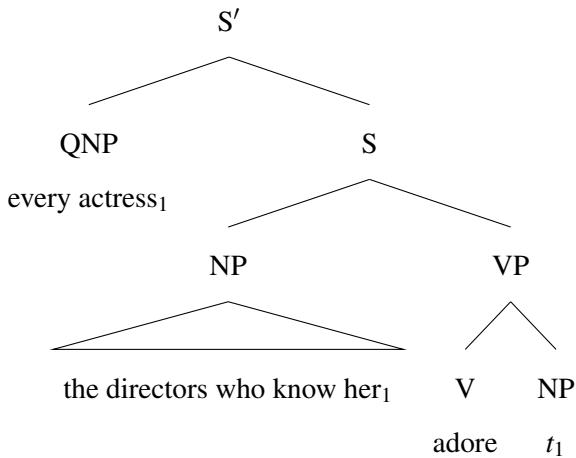
Semantic accounts of CBT's indexing system generally attribute the anomalies in (6)-(7) to a peculiar property of quantification in natural language. There are various proposals

regarding this issue in the literature, but an influential one is based on the Bijection Principle of Koopman and Sportiche (1983, pg. 146). The particular formulation of this principle need not concern us here, but one of its corollaries is relevant to the issue at hand:
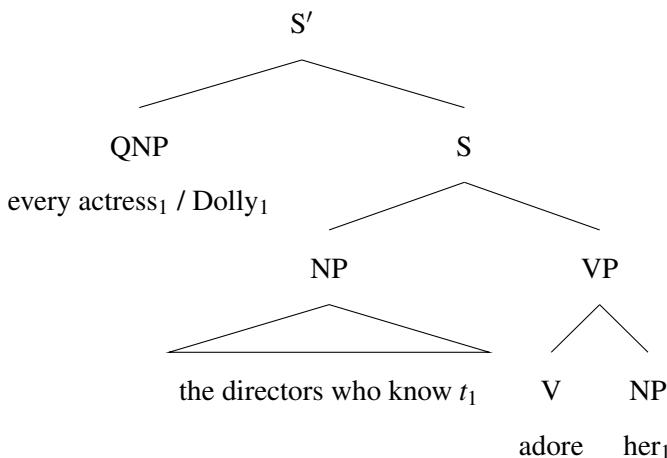
(9)   Constituents in A′-positions must directly syn-bind one and only one variable, where

    a.   *β directly syn-binds α iff* there is no *γ* that is syn-bound by *β* and syn-binds *α*.

Since the specifier of S′ counts as an A′-position, (9) entails that natural language quantifiers can directly syn-bind *at most* one variable. This is precisely what rules out (6) and the sloppy reading for (7). The LFs for these examples are summarized in (10):

(10)   a.

$$
\begin{array}{c}
\text{S}'\\
\diagup\quad\diagdown\\
\text{QNP}\qquad\text{S}\\
\text{every actress}_1\qquad\diagup\quad\diagdown\\
\text{NP}\qquad\qquad\text{VP}\\
\text{the directors who know her}_1\quad\text{V}\quad\text{NP}\\
\text{adore}\quad t_1
\end{array}
$$

b.

$$
\begin{array}{c}
\text{S}'\\
\diagup\quad\diagdown\\
\text{QNP}\qquad\text{S}\\
\text{every actress}_1\;/\;\text{Dolly}_1\qquad\diagup\quad\diagdown\\
\text{NP}\qquad\qquad\text{VP}\\
\text{the directors who know }t_1\quad\text{V}\quad\text{NP}\\
\text{adore}\quad\text{her}_1
\end{array}
$$
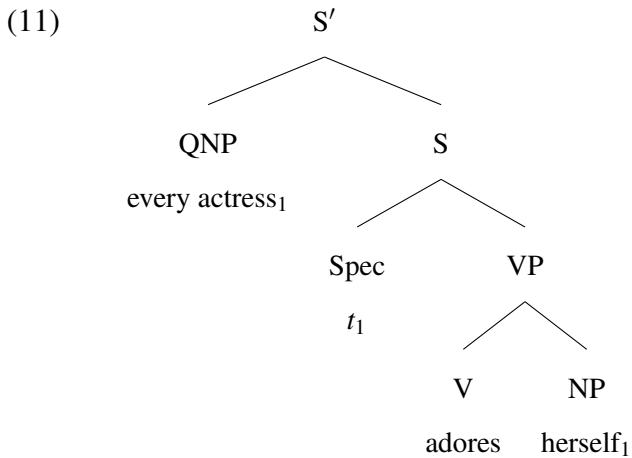
The QNPs in (10) directly syn-bind both their traces and the pronoun *her*. In a licit quantificational structure like (11), a QNP can only directly syn-bind its trace, as dictated by (9).

(11)

```
                        S′
                   ╱         ╲
              QNP               S
         every actress₁      ╱     ╲
                          Spec      VP
                           t₁     ╱    ╲
                                 V      NP
                              adores  herself₁
```

Coreference is allowed in (8) – as opposed to variable-binding (cf. (7)) – because referential antecedents have the option of staying *in situ*, thus avoiding a violation of (9). The Bijection Principle only rules out structures with quantifiers – in particular, those where the traces left behind by Quantifier Raising fail to c-command the relevant pronouns (cf. (10)).

Note, however, that this whole story misses the straightforward generalization that variable binding is only possible when an antecedent (regardless of its semantics) c-commands a pronoun in surface structure. This simple summary of the facts cannot be incorporated into the CBT precisely due to the adoption of an ambiguous indexing system (BÜRING, 2004).

The situation can be put as follows. The ambiguity of coindexing motivates the appeal to a covert level of structure where the differences between coreference and variable binding are transparently "legible" by the interpretive system. But this covert level (LF) is one where the surface generalization that holds of most cases of variable binding cannot be stated directly – hence the need to invoke the theory-laden principle in (9). CBT's ambiguous indexing system requires the hypothesis of obligatory quantifier raising, which, in turn, makes it impossible to formulate the c-command constraint on variable binding.

But why is an inherently ambiguous indexing system adopted by the CBT in the first place? The reasons, Grodzinsky and Reinhart (1993) argue, are mostly historical. Throughout the intellectual development of the various theories of anaphora which led to the CBT, coreference was taken to be the central empirical problem.[2] It was by means of conditions *on indexing* – e.g. the Condition C in (12) (CHOMSKY, 1981, pg. 188) – that the possibility of coreference in (13) and the impossibility of coreference in (14) were originally explained.

---

[2]   This is arguably due to the fact that the first accounts of anaphora in MGG were transformational theories like Lees and Klima's (1963). As we saw, these proposals were outright incapable of handling variable binding. Coreference persists as a primary concern even in Chomsky's (1993) indexless Binding Theory, where the Binding Conditions are stated as interface constraints governing the conditions where coreference is admissible.

(12)   CONDITION C: R(eferential)-expressions cannot be c-commanded by coindexed NPs.

(13)   a.   A party without Dolly$_3$ annoys her$_3$.

        b.   Everybody who knows her$_5$ says Dolly$_5$ is a fabulous person.

(14)   a.   *He$_6$ forced Peter Smith to hire John$_6$'s girlfriend in his lab.

        b.   *He$_1$ said there was a man who was trying to give Otto$_1$ directions.

However, the idea that syntactic constraints on indexing can determine these facts is simply wrong. Relations defined over phrase-structure configurations are too categorical to explain which coreference relations are judged acceptable by speakers. For instance, simply providing richer contextual support for examples like (14) often makes the non-coreference judgments disappear. The example in (15a) comes from Schlenker (2005a, pg. 387) and (15b) is adapted from Dubinsky and Hamilton (1998, pg. 687):

(15)   a.   *John Smith* was so devoid of any moral sense that *he* forced Peter Smith to hire
             John's girlfriend in his lab.

        b.   *Otto* is a jerk. *He* ran over a man who was trying to give *Otto* directions.

If (non-)coreference judgments shift as a result of a manipulation of context, they are more likely to be the yield of context-sensitive factors (e.g. processing, pragmatic preferences) than of narrow syntax (REINHART, 1983; VARASCHIN; CULICOVER; WINKLER, in press). Many of the effects that are attributed to a grammatical non-coreference rule like (12) can be handled by the interface strategy in (16), which Reinhart (2006) justifies on processing grounds.

(16)   RULE I (GRODZINSKY; REINHART, 1993, pg. 79):
        NP A cannot corefer with NP B if replacing A with C, C a variable A-bound by B,
        yields an indistinguishable interpretation.

However, as soon as we abandon the urge to posit non-coreference rules like (12), we no longer need to encode coreference by means of a syntactic device. This suggests the following hypotheses, which lie at the heart of Reinhart's (1983) general theory of anaphora:

(17)   REINHART'S (1983) HYPOTHESES:

        a.   Narrow syntax is only concerned with variable binding.

        b.   Coreference is the result of discourse-level processes of reference resolution.

There are two ways to implement (17). The more conservative one (which the RCT adopts) is to keep Chomsky's (1981) theory of free indexing, but to state the translation procedure between surface structure and semantic structure in a way that the sole interpretation for coindexing is the bound-variable one. Reinhart (1983, pg. 160) proposes (18) (where, for any string Φ and any NP β which occupies a non-A′-position in Φ, Φ^{β}/x is the result of replacing β and all occurrences of pronouns coindexed with and c-commanded by β in Φ by x):

(18) $[_{S'} \Phi] \Rightarrow [_{S'} \beta(\lambda x.\Phi^{\beta}/x)]$

The procedure in (18) is one among a variety rules for mapping syntax to semantic structure. It has the effect of λ-abstracting over the semantics of an antecedent and all of the coindexed pronouns it c-commands. In the resulting representation, the pronouns that the antecedent *syntactically* binds end up being *semantically* A-bound by it as well, according to the definition of A-binding repeated below. On this view, syn-binding (i.e. coindexing under c-command) becomes isomorphic to A-binding. The application of (18) is illustrated in (20).

(19)   A(RGUMENT)-BINDING

          β A-binds α iff β is the sister of a λ-predicate whose operator logically binds α.

(20)   a.   Only Bobby$_3$ says that he$_3$ is happy. $\Rightarrow$ Only Bobby ($\lambda x.$ x says (x is happy))

          b.   Every actor$_5$ says that he$_5$ is happy. $\Rightarrow$ Every actor ($\lambda x.$ x says (x is happy))

Reinhart's theory only allows coindexing to be interpreted as variable binding, in accordance with (18). This means that coindexing which is not coupled with c-command (cf. (13) above) becomes literally uninterpretable. Coreference needs to be captured by a separate (non-syntactic) mechanism, such as the contextual assignment of identical values to NPs with different indices (GRODZINSKY; REINHART, 1993, pg. 77). We can get the coreference reading for (21) if we assume that the context involves an assignment function g such that $g(1) = g(7)$.[3]

(21)   Only Bobby$_1$ says that he$_7$ is happy.

The assignment of the same reference to the two NPs in (21) is no different from what would happen if each of these NPs occupied a different sentence in discourse, as in (22).

---

[3]   Technically, assignment functions are functions from the set of natural numbers (from which indicies are drawn) to individuals. We can think of an assignment as part of what *defines* a context (BÜRING, 2005, pg. 27). A sentence containing a free pronoun means different things in different contexts. In the context defined by an assignment g' such that $\langle 7, \text{Curly} \rangle \in g'$, the sentence *He$_7$ sings* expresses the proposition that Curly sings.

(22)   Bobby$_1$ is here. He$_7$ seems happy.   (where $g(1) = g(7)$)

That there is no parallel coreference reading for (20b) is a consequence of the fact that quantifiers never refer to individuals (and, hence, they cannot receive their interpretation from assignment functions in the usual sense):

(23)   *Every actor$_1$ is here. He$_7$ seems happy.   ($g(1) = g(7)$ is unspecified)

The only way a pronoun can be coconstrued with a quantified NP is if it is interpreted as a bound-variable – which, according to (18), can only happen when the pronoun is syntactically bound (i.e. coindexed under c-command).

There is, however, a more radical implementation of (17) which abandons the use of indices (and the derivative notion of *syntactic binding*) altogether. Instead of stating constraints on A-binding "indirectly" by means of constraints on coindexing in narrow syntax, we can simply make them specify the conditions under which particular kinds of pronouns can be placed in correspondence with A-bound variables in semantic structure. In this spirit, CBT's Condition A (repeated as (24)), can be recast informally as the correspondence rule in (25)[4]

(24)   CONDITION A: An anaphor is coindexed to a locally c-commanding NP.

(25)   CONDITION A (INDEXLESS VERSION): An anaphor α must be interpreted as an A-bound variable, where the β that A-binds α must be local with respect to α in syntax.

That this is in fact a possibility was anticipated by Reinhart (1983, pg. 160) herself:

> Nothing in my analysis hinges on the existence of a separate coindexing device. The translation procedure for pronouns can be stated directly on non-coindexed surface structures, λ-abstracting on a given NP and translating all pronouns c-commanded by this NP (under the appropriate auxiliary conditions [specified by an appropriate Binding Theory]). The crucial point in my analysis is that what is needed in the grammar to account for anaphora is a mechanism determining when a pronoun can be translated as a bound variable.

Reinhart's approach – whether in the conservative version assuming indexing or in the indexless version – makes it possible to envision a much more surface-oriented theory of variable binding in natural language than approaches based on the CBT. This point will be particularly relevant when I sketch the constructional reduction of the RCT in Chapters 5-7.

---

[4]   This opens up the possibility of assuming a variable-free semantics, which dispenses with the idea that that pronouns start out as assignment-dependent (i.e. free) and become bound in the course of derivations. Szabolcsi (1987, 1992) works out an alternative which explicitly takes Reinhart's hypotheses in this direction.

2.1.2 Variable binding without c-command?

The conception of the syntax-semantics interface outlined above presents a lucid picture of what a linguistic theory of anaphora might look like. From the point of view of the grammar *per se* (i.e. the context-independent system of knowledge internalized by the speaker), there are two sets of conditions governing the interpretation of pronouns: (i) a general condition determining when pronouns can be interpreted as bound variables and (ii) specific conditions regulating the environments where particular types of pronouns (e.g. anaphors, pronominals and reflexives) can be so interpreted. The first of these is what Reinhart (1983) calls the BOUND ANAPHORA CONDITION. The second is the subject of theories like CBT and RCT.

The translation procedure in (18) incorporates a specific proposal regarding (i), according to which variable binding is determined by c-command in surface structure. This provides a straightforward explanation for why quantified antecedents and sloppy readings are impossible in (6), repeated below as (26):

(26) a. *The directors who know *her* adore *every actress*.

  b. *The directors who know *every actress* adore *her*.

There are, however, well-known counterexamples to the generalization that A-binding can only take place from a c-commanding position in surface structure. These typically involve binding out of possessors, nominal arguments and modifiers of quantificational NPs (HIGGINBOTHAM, 1980; MAY, 1985; KAYNE, 1994; BARKER, 2012):

(27) a. [*Every boy*'s mother] thinks *he* is a genius.

  b. [The fate of *every individual*] is decided by *his* inner ego.

  c. [Someone from *every city*] hates *it*.

However, there is some reason to believe that the pronouns in (27) are not coconstrued with their antecedents in the same sense as typical pronouns A-bound under c-command:

(28) a. *Every boy* thinks *he* is a genius.

  b. *Every individual* is guided by *his* inner ego.

  c. *Every city* has inhabitants who hate *it*.

To see why (27) and (28) may not be the same kinds of beasts, it is useful to examine

how German demonstrative pronouns (DEM) (WILTSCHKO, 1998; PATEL-GROSZ; GROSZ, 2010; HINTERWIMMER; BROCHER, 2018) and subject pronominals in Brazilian Portuguese (BP) (MONTALBETTI, 1984; MENUZZI, 2003; MÜLLER, 2003) behave in similar contexts. What is peculiar about these expressions is that they display a general dispreference for bound-variable interpretations, as the examples below illustrate:[5]

(29) a. **Jeder Jungen* glaubt, dass *dieser* ohne   Fehler ist.
       every boy   thinks   that   DEM   without flaws   is
       '*Every boy* thinks that *he* is without flaws.'

    b. **Jeder Mann* behauptet, dass *der*   intelligent ist.
       every man   claims   that   DEM intelligent is
       '*Every man* claims that *he* is intelligent.'

(30) a. ??*Nenhuma atriz*   disse que *ela* deveria ganhar o   Oscar.
       no   actress said   that she should   win   the Oscar
       '*No actress* said that *she* should win the Oscar.'

    b. ??*Qualquer médico* sabe   que *ele* não pode viajar agora.
       any   doctor   knows that he not can   travel now
       '*Any doctor* knows that *he* cannot travel now.'

Though judgments are less clear, Büring (2005, pg. 123) and Krifka (2018, pg. 44) report something similar holds for English epithets, as (31) illustrates:[6]

(31) a. *Every player* ran over a man who was trying to give {*him* / **the jerk*} directions.

    b. It was said of *every player* that {*he* / **the idiot*} lost a hundred dollars on the slots.

Despite resisting bound-variable readings in general, overt subject pronominals in BP, German demonstratives and English epithets can all be coconstrued with non-c-commanding QNPs in contexts parallel to (27). The data below are adapted from Hinterwimmer (2015, pg. 67), Krifka (2018, pg. 45), Menuzzi and Lobo (2016, pg. 351) and Koster (1987, pg. 97)

---

[5] According to Montalbetti (1984), the anti-binding effect in BP is a consequence of the Overt Pronoun Constraint, which says that, when a null form is available (e.g. a *pro* subject), overt pronouns cannot behave as bound variables. He gives the following BP example to illustrate (MONTALBETTI, 1984, pg. 191):

(i)   *Ninguém* acha que {*pro* / **ele*} é inteligente.

In cases like (30), where more referential material is added to the QNPs, A-binding of overt subject pronouns is not strictly impossible like it is in (i), but merely disfavored (MENUZZI, 2003).

[6] In addition to not being bindable by QNPs, epithets are also subject to an ANTI-LOGOPHORICITY restriction: i.e. they cannot refer to a discourse referent who counts as perspective bearer in their discourse context (DUBINSKY; HAMILTON, 1998). This condition is satisfied in (31), as (ia) shows:

(i)   a. *John* ran over a man who was trying to give *the jerk* directions.
       b. It was said of *John* that *the idiot* lost a thousand dollars on the slots.

(32)   a.   [Die Mutter *jedes   Jungen*] glaubt, dass *dieser* ohne   Fehler ist.
            the mother of.every boy   thinks that DEM   without flaws   is
            '*Every boy*'s mother thinks that *he* is without flaws.'

       b.   [Mindestens eine Person aus   *jeder Stadt*] hasst *diese*.
            at.least   one person from every city   hates DEM
            'Someone from *every city* hates *it*.'

(33)   a.   [O pai   de *nenhuma atriz*] disse que *ela* deveria ganhar o   Oscar.
            the father of no   actress said that she should win   the Oscar
            '*No actress*' father said that *she* should win the Oscar.'

       b.   [A mãe   de *qualquer médico*] sabe   que *ele* não pode viajar agora.
            the mother of any   doctor   knows that he not can   travel now
            '*Any doctor*'s mother knows that *he* cannot travel now.'

(34)   a.   [Someone from *every city*] hates *the place*.

       b.   [The fate of *every dictator*] is determined by *the jerk*'s actions.

These facts lend plausibility to the conclusion that coconstrual without c-command is fundamentally different from coconstrual with c-command. Arguably, only the latter corresponds to variable binding in semantics, as Reinhart (1983) proposes.[7] If this is indeed the case, pronouns that are coconstrued with non-c-commanded QNPs are not strictly A-bound by their antecedents – at least not in the same sense as those that are c-commanded by QNPs. This is why classes of NPs that disfavor bound-variable readings (e.g. German demonstratives, BP subject pronouns and English epithets) are acceptable in contexts like (32)–(34).

Büring (2004), following a suggestion by Bach and Partee (1980), offers a formal semantic account of non-c-command binding which preserves the basic Reinhartian picture supported by the data above. He argues that pronouns that are apparently A-bound by non c-commanding QNPs are, in fact, interpreted as disguised definite descriptions containing a free relation variable *R*. According to this proposal, (35a) is analyzed as having the semantic structure in (35b), where the content of the pronoun is underlined. In this case, *R* is instantiated as the *son-of* relation, which maps a mother to her son, giving rise to the reading sketched in (35c):

(35)   a.   [*Every boy*'s mother] thinks *he* is a genius.

       b.   Every boy's mother ($\lambda x.\ x$ thinks the $R(x)$ is a genius)

       c.   *Every boy's mother* thinks *her* son is a genius.

Note that, on this account, it is the whole container NP *every boy's mother* that does the

---

A-binding, rather than the embedded QNP *every boy*. This proposal also carries over to other apparent counterexamples to the c-command generalization in (27), as (36) exemplifies:

(36)   a.   [Someone from *every city*] hates *it*.

       b.   Someone from every city (λx. *x* hates the $R(x)$)

       c.   [*Someone from every city*] hates the city *they* are from.

Büring predicts that apparent cases of A-binding without c-command will only emerge when the host NP that contains the apparent binder carries its own quantificational force, which can in turn bind the variable contained in the descriptive content of the pronoun. This prediction is in fact borne out. In (36), the quantificational status of the host DP is quite explicit, whereas in an ill-formed case such as *[The directors who know **every actress**] adore **her***, there is no quantification over directors. In (35a), there is quantification over mothers, albeit implicit: the sentence concerns *boy-mother* pairs, rather than a single person who is the mother of every boy. Consider Reuland's (2011, pg. 75) interesting example below:

(37)   [The owner of *every car* in the street] should move *it* on Mondays.

The subject NP in (37) is ambiguous between a referential reading where there is a single person who owns every car in the street and a reading where each car has a possibly different owner. Coconstrual between *every car* and *it* is only possible under the latter, which implies a quantification over owners, depicted in the semantic structure in (38) (where *R=car-of*):

(38)   The owner of *every car* in the street (λx. *x* should move the $R(x)$ on Mondays).

Büring's (2004) theory effectively reduces the appearance of A-binding in non-c-command configurations to the e-type strategy independently postulated for handling different varieties of donkey anaphora (COOPER, 1979; EVANS, 1980; HEIM; KRATZER, 1998). In a typical case of donkey anaphora, apparent non-c-command A-binding is effected by a quantificational indefinite NP. A paradigm example is given in (39), along with Büring's (2004) proposed analysis. In that case, the free *R* variable is instantiated by the *donkey-of* relation.

(39)   a.   [Every farmer who owns *a donkey*] loves *it*

       b.   Every farmer who owns a donkey (λx. *x* loves the $R(x)$)

       c.   [*Every farmer who owns a donkey*] loves the donkey that *he* owns.

Treating apparent non-c-command A-binding on a par with donkey anaphora also explains why German demonstratives, BP subject pronouns and English epithets can adequately replace standard donkey pronouns (PATEL-GROSZ; GROSZ, 2010):

(40)   a.   *Jeder Bauer*, der einen Esel   besitzt, liebt *diesen*.
            every farmer who a   donkey has   loves DEM

       b.   Todo fazendeiro que tem *um burro*   diz que *ele* é lindo.
            every farmer   who has a   donkey says that he is beautiful
            'Every farmer who owns *a donkey* says that *he* is beautiful.'

       c.   Every farmer who owns *a donkey* loves *the poor animal*.

There is certainly much more to be said for and against Büring's (2004) proposal (see Bruening (2014) for an opposing view). The discussion here was kept largely informal because my purpose was merely to show that there is a plausible enough way to retain Reinhart's (1983) surface-oriented theory of variable binding in light of potential counterexamples. My main point was that we do not need to resort to complex derivational devices like Quantifier Raising to explain when variable binding is possible: the surface structure of expressions will do.[8]

The issue I turn to now are the precise conditions under which different kinds of pronouns are construed as A-bound variables in semantic structure. In Chapter 1, I argued that the conditions proposed within the CBT fail to deliver a descriptively adequate account of the facts about the distribution and interpretation of anaphors and pronominals. I argue below that the reflexivity-and-chains theory (RCT) of Reinhart and Reuland (1993) does a much better job in this, even though it perpetuates some problems from earlier approaches.

---

[8]   For my purposes, the surface-oriented nature of Reinhart's theory is more important than the specific details of her proposal. It may well be that an alternative to c-command, such as o-command (POLLARD; SAG, 1994; HUKARI; LEVINE, 1995; WALKER, 2011), whose simplified recursive definition is given in (i), works better as a generalization for the surface pattern that licenses A-binding in semantics. This seems to be correct if we look at data like (ii) (BALTIN; POSTAL, 1996; BARKER, 2012; PESETSKY, 1995):

(i)   A syntactic object β O-COMMANDS a syntactic object α *iff* β is less oblique than α or β o-commands a γ which dominates α, where

       a. β is LESS OBLIQUE than α *iff* β precedes α on the following hierarchy:
            subject $\prec$ indirect object $\prec$ direct object $\prec$ prepositional goal $\prec$ prepositional theme $\prec$ adjunct

(ii)   a.   A book was given [to *every boy*] by *his* mother.
       b.   Amy talked [to *every student*] about *his* grades.
       c.   [In *his* own mind], *every actor* is the most important person in the world.

In virtue of the branching structure (of the PPs in (iia) and (iib) and of the A′ attachment in (iic)), the pronouns in (ii) are o-commanded, but not c-commanded by their antecedents: in (iia), the antecedent is a prepositional goal and the pronoun is an adjunct; in (iib), the antecedent is a prepositional goal and the pronoun is a prepositional theme; in (iic), the antecedent is a subject and the pronoun is part of an adjunct. Note, however, that this o-command alternative also needs to resort to something like Büring's (2004) e-type strategy to handle cases of binding out of NPs since the antecedents in such cases do not o-command the pronouns.

2.2   THE REFLEXIVITY-AND-CHAINS THEORY

In this section, I give a detailed summary of Reinhart and Reuland's (1991, 1992, 1993, 1995) reflexivity-and-chains theory of anaphora (RCT). I focus here mostly on positive aspects of the theory, postponing more severe criticisms of the approach until Chapter 3. I will ultimately argue that the RCT is correct in many respects (e.g. its typology of pronouns and its emphasis on predicate-argument structure) but wrong in others (e.g. its claim to universality and its configurational residues).

In addition to its empirical flaws, the CBT account of anaphora discussed in Chapter 1 also carries some degree of conceptual arbitrariness. Though concepts like c-command and government are, to a large extent, independently motivated (i.e. they are invoked in other parts of the grammar), notions like SUBJECT, governing category and syn(tactic)-binding don't seem quite as virtuous. Moreover, the relations between these theoretical constructs and the inherent properties of NPs also come across as largely stipulative in the overall CBT framework. For instance, there is no property of anaphors (*qua* referentially defective NPs) that explains why they must be bound in the smallest domain containing their governors and accessible SUBJECTS. Likewise, there is no "deep" reason why pronominals cannot be bound in similar environments. These facts are just stipulated by modular principles (Conditions A and B), which were (at the time of their formulation) thought to be constitutive of Universal Grammar (UG).[9]

Reinhart and Reuland's (1991, 1992, 1993, 1995) reflexivity-and-chains theory (RCT) was designed to overcome these conceptual shortcomings. Instead of simply stipulating Binding Conditions as irreducible principles, the RCT tries to explain and represent facts about anaphoric dependencies in terms of constraints and relations that are independently available in other parts of the grammar. As in other post-CBT accounts of anaphora, there is, strictly speaking, no Binding Theory module in RCT's architecture. The Binding Conditions are viewed as hypostases at best and the effects they purport to explain are attributed to the interaction of different subsystems. In addition, the RCT also tries to establish a closer connection between the coconstrual possibilities of NPs and their intrinsic morphosyntactic make-up.

In light of these goals, it becomes of crucial importance to determine precisely which

---

[9]   Subsequent work recognized the CBT as dubious in light of the minimalist pursuit of a leaner and more biologically realistic UG (HORNSTEIN, 2001, ZWART, 2002; HICKS, 2009, ROORYCK, J.; VAN WYNGAERD, 2011, i.a.). Nonetheless, one can raise similar concerns about Charnavel's (2019) recent attempt to recast the CBT in phase-theoretic parlance (CHOMSKY, 2001). Exchanging outdated terms (e.g. "governing category", "empty PRO subject") for more contemporary jargon (e.g. "spell-out domain", "null logophoric operator") does not remove the conceptual arbitrariness of the general framework.

internal properties of pronouns are significant for their behavior with respect to anaphora. The RCT assumes the typology of pronouns presented in Chapter 1, which is stated in terms of the properties of referential independence, now represented as R, and reflexivizing function:

	Reflexive Anaphor	Non-reflexive Anaphor	Pronominal
R(eferential Independence)	-	-	+
Reflexivizing Function	+	-	-

Table 2 – The typology of pronouns (adapted from Reinhart and Reuland (1993, pg. 659))

As we saw in Chapter 1, reflexivizing function is the persistent capacity to impose identity between two thematic arguments of a predicate. On a first approximation, we can think of this as an inherent semantic property of lexical items, which can be modeled by attributing to them a semantics consisting minimally of the structure depicted in (41) (BACH; PARTEE, 1980; SZABOLCSI, 1987; KEENAN, 1988; SCHLENKER, 2005b; SPATHAS, 2010).[10]

(41) $\lambda P_{\langle e, \langle e,t \rangle \rangle}.\lambda x.P(x)(x)$

This approach treats reflexives as arity reducers: the job of (41) is basically to take a binary predicate and turn it into a unary one, binding two of its arguments to the same λ-operator. The application of an item bearing a reflexivizing function in this sense to a predicate is shown in (42). In the resulting structure, the two thematic arguments of the predicate are identified.

(42) $[\![admires\ himself]\!] = \lambda P_{\langle e, \langle e,t \rangle \rangle}\lambda x.P(x)(x)(\lambda y.\lambda z.\ z\ admires\ y)$
          $= \lambda x.\lambda y.\lambda z.\ z\ admires\ y(x)(x) = \lambda x.\ x\ admires\ x$

The R property can be intuitively characterized as the capacity to support reference without relying on a previously established referent in the discourse. In this sense, English pronominals are +R and reflexives are -R, as we saw in Chapter 1. We will see below that the RCT actually looks deeper into this and proposes that R can be reduced to ϕ-feature specification.

The central idea which defines the RCT is that each of these properties is governed by a different module of grammar. Reflexivizing function is governed by REFLEXIVITY THEORY and R is governed by CHAIN THEORY, which also regulates A-movement in general (passives, raising, etc.). The fact that the distribution and interpretation of pronouns is governed by these two sets of principles is what gives the RCT its name. I will go over each of these below.

---

[10] The λ-term in (41) is what is known as a duplicator function within combinatory logic (SZABOLCSI, 1987, 1992; STEEDMAN, 1988). In order to handle the application of reflexives to ternary relations and object raising predicates, we need a a polymorphic extension of (41), as in Szabolcsi (1987).

2.2.1   Reflexivity Theory

The Reflexivity Theory is concerned with the morpho-syntactic encoding of reflexive predicates – i.e. predicates whose coarguments are coconstrued. If we assume the Reinhartian picture outlined in Section 2.1, this coconstrual relation can only be grammatically relevant if it corresponds to A-binding. The notion of a reflexive predicate can be defined accordingly as follows (REULAND, 2011, pg. 81):

(43)   A predicate is REFLEXIVE iff two of its arguments are bound by the same λ-operator.

Given the translation procedure for indexing summarized in (18) above, both of the sentences in (44) correspond to the reflexive predicate in (44c) – though (44b) is ill-formed.

(44)   a.   Rose$_5$ despises herself$_5$.

      b.   *Rose$_2$ despises her$_2$.

      c.   Rose ($\lambda x.\ x$ despises x)

The situation is less clear for (45a)-(47a). In (45a) it is not obvious whether the pronominal is part of the predicate headed by *over* or of the one headed by *pull*. Similar uncertainties hover over the reflexive anaphor in (46a) and the non-reflexive Dutch anaphor *zich* in (47a):

(45)   a.   Rose$_2$ pulled the blanket over her$_2$.   (HESTVIK, 1991, pg. 462)

      b.   Rose ($\lambda x.\ (x$ pulled the blanket) & (the blanket ($\lambda y.\ y$ is over x)))

(46)   a.   Rose$_7$ praised (both) Herbie and herself$_7$.   (REULAND, 2001, pg. 487)

      b.   Rose ($\lambda x.\ (x$ praised Herbie) & (x praised x))

(47)   a.   Rose$_2$ voelde zich$_2$ wegglijden.
           Rose felt   SE   slip-away.   (REULAND, 2011, pg. 118)
      b.   Rose ($\lambda x.\ x$ felt (x slip away))

The reason why these cases are more dubious than (44) is because the notion of *predicate* is fundamentally ambiguous. In order to remedy this, the RCT makes the following distinction between SYNTACTIC and SEMANTIC predicates:

(48)   a.   The SYNTACTIC PREDICATE formed of (a head) P is P, all of the projections that realize grammatical functions associated with P (i.e. the SYNTACTIC ARGUMENTS of P), and an external argument of P (i.e. the SUBJECT of P).[11]

---

[11]   The definition of syntactic predicate in (48) corresponds roughly to Chomsky's (1986b) notion of Complete

b.   The SEMANTIC PREDICATE formed of P is P and all arguments that realize a semantic role associated with P (i.e. the SEMANTIC ARGUMENTS of P).

The only syntactic predicate in (45a) is the one headed by *pull*, whose syntactic arguments are the subject *Rose*, the object *the blanket* and the PP *over her*. This predicate assigns a grammatical function to the PP as a whole (*as per* the analysis in Marantz (1984)), so it does not include the pronominal *per se*. As a consequence, it is also not reflexive, according to (43). The PP does not qualify as a syntactic predicate on its own because it lacks a subject. It does, however qualify as a *semantic* predicate: namely, a binary spatial relation encoded by *over*. One of the semantic arguments of this relation is controlled by the object of the verb (*the blanket*) and the other is a bound variable that corresponds to the pronominal.

The syntactic predicate in (46a) consists in *Rose*, *praise* and the conjoined NP. The reflexive by itself does not count as one of its syntactic arguments. This means that the syntactic predicate is not reflexive. However, there is a reflexive *semantic* predicate involved here, since the semantic structure of (46a) is (46b), where the variables that realize the two semantic roles of *praise* in the second conjunct are bound by the same λ-operator.[12]

Just the opposite happens in (47a). The semantic and syntactic predicates therein are two: one headed by *voelen* ('feel') and another headed by *wegglijden* ('slip away'). The reflexive is a syntactic argument of both, because it realizes the subject grammatical function of the lower predicate and the object function of the higher one. However, it is only a semantic argument of *wegglijden*. Since this predicate is both syntactically and semantically monovalent (and, thus, irreflexive), only the syntactic predicate headed by *voelen* comes out as reflexive.

On the basis of these notions, the RCT posits the following conditions:

(49)   REFLEXIVITY CONDITION A (RCA) (cf. (43))

         A reflexive-marked syntactic predicate is reflexive.

(50)   REFLEXIVITY CONDITION B (RCB) (cf. (43))

         A reflexive semantic predicate is reflexive-marked.

Functional Complex, as Reinhart and Reuland (1992) recognize. This is not quite the definition proposed in Reinhart and Reuland (1993, pg. 678), but I stick with it for convenience, because it allows for a smoother transition into the constructional approach I will ultimately adopt in Chapters 5-7.

[12]   A discussion of how the grammar assigns (46b) as the semantic structure of (46a) by type-driven interpretation is delayed until Chapter 5. For now, it suffices to note that the duplicator meaning in (41), which can be taken to define the reflexivizing property, can be conjoined (by generalized conjunction) with a type-lifted meaning for Herbie ($\lambda P.\lambda x.P(\text{herbie})(x)$). The resulting meaning term ($\lambda P.\lambda x.P(\text{herbie})(x)\&P(x)(x)$) can then pick up the meaning of *praise*, leading to the predicate in (46b) (cf. Morrill (2011, pg. 47) for a concrete proposal).

(51)   A predicate (formed of P) is REFLEXIVE-MARKED iff either P is lexically reflexive or
         one of P's arguments is a reflexive pronoun.

Unlike the Binding Conditions of the CBT, RCA and RCB are not constraints on pronouns, but on predicates. The motivation for this view – which is shared with HPSG's binding theory (POLLARD; SAG, 1992, 1994) – is the observation that the distribution of reflexive-marking devices is not (only) sensitive to configurational relations but (also) to properties of predicate-argument structure. Reinhart and Reuland (1993) argue that most of the empirical shortcomings of the CBT – for instance, the inability to deal with discursive reflexives, non-reflexive anaphors and locally bound pronominals – can be solved by assuming the predicate-centered outlook. The first of these is taken up by RCA and the other two by RCB.

Let us start with simple illustrations. Consider how RCA handles the following:

(52)   a. $Bobby_2$ criticized $himself_2$.

               Bobby ($\lambda x.$ x criticized x)

         b. $Bobby_9$ relies on $himself_9$.

               Bobby ($\lambda x.$ x relies on x)

         c. $Bobby_1$ appears to $himself_1$ to be a genius.

               Bobby ($\lambda x.$ x appear to x (x is a genius))

         d. $Bobby_1$ forgot PRO_1 to defend $himself_1$.

               Bobby ($\lambda x.$ x forgot (x defend x))

In (52a), the syntactic predicate headed by *criticize* is reflexive-marked (i.e. one of its arguments is a reflexive pronoun). Since this predicate is reflexive, RCA is satisfied. The same reasoning applies to (52b). Note that, unlike in (45a), the complement of the preposition in this particular case *does* form a syntactic predicate with the main verb – it realizes the verb's oblique grammatical function.[13] In (52c), *Bobby* is a syntactic argument of *appear*, but not a semantic

---

[13]   There are several ways to distinguish P-complements that form syntactic predicates with the verb from those
       that do not (cf. (45a)). The former, for instance, are far more acceptable in pseudo-passive structures (cf. (i))
       and involve prepositions that cannot be contrastively stressed (cf. (ii)) (CULICOVER, 1976; FROUD, 2001):

(i)   a.   Bobby was relied on by everyone.
         b.   *Rose was hid behind by the children. (cf. The children hid behind Rose)
(ii)   a.   *Bobby relies ON his friends.
         b.   Rose pulled the blanket OVER the children.

       Moreover, prepositions whose complements realize grammatical functions of the verb (i.e. *functional* prepositions) are semantically empty and rigidly determined by the verb. This is what leads some proposals within
       MGG to view them as mere case markers (KAYNE, 1975; ROORYCK, 1996), as opposed to lexical prepo-

one. Since RCA is stated in terms of syntactic predicates (and the syntactic predicate headed by *appear* is in fact reflexive), (52c) is also correctly predicted to be grammatical. Lastly, in (52d), *defend* forms a reflexive-marked syntactic predicate with a covert PRO subject. Given that this predicate is in fact reflexive under the depicted reading, RCA is satisfied.

The examples in (53), however, are *not* licensed by RCA. In all of these cases, reflexives are arguments of syntactic predicates that fail to be reflexive. For instance, the syntactic predicate headed by *love* in (53a) is not one whose coarguments are bound by the same λ-operator.

(53)   a.   *$Bobby_6$ told me that Marta loves $himself_6$.

Bobby ($\lambda x. \; x$ told me (Marta loves x))

b.   *$Bobby_2$'s friends admire $himself_7$. (under an assignment g, where $g(2) = g(7)$)

Bobby ($\lambda x. \; x$'s friends admire y) & y =Bobby

c.   *The $actress_5$ defends $himself_1$.

The actress ($\lambda x. \; x$ defends y)

Note, furthermore, that, for all of the well-formed cases in (52), the definition of reflexivity in (43) requires reflexive pronouns to be A-bound in semantics. In fact, the RCT as a whole excludes coreference interpretations for any reflexive that serves as an argument of a syntactic predicate, as we see in (54). The CBT makes a different prediction. Since its Binding Conditions are stated as constraints on indices (which are, as we saw, interpretively ambiguous), the CBT expects the reflexives in (54) to receive both coreferential and A-bound readings.

(54)   a. $Bobby_6$ relies on $himself_6$, and so does $Paul_3$.

Bobby ($\lambda x. \; x$ relies on x) & Paul ($\lambda x. \; x$ relies on x)   (sloppy reading only)

b.   Only $Bobby_2$ criticizes $himself_2$.

Only Bobby ($\lambda x. \; x$ criticizes x)

The data above strongly favor the RCT account (but see Section 3.3 for some caveats). The second conjunct in (54a) can mean that Paul relies on Paul, but not that Paul relies on *Bobby*. (54b) could be true in a scenario where other people also criticize Bobby, but not in one where other people are also self-critical. As we saw in Chapter 1, both of these readings are indications that the anaphor is construed as a bound variable, as RCA predicts.

The idea that RCA-abiding reflexives are necessarily interpreted as bound variables allows

---

sitions like *behind*, which assign case and semantic roles of their own. This distinction is independent of the argument/adjunct status of the PP itself: the PPs in (45a) and (52b) both qualify as arguments.

us to solve a puzzle raised by Kayne (1994, pg. 25). The contrast in (55) is a problem for the CBT because it requires the definition of binding to be different for different kinds of pronouns: the pronominal must be bound by the possessive, but the reflexive must not.

(55)   *Every girl*'s father loves {*her* / *herself}.

       Every girl's father ($\lambda x.$ x loves $\underline{\text{the } R(x)}$)

The RCT provides an immediate explanation for this contrast. Since, according to Büring (2004), instances of apparent A-binding without c-command are, in fact, cases where a pronoun is construed as a covert description, (55) counts as a violation of RCA. The syntactic predicate headed by *love* is reflexive-marked, but it is not reflexive: one of its syntactic arguments corresponds to a bound variable x, while the other corresponds to the description that is associated with the pronoun (i.e. $\underline{\text{the } R(x)}$). If the predicate headed by *love* was genuinely reflexive, as RCA requires, we would get something like (56), which is the semantic structure assigned to ***Every girl's father** loves **himself***.

(56)   Every girl's father ($\lambda x.$ x loves x)

Consider now the work that RCB is supposed to do for some of the cases I mentioned above. The examples in (57), which are analogous to (52), are ruled out because all of them express reflexive semantic predicates that are not reflexive-marked.

(57)   a. *Bobby$_2$ criticized him$_2$.

             Bobby ($\lambda x.$ x criticized x)

         b. *Bobby$_9$ relies on him$_9$.

             Bobby ($\lambda x.$ x relies on x)

         c. *Bobby$_1$ forgot PRO$_1$ to defend him$_1$.

             Bobby ($\lambda x.$ x forgot (x defend x))

The examples in (58), in turn, are all allowed by RCB because they don't express reflexive semantic predicates at all.[14] Though the λ-term assigned as semantic representation for (58a)

---

[14] A separate principle is needed to rule out cases where coarguments of a semantic predicate corefer without being A-bound. These can be represented with RCT's indexing system as in (i):

(i)   *Bobby$_8$ criticized him$_1$. (under an assignment g, where $g(8) = g(1)$)
          Bobby ($\lambda x.$ x criticizes y) & $y =$ Bobby

     Example (i) does not violate the RCB stated in (50) because its semantic predicate is not reflexive according to

is technically reflexive according to (43), it does not correspond to a semantic predicate, which is restricted to the domain of a head and its immediate semantic arguments (cf. (48b)).

(58) a. Bobby$_6$ told Susan that Marta loves him$_6$.

Bobby ($\lambda x.$ x told me (Marta loves x))

b. Bobby$_2$'s friends admire him$_7$. (under an assignment g, where $g(2) = g(7)$)

Bobby ($\lambda x.$ x's friends admire y) & $y =$ Bobby

c. The actress$_5$ defends him$_1$.

The actress ($\lambda x.$ x defends y)

For most of the data examined above – all except the semantic facts in (54)-(55) – the RCT makes the same predictions as the CBT. The RCT begins to show its distinctive character when we examine contexts where both reflexives and pronominals are acceptable. These are precisely the cases that posed problems for the complementarity assumption of the CBT in Chapter 1:

(59) a. Rose$_8$ said that the NYT published that picture of {her$_8$ / herself$_8$}.

Rose ($\lambda x.$ x said (the NYT published that (picture of x)))

b. Rose$_2$ pulled the blanket over {her$_2$ / herself$_2$}.

Rose ($\lambda x.$ (x pulled the blanket) & (the blanket ($\lambda y.$ y is over x)))

c. Curly$_4$ saw a snake behind {him$_4$ / himself$_4$}.

Rose ($\lambda x.$ (x saw a snake) & (the snake ($\lambda y.$ y is behind x)))

d. Susan$_8$ said that laymen admire physicists like {her$_8$ / herself$_8$}.

Joanne ($\lambda x.$ x said (laymen admire physicists (like x)))

e. Rose$_6$ said that Herb invited Louise and {her$_8$ / herself$_6$} for lunch.

Rose ($\lambda x.$ x said ((Herb invited Louise for lunch) & (Herb invited x for lunch))

f. It is {him$_2$ / himself$_2$} that Curly$_1$ claims that Laurie loves. ($g(2) = g(1)$)

Curly ($\lambda x.$ x claims (Laurie loves y)) $y =$ Curly[15]

Pronominals are allowed by RCB in all of the structures above because none of them contains a reflexive semantic predicate in need of reflexive-marking. Reflexives are *also* predicted

---

the definition in (43) (which entails A-binding). Fortunately, however, the principle the RCT can use to rule out coargument coreference in cases like these is precisely the same one which is independently needed to handle most Condition C effects: the extra-grammatical Rule I of Grodzinsky and Reinhart (1993) (cf. (16) above).

[15] The semantic structures in (59) assume Reinhart's (1983) interpretive procedure motivated in Section 2.1, whereby A-binding is only obtained under c-command. It is possible to derive an interpretative effect similar to A-binding for the examples like (59f) by indirect means, as discussed in the connectivity literature (JACOBSON, 1994; STERNEFELD, 2000; SCHLENKER, 2003a). I will not address this issue here.

to be acceptable here, but for a peculiar reason: they do not fall under the scope of RCA.

The reflexives in (59a)-(59c) are not arguments of syntactic predicates because the heads whose grammatical functions they realize (namely, *picture*, *description*, *over* and *near*) do not have subjects (cf. (48a)).[16] In (59d)-(59f), the reflexive pronouns are not syntactic arguments at all, as they do not realize a grammatical function of any predicate. In the conjunction and NP adjunct structures in (59d)-(59e), the reflexives are merely *parts* of the syntactic arguments of the verbs *admire* and *invite*. In (59f), the reflexive occupies a non-argument position.

The upshot is that, in the examples in (59), there is no reflexive-marking of a syntactic predicate in the first place. Reflexives in these contexts are, therefore, EXEMPT from RCA and free to be determined by discourse factors like empathy, along the lines discussed in Chapter 1. I argue in Section 3.3 that the concept of exemption is the most fragile part of the RCT.

The case of binding out of locative PPs in (59b)-(59c) merits special attention. It is important to see that these PPs function as independent semantic predicates, in contrast to the ones in (60), whose complements integrate the semantic predicate of the verb. This is why RCB licenses pronominals in the former, but not in the latter:

(60)   a.   Bobby$_5$ relies on {*him$_6$ / himself$_6$}.

            Bobby ($\lambda x.$ x relies on x)

    b.   Bobby$_5$ talks to {*him$_6$ / himself$_6$}.

            Bobby ($\lambda x.$ x talks to x)

There are, however, cases where locative PPs do encode reflexive semantic predicates. In such cases RCB correctly demands reflexive marking, as we see in (61), which I previously discussed in connection to the Small Clause analysis of PPs in Chapter 1.

(61)   a.   Bobby$_5$ rolled the carpet$_6$ over {*it$_6$ / itself$_6$}.

            Bobby ($\lambda x.$ (x rolled the carpet) & the carpet ($\lambda y.$ y is over y))

---

[16] The requirement that syntactic domains for reflexivization have *subjects* (and not merely higher coarguments) is what distinguishes the RCT from HPSG's Binding Theory (POLLARD; SAG, 1994). The latter predicts both sentences in (i) to be equally unacceptable, while the RCT (assuming genitives are subjects of NPs) says (ia) should be worse than (ib). The judgments I gathered support the RCT (see also Menuzzi (1999, pg. 138)):

(i)   a.   *?*Tony* said that [Maria's letters about *himself*] were really frequent.
      b.   ?*Tony* said that [letters to Maria about *himself*] were really frequent.

However, experimental evidence in Asudeh and Keller (2001) and Runner and Kaiser (2005) argues against both analyses. What these studies show is that, in suitable contexts, speakers accept coconstrual even in cases like (ia). We can account for this keeping the RCT unchanged simply by assuming that genitives are not in fact subjects of NPs. This would entail that nominal projections are *never* syntactic predicates.

    b.   Julie$_1$ stepped on {*her$_1$ / herself$_1$}.

          Jule (λx. (x stepped) & (x is on x))

Despite being required by RCB, the reflexives in (61) are exempt from RCA for the same reason as those in (59b)-(59c): they are not arguments of syntactic predicates. The mismatch between syntactic and semantic predicates makes (61) a context where reflexives are both exempt and obligatory. It is the recognition of this mismatch which allows the RCT to overcome the problems the CBT faced in accounting for the distribution of pronouns within PPs.

The RCT account also has the virtue of eliminating a longstanding puzzle for theories of anaphora within MGG: the status of reflexives which undergo A′-movement.

(62)   [Which pictures of *herself*]$_6$ does Max think *Lucie* likes t_6?

What the CBT needs to say here is that there is a full copy of the moved *wh*-phrase which is locally c-commanded by *Lucie* in (62). Otherwise, Condition A of the CBT would be violated (CHOMSKY, 1977, 1993; MAY, 1977; HIGGINBOTHAM, 1983; BARSS, 1986). However, if this were true, we would need to see a contrast between (62) and (63), since the latter should be ungrammatical. This does not seem to be correct (POLLARD; SAG, 1994, pg. 248):

(63)   [Which picture of *himself*]$_2$ does *Max* think Lucie likes t_2?

Since reflexives within (subjectless) NPs are not restricted by locality in the RCT, there is no need to resort to any special reconstruction mechanism to explain (62)-(63). The reflexives therein are discursive occurrences. As such, they simply refer to whoever is construed as the empathy loci in context, regardless of structural factors.[17]

A potential objection to this kind of approach are cases like (64a), where a reflexive is

---

[17] Unlike what happens with NPs, when the fronted *wh*-phrase contains an AP, the reflexive must be coconstrued with the embedded subject, as Reinhart and Reuland's (1993, pg. 684) contrast in (i) illustrates:

(i)   a.   [How proud of *herself*]$_7$ does Max think *Lucy* is t_7?
        b.   *[How proud of *himself*]$_7$ does *Max* think Lucy is t_7?

Following Koopman and Sportiche (1991), Reinhart and Reuland argue that this happens because the internal structure of the moved AP contains, as its subject, a trace coindexed with *Lucy* — the assumption being that *Lucy* is base-generated as an external argument inside the AP and only gets moved to the subject the inflected verb later in order to satisfy the EPP. The full structure of (i) is, therefore, something like (ii):

(ii)   [How t_7 proud of {herself$_7$ / *himself$_2$}]$_1$ does Max$_2$ think Lucy$_7$ is t_1?

Since the AP headed by *proud* contains a subject (namely, t_7), in contrast to the NPs in (62)-(63), it does qualify as a syntactic predicate. A reflexive argument in (ii) only satisfies RCA if it is coindexed with this subject, giving rise to a reflexive syntactic predicate.

obligatory – in contrast with (59), where both reflexives and pronominals are acceptable.

(64) a. Rose$_9$ took a picture of {herself$_9$ / *her$_9$}.

  b. Rose ($\lambda x.\ x$ took-picture-of x)

The pattern in (64a) can be dealt with, however, if we assume the logical syntax outlined in (64b), according to which the semantics of *picture* is incorporated to the semantics of the light-verb *take* – as in the analysis of Culicover and Jackendoff (2005, chap. 6). Since the resulting predicate is semantically reflexive, using a pronominal would violate RCB (see Reinhart and Reuland (1993, pg. 685) for a slightly different implementation of the same idea).

I leave open the issue of whether the semantic incorporation in (64a) accompanies a syntactic incorporation of the N-complement into the syntactic predicate headed by *take*, as in Culicover and Jackendoff (2005). The fact that the complement of *picture* can be passivized (cf. (65a), from the internet) seems to suggest that this syntactic incorporation in fact occurs. This contrasts with regular complements of non-incorporated *picture* Ns (cf. (65b)).

(65) a. "My cosplay as a 'female' was taken pictures of by so many people!!!"

  b. *My cosplay as a 'female' was seen pictures of by so many people!!!

If the N-complement counts as part of the syntactic predicate headed by *take* (64a), this predicate is reflexive-marked. If not, (64a) would be another case where the reflexive is exempt from RCA *and* in complementary distribution with pronominals (the latter property due to RCB).

There is one class of structures where the asymmetry between semantic predicates (governed by RCB) and syntactic ones (governed by RCA) clearly yields this curious result, as it also did for the locative PPs in (61). These are examples involving conjunctions, such as (66):

(66) a. Bobby$_7$ praised both Marta and {himself$_7$ / *him$_7$}.

  Bobby ($\lambda x.\ (x$ praised Marta) & (x praised x))

RCB excludes pronominals in (66) (under a distributive interpretation) because the semantic structure of this example includes a reflexive predicate (i.e. *x praised x*)) which is not appropriately licensed by reflexive-marking in syntax. This is precisely why Reinhart and Reuland (1993, pg. 675) state RCB in terms of semantic representations, instead of syntactic ones. (The CBT also fails to predict this effect due to its sole concern with syntactic coindexing.)

Since the reflexive in (66) is not a syntactic argument, it does not reflexive-mark a SYN-TACTIC predicate. This is what makes it exempt from RCA in the RCT. Note, however, that the reflexive does reflexive-mark the SEMANTIC predicate linked to *praise* (i.e. it is one of its semantic arguments), so RCB is not violated if a reflexive is used in (66). Thus (66) is a context where exempt reflexives and pronominals are in complementary distribution.

Stating RCB as a condition on semantic predicates also allows the RCT to overcome another source of empirical embarrassment for the CBT I mentioned in Chapter 1: the distribution of non-reflexive anaphors. Consider the Dutch data in (68), drawn from Reinhart and Reuland (1993, pg. 665-666; 711). These cases were tricky for the CBT because *zich* behaves like the pronominal *hem* in contexts like (67) and like the reflexive *zichzelf* in contexts like (68).

(67)   a.   *Max* haat {*zichzelf* / **zich* / **hem*}.
            Max hates {REFL   / SE   / PRON}

       b.   *Max* praat   met {*zichzelf* / **zich* / **hem*}
            Max speaks with {REFL   / SE   / PRON}

       c.   *Henk* overreedde {*zichzelf* / **zich* / **hem*} te zingen.
            Hank persuaded {REFL   / SE   / PRON} to sing

(68)   a.   *Max* schaamt {*??zichzelf* / *zich* / **hem*}.
            Max shames {REFL   / SE / PRON}
            'Max is ashamed'

       b.   *Max* wast   {*zichzelf* / *zich* / **hem*}.
            Max washes {REFL   / SE / PRON}

       c.   *Henk* hoorde {*zichzelf* / *zich* / **hem*} zingen.
            Henk heard {REFL   / SE / PRON} sing.

These patterns come out more naturally in the RCT. Since *zich* is a NON-REFLEXIVE anaphor, RCB rules it out in contexts where a reflexive semantic predicate is formed: i.e. precisely those in (67). In these cases, reflexive-marking is obligatory and *zich* should behave just like a pronominal, given that both of these forms lack a reflexivizing function (cf. Table 2).

In contexts like (68), what we have in turn are predicates that are either lexically reflexive (e.g. (68b)) or not semantically reflexive at all (the object-raising case in (68c) and possibly (68a)). RCB does not require reflexive-marking by a reflexive argument in either of these scenarios. Therefore a non-reflexive anaphor like *zich* is acceptable.

Moreover, since non-reflexive anaphors are not subject to RCA, they carry no necessary locality requirement, as they did within the CBT. As Reuland (2011, pg. 291-292) observes, *zich* can in fact function as LONG DISTANCE ANAPHOR in restricted contexts such as non-object

positions within causative or perception verb infinitival complements:

(69)   a.   *Jan* hoorde Marie een lied voor {*zichzelf / *zich* / *hem*} fIuiten.
             Jan heard Marie a   song for {REFL   / SE / PRON} whistle
             '*John* heard Marie sing a song for *him*.'

       b.   *Jan* liet Marie {*zichzelf / *zich* / *hem*} een boek brengen.
             Jan let Marie {REFL   / SE / PRON} a   book bring
             '*John* let Mary bring *him* a book.'

In other languages, the long distance use of non-reflexive anaphors is less restricted, as we see in the following Icelandic and European Portuguese examples from Chapter 1:

(70)   a.   *Jón* segir   [að María elski   *sig*].
             John say.PRES.IND that Mary love.PRES.SBJV SE
             '*John* says that Mary loves *him*.'

       b.   *Maria* soube do   João [que alguém   tinha falado mal de *si*].
             Maria knew from.the João that somebody had   spoken bad of SE
             '*Maria* heard from João that somebody spoke badly of *her*.'

Though not a non-reflexive anaphor proper, the Frisian pronominal *hem* manifests the same pattern observed for Dutch *zich* in (67)-(68) (REULAND; REINHART, 1995):

(71)   a.   *Willem* bewûnderet {*himsels* / *him}.
             Willem admires   {REFL   / PRON}

       b.   Max pratet mei {*himsels* / *him}.
             Max speaks with {REFL   / PRON}

(72)   a.   *Willem* skammet {??*himsels* / *him*}.
             William shames {REFL   / PRON}
             'William is ashamed'

       b.   *Jack* wasket {*himsels* / *him*}.
              Jack washed {REFL   / PRON}

       c.   *Jan* seach [{*himsels* / *him*} in 'e film de partij winnen].
             Jan saw   {REFL   / PRON} in the film the match win
             '*John* saw *himself* win the match in the film.'

As we saw in Chapter 1, these facts are puzzling for the CBT: the same item seems to be subject to the local disjointness effects typical of pronominals in some contexts (cf. (71)) but not in others (cf. (72)). Since RCB is sensitive to inherent properties of predicates, the puzzle is largely eliminated. A reflexive pronoun is mandatory in (71) because the predicates therein are semantically reflexive and are not reflexive-marked in the lexicon. The opposite happens

in (72): Reuland and Reinhart (1995) argue that (72a)-(72b) are lexically reflexive-marked and that and (72c) is a case where there is no reflexive semantic predicate in the first place.

Other aspects of the RCT account of the distribution of pronouns in (67)-(72) seem less obvious and merit further discussion. As we've seen, proponents of the RCT claim that the predicates headed by *schaamt/skammet* ('be ashamed') in (68a)/(72a) and *wassen/wasket* ('wash') in (68b)/(72b) are INTRINSICALLY REFLEXIVE – i.e. they are "marked [as reflexive] in the lexicon" (REINHART; REULAND, 1993, pg. 662). In Chapter 3 I argue that this concept of lexical reflexivity is suspicious. One of its problems is that it lumps together predicates that differ from each other when it comes to the possibility of accepting a reflexive: *zichzelf* and *himsels* are very marginal in (68a)/(72a) but completely acceptable in (68b)/(72b).

The RCT handles this difference by stipulating that verbs like *wassen/wasket* ('wash') are listed twice in the lexicon: in their reflexive entries, they allow the non-reflexives *zich/him* and exclude *zichzelf/himsels*; in their transitive entries, a reflexive is required to comply with RCB. The verbs *shaamt/skammet* ('shame'), on the other hand, only have a reflexive entry. Using this intrinsically reflexive entry with a syntactic reflexive like *zichzelf/himsels* thus results in redundant reflexive marking, which is penalized (REINHART; REULAND, 1993, pg. 667).

The idea that certain predicates can be reflexive marked in the lexicon is abundantly employed in more recent versions of the RCT to explain languages in which local binding of non-reflexives is freely allowed. The data in (73) illustrates the phenomena with Traditional Jambi Malay pronominals (COLE; HERMON; YANTI, 2015, pg. 148) and (74) with the non-reflexive SE anaphor in European Portuguese (MENUZZI; LOBO, 2016, pg. 344).

(73)   Yanti pikir [*Arna* cinto *dio*].

         Yanti thinks Arna loves PRON

         'Yanti thinks *Arna* loves *her(self)*.

(74)   Pedro acredita [que *Roberto* torce por *si*].

         Pedro believes that Roberto roots for SE

         'Pedro believes that *Roberto* roots for *himself*.'

In both (73) an (74) we see reflexive semantic predicates that are not reflexive marked in syntax (i.e. that do not take a reflexive as an argument). From the point of view of RCB, it does not make a difference whether the non-reflexive argument is a pronominal or an anaphor – this is exactly what allows the RCT to predict the Dutch and Frisian contrasts in (67)-(68) and

(71)-(72). RCB only cares cares about the reflexive-marking property which defines reflexives.

Therefore, the only way avoid the conclusion that (73) and (74) violate RCB is to claim that the predicates in these examples are reflexive marked in the lexicon. This is the tack that Reuland (2011, 2016) takes. Following Reinhart and Siloni (2005), he argues that what allows non-reflexive pronouns in contexts like (73) an (74) the presence of a generalized LEXICAL BUNDLING parameter. What lexical bundling does is combine two of a predicate's semantic roles into one complex role while also reducing the arity of the original predicate. In Dutch and Frisian, bundling is restricted to grooming predicates like *wassen/wasket* ('wash'). In Traditional Jambi Malay and European Portuguese, however, bundling applies freely. This guarantees that all transitive verbs in these language have intrinsically reflexive variants.[18]

Though some of the RCT analyses sketched above may not be entirely convincing, they serve to illustrate that the RCT makes a considerable effort to overcome the empirical deficiencies of the CBT. In fact, the RCT successfully predicts the possibility of discursive reflexives, the distribution of non-reflexive anaphors, the existence of locally bound pronominals and the behavior of pronouns within PPs – all of which escaped the grasp of the CBT.

But there is still one set of facts that the CBT predicts and that the principles of the RCT stated so far are silent about: namely, the unavailability of locally bound pronominals in (68). From the point of view of the CBT, pronominals cannot be used in these structures because they would be locally c-commanded by their antecedents, violating Condition B. RCB of the Reflexivity Theory, in its turn, says nothing about these cases. Since the semantic predicates in (68) are either reflexive-marked in the lexicon or not reflexive at all, no reflexive marking in syntax should be necessary. Hence, as far as RCB is concerned, both *zich* and *hem* (which are not reflexives) should be fine. But only the former is acceptable. This becomes especially puzzling when we compare (68) to the Frisian examples in (72), where locally bound pronominals are in fact acceptable. The parts of the RCT presented thus far do not explain this difference.

The dilemma is this: since the Reinhart and Reuland (1993) state RCB in terms of semantic representations, RCB loses the capacity to explain purely syntactic anti-locality effects that fell under Condition B of the CBT. This also crops up in the following English examples:

---

[18] On the face of it, this seems to make RCB unfalsifiable. However, Reuland (2011) suggests that bundling has morphosyntactic reflexes. In Romance languages, SE anaphors serve as overt markers of lexical bundling (BAAUW; DELFITTO, 2005). For Traditional Jambi Malay, Reuland (2016) hypothesizes that bundling may be related to the absence of overt voice morphology (but see Yanti, Cole and Hermon (2017)).

(75)   a.   Bobby$_2$ believes {*him$_2$ / himself$_2$} to be a genius.

            Bobby ($\lambda x.$ x believes (x is a genius))

       b.   Bobby$_9$ appears to {*him$_9$ / himself$_9$} to be a genius.

            Bobby ($\lambda x.$ appear to x (x is a genius))

The semantic predicates in (75) are not reflexive, so nothing needs to be reflexive marked. Therefore, RCB, by itself, does not explain why a pronominal is not allowed in these structures. These and other purely syntactic residues CBT's Binding Conditions are subsumed under the second part of the RCT, to which I will now turn: the generalized theory of A-chains.

2.2.2   Chain Theory

The principle to which the RCT attributes the unavailability of pronominals in (68) and (75) is not a constraint on predicates but on CHAINS – i.e. the same sort of formal object employed to represent movement dependencies. The suggestion that anaphora and movement are related is, of course, not new. It was one of the founding themes of the Conditions Framework of Chomsky (1973) and also an important feature of the binding theoretic account of the distribution of NP traces in Chomsky (1981) and Aoun (1986).

The Chain Theory is an attempt to restore what was fundamentally right about these proposals: namely, the generalization that, in the syntactic domain where an NP can (syntactically) bind its trace, an NP can never bind a non-anaphor (i.e a +R element), as (76)-(78) illustrate.

(76)   a.   Felix$_3$ was fired t_3.

       b.   Felix$_1$ fired {himself$_1$ / *him$_1$}.

(77)   a.   He$_4$ is believed [t_4 to be smart].

       b.   He$_7$ believes [{himself$_7$ / *him$_7$} to be smart].

(78)   a.   Felix$_8$ was [t_8 expected to be considered [t_8 smart]].

       b.   Felix$_2$ expects [{himself$_2$ / *him$_2$} to consider [{himself$_2$ / *him$_2$} smart]].

The generalization stated above can only be expressed in a principle of grammar if the syntactically bound NPs in (76)-(78) are regarded as objects of the same type. Chomsky (1981) did this by classifying NP traces as anaphors. The RCT adopts a slightly different tactic and extends the concept of chain, originally designed to represent movement, to include sequences of coconstrued NPs as well. A somewhat simplified version of Reinhart and Reuland's (1993, pg. 693) definition is given in (79):

(79)   C is a CHAIN *iff* C is the maximal sequence of links $(\alpha_1, ..., \alpha_n)$ such that, for all j, α_j antecedent governs α_{j+1}, where δ ANTECEDENT GOVERNS γ *iff*:

    a. δ and γ are coindexed;

    b. δ c-commands γ, and

    c.   there is no barrier between δ and γ.

The definition above does not require chains to include only one θ-marked position, as in Chomsky (1986a). According to (79), a chain is simply a sequence of positions where each link (except for the head link) satisfies antecedent government. In this sense, chains characterize not only dependencies formed by movement (the a-cases in (76)-(78)), but also purely representational dependencies formed by way of A-binding (the b-cases in (76)-(78).).

Unifying movement and anaphoric relations under the concept of chain allows the patterns in (76) to be explained by a single principle which is neither specific to movement nor to anaphora. With this in mind, Reuland and Reinhart (1995, pg. 255) propose the following condition on A-chains (i.e. chains which are headed by argument positions):

(80)   CHAIN CONDITION (CC)

    A maximal A-chain $(\alpha_1, ..., \alpha_n)$ contains exactly one link – α_1 – which is +R.

To see how the CC accounts for (76)–(78) and the data mentioned in the end of the last section, it is useful to disentangle it into two separate demands (MENUZZI, 1999, pg. 45):

(81)   a.   CHAIN VISIBILITY: The head of an A-chain must be +R.

    b.   CHAIN ECONOMY: Non-head links of A-chains must be -R.

According to (79), the sequences of coindexed NPs in (76)-(78) all count as A-chains. Chain Economy determines that only anaphors and NP traces are allowed to tail the chains in these structures, because only they are -R (i.e. referentially dependent). Pronominals, in turn, cannot be used in (76b)/(77b)/(78b), because doing so would result in an A-chain where at least one non-head link is +R. The pronominal in (76b) is also ruled out by the RCB of the Reflexivity Theory ((50) above), since it does not reflexive-mark a reflexive semantic predicate.

An explanation in terms of Chain Economy is also applicable to the Dutch and English examples I brought up near the end of the last section, which I partly repeat below.

(82) a.   Bobby$_2$ believes {himself$_2$ / *him$_2$} to be a genius.

b.   Bobby$_9$ appears to {himself$_9$ / *him$_9$} to be a genius.

(83) a.   Max$_4$ wast   {zich$_4$ / zichzelf$_4$ / *hem$_4$}.
Max washes {SE   / REFL   / PRON}

b.   Henk$_3$ hoorde {zich$_3$ / zichzelf$_3$ / *hem$_3$} zingen.
Henk heard   {SE   / REFL   / PRON} sing.

In these cases, using a pronominal would yield an ill-formed A-chain whose tail is +R. Only -R elements like *himself*, *zich* or *zichzelf* are allowed in these positions. This means that, even though *zich* is grouped with pronominals with respect to RCB (since both lack a reflexivizing function), it is grouped with reflexives with respect to CC (since it is also an anaphor).[19]

The data in (82)-(83) shows that, even when RCB is not at stake, anaphors and pronominals are in complementary distribution in the tail position of chains. Non-complementarity thus only truly emerges when no chain or reflexive predicate is formed. This is what happens when pronouns are arguments of locative Ps in English and Dutch (REULAND, 2011, pg. 279):[20]

(84) a.   Curly $_1$ saw a snake near {him$_1$ / himself$_1$}.

b.   Curly$_5$ zag een slang naast {hem$_5$ / zich$_5$}.
Curly saw a   snake near {PRON / SE}

The Chain Theory also proves useful to derive some other effects captured by the CBT which the Reflexivity Theory, in and of itself, cannot anticipate. Since the concept of reflexive-marking is essentially non-configurational (i.e. it merely requires that a reflexive be an *argument* of a predicate), neither RCA nor RCB alone can predict the unacceptability of (85) (BARSS; LASNIK, 1986; LARSON, 1988; JACKENDOFF, 1990a).

---

[19] As non-reflexive anaphors, SE forms like *zich* are *only* subject to the CC. Reflexive anaphors, however, are subject to both the CC and RCA. The reflexives in in (83) and (82b) do satisfy RCA because the syntactic predicates they reflexive mark are indeed reflexive. The situation for (82a) is more complicated, as Reinhart and Reuland (1993, pg. 707) recognize. Since *himself* is assigned a grammatical function by both *expect* and *be a genius* (i.e. it is the object of the former and the subject of the latter), *both* of these predicates are reflexive-marked. RCA thus requires both of them to be reflexive – a condition which is only met for the predicate headed by *expect*. Reinhart and Reuland solution is to claim that, in such cases, the object-raising verb and the embedded predicate are amalgamated to form a single complex predicate in LF, as in (i):

(i)   Bobby$_3$ [[believes [to be a genius]$_1$] t_1 himself$_3$].

If the relevant syntactic predicate that *himself* reflexive marks in (82b) is the single complex predicate headed by *believe to be a genius*, RCA is satisfied. In Chapter 5, I propose an alternative treatment for reflexives which does not need to appeal to covert operations.

[20] Following Chomsky (1986a), Reinhart and Reuland (1993) argue that locative Ps create a minimality barrier because they introduce an independent thematic domain.

(85) a. *Himself$_3$ recommended Bobby$_3$.

b. *Marta showed himself$_7$ to Bobby$_7$

c. *Bobby$_5$ said [that Marta showed himself$_5$ to himself$_5$].

d. *Joanne can't imagine [himself$_1$ praising Bobby$_1$].

e. *Marta$_2$ said [that herself$_2$ defended herself$_2$].

In each of these examples, a reflexive-marked syntactic predicate is reflexive (*as per* RCA) and a reflexive semantic predicate is reflexive marked (*as per* RCB). The problem is, of course, that reflexives in these structures are in the wrong position. However, unlike the Binding Conditions of the CBT, neither RCA nor RCB say anything about the positions reflexives have to occupy (e.g. that they should be preceded or c-commanded by their antecedents).

In the RCT, the explanation for these structural residues of the CBT is taken up by Chain Visibility: in all of the examples in (85), what we have are A-chains whose heads are -R.[21] In this sense, these examples violate the CC for the same reason as (86).

(86) *Harold$_9$ seems that [t_9 was kissed t_9 by Marian].

Moreover, since the presence of a single reflexive argument suffices to reflexive mark a predicate once and for all, RCB cannot not explain why a pronominal is not acceptable in the English and Dutch three-place predicates below (REINHART; REULAND, 1993, pg. 691):

(87) a. He$_6$ (accidentally) assigned {*him$_6$ / himself$_6$} to himself$_6$.

b. Henk$_3$ wees   {*hem$_3$ / zich$_3$} aan zichzelf$_3$ toe.
    Henk assigned {PRON / SE}   to REFL   PARTICLE

The predicates in (87) are reflexive marked, as required by RCB. The reason why an anaphor is nonetheless obligatory as the intermediate arguments in (87) is Chain Economy: pronominals, *qua* +R elements, are not allowed in non-head positions of chains.

All in all, the Chain Theory is a concession to the thesis that some facts about the distribution pronouns require reference to purely hierarchical constraints. However, the RCT attempts to motivate these constraints by relating them to conditions which are independently required to explain A-movement. This is made possible by adopting a purely representational definition of chains. The unification of A-movement and A-binding under a generalized conception of

---

[21] Examples (85a), (85b) and (85d) violate Chain Economy in addition to Chain Visibility because they also contain a +R NPs that serve as tails of their respective A-chains.

chains is also proposed, within a different (more derivational) framework, by Lidz and Idsardi (1998), Hornstein (2001) and Boeckx, Hornstein and Nunes (2007).

The Reflexivity part of the RCT determines the distribution of pronouns according to whether they carry the reflexivizing function property. The relevant distinction for RCA and RCB is, accordingly, the one between reflexives and non-reflexives. The Chain Theory, on the other hand, regulates the distribution of NPs on the basis of the R property. This means that the CC only cares about the distinction between anaphors (-R) and non-anaphors (+R).

Up to now, I have been assuming that R is equivalent to the semantic property of referential independence. However, this way of seeing things makes the Chain Economy part of the CC seem conceptually suspicious. On the one hand, it looks like an attempt to sneak into the RCT a *doppelgänger* of CBT's Condition B: both principles end up being constraints against the local binding of non-anaphors (SAFIR, 2004, pg. 18). On the other hand, as we already saw for (76b) above, most of the cases excluded by the Chain Economy are also excluded, on independent grounds, by RCB of the Reflexivity Theory.

The rationale for keeping Chain Economy and RCB as independent principles is largely based on cases where there is a mismatch between syntactic and semantic predicates: structures like conjoined NPs (cf. (88a)) and raising predicates (cf. (88b)).

(88)   a.   *Bobby$_7$ praised both Marta and him$_7$.

          b.   *Bobby$_2$ believes him$_2$ to be a genius.

Pronominals can only be excluded from (88a) by something like RCB: a principle which is sensitive to reflexivity in *semantic* structure. To rule out (88b), on the other hand, we need something like the CC: a purely syntactic principle disallowing local antecedents for +R elements, regardless of whether reflexive semantic predicates are formed or not. Though this is decent motivation for keeping the principles separate, the partial redundancy between the CC and RCB across most of the data does look theoretically unwholesome, to say the least.

Furthermore, given that the CC as stated so far mimics closely the effects of CBT's Condition B, it stumbles upon many of the same empirical difficulties as the latter. Locally bound pronominals in Frisian and Traditional Jambi Malay, for instance, continue to pose a problem (REULAND; REINHART, 1995; COLE; HERMON; YANTI, 2015):

(89)   a.   *Jack* wasket *him*.
               Jack washed PRON

      b.   *Jan* seach [*him* in 'e film de partij winnen].
               Jan saw   PRON in the film the match win
               '*John* saw *him(self)* win the match in the film.'

(90)   a.   *Yanti* neŋoʔ *dio*   di tipi.
               Yanti saw   PRON on TV

      b.   Yanti pikir   [*Arna* cinto *dio*].
               Yanti thinks Arna loves PRON

In (89)/(90), we have what appear to be a +R elements (i.e. pronominals) occupying non-head position of an A-chain. This would be a violation of Chain Economy.

In order to give the CC more conceptual justification and predictive force, avoiding the problems raised by (89)/(90), the RCT argues that it is necessary to reduce R to a purely syntactic property. This property should be, at once, one that underlies an NP's capacity for independent reference and its behavior with respect to the formation of A-chains. Reinhart and Reuland (1993, pg. 697) pose the issue as follows:

> It is not the case that referential properties of NPs miraculously restrict their syntactic behavior; rather, some independent syntactic properties of NPs determine how they can be used to refer. Thus, R should be a purely syntactic property. Having this property is a necessary condition for an expression to function as an independent argument, but R itself does not have anything to do with reference. What could it be?

Following Bouchard (1984), Reinhart and Reuland (1993) propose that R – which is treated as a primitive property in many theories (CHOMSKY, 1981; BURZIO, 1998; SAFIR, 2004) – is in fact determined by the ϕ-feature specification of NPs:

(91)   An NP is +R iff it carries a full specification for ϕ-features (including structural Case).

The hypothesis above gives the RCT a strategy for tackling facts concerning locally bound pronominals without abandoning the CC: locally bound pronominals should be possible only if the pronominals in question are deficient in ϕ-features – i.e. if they have a featural profile akin to that of an anaphor. Proponents of the RCT argue that this is precisely what happens in Frisian and Traditional Jambi Malay (REULAND; REINHART; 1995; REULAND 2011, 2016).

Consider the latter case first. According to Reuland (2016), like the Dutch anaphor *zich*, pronominals in Traditional Jambi Malay lack a number distinction. This situation contrasts with the one found in the dialect of Jambi spoken in Jambi City, known as Jambi City Malay.

(92) a. Dio   kagum samo Budi.   (Traditional Jambi Malay)
         PRON admire with Budi
         '{She / he / they} admire Budi.'

     b. Dio   nengok Eko di sekolah.   (Jambi City Malay)
         PRON.SG saw   Eko in school
         '{He / she / *they} see Eko in school.'

The Traditional Jambi Malay *dio* pronominal in (92a) is ambiguous in a way that the Jambi City Malay variant in (92b) is not. This indicates that pronominals in Traditional Jambi Malay are poorer in ϕ-features than the ones in Jambi City Malay: the former lack a number feature which is present in the latter. If (91) is correct, this contrast regarding number specification should correlate with the (im)possibility of local binding. This is in fact what we find:

(93) a. *Budi* mukul *dio*.   (Traditional Jambi Malay)
         Budi hit   him.

     b. **Eko* muji   *dio*.   (Jambi City Malay)
         Eko praise him.

The binding difference between the two variants of Jambi comes out as a byproduct of a lexical difference in ϕ-feature specification. Local binding of pronominals in (93a) does not violate the CC because *dio* is ϕ-deficient in Traditional Jambi Malay. Since Jambi City Malay *dio* is more fully specified, Chain Economy rules out local binding in (93b)[22]

Building on earlier work by Hoekstra (1994), Reuland and Reinhart (1995) propose an explanation along similar lines for the Frisian data in (89). According to them, the Frisian pronominal *hem* is specified for inherent Case and unspecified for structural Case. Assuming that only the latter counts as a ϕ-feature for the purposes of chain formation, *hem* comes out as ϕ-deficient, and, thus, as -R. This means that (89) does not imply a violation of the CC. [23]

---

[22] Note, however, that pronominals in Jambi City Malay are also unspecified for gender. Reuland (2016) does not provide any rationale for why lack of specification for gender does not suffice to count a form as -R.

[23] Van Gelderen (2000) contends that this is also why locally bound pronominals were allowed in Old English, as the following attested data illustrates (see also Reuland (2011, pg. 284)):

(i) a. *hine*   *he* bewerað mid wæpnum
         him(self) he defended with weapons
         '*he* defended *himself* with weapons'   (KÖNIG; SIEMUND, 2000a, pg. 44)

     b. hweðer *he hine* gefreclsian wolde
         whether he him set-free   would
         'whether *he* would set *him* free'   (KEENAN, 2002, pg. 331)

     c. *he hine* to guðe gegyred hæfde
         he him to battle girded   had
         '*he* had made *himself* ready for battle'   (VAN GELDEREN, 2000, pg. 36)
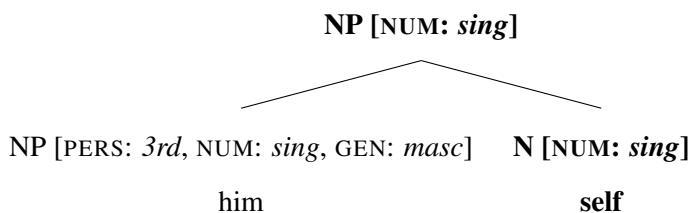
The transition from Old and Middle English to Modern English is attributed to the gradual loss of inherent

The characterization of -R forms as ϕ-deficient also neatly accounts for the behavior of the Dutch pronouns *zich* and *zichzelf*, which are only specified for 3rd person. English reflexives, however, present a challenge. Forms like *himself* and *herself* seem to be fully specified for number, person and gender and case, just like their pronominal counterparts *him* and *her* (BURZIO, 1991, pg. 96). Nonetheless, they do not trigger any kind of Chain Economy effects.

Reuland and Reinhart (1995) try to resolve this apparent contradiction by stipulating that, since English reflexives lack a case contrast (i.e. there are no forms *heself*, *theyselves*, etc.), they are unspecified for Case, much like the Frisian pronominals. ϕ-deficiency follows if case counts as a ϕ-feature, as in Chomsky (1981). Anagnostopoulou and Everaert (1999) argue, however, that the deficient case paradigm of English reflexives (i.e. the absence of a nominative/accusative opposition) can hardly be an argument for anything – it is, rather, an *explanandum* in its own right.[24]

A more promising line of reasoning says that the ϕ-deficiency of English reflexives derives, not from their inherent features, but from their morphosyntactic composition. Anagnostopoulou and Everaert (1999) and Menuzzi (1999), following Helke (1970), argue that English reflexive forms are compounds which are headed by (and, thus, inherit their ϕ-deficiency from) the *self* morpheme, which is only specified for number (see also Chomsky (1981, pg. 102)). Building on Postal (1966b), these authors claim that the pronominal part of English reflexives is a specifier-like element, as illustrated in (94). Note that the whole reflexive form inherits its poor ϕ-feature profile from its head, indicated in boldface:[25]

(94)   **NP [NUM: *sing*]**

          NP [PERS: *3rd*, NUM: *sing*, GEN: *masc*]   **N [NUM: *sing*]**

                    him   **self**

The representation in (94) treats forms such as *himself* as an instance of what Faltz (1985, pg. 29) calls HEAD REFLEXIVES. It goes against a large body of work which treats English

---

Case, which is assumed to imply the emergence of structural Case. Since Modern English pronominals became specified for structural Case, they came to qualify as +R and trigger Chain Economy effects.

[24] A similar reification of the facts is also advanced in Pollard and Sag's (1992, pg. 290) Binding Theory as an explanation for the ungrammaticality of subject reflexives in *Himself loves John* (cf. (85a)).

[25] The idea that English reflexives are morphosyntactically complex has a long history – hints of it are found as far back as Chomsky (1955, chap. 9) and Langacker (1969). The approaches mentioned here differ in subtle details of implementation which I shall gloss over. What they share is the basic intuition that the pronominal part of reflexives may be fully specified for ϕ-features without the whole composite structure being specified.

reflexives as undecomposable units (e.g. Pollard and Sag (1992), Safir (2004), Hicks (2009), Rooryck and Vanden Wyngaerd (2011), Charnavel (2019)).

There is, however, good positive evidence in favor of the view that English reflexive anaphors are morphosyntactically complex. Note that *self* can be used as a regular independent noun in structures like (95) (taken from the internet). In (95a) in particular, *self* incontrovertibly heads its own nominative NP, which shows that *self* is neutral with respect to case, given that it has the same form in both nominative and accusative.

(95) a. "Mary's self is celebrated because of her contributions to her country."

  b. "I don't even know my damn self."

  c. "You must understand that we should love our own selves."

  d. "How to get your daughter to accept her amazing self."

The examples in (95b)-(95d) are interesting because they contain the same possessive forms that are found in standard reflexives (*myself*, *herself*, *ourselves*). As a matter of fact, there is no reason to see these cases as substantially different from standard reflexives – all of them encode reflexive predicates. The observation that (some) reflexives are prefixed by possessives again points to the conclusion that the *self* morpheme they contain is a regular noun.

The fact that both the pronominal part of reflexives as well as the *self* morpheme can be contrastively stressed independently of each other also favors a decompositional analysis. As Sauerland (2008) notes, some contexts favor focus on *self* (e.g. (96)) while others favor focus on the pronominal (e.g. (97)). (Jacobson (2000) and Spathas (2010) give similar examples.)

(96) a. ??John likes Mary, but $Bill_7$ likes $HIMself_7$.

  b. John likes Mary, but $Bill_4$ likes $himSELF_4$.

(97) a. $John_9$ likes $himself_9$ and $Bill_1$ likes $HIMself_1$.

  b. ??$John_2$ likes $himself_2$ and $Bill_6$ likes $himSELF_6$.

Moreover, as Stowell (1996) observes, in headline-style English, determiners and other specifier-like elements tend to be omitted. In this respect, the pronominal part of the reflexive behaves just like an ordinary syntactic specifier, as the data in (98) shows. If *himself* was an undecomposable unit (as in most approaches to reflexives), (98c) should be perfectly acceptable in headline-English (ANAGNOSTOPOULOU; EVERAERT, 1999, pg. 114).

(98)   a.   Clinton$_5$ describes himself$_5$ as smarter than his$_5$ wife.

         b.   Clinton$_1$ describes self$_1$ as smarter than wife.

         c.   ??Clinton$_9$ describes himself$_9$ as smarter than wife.

Another (perhaps marginal) argument which supports an analysis in which reflexives are headed by *self* is the following: If reflexives were headed by the pronominal part of their morphology, we would expect (99a) to be slightly less deviant than (99b), since verb agreement would be established on the basis of the second person of the subject. Though both sentences are clearly ungrammatical, the opposite seems to be the case. This suggests that the agreement features that project to the whole reflexive NP are not those associated with *you*, but with *self* (ANAGNOSTOPOULOU; EVERAERT, 1999, pg. 113).

(99)   a.   *You$_4$ think yourself$_4$ work too hard.

         b.   *?You$_8$ think yourself$_8$ works too hard.
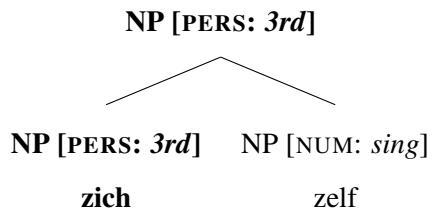
Furthermore, as Menuzzi (1999, pg. 241) points out, the use of reflexive anaphors in object position implies that such forms inherit their distributional properties from the *self* morpheme, rather than from the pronoun they contain. Otherwise, forms prefixed by possessive pronominals, such as *myself*, *herself* and *ourselves*, could only occur only as possessives. This also favors the conclusion that *self* is the head of reflexives, as proposed in (94).

The analysis proposed here entails that *self* does not, on its own, constitute a phrasal projection (i.e. an NP). This also explains why in Modern English, only reflexive forms as a whole exhibit an adnominal emphatic use (e.g. *Sue herself sang* vs. **Sue self sang*). This contrasts with what we find in other Germanic languages (SIEMUND, 2002). The adnominal use of bare *self*-forms is illustrated with the German, Dutch and Old English examples below:
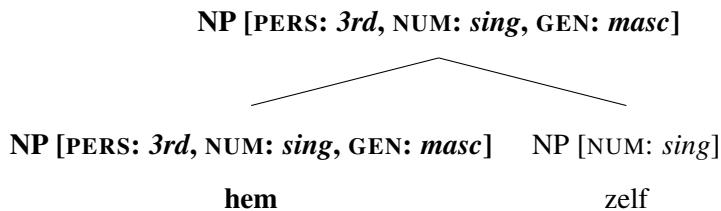
(100)   a.   Lizzy barbierte den Vater selbst.
               Lizzy shaved   the father SELF
               'Lizzy shaved the father himself.'   (EDMONDSON; PLANK, 1978, pg. 374)

         b.   Betty waste   Theo zelf.
               Betty washed Theo SELF
               'Betty washed Theo himself.'   (GEURTS, 2004, pg. 5)

         c.   Wolde self   cyning symbel þicgan.
               wanted SELF king   feast   consume
               'The king himself wanted to eat.'   (VAN GELDEREN, 2000, pg. 40)

These examples suggest that *self*-forms in other Germanic languages are full phrasal projections that attach to their host NPs as adjuncts, forming [NP NP] structures. This yields interesting consequences for *self*-like elements that attach to anaphor and pronominal NPs. For instance, the underlying morphosyntactic structures of the Dutch pronouns *zichzelf* and *hemzelf* should be something along the lines of (101). Both (101a) and (101b) contrast with (94) in that the complex forms are headed by (and inherit their ϕ-feature profile from) the host pronoun NP instead of from *zelf*. Boldface represents the sequence of projected heads:

(101) a.

NP [PERS: *3rd*]

NP [PERS: *3rd*]   NP [NUM: *sing*]
zich   zelf

  b.

NP [PERS: *3rd*, NUM: *sing*, GEN: *masc*]

NP [PERS: *3rd*, NUM: *sing*, GEN: *masc*]   NP [NUM: *sing*]
hem   zelf

This analysis predicts that *hemzelf*, notwithstanding its surface similarity to the English reflexive *himself*, should be ruled out by Chain Economy in local binding environments for the same reason as the simple pronominal *hem* is. This is in fact what we find:

(102) a. Jan$_4$ haat {*hemzelf$_4$ / *hem$_4$ / zichzelf$_4$}.
           Jan hates {PRON-SELF / PRON / REFL}

  b. Jan$_2$ schoot op {*hemzelf$_2$ / *hem$_2$ / zichzelf$_2$}.
      Jan shot   at {PRON-SELF / PRON / REFL}

(MENUZZI, 1999, pg. 47)

There are, in sum, fair reasons to view complex *self*-forms (in English and other Germanic languages) as morphosyntactically composite. If we assume that the ϕ-features of heads are passed up to their maximal projections (as is necessary to account for the fact that NPs like *the boys* are plural), this view yields interesting typological predictions, corroborating a fundamental thesis of the RCT: namely, the idea that the morphosyntactic make-up of NPs determines their behavior with respect to anaphora.

2.3 CONCLUSION

In the present chapter, I have argued that the reflexivity-and-chains theory (RCT) of Reinhart and Reuland (1991, 1992, 1993, 1995) is, overall, a more promising approach to anaphora than the two earlier proposals discussed in Chapter 1: the transformational theory of Lees and Klima (1963) and the Classical Binding Theory (CBT) of Chomsky (1981, 1986b).

On a more abstract level, the RCT is based on a much more solid understanding of the relationship between syntax and semantics, which is built into Reinhart's (1983) theory of bound variable anaphora. The specific empirical virtues of the RCT stem from the fact that it takes into account the specific properties of predicates (captured by the Reflexivity Theory) as well the relation between syntactic dependencies and the morphosyntactic constitution of NPs (captured by Chain Theory). Moreover, insofar as the RCT does not essentially rely on a thematic hierarchy or on intricate phrase-structure-related notions of licensing, it also opens the possibility of greatly simplifying the structures that a theory of anaphora requires from syntax – a property which makes it particularly useful as a kind of stepping stone for the constructional reduction undertaken in Chapters 5-7.

The next chapter goes over many of the points discussed in this chapter and shows that, in spite of all of its virtues, the RCT presents conceptual and empirical flaws that call for an approach to anaphora which is different from the ones we have seen so far.

3 PERSISTENT PROBLEMS

Reinhart and Reuland's (1991, 1992, 1993) reflexivity-and-chains theory (RCT) represents what is arguably the last major breakthrough in the study of anaphora within theoretical linguistics, averting many of the pitfalls that previous approaches fell into. The Reflexivity part of the RCT explains facts concerning discursive reflexives, locally bound pronominals within PPs and the distribution of non-reflexive anaphors. Chain Theory, in turn, (particularly, the reduction of referential (in)dependence to ϕ-feature specification) explains cases of locally bound pronominals in languages such as Frisian and Traditional Jambi Malay.

In spite of these advantages, the RCT faces some empirical difficulties – some of which are shared with previous approaches to anaphora. First, there is a set of familiar problems that gravitate around the formulation of the Chain Theory. For instance, since the Chain Economy part of the Chain Condition countenances the generalization expressed by Condition B of the CBT, it struggles with data concerning locally bound pronominals in languages like Brazilian Portuguese. Second, RCT's notion of REFLEXIVE MARKING does not adequately capture the fact that some reflexive predicates are liberated from the requirement of having overt reflexive arguments. Third, the RCT goes amiss when when it tries to push the Two Reflexives Hypothesis (TRH) as a solution to the problem of discursive reflexives. By letting syntax define when reflexives can be interpreted in accordance to discursive constraints, the RCT makes wrong predictions about the distribution of non-local reflexives in English.

The purpose of this chapter is to offer a diagnostic of these problems. By so doing, I do not wish to imply that the RCT does not incorporate other mistaken assumptions that further compromise its empirical adequacy. More general critiques of the RCT framework can be found in Menuzzi (1999), Safir (2004), Rooryck and Vanden Wyngaerd (2011) and Charnavel (2019). The choices I make here reflect my interest in discussing solely the problems for which I feel capable of offering a more adequate solution – a task to be undertaken in Chapters 5–7.

Section 3.1 goes over some of the conceptual and empirical difficulties faced by the RCT's Chain Theory – e.g. the status of Chain Visibility, Chain Economy and the role of indices. I ultimately argue that the Chain Theory is unnecessary for the theory of anaphora. Section 3.2 discusses RCT's problematic notion of reflexive-marking. Section 3.3 closes the chapter discussing the problems with the RCT's treatment of reflexives – in particular, the difficulties it runs into by adopting what I call the Two Reflexives Hypothesis (TRH).

3.1 PROBLEMS WITH THE CHAIN THEORY

In this section, I examine a set of interrelated problems faced by the Chain Theory component of the RCT. For convenience, I repeat the relevant definitions below:

(1) C is a CHAIN *iff* C is the maximal sequence of links $(\alpha_1, ..., \alpha_n)$ such that, for all j, α_j antecedent governs α_{j+1}, where δ ANTECEDENT GOVERNS γ *iff*:

    a. δ and γ are coindexed;

    b. δ c-commands γ, and

    c.   there is no barrier between δ and γ.

(2) CHAIN CONDITION (CC)

A maximal A-chain $(\alpha_1, ..., \alpha_n)$ contains exactly one link – α_1 – which is +R.

Section 3.1.1 starts by tackling the dubious role of CHAIN VISIBILITY (i.e. the requirement that heads of chains be +R). Section 3.1.2 does the same for CHAIN ECONOMY (i.e. the requirement that non-head links of chains be -R), also pointing out how RCT's understanding of referential independence (R) is too narrow to capture the behavior of complex emphatic forms like *ele mesmo* ('him same') in Brazilian Portuguese. Section 3.1.3 discusses the conceptual and empirical troubles implicated in the CC's essential use of syntactic indices and points out some difficulties for the most popular alternative to indexing – namely, the one that uses identity of ϕ-features as a syntactic surrogate for semantic identity (POLLARD; SAG, 1994; CHOMSKY, 2007, 2008; HEINAT, 2006; REULAND, 2011, ROORYCK; VANDEN WYNGAERD, 2011). Ultimately, the Chain Theory part of the RCT will be shown to be unnecessary.

3.1.1 Chain Visibility

In Chapter 2, we saw that the CC imposes two logically independent constraints on A-chains. These constraints are stated in (3) using the terminology of Menuzzi (1999, pg. 45):

(3)   a.   CHAIN VISIBILITY: The head of an A-chain must be +R.

    b.   CHAIN ECONOMY: Non-head links of A-chains must be -R.

My previous discussion focused mostly on Chain Economy, which is responsible for handling the syntactic remnants of CBT's Condition B. Chain Visibility was invoked solely to account for non-c-commanded anaphors with local antecedents, such as the ones we find in (4):

(4) a. **Bobby* said [that Marta showed *himself* to *himself*].

   b. **John* said [that *himself* defended *himself*].

However, in the current system, Chain Visibility has broader consequences than what (4) may suggest. It also excludes non-local binding of reflexive anaphors in (5):

(5) a. **Harold* said [$_{S'}$ that Marian loves *himself*].

   b. **[$_{NP}$The members of *Harold*'s famous marching band] admire *himself*.

In (5a), the pair (Harold, himself) does not form an A-chain due to the presence of an S′ barrier. In (5b), the lack of c-command between *Harold* and *himself* also prevents the formation of an A-chain (as per (1b)). This means that, in both of these cases, *himself* actually forms an A-chain on its own (i.e. a SINGLETON A-CHAIN), where it serves as both the tail and the head. Since *himself* is -R, this is blocked by Chain Visibility.

These examples illustrate a surprising consequence of Chain Theory as I have stated it so far: given the definitions in (1)-(2), Chain Visibility winds up *entailing* a version of CBT's Condition A. The idea is this. Whenever an anaphor lacks a local c-commanding antecedent, a -R singleton A-chain is formed and a violation of Chain Visibility ensues. In order to avoid the formation of a -R singleton A-chain, anaphors are, therefore, required to be locally bound – i.e. to have a +R antecedent that functions as the head for the local A-chains that they tail.

This makes the CC vulnerable to the many of same empirical criticisms that befell Condition A of the CBT in Chapter 1. To begin, every occurrence of a non-locally bound discursive reflexive turns out to be a counterexample to Chain Visibility. In all of the examples in (6), for instance, we have -R NPs that function as the heads of singleton A-chains:

(6) a. The description of *herself* in the paper really annoyed *Annie*.

   b. *Joanne* said that Bobby invited Larry and *herself* for lunch.

   c. It is *himself* that *Curly* claims that Laurie loves.

The long-distance use of non-reflexive anaphors threatens the validity of Chain Visibility for a similar reason. Recall from Chapter 1 that the Icelandic anaphor *sig* can be bound across an arbitrary number of finite subjunctive (SBJV) clauses (ANDERSON, 1986, pg. 66-67):

(7) *Jón* segir [að María viti   [að Ólafur vilji   [að Billi meiði   *sig*]]].
    John says that Mary knows.SBJV that Olaf   wants.SBJV that Bill hurts.SBJV SE
    '*John* says that Mary knows Olaf wants Bill to hurt *him*.'

Unless we are willing to resort to some *ad hoc* ancillary hypothesis – e.g. claim that finite subjunctives are not barriers in Icelandic and, therefore, that (Jón, sig) forms a licit A-chain in (18), or that *sig* is not really -R – Chain Visibility will incorrectly rule out cases like (18).

Moreover, under the reasonable assumption that expletives are -R, Chain Visibility is also contradicted by the following examples from Postal and Pullum (1988, pg. 642), where the dummy *it* occupies an subcategorized position governed by a verb:

(8)   a.   He never gave it a thought that Bolshies are human beings.

       b.   I resent it greatly that you didn't call me.

       c.   I didn't suspect it for a moment that you would fail.

As Reinhart and Reuland (1993, pg. 702) recognize, the only way to avoid these kinds of problems is by preventing the CC from applying to singleton chains. One way of doing this is by redefining the general notion of chain so as to require the presence of at least two links:[1]

(9)   C is a CHAIN *iff* C is the maximal sequence of links $(\alpha_1, ..., \alpha_n)$, $n \geq 2$, such that, for all j, α_j antecedent governs α_{j+1}.

Since (9) does not treat singleton NPs as chains, the structures in (6)-(8) are no longer violations of Chain Visibility: i.e. they no longer contain a -R element as a head of a CHAIN. On this picture, the main empirical duty of Chain Visibility is to rule out cases like (4), which involve A-chains with at least two -R links – one occupying the head and the other the tail.

By virtue of (9), therefore, Chain Theory only constrains -R elements that have local antecedents (i.e. those that integrate genuine $n \geq 2$-membered chains). For the purposes of the theory of anaphora, therefore, Chain Visibility becomes equivalent to the statement in (10):

(10)   CHAIN VISIBILITY (PARAPHRASE)

        An anaphor that has a local antecedent must be bound.

---

[1]   As it stands, this move looks entirely stipulative. However, further motivation comes from assuming that the primary mechanism for anaphoric chain formation in syntax is the valuation of unvalued ϕ-features by a (local) Agree operation (PESETSKY; TORREGO, 2007). This is what Reuland (2011) proposes in his minimalist construal of the RCT. In fact, (9) captures the gist of Reuland's (2011, pg. 151) notion of a CHECKING CHAIN:

(i) (α, β) form a CHECKING CHAIN if (a) β's ϕ-features have been valued by α; and (b) (α, β) meets standard conditions on chains such as uniformity, c-command and locality.

       Given that agreement is, minimally, a binary relation, the $n \geq 2$ requirement in (9) winds up not being as artificial as it first looked. I discuss proposals to encode coconstrual relations by means of agreement in Section 3.1.3. Note that excluding singletons from the definition of chain, as (i) and (9) do, requires restating other grammatical principles that invoke the notion of chain, such as the θ-Criterion (CHOMSKY, 1981, pg. 335).

The cases in (4) are still excluded by (10) because they contain an anaphor which has a local antecedent but is not bound: namely, the head of the chain. Notice, however, that (10) implies that there is no syntactic requirement that anaphors, *qua* -R NPs, have to meet in *every context* (REULAND, 2011, pg. 144). Anaphors are neither obliged to be syntactically bound (lacking a local antecedent exempts them from this, as in (6)–(8)) nor to have local antecedents.

This last point is significant because it eliminates a redundancy that was implicit in the previous version of the RCT. It is not necessary to appeal to Chain Visibility to derive the impossibility of non-local binding for (5) (repeated below as (11)) because these structures are independently ruled out by Reflexivity Condition A (RCA) – which, for present purposes, I take to be equivalent to Reinhart and Reuland's (1991, pg. 292) (12):

(11)   a.   **Harold said [$_{S'}$ that Marian loves *himself*].

   b.   **[$_{NP}$The members of *Harold*'s famous marching band] admire *himself*.

(12)   REFLEXIVITY CONDITION A (RCA)

   A reflexive argument of a syntactic predicate P makes P reflexive.

The reflexives in (11) are not excluded by Chain Visibility construed as in (10) because they do not have local antecedents. RCA alone is responsible for filtering them out.

However, simply preventing Chain Visibility from applying to singleton NPs (in order to allow for long-distance anaphors) while putting nothing in its place may not be the best solution. To see why, note that even though the previous construal Chain Visibility was redundant with RCA in ruling out non-locally bound REFLEXIVE anaphors in argument positions, it was not redundant when it comes to ruling out NON-REFLEXIVE anaphors in similar environments.

As Manzini and Wexler (1987) and Menuzzi (1999) note, most non-reflexive anaphors are subject to *some kind* of locality requirement – even if these are typically less stringent than the those that affect reflexive anaphors. To take a previous example, the Icelandic anaphor *sig* – which happens to be a particularly unrestricted non-reflexive anaphor – cannot be bound across an indicative (IND) clause boundary (HYAMS; SIGURJÓNSDÓTTIR, 1990, pg. 62):

(13)   **Jón veit   [að Pétur rakar   *sig*].
   John knows that Peter shaves.IND SE

Moreover, locality constraints for non-reflexive anaphors vary widely from language to language (cf. Dalrymple (1993) for a survey). We see variation even among closely related

Germanic languages. Icelandic, as we saw, allows binding of *sig* out of control infinitives and finite clauses, as (14) shows (HAIDER; OLSEN; VIKNER, 1995, pg. 14):

(14)   a.   *Petur* leyfði   [mér að raka *sig*].
             Peter allowed me   to shave SE

        b.   *María* heldur   [að ég elski   *sig*].
             Mary believes that I love.SBJV SE

Danish allows binding of its non-reflexive anaphor *sig* out of control infinitives, but not out of finite clauses (THRAINSSON, 1991; HAIDER; OLSEN; VIKNER, 1995, pg. 14):

(15)   a.   *Peter* tillod   [mig at barbere *sig*]
             Peter allowed me   to shave   SE

        b.   *Marie* tror   [at jeg elsker *sig*].
             Mary believes that I   love   SE

Dutch, by contrast, does not allow binding of *zich* in either of these environments as (16) illustrates (ROORYCK; VANDEN WYNGAERD, 2011, pg.166):

(16)   a.   *Jan* vroeg mij om   [voor *zich* te zorgen].
             John asked me COMP for   SE   to take.care.
             '*John* asked me to take care of *him*.'

        b.   *Jan* dacht   [dat ik voor *zich* zou   zorgen].
             John thought that I for SE   would take.care
             '*John* thought that I would take care of *him*.'

Since non-reflexive SE anaphors are not subject to RCA, there is nothing in the RCT as we have it now that would lead us to expect the non-local binding of these forms to be subject to syntactic constraints of any sort – let alone *cross-linguistically variable* constraints. If Chain Visibility only applies to anaphors with local antecedents, excluding singleton NPs (as in (10)), it can never be relied on to predict illicit non-local bindings of anaphors like the ones we see above. The previous Chain Visibility does not help either, because it indiscriminately rules out every instance of a non-locally bound anaphor in every language.

We are, therefore, faced with a dilemma. On the one hand, including singleton NPs into the definition of chains turns the RCT into a partly redundant and excessively rigid theory where long-distance anaphors are always impossible. On the other hand, excluding singleton NPs from the definition of chains, as we proposed above, turns the RCT into what appears to be an excessively unconstrained theory, where non-reflexive anaphors are *always* allowed to find

non-local antecedents – a theory where the notion of a *local binding* requirement plays no role.

Most proponents of the RCT bite the bullet and stick with the second of these alternatives (REINHART; REULAND, 1991; SIGURJÓNSDÓTTIR; HYAMS, 1992; REULAND, 2011). As a result, they are forced to appeal to independent syntactic or discourse factors to explain the different kinds of locality effects we see in connection to non-reflexive anaphors in various languages. This is a heavy burden to meet – one that may lead to excessive proliferation of syntactic machinery and dubious claims about constraints on discourse anaphora.[2]

As Fox (1993) notes, even reflexives are not completely immune to the problems that result from suspending the effects of Chain Visibility for singleton NPs. Although a locality requirement for argument reflexives is guaranteed by RCA, in a system where Chain Visibility is reduced to (10), there is nothing to account for the ungrammaticality of (17):[3]

(17) *Bobby* said [that *himself* shaved].

Bobby ($\lambda x.\ x$ said (x shaved x))

Since the reflexive in (17) is not part of an A-chain (it does not have a local antecedent),

---

[2] Sigurjónsdóttir and Hyams (1992), for example, argue that the contrast between long-distance binding out of indicative and subjunctive clauses in Icelandic (cf. (13)-(14b)) is entirely due discourse constraints. They start by assuming (as Reuland (2011, chap. 8) also does) that finite tense blocks the formation of an A-chain between *sig* and its antecedent. When this happens, they argue, *sig* must be interpreted logophorically. The crucial claim then is that the subjunctive mood provides a suitable logophoric antecedent, while the indicative does not. The problem with this proposal is that it is not clear whether the indicative/subjunctive opposition has this effect in other languages that have non-reflexive anaphors. Latin, for instance, seems to allow long-distance binding of the non-reflexive anaphor *sibi* out of both kinds of tensed clauses (BENEDICTO, 1991, pg. 174-175):

(i) a. *Epaminondas*   ei   [qui   *sibi*   ex lege   praetor   successerat]   exercitum
         Epaminondas.NOM him.DAT that.NOM SE.DAT by law.ABL praetor.NOM succeeded.IND army.ACC
         non tradidit.
         not transferred
         '*Epaminondas* didn't transfer the army to the one that succeeded *him* as a praetor following the law.'
     b. *Ariouistus*   respondit magnam Caesarem iniuriam facere   [qui   suo aduentu
         Ariovistus.NOM answered big.ACC Caesar.ACC injury.ACC make.INF who.NOM his arrival.ABL
         uectigalia   *sibi*   deteriora faceret].
         income.ACC SE.DAT worse   made.SBJV
         '*Ariovistus* answered that Caesar was doing him a serious injury, for his advance was damaging *his* revenues.'

In order to explain the difference between Latin and Icelandic, Sigurjónsdóttir and Hyams (1992) would need to claim that the discourse constraints block the long-distance use of *sig* in indicatives are absent in Latin. This is implausible, because it is unlikely that discourse constraints of this sort vary from language to language.

[3] As for the other instances of illicit bindings in (13)-(16), proponents of the RCT who stick to the version of Chain Visibility summarized in (10) have to appeal to an independent principle in order to explain why (17) is impossible. Reuland (2011, pg. 259-265) argues that what precludes (17) is the fact that *himself*, as a ϕ-deficient element, cannot check the uninterpretable agreement features on T(ense), causing the derivation to crash. In fact, he claims that this is what prohibits subject anaphors in all languages that do subject-verb agreement. Preminger (2019) shows that this generalization (often called the Anaphor Agreement Effect) happens to be incorrect (see also Section 3.1.2). So the fact that Reuland (2011) derives it actually counts against his theory.

Chain Visibility has nothing to say about it. However, here we cannot appeal to the RCA as we did for (11) because the reflexive is an argument of the syntactic predicate headed by *shave*, which happens to be reflexive (i.e. its arguments are bound by the same λ-operator).

The simplest way to exclude (17) is to build a hierarchical binding requirement for reflexive anaphors into the RCA itself. That is, in addition to saying that a reflexive that occupies the argument position of a syntactic predicate has to make its predicate reflexive, as in (12), we need to add a restriction that determines who can be a binder for a reflexive, as in (18):

(18)   REVISED RCA

A reflexive argument of a syntactic predicate P is bound by a higher argument of P.[4]

This version of the RCA correctly blocks (17) because the reflexive therein is not bound by a higher argument of *shave* (given that there is none). But (18) also has a surprising consequence: it correctly prevents the illicit structures in (4), which were the initial motivations for Chain Visibility in the first place. These examples, repeated below as (19), each contain at least one reflexive argument which is not bound by a higher argument of their syntactic predicates:

(19)   a.   *Bobby* said [that Marta showed *himself* to *himself*].

      b.   *John* said [that *himself* defended *himself*].

One of the reasons why the RCA in (18) is able to take over the explanatory role of the Chain Visibility in (10) because the concept of BARRIER (which is part of the definition of chains) and the concept of SYNTACTIC PREDICATE (which is part of RCA) pick out the same kinds of domains when it comes to anaphora. This uncovers a hitherto unnoticed redundancy within the RCT: it superfluously appeals to two separate notions of locality, with roughly the same effects. If we keep only the notion of syntactic predicate as our local domain for reflexives, Chain Visibility can be eliminated, with its empirical effects fully absorbed by the RCA.

Moreover, insofar as we are willing to build into the RCA a specific syntactic requirement about who can count as a binder for reflexive anaphors, there is nothing stopping us from doing something similar for NON-REFLEXIVE anaphors. That is, with Chain Visibility out of the

---

[4]   This statement of RCA is close to the one in Reinhart and Reuland's (1992) early version of the RCT. I use "higher" as a cover term for any relation that defines a total order between coarguments of a predicate (c-command, o-command, etc.). As usual, "binding" refers to A-binding as defined in Chapter 1. Note that (18) actually entails the previous formulation of the RCA in (12); if a reflexive is bound by a higher argument of a predicate P, then P is reflexive. What (18) adds to (12) is the specific syntactic demand that whoever winds up binding the reflexive must be a higher argument of the same syntactic predicate where the reflexive is bound. This also solves the problem with RCT's treatment of reflexives in raising-to-object structures, briefly noted in Chapter 2.

picture, we could define different locality requirements for different anaphors on a language-by-language basis, without worrying too much about repercussions for the general treatment of chains or even for the description of anaphors in other languages. This would give us a useful way of addressing the issues posed by (13)-(16).

Discussing the cross-linguistic facts that bear on this last suggestion would take me well-beyond the scope of this Thesis (see Manzini and Wexler (1987) and Dalrymple (1993) for concrete proposals). My point here was merely to show that Chain Visibility creates more problems than it solves and that there are no relevant problems that Chain Visibility solves that cannot be solved by a suitably reformulated RCA. This means that we can safely discard Chain Visibility and focus on what really is the conceptual core of RCT's Chain Theory: Chain Economy. This is the issue I turn to next.

3.1.2 Chain Economy and referential (in)dependence

As we saw in Chapter 2, the main reason why Reinhart and Reuland (1993) invoked the Chain Theory in the first place was the existence of purely syntactic anti-locality effects involving pronominals, which do not follow from their semantically-based Reflexivity Condition B defined in (20). Paradigm illustrations of such anti-locality effects are given in (21).

(20)   REFLEXIVITY CONDITION B (RCB)

A reflexive semantic predicate is reflexive-marked.

a.   The SEMANTIC PREDICATE formed of P is P and all arguments that realize a semantic role associated with P (i.e. the SEMANTIC ARGUMENTS of P).

b.   A predicate (formed of P) is REFLEXIVE-MARKED iff either P is lexically reflexive or one of P's arguments has a reflexivizing function.

(21)   a.   *$Bobby$ believes him to be a genius.

Bobby ($\lambda x.\ x$ believes (x is a genius))

b.   *$Bobby$ appears to him to be a genius.

Bobby ($\lambda x.$ appears to x (x is a genius))

The ungrammaticality of the examples in (21) does not follow from RCB because they do not contain reflexive semantic predicates. What prohibits these structures is the the part of the CC I have been calling Chain Economy:

(22)   CHAIN ECONOMY: Non-head links of A-chains must be -R (i.e. they must not be +R).

The structures in (21) are filtered out by Chain Economy because they contain A-chains that are tailed by a +R element – namely, a pronominal.

This is exactly the same prediction that the syntactic Condition B of the CBT would make. In fact, a version of the latter is entailed by Chain Economy, since every non-head link of an A-chain is, by definition, syntactically bound (c-commanded by a coindexed NP) in a kind of local domain – the domain of the closest barrier. The parallel is made explicit in (23):

(23) a. CHAIN ECONOMY (PARAPHRASE)

A +R element must not be syntactically bound in a local domain.

b. CONDITION B OF THE CBT

A pronominal must not be syntactically bound in a local domain.

As Safir (1997) notes, assessing the adequacy of a version of Condition B stated in terms of barriers and of Chain Economy will depend on the same sorts of considerations. For example, given that pronominals can be bound in (24), both principles would need to say that picture NPs count as barriers in English:

(24) *Curly* saw [a picture of *him*].

Therefore, any substantial difference that may exist between the two principles in (23) will hinge solely on how one understands the property of referential (in)dependence (R).

In principle, the concept of an +R element is broader than that of a pronominal, given that the former also includes non-pronoun NPs such as names and definite descriptions. However, Chain Economy is not motivated by its role in explaining the distribution of referring expressions of this sort – this is the responsibility of whatever handles the empirical effects of Condition C of the CBT (cf. Varaschin, Culicover and Winkler (in press) for a proposal). Since non-pronouns have to be excluded whenever they are bound at all (regardless of locality), Chain Economy is redundant in ruling out cases where non-pronouns are bound in a local domain.[5]

Whatever motivation one can give for Chain Economy will, therefore, come exclusively from its ability to derive distributional facts concerning +R pronouns (i.e. NPs that play role of

---

[5] That is, even though Chain Economy is capable of predicting the failure of coconstrual in (ia), only something like Condition C can extend this prediction over non-local domains, such as the one in (ib)

(i) a. **He* likes *Bobby*
     b. **He* refuses to talk to the woman [who wants to marry *Bobby*].

This is why Chain Economy is redundant as far as the prediction in (ia) is concerned.

variables in semantic structure). However, this matches exactly the definition of pronominals I have been assuming since Chapter 1. This means that, for the purposes of this Thesis, Chain Economy and a barriers version of Condition B of the CBT are one and the same principle.

This, however, is not quite the way the RCT sees things. Reuland (2011, 2016) argues that the notion of a +R element, understood as an NP which is fully specified for ϕ-features, is more well motivated than that of a pronominal, because it is capable of making correct predictions for the Frisian and Traditional Jambi Malay facts mentioned in previous chapters:

(25)   *Jack* wasket *him.*   (Frisian)
        Jack washed PRON
        '*Jack* washed *him(self).*'

(26)   *Yanti* neŋoʔ *dio*   di tipi.   (Traditional Jambi Malay)
        Yanti saw   PRON on TV
        '*Yanti* saw *her(self)* on TV.'

The claim is that Frisian *him* and Jambi *dio* are -R (i.e. ϕ-deficient, for the reasons mentioned in Chapter 2), but still pronominals nonetheless, given that they can be non-locally bound and deictically refer to a sentence-external entity. On this conception, Chain Economy fares better than the barriers-based Condition B, because (25)-(26) violate the latter, but not the former.

The account of (25)-(26) forces upon the RCT an implicit conceptual shift: the dissociation of ϕ-feature specification from the actual *semantic* property of referential independence, in terms of which pronominals are defined. This reduces Chain Economy to a claim which is weaker than what I have been assuming: it no longer carries any implications for referentially independent NPs (given that *him* and *dio* are referentially independent), but only for NPs that are fully specified for ϕ-features. In other words, if the explanation of (25)-(26) is to go through, the RCT has to abandon the thesis that ϕ-feature specification determines referential (in)dependence, which was explicitly embraced by Reinhart and Reuland (1993, pg. 697).

There are, in fact, independent reasons for doubting that ϕ-feature specification fully determines the referential properties of NPs. Just as there are examples of ϕ-deficient NPs that are not referentially deficient (cf. (25)-(26)), there are also examples of referentially deficient NPs that are clearly not ϕ-deficient. An interesting example are complex emphatic forms such as *ele mesmo/ela mesma* ('him same'/'her same') in Brazilian Portuguese (BP).

The internal morphosyntactic constitution of these forms is pretty unremarkable: they consist of a personal pronoun (e.g. *ele*, for 3rd person masculine singular) attached to an emphatic modifier, which has to agree in number and gender with the pronoun (e.g. *mesmo*, for

masculine singular). At least in the 3rd person, the pronoun and emphatic parts of these complex forms are fully specified for number, person and gender, as the following Table illustrates:

	SING		PLUR	
	MASC	FEM	MASC	FEM
1st	eu mesmo	eu mesma	nós mesmos	nós mesmas
2nd	tu mesmo	tu mesma	vocês mesmos	vocês mesmas
3rd	ele mesmo	ela mesma	eles mesmos	elas mesmas

Table 3 – Complex Emphatic Pronouns in BP

Moreover, there is good evidence that the ϕ-features of the pronominal and emphatic dependents are projected up to the phrasal NP level. The examples in (27) show that 3rd person feminine plural *elas mesmas* triggers necessary 3rd person plural agreement with the verb and also plural feminine agreement with adjectives in raising structures:

(27)   As atrizes trabalharam muito. ('The actresses worked a lot.')

   a.   Elas mesmas {escrev -eram / *-eu / *-vi} a   peça.
          They same   {wrote -3PL / -3SG / -1SG} the play
          'They themselves wrote the play.'

   b.   Mas elas mesmas nem parecem   {cansad -as   / *-os   / *-a}.
          But they same   not seem.3PL {tired -PL.FEM / -PL.MASC / -SG.FEM}
          'But they themselves do not seem tired.'

Despite their full ϕ-specification these complex emphatic forms, much like the anaphoric clitic *se* and its oblique tonic counterpart *si*, fail the deictic test for referential independence:

(28)   [Context: Joana and Roberto are sitting in a bar. Joana spots her husband, Lauro, in the dance floor. She decides to call Roberto's attention to Lauro.]

   a.   *Se olha! / *Olha pra si!   [pointing towards Lauro]
         SE look! / look   at SE

   b.   *Olha pra ele mesmo!   [pointing towards Lauro]
         look at   him same

These facts suggest that the forms in (17) are genuine ANAPHORS – albeit for different reasons. Forms like *se/si*, like the usual SE forms in Romance languages, are anaphoric for the traditional reasons expected by the RCT: they are poorly specified for ϕ-features.

The anaphoric status of the complex BP emphatics in Table 3, however, has to be attributed to some other cause. A plausible alternative is that the prosody associated with the internal structure of these complex forms has consequences for their information structure sta-

tus, as Zribi-Hertz (1995) suggests for their French counterparts *lui-même/elle-même* and the English complex possessives *his own/her own*. The reasoning goes as follows. Adjunction of an emphatic modifier (e.g. *mesmo*, *même*, *own*) triggers deaccenting on the pronoun part of the complex form (e.g. *ele*, *lui*, *his*) – as per Chomsky and Halle's (1968) Nuclear Stress Rule. Deaccenting is, in turn, associated with a constant information structure effect: it signals that the referent of the deaccented NP must be GIVEN in the discourse context (ROCHEMONT, 2016). This givenness requirement, which is caused by independent properties of the prosodic environment, forces the pronoun to find a linguistically antecedent nearby, thereby depriving it of its referentially independent status.[6]

The case of complex BP emphatic pronouns demonstrates that there is much more to referential (in)dependence than mere ϕ-feature specification. Proponents of the RCT are, therefore, justified in abandoning the view that ϕ-feature specification determines referentiality. This allows them to construct a weaker version of Chain Economy along the lines mentioned above – one that rules out ϕ-feature specified NPs in the non-head positions of A-chains, but does not really care about their status *vis-à-vis* referential (in)dependence.

However, even this weaker version of Chain Economy runs into severe empirical problems. As Safir (1997) notes, quantifiers like *someone*, *anyone* and *everyone* and proper names like *Alex*, *Jamie* and *Drew* are all unspecified for gender and Case. Nonetheless, they do not count as anaphors for the purposes of Chain Theory – i.e. Chain Visibility is not supposed to exclude them from head positions of A-chains (e.g. *Someone$_1$ was murdered t$_1$*).

More directly to the point, there are also cases of pronominals that are clearly fully specified for ϕ-features and, nonetheless, can be locally bound (i.e. that appear in non-head positions of A-chains). We find possible examples of this even in English. Recall from Chapter 1 that pronominals can be bound out of locative PPs, regardless of whether these PPs are arguments (cf. (29)) or adjuncts (cf. (30)) (HESTVIK, 1991):

(29)   a.   *Curly* saw a snake [behind *him*].

          b.   *Harold* found a copy of Lost Illusions [over *him*].

---

[6]   On that note, it is relevant to observe that, despite being anaphors, BP emphatics can be used in subject position (cf. (27)). This contradicts the aforementioned Anaphor Agreement Effect (RIZZI, 1990; WOOLFORD, 1999) which Reuland (2011) attempts to derive within his revised version of the RCT:

(i)   ANAPHOR AGREEMENT EFFECT: Anaphors do not occur in syntactic positions construed with agreement.

Insofar as the Anaphor Agreement Effect is wrong (abundant further evidence for this is given in Preminger (2019)), any general theory of anaphora which derives it faces empirical difficulties.

(30)   a.   *Curly* looked [behind *him*].

     b.   *Marian* pulled a blanket [over *her*].

The problem is that, in contrast locative PP adjuncts (cf. (31)), locative PP arguments do not seem to be barriers for chain formation, given the possibility of movement in (32):

(31)   a.   *Who$_7$ did Curly see a snake [behind t_7]?

     b.   *Which librarian$_1$ did Harold find a copy of Lost Illusions [over t_1]?

(32)   a.   Who$_7$ did Curly look [behind t_7]?

     b.   Which books$_3$ did Marian pull a blanket [over t_3]?

It seems, therefore, that examples like (29) are violations of Chain Economy: they contain a pronoun which is fully specified for ϕ-features but which, nonetheless, is bound within the domain of the closest barrier (MENUZZI, 1999, pg. 156).

This was basically the same problem Condition B of the CBT encountered in Chapter 1. There is, however, a potential way out of it which is suggested by my reformulation of Chain Visibility in (18). Instead of stating the anti-locality effect for +R elements in terms of the concept of chain, we can use the concept of SYNTACTIC PREDICATE, which is independently required for the statement of the RCA. This gives us the following principle:

(33)   SYNTACTIC CONDITION B

     A +R argument of a syntactic predicate P is not bound by any higher argument of P.

(34)   The SYNTACTIC PREDICATE formed of (a head) P is P, all of the projections that realize grammatical functions associated with P, and the subject of P.

The PPs in (29)/(30) do not count as syntactic predicates because they lack subjects. This means that the pronominals therein are *exempt* from the anti-locality effect enforced by (69) for the same reason as reflexives in these positions would be exempt from RCA. If Chain Economy is replaced by (69), (29)/(30) are, thus, no longer a threat to the RCT. This move enables us to purge the configurational notion of chain from the RCT, thereby eliminating the redundancy of referring to two separate syntactic notions of locality within our theory of anaphora.[7]

---

[7]   Note that (69) does not eliminate the need for a *semantic* notion of locality, the SEMANTIC PREDICATE, which is required for the statement of the Reflexivity Condition B in (20). This condition is still needed to capture the impossibility of binding involving pronominals that are not arguments of syntactic predicates:

(i)   a.   *Bobby* praised both Marta and *him*.
     b.   *Curly twisted *Laurie* around *her*.

However, even though (69) is generally adequate as an account of English pronominals (and something along its lines will indeed be proposed in Chapter 7), it still fails remarkably when we attempt to extend it to other languages, such as certain dialects of Brazilian Portuguese (BP) (GALVES, 1986; LACERDA; OLIVEIRA; LEITÃO, 2014; CARVALHO, 2019):

(35)   a.   *O Paulo* viu *ele* no espelho.
the Paulo saw PRON in-the mirror
'*Paulo* saw *him(self)* in the mirror.'

    b.   *A Joana* esqueceu de incluir *ela* na lista de convidados.
the Joana forgot to include PRON in-the list of guests
'*Joana* forgot to include *her(self)* in the guest list.'

    c.   *O Pedro* não reconheceu *ele* na foto.
the Pedro not recognized PRON in-the photo
'*Pedro* did not recognize *him(self)* in the photo.'

As we will see in detail in Chapter 7, the BP pronouns *ele/ela* are +R on all accounts: they are referentially independent in the original semantic sense as well as fully specified for ϕ-features. Nonetheless, they can be arguments of syntactic predicates and bound by higher coarguments in contexts like (35). There is simply no coherent statement of a universal syntactic anti-locality effect on +R elements – be it Chain Economy, CBT's Condition B or the Syntactic Condition B in (69) – that can be made compatible with data like (35).

In conclusion, it seems that some principle is in fact necessary to account for purely syntactic anti-locality effects that do not fall under Reflexivity Condition B. However, whatever explains these particular effects must be flexible enough to capture the possibility of local pronominal binding for English locative PPs and BP in general. Chain Economy fails on both grounds, regardless of how one conceptualizes the R property that it invokes.

3.1.3 Indices and their alternatives

The RCT of Reinhart and Reuland (1993) retains indices as genuine syntactic objects, despite the fact that they are not necessary for the statement of anaphoric binding constraints in Reinhart's (1983) interpretive system, as we saw in Chapter 2. This is a direct consequence of the Chain Condition (CC). Recall that the CC is supposed to apply to syntactic dependencies formed by coconstrual as well as by A-movement (e.g. passive, raising). These do not encode the same kinds of interpretive effects. So, for example, while the coconstrual dependency in (36a) corresponds to the A-binding structure to its right, the movement one in (36b) does not.

(36)   a. $Felix_1$ fired $himself_1$. $\Rightarrow$ Felix ($\lambda x.\ x$ fired x)

       b. $Felix_3$ was fired t_3. $\nRightarrow$ Felix ($\lambda x.\ x$ was fired x)

Since the dependencies regulated by the CC cannot be individuated in terms of a uniform semantics (e.g. A-binding), a syntactic surrogate must be invoked: the concept of an A-chain, which neutralizes the differences in (36). The semantic heterogeneity of the objects that the CC applies to is, thus, what forces indices upon a system that would otherwise not need them.

This creates conceptual and empirical problems for the RCT. First, it makes the RCT incompatible with the current practices of the Minimalist Program, which eliminated the use of indices on the grounds of optimization. Reuland (2011, pg. 55) summarizes this basic point:[8]

> [The Minimalist Program] proposes that the computational system of human language reflects the combinatorial properties of a purely morpholexical vocabulary. Furthermore, its guiding hypothesis is that [the computational system of human language] is an optimal solution for a system pairing form and interpretation that is to meet the specific conditions imposed by the human systems of thought and perception/articulation. Such an optimal system should meet the condition of *inclusiveness*: any structure formed by the computation is constituted of elements already present in the lexical items selected. No new objects such as indices are added in the course of the derivation. Hence, indices [...] have no place in syntax, unless coindexing is really morphosyntactically expressed, which is not the case in any language we know of.

In addition to these abstract considerations, the demise of traditional indices is also motivated on empirical grounds (FOX, 1998; HEIM, 1998; SAFIR, 2004; REINHART, 2006). Since coindexing is essentially a transitive relation, it has difficulties in encoding the non-transitivity of semantic A-binding – a point originally raised by Higginbotham (1983). That is, if α is coindexed with β and β is coindexed with γ, α is also coindexed with γ. However, if α A-binds β and β A-binds γ, it is not necessarily true that α A-binds γ.

For most cases, this is a distinction without a difference. But when we have two potential A-binders for a given variable, the disadvantages of transitive coindexing representations come to light. This is precisely what happens in Safir's (2004, pg.106) example in (37):

---

[8]   In more recent years, the condition of inclusiveness, which establishes the rationale for eliminating indices, came to be viewed as a corollary of the so called Strong Minimalist Thesis: the idea that syntactic objects are constructed by a single structure-building operation known as Merge, as in the recursive definition in (i).

(i)   a.   If α is a lexical item, then α is a syntactic object.
         b.   If α and β are syntactic objects, then Merge(α, β) is a syntactic object.
         c.   For any syntactic objects α, β, Merge(α, β)=$\{\alpha, \beta\}$.

   If the only operation in the grammar is Merge and syntactic indices are not lexical items (or parts thereof), inclusiveness simply follows as a theorem because Merge, as defined in (ic), does not add any new structure to its inputs other than the fact that they constitute a set (COLLINS; STABLER, 2016; HORNSTEIN, 2018).

(37)   a.   *Every woman* thought that only *she* voted for *her*.

        Every woman ($\lambda x.\ x$ thought (only x ($\lambda y.\ y$ voted for x)))

    b.   *Every woman* thought that only *she* voted for *herself*.

        Every woman ($\lambda x.\ x$ thought (only x ($\lambda y.\ y$ voted for y)))

In (37a), we get a transitive A-binding relation between the three coconstrued NPs: *every woman* A-binds *she*, *she* A-binds *her* and *every woman* also A-binds *her*, as expected. This does not happen in (37b): even though *every woman* A-binds *she* and *she* A-binds *herself*, *every woman* does not A-bind *herself*. In virtue of the presence of the focus-sensitive operator *only*, the readings for (37a) and (37b) are actually truth-conditionally distinct: the latter entails that every woman thought that no other woman was a self-voter, while the former entails that every woman thought that she got a sum total of one vote (her own). So, a situation where a woman thought that she got more than one vote is compatible with (37b), but not with (37a).

Since the RCT relies on transitive coindexing to serve as a syntactic proxy for A-binding, it is unable distinguish these two readings. Moreover, (37a) would be wrongly excluded by the CC, because it could only be attained by the indexing in (38), which violates Chain Economy: the +R element *her* functions a the tail link of the A-chain headed by *she*.

(38)   Every woman$_8$ thought that only she$_8$ voted for her$_8$.

The reason why (37a) is acceptable, according to Heim (1998) and Reinhart (2006), is because the pronouns, though covalued, are not actually in a local A-binding relation with each other – they are both independently cobound by the non-local quantifier subject. Simple interpretive theories based on coindexing have no way to capture this special kind of cobinding relation and, therefore, wind up prohibiting (37a) in virtue of whatever principle handles anti-locality effects involving pronominals (Condition B of the CBT, Chain Economy, etc.).

These problems are good reasons for abandoning traditional indices. With the demise of indices, a popular alternative has been to use the identity of ϕ-features established by agreement as a way of encoding A-binding relations in syntax (HEINAT, 2006; CHOMSKY, 2007, 2008; REULAND, 2011; KRATZER, 2009; ROORYCK; VANDEN WYNGAERD, 2011).

There is some independent conceptual appeal to this idea: structures like those in (39) can be ruled out by claiming a failure of agreement, just as we see in cases like (40).

(39)   **The guy who lives next door* admires *themselves*.

(40)   *The farmer and the cowboy is friend. (cf. The farmer and the cowboy are friends.)

The way agreement-based theories are implemented typically involves saying that bound pronouns come from the lexicon with unvalued ϕ-features. It is assumed that, in order to be interpretable, unvalued features must be valued in the course of the derivation. An element bearing unvalued features acts, therefore, as a PROBE that searches within a certain domain for an element bearing a value for the features it lacks (the GOAL). The syntactic operation AGREE "copies the value of the valued feature onto the valueless one thereby ensuring that the syntactic object bearing the feature is interpretable by the interfaces" (HICKS, 2009, pg. 107).[9]

On this approach, the structure in (39) is ill-formed because the reflexive contains ϕ-features that could not have been copied from the subject by Agree – and, since the reflexive is assumed to be poor in ϕ-features, there is nowhere else its features could have come from. This also solves the more abstract problem noted in connection to (37) because the valuation relation correctly models the non-transitive nature of A-binding: if α values the features of β, thereby allowing β to value the features of γ, it is not the case that α itself has valued the features of γ.

But this use of agreement as a syntactic surrogate for A-binding runs into obvious problems. An immediate threat comes from data like (41), from Levine (2010, pg. 275):

(41)   I know *someone* who thinks *they* are the greatest thing since sliced bread.

The the quantified NP and the pronoun in (41) stand in an A-binding relationship. However, contrary to expectations, their ϕ-features do not match (*someone* is 3SG, *they* is 3PL), as indicated by the discrepant agreement on the verbs *thinks* and *are* (see also Conrod (2018)).

Though these agreement mismatches are somewhat marginal in English, they can be abundantly found in BP (MENUZZI, 1999). BP has two pronominals that express the semantics of 1st person plural: the canonical Romance 1PL pronoun *nós* (with its clitic counterpart *nos*) and the expression *a gente* (literally, 'the people'), which functions morphosyntactically like a 3rd person singular form, as indicated by verbal agreement. What is relevant here is that we can get A-binding of *a gente* by *nós/nos* and vice versa, in spite of the formal agreement mismatch between these forms. The following examples illustrate (for clarity, I gloss the pronouns in

---

9   Agreement-based theories come in different varieties and my purpose here is not to get into details, but merely to present general difficulties that all of them face (see also Preminger (2019)). There are also non-derivational variants of agreement-based theories, like the one in Pollard and Sag (1994). These, however, run into the same problem for modeling the non-transitivity of A-binding as the traditional theories based on coindexing (cf. (37)). The reason is that structure-sharing of ϕ-features alone, without a derivational mechanism that establishes an asymmetry between an agreement probe and an agreement goal, is *also* a transitive relation.

accordance to their morphosyntactic ϕ-feature specifications, instead of their semantics):

(42) a. *Nós* achamos que o Paulo já viu *a gente* na TV.
PRON.1PL think.1PL that the Paulo already saw PRON.3SG on-the TV
'*We* think that Paulo has already seen *us* on TV.'

b. *A gente* acha que o Paulo já *nos* viu na TV.
PRON.3SG think.3SG that the Paulo already CL.PRON.1PL saw on-the TV
'*We* think that Paulo has already seen *us* on TV.'

These examples demonstrate that ϕ-feature matching as a result of Agree is not a necessary condition for A-binding to take place. However, these examples are not objections to Reuland's (2001, 2011) specific agreement-based reconstruction of the CC. On his account, only anaphors that form an A-chain require ϕ-feature agreement in order to express A-binding.[10] This is because only anaphors are assumed to be unvalued for ϕ-features, as we saw. Following Menuzzi (1999), we can state this more restricted agreement requirement as follows:

(43)   AGREEMENT CONDITION ON A-CHAINS

An anaphor α agrees with an antecedent β's ϕ-features only if (α, β) form an A-chain.

The reason why (41)-(42) are not objections to Reuland and Menuzzi is because the A-binding relationships therein do not involve the formation of an A-chain, which presupposes locality. But even the more restricted agreement requirement expressed in (43) encounters numerous counterexamples. In English, one mainly sees these in cases where a subject that triggers morphosyntactic singular agreement A-binds a plural reflexive anaphor. The following examples are all naturally occurring data collected from the internet:

(44) a. "Probably *every couple* thinks of *themselves* that way [. . .]."

b. "[M]aybe the reason you like her so much is because she looks like you and *everyone* admires *themselves*."

c. "*I*'m really proud of *ourselves*."

d. "*I*'m angry at *ourselves* about that loss to Oregon State."

Contrary to what Menuzzi (1999) and Reuland (2011) claim, we also find counterexamples to (43) in some dialects of BP: the 1PL pronoun forms *nós/nos* can locally A-bind the 3SG

---

[10]   The non-local binding of pronominals is taken to be established on purely semantic grounds. The difference between (37a) and (37b) above is captured in this theory by the assumption that only the latter contains a chain dependency formed by Agree. That is, only in (37b) are the features of *she* and the A-bound pronoun actually token-identical. In (37a), there is only type-identity of features, and, hence, no chain formation.

forms *a gente*/*se* and vice versa. The examples in (45) come from Brito and Sedrins (2017) and the ones in (46)-(47) are gathered from the internet. Pereira (2007, pg. 242) found data similar to (45a)/(46a)/(47a) in spoken corpora from the city of São Paulo:

(45) a. *Nós   se   vimos   no   espelho.*
          PRON.1PL REFL.3SG saw.1PL in-the mirror
          '*We* saw *ourselves* in the mirror.'

      b. *A gente   nos   viu   no   espelho.*
          PRON.3SG REFL.1PL saw.3SG in-the mirror
          '*We* saw *ourselves* in the mirror.'

(46) a. "*Nós   se   conhecemos em um trabalho de jovens.*"
          PRON.1PL REFL.3SG met.1PL   in a work of youngsters
          '*We* met *each other* during youth work. '

      b. "Foi assim que *a gente   nos   conheceu.*"
          was thus that PRON.3SG REFL.1PL met.3SG
          'This was how *we* met *each other*.'

(47) a. "*Nós   se   casamos   em 56 [...].*"
          PRON.3SG REFL.1PL married.3SG in 56
          'We got married in 1956.'

      b. "*A gente   nos   casou   logo no   começo   [...].*"
          PRON.1PL REFL.3SG married.1PL soon in-the beginning
          'We got married right at the beginning. '

It is probable that the a and b-cases in (45)-(47) reflect different dialects of BP. In my own personal judgment, the cases where *a gente* A-binds *nos* are considerably degraded, while the ones where *nós* binds *se* are much more natural. Regardless of whatever sociolinguistic issues may be involved here, any instance of containing a local agreement mismatch suffices to undermine the claim that agreement is necessary for local A-binding, as (43) implies.

All of the data above point to the conclusion that are no well-motivated formal objects that can serve as suitable syntactic proxies for the semantic A-binding relation. This makes it impossible to state the CC, which requires A-binding and A-movement to be unified in terms of a unitary syntactic object. Since Chain Theory is the only part of the RCT that requires the employment of syntactic proxies for the A-binding (e.g. indices, ϕ-features), all of the problems above can be avoided if we simply abandon it. Not much is lost since, as we have already seen above, the effects of Chain Visibility and Chain Economy can be incorporated into other parts of the theory. This leads us to a theory of anaphora fully situated at the interface, where binding principles take the form of correspondence rules between narrow syntax and semantic structure, as independently suggested by Jackendoff (1990b, chap. 3) and Reinhart (2006, chap. 4).

3.2 ON REFLEXIVITY CONDITION B AND THE NOTION OF REFLEXIVE-MARKING

One of the major innovations of the RCT, in comparison to earlier theories of anaphora, was the introduction of a binding principle which is sensitive to properties of semantic predicates. This is the Reflexivity Condition B (RCB), whose definition I repeat below:

(48)   REFLEXIVITY CONDITION B (RCB)

A reflexive semantic predicate is reflexive-marked.

a.   The SEMANTIC PREDICATE formed of P is P and all arguments that realize a semantic role associated with P (i.e. the SEMANTIC ARGUMENTS of P).

b.   A predicate (formed of P) is REFLEXIVE-MARKED iff either P is lexically reflexive or one of P's arguments is a reflexive pronoun.

In this section, I argue that, while the RCB is largely correct, the concept of reflexive marking that it invokes is doubly problematic. First, it embodies a suspicious notion of LEXICAL REFLEXIVE MARKING, which does not adequately capture why some predicates are liberated from the requirement of having to be overtly reflexive marked in syntax. Second, the definition of syntactic reflexive marking itself (i.e. marking by a reflexive pronoun argument) yields contradictory consequences when we look at the behavior of BP emphatic pronouns.

As we saw in Chapter 2, one of the phenomena RCB provides an explanation for are the contrasts between reflexive anaphors and non-reflexive forms (SE anaphors and pronominals) in structures conveying reflexive semantic predicates, like the Dutch examples in (49). Since *zich* and *hem* do not have a reflexivizing function, only the dedicated reflexive form *zichzelf* is able to provide the appropriate reflexive-marking required by RCB.

(49)   a.   *Max* haat   {*zichzelf* / *zich* / *hem*}.
             Max hates {REFL   / SE   / PRON}

b.   *Max* praat   met {*zichzelf* / *zich* / *hem*}.
       Max speaks with {REFL   / SE   / PRON}

The data above illustrate a universal tendency that is attested across the world's languages: semantic reflexivity is typically marked by means of morphosyntactic devices that make reflexive structures more complex overall than those that convey semantic disjointness (COMRIE, 1999; LEVINSON, 2000; HASPELMATH, 2008). This can be expressed in terms of the following universal of reflexive marking:

(50)   REFLEXIVE MARKING UNIVERSAL 1

     Forms that signal reflexivity are at least as complex than forms that do not.

The notion of complexity referred to in (50) is meant to cover all levels of linguistic form: phonology, morphology and syntax.[11] Table 4, taken from Haspelmath (2008, pg. 48) illustrates further examples that confirm this general tendency:

	Reflexive Forms	Non-reflexive Forms
Dutch	zichzelf	zich / hem
English	herself	her
Greek	ton eaftó tu	ton
Hebrew	(et) ʔacmo	oto
Turkish	kendini	onu
Oriya	nijaku	taaku
Lezgian	wič	am
Japanese	zibun	∅
Chinese	ziji	tā
German	sich	ihn
French	se	le

Table 4 – Reflexive Forms *vs.* Non-reflexive Forms

The first nine languages provide positive evidence for the universal in (50) because their reflexive forms are longer/more morphologically complex than their non-reflexive forms. German and French do not confirm (50), but they also do not contradict it.

The RCT, in and of itself, does not directly predict this typological pattern. But we can make it do so simply by building into the RCT the idea that formal complexity of a pronoun (on some linguistic level) is associated with its capacity to function as a syntactic reflexive marker.[12] With this association in hand, RCB correctly derives (50) and the asymmetries in Table 4.

However, the data in (51), also from Chapter 2, presents a slightly more complex picture. The reflexive marking requirement that yields the judgments in (49) seems to be absent here:

(51)   a.   *Max* schaamt {??*zichzelf* / *zich* / **hem*}.
            Max shames {REFL   / SE   / PRON}
            'Max is ashamed'

---

[11] What unifies complexity in all of these levels is arguably some general notion of *markedness as difficulty* (HASPELMATH, 2006). A reasonable conjecture is that a form ρ counts as more complex than ψ to the extent that producing and interpreting ρ imposes more processing effort than ψ (GIVÓN, 1991; WURZEL, 1998; HAWKINS, 2004; CULICOVER, 2013b). More on this in Chapter 5.

[12] Though this assumption is not a part of the RCT's core, it is suggested in Reinhart and Reuland's (1993, pg. 658) characterization of non-reflexive anaphors as as SIMPLEX anaphors and of reflexive anaphors as COMPLEX anaphors. The connection between formal complexity and the property of being a reflexive pronoun is further explored in recent RCT work, such as Volkova and Reuland (2014) and Reuland, Wong and Everaert (2020).

  b. *Max* wast {*zichzelf* / *zich* / **hem*}.
    Max washes {REFL / SE / PRON}

In order to avoid saying that the structures with *zich* in (51) are violations of RCB, Reinhart and Reuland (1993, pg. 666) claim that the predicates headed by *schaamt* ('be ashamed') and *wassen* ('wash') are are "intrinsically reflexive". Reinhart and Reuland (1993, pg. 662) take this to mean that such predicates are "marked [as reflexive] in the lexicon", and therefore, do not require an overt reflexive argument in syntax in order to be reflexive marked.

Note, however, that this concept of lexical reflexive-marking lumps together two classes of predicates that differ from each other when it comes to the possibility of taking a reflexive pronoun as an argument: *zichzelf* is very marginal in (51a) but completely acceptable in (51b). The RCT handles this difference by stipulating that verbs like *wassen* ('wash') are doubly listed in the lexicon. In their intrinsically reflexive entries, they allow the non-reflexive *zich* and exclude *zichzelf*; in their ordinary transitive entries, a reflexive is required to comply with RCB. Verbs like *shaamt* ('shame'), on the other hand, only have an intrinsically reflexive entry. Reinhart and Reuland (1993, pg. 667) claim that combining such verbs with a syntactic reflexive like *zichzelf* violates economy, insofar as it implies redundant reflexive marking.

Lexical reflexivity is also used to explain why English verbs like *shave* and *undress* allow reflexive readings in the absence of an overt reflexive argument, as (52) shows:

(52) a. Bobby shaved.

   b. Bobby undressed.

(53) a. *Bobby* shaved *himself*.

   b. *Bobby* undressed *himself*.

The sentences in (52) convey the same reflexive semantic predicates as those in (53). For the RCT, this can only mean that *shave* and *dress*, like *wassen* in (51b), must have have an intrinsically reflexive lexical entry in addition to a non-reflexive one, which is the one used in (53). Otherwise, RCB would mistakenly predict reflexive readings to be unacceptable for (52).

There are numerous problems with this account. To begin, the assumption that RCB applies to (51a) is strange because it implies that the semantic predicate therein is in fact reflexive. It is much more straightforward to assume that *shaamt* simply belongs to the class of REFLEXIVE INTRANSITIVE VERBS, in the sense of Lees and Klima (1963, pg. 25). English exemplars of this verb class include *behave*, *perjure* and *absent*. What all of these verbs have in common

is the fact that they are, at once, *syntactically transitive* and *semantically monadic*.

The fact that these verbs only have one semantic argument can be shown on the basis of various tests: e.g. they can't be conjoined with semantically transitive verbs (cf. (54)) and their objects can't be replaced by non-reflexive NPs (cf. (55)):

(54) a. **Max* haat en   schaamt *zich*
           Max haat and shames   SE
      b. **Bobby* praised and behaved *himself*.

(55) a. *Max schaamt Marie.
           Max shames Marie
      b. *John behaved Mary.   (SAFIR, 2004, pg. 128):

If *shaamt* is a semantically monadic verb, using a genuine reflexive in (51a) counts as a violation of the θ-criterion (or whatever subsumes it): the complement of *shaamt* is a non-thematic position and *zichzelf* can only be used as a thematic argument. Since there is no reflexive predicate in the semantics of (51a), RCB does not require any specific reflexive marking and the anaphor *zich* is used as a kind of dummy argument.[13]

Moreover, simply stipulating two separate lexical entries for verbs like *wassen* ('wash'), *shave* and *undress* (cf. (51b) and (53)) does not explain why the reflexive-marking requirement is dropped for similar predicates in *other* languages as well. That is, it is not only in Dutch and English that verbs denoting typically self-directed actions (e.g. *wash*, *shave* and *undress*) are able to receive reflexive readings without the complex marking that is used to signal reflexivity for other transitive verbs (e.g. *kiss*, *hate*, *criticize*).

The concept of lexical reflexive-marking prevents the RCT from deriving this widely attested typological pattern, which can be stated as the following universal of reflexive marking (FALTZ, 1985; HAIMAN, 1983; KÖNIG; SIEMUND, 2000b; HASPELMATH, 2008):

(56) REFLEXIVE MARKING UNIVERSAL 2

      Forms used to signal reflexivity for prototypically non-reflexive verbs are at least as complex as the forms used to signal reflexivity for prototypically reflexive verbs.

The notion of prototypical (non-)reflexivity comes from the functionalist work on anaphora (HAIMAN, 1983; FALTZ, 1985; LEVINSON, 1991; COMRIE, 1999; ARIEL, 2008). It is based on the intuition that reflexive interpretations are more natural for some predicates (e.g.

---

[13] In the case of English reflexive intransitive verbs, the forms that are lexically specified as the syntactic dummy arguments are homophonous with reflexives: e.g. *Bobby {behaves / perjures / absents} himself*.

wash, *shave* and *undress*) than for others (e.g. *kiss*, *hate*, *criticize*). There is debate about what causes this, but it seems likely that prototypical (non-)reflexivity, like other kinds of prototypes, is determined by inductive regularities in speakers' experience of the world (LEVINSON, 2000; KÖNIG; SIEMUND, 2000b). In most cultures, people experience more often self-directed instances of actions like *shaving* than of actions like *kissing*.[14] Haspelmath (2008, pg. 44) provides the following illustrations of (56) in English and other languages:

	Non-prototypically Reflexive		Prototypically Reflexive	
English	hate **herself**		shave ∅	
Russian	nenavidet' **sebja**	'hate oneself'	myt'-**sja**	'wash'
Hungarian	utálja **mag-á-t**	'hates herself'	borotvál-**koz**	'shave'
Greek	aghapái **ton eaftó tu**	'loves himself'	dín-**ete**	'dresses'
Turkish	**kendini** sev-iyor	'loves himself'	yıka-**n**-iyor	'washes'
Dutch	haat **zichzelf**	'hates himself'	wast **zich**	'washes'
Frisian	hearde **himsels**	'hates himself'	wasket **him**	'washed'
Jamul Tiipay	**naynaach mat**-aaxway	'killed himself'	**mat**-sxwan	'scratched'

Table 5 – Prototypically reflexive *vs.* Non-prototypically reflexive marking forms

If intrinsic reflexivity is seen as an idiosyncratic property of lexical items, as it is in the RCT, we cannot account for why only a specific class of predicates (prototypically reflexive ones) can be reflexive in the absence of complex reflexive marking. This means that we cannot not rule out the existence of a hypothetical language where verbs like *kill* and *hear* are lexically reflexive and verbs like *shave* and *dress* are not. Such a language has never been attested.

Moreover, the assumption that intrinsic reflexivity is a property of individual lexical items (also implicit in the concept of *lexical reflexive marking*) is inadequate because it fails to take into account the fact that exemption from the overt reflexive marking requirement is sensitive to compositional factors. We can see this by examining the following BP data ((58) is adapted from adapted from Menuzzi (1999, pg. 145), who makes a similar point):

(57)   a.   *O Pedro* não reconheceu {*ele / *ele mesmo*} na festa.
              the Pedro not recognized {PRON / PRON} same in-the party
              '*Pedro* did not recognize *him(self)* in the party.'

       b.   *O Pedro* não reconheceu {*ele* / *ele mesmo*} na foto.
              the Pedro not recognized {PRON / PRON same} in-the photo
              '*Pedro* did not recognize *him(self)* in the photo.'

---

14   These inductive regularities gathered from world knowledge are plausibly reflected in frequency of reflexive use in language: given a large enough corpus of utterances, prototypically reflexive predicates like *shave* will occur more often with reflexive interpretations than predicates like *kiss*, which are not prototypically reflexive. See Ariel (2008) and Haspelmath (2008) for some frequency counts that confirm this prediction.

(58)   a.   Nessa carta *o Jô* fica falando {*d*ele* / d*ele mesmo*}.
               in-this letter the Jô stays talking {of-PRON / of-PRON same}
               'In this letter *Jô* keeps talking about *him(self)*.'

        b.   Nessa carta *o Jô* fica falando {d*ele* / d*ele mesmo*} o tempo todo.
               in-this letter the Jô stays talking {of-PRON / of-PRON same} the time all
               'In this letter *Jô* keeps talking about *him(self)* all the time.'

The lexical heads that select the pronominals (the verbs *reconheceu* and *falar*) are kept constant within each of the contrasts above. The only thing that changes is the presence of other elements in the sentence. This change has the effect of making the scenarios described in (57b)/(58b) more plausibly reflexive than those in (57a)/(58a): e.g. recognizing oneself in a photo is a more common type of situation than recognizing oneself in a party.

The prototypicality of the situation as a whole (and not just of the event denoted by the verb) seems to have an effect on the reflexive marking requirement enforced by RCB: special marking by the emphatic *mesmo* is required in the prototypically non-reflexive situations in (57a)/(58a), but not for the more prototypical situations in (57b)/(58b). Since, as we saw in section 3.1.2, pronominals in BP do not exhibit the syntactic anti-locality effect captured by Chain Economy, they are actually fully acceptable in (57b)/(58b), despite being locally bound.

Geurts (2004) reports that the non-reflexive Dutch anaphor *zich* is subject to a similar effect in the following examples:

(59)   a.   *De zuster* diende   {*zich / *zichzelf*} opium toe.
               the nurse administered {SE / REFL} opium PARTICLE
               '*The nurse* administered *herself* opium.'

        b.   *Betty* dient   {*zich* / *zichzelf*} weer eens opium toe.
               Betty administers {SE / REFL} yet again opium PARTICLE
               '*Betty* administers *herself* opium yet again.'

If we were to explain the optionality of reflexive marking in (59b) in terms of lexical reflexivity, we would have to say that *dient* ('administers') has two lexical entries: an inherently reflexive entry that licenses *zich* and a transitive entry that requires *zichzelf*. This is what Reinhart and Reuland propose for *wast* ('wash') in (51b). The problem, in this case, is that this dual-entry account does not not explain why the reflexive is obligatory in (59a). What seems to be going on is that the reflexive marking requirement is dropped if the situation described by the sentence as a whole (not just the verb) is portrayed as habitual or common. In the case of (59b), this habitual meaning is conveyed by adding the temporal adjunct *weer eens* ('yet again').

A proper explanation for the optionality of reflexive marking requires a reformulation of RCB in pragmatic terms, as anticipated in the work of Farmer and Harnish (1987) and Levinson (2000). This is a task I undertake in Chapter 7. A pre-condition for a pragmatic reduction of RCB is to have it apply only to prototypically non-reflexive semantic predicates, as in (60):

(60)   REVISED RCB

A prototypically non-reflexive semantic predicate is reflexive-marked.

(61)   A predicate is REFLEXIVE-MARKED iff one of its arguments is a reflexive pronoun.

The formulation in (60) allows us to eliminate lexical reflexive-marking as a subkind of reflexive-marking, since the cases previously treated in such terms are better understood as instances of prototypical reflexivity. On the account suggested here, therefore, structures like *Bobby shaved* or *Max wast zich* ('Max washed') neither obey nor violate the RCB – they are exempt from it. That is, they simply do not contain the kinds of predicates RCB applies to.

Excluding lexical reflexivity from the definition reflexive-marking thus allows us to simplify and expand the empirical coverage of RCB. However, it also creates a problem for Reflexivity Condition A, whose classical statement also calls upon the notion of reflexive-marking:

(62)   REFLEXIVITY CONDITION A (RCA)

A reflexive-marked syntactic predicate is reflexive.

In order to see why the simpler definition of reflexive marking in (61) creates problems for RCA, we consider again the system of anaphora of BP. Note first that, in order for a predicate to be reflexive-marked in the sense of (61), it suffices that one of its arguments be the kind of expression that can signal semantic reflexivity across a consistent range of diverse predicates in a language – i.e. a reflexive pronoun, according to the definition in Chapter 1. In BP, both the SE clitic and the complex emphatic pronouns mentioned in section 3.1.2 seem to meet this requirement, in contrast to the pronominal *ele*, which cannot be used to signal reflexivity with prototypically non-reflexive predicates (cf. (57)-(58) above).[15]

---

[15] As we saw in Table 4, reflexives are typically expressions that are more complex than the forms that signal semantic disjointness for most predicates. One might wonder whether BP presents a counterexample to this universal tendency, insofar as the SE clitics seem to be simpler than the non-reflexive pronominals *ele/ela*. Even though this may well be the case on a purely phonological level, there are other levels of analysis where SE might actually count as more complex pronominals. For instance, it is plausible to assume that, since BP has been progressively losing its clitics (CYRINO, 2003; CARVALHO; CALINDRO, 2018), the use of clitic SE implies a morphological strategy which is highly marked in BP.

(63) a. *O Roberto se* barbeou.
    the Roberto SE shaved

    '*Roberto* shaved *himself*.'

  b. *O Roberto se* odeia.
    the Roberto SE hates

    '*Roberto* hates *himself*.'

(64) a. *O Roberto* barbeou *ele mesmo*.
    the Roberto shaved PRON same

  b. *O Roberto* odeia *ele mesmo*.
    the Roberto hates PRON same

However, emphatics, unlike the SE clitics, are not used *exclusively* to signal reflexivity: they can also be coconstrued with non-c-commanding and non-local antecedents, merely signaling focus (VIEIRA, 2015). The examples below are adapted from Grolla (2011, pg. 78-80):

(65) a. [A atitude da *Amy*] prejudicou *ela mesma*.
    the attitude of-the Amy damaged PRON same

    '*Amy*'s attitude damaged *her(self)*'

  b. [A força da *Marta*] está dentro d*ela mesma*.
    the force of-the Marta is inside of-PRON same

    '*Marta*'s strength is inside *her(self)*'

(66) a. *O Roberto* disse pra Joana [que *ele mesmo* comeu o bolo].
    the Roberto said to-the Joana that PRON same ate the cake

    '*Roberto* said to Joana that *he himself* ate the cake.'

  b. *O Roberto* não reclama quando seus alunos atrasam para uma reunião.
    the Roberto not complains when his students be.late for a meeting.
    [*Ele mesmo* faz isso com frequência].
    PRON same does this with frequency.

    '*Roberto* doesn't complain when his students are late for a meeting. *He himself* does that frequently.'

These facts present us with a paradox. In contexts like (64), the complex emphatic pronouns seem to function as reflexives for the purposes of RCB: they fulfill the duty of reflexive-marking semantic predicates that are reflexive. However, in (65)-(66), these very same expressions do not function like reflexives for the purposes of RCA: if they did, (65)-(66) would contain reflexive-marked syntactic predicates that are not actually reflexive, violating RCA. In other words, the concept of reflexive-marking in (61) has to meet contradictory demands for RCA and RCB: it has to encompass the configurations where emphatic pronouns appear as

arguments for the latter (cf. (64)), but exclude them for the former (cf. (65)-(66)).

The solution to this puzzle involves rejecting the view that the RCA is a universal principle that applies equally to syntactic predicates in all languages. In Chapter 5 I propose that: (i) RCA should be stated as a constraint on reflexive pronouns rather than on syntactic predicates (as I already did in (18) above); and (ii) RCA only applies to pronouns whose reflexive status is fully grammaticalized. This has not yet happened for BP emphatics.[16]

Note that, on a system of this sort, we lose the symmetry that existed between RCA and RCB in Reinhart and Reuland's (1993) version of the RCT. In fact, given the results gathered in the previous sections, the principle that does the work of RCA winds up being symmetric to the principle that replaces Chain Economy – what I called SYNTACTIC CONDITION B in section 3.1.2. Both the Revised RCA and Syntactic Condition B regulate the binding possibilities of different kinds of pronouns within the domain of the syntactic predicate. In this sense, the system we have at this point is actually closer in its overall structure to the Classical Binding Theory (CBT) of Chomsky (1986b) than to Reinhart and Reuland's proposal. However, my Revised RCA still differs from CBT's Condition A in that it is based on predicate-argument structure, rather than on governing categories, phases or other configurational notions. A partial summary of the version of the RCT we arrived at is given below:

(67)   REVISED RCA (VERSION 2)

A grammaticalized reflexive which is an argument of a syntactic predicate P is bound by a higher argument of P.

(68)   REVISED RCB

A prototypically non-reflexive semantic predicate is reflexive-marked.

(69)   SYNTACTIC CONDITION B

A +R argument of a syntactic predicate P is not bound by any higher argument of P.

The concept of reflexive marking adduced in (68) still refers to the property of having a reflexive pronoun as an argument, as per (61) above. The major difference is that RCA is no longer stated in terms of this concept, since it is only supposed to apply to *grammaticalized*

---

[16] I still want to say that BP emphatic pronouns are reflexives because they do carry a reflexivizing function – i.e. they are capable of signaling reflexivity for predicates of all sorts. The reflexivizing function of emphatics in BP clearly derives from their marked status with respect to bare pronominals: *ele mesmo* is both phonologically as well as morphosyntactically more complex than *ele*. A similar point could be made for pronominals accompanied by the *self* modifier in Old English, which served double duty as markers of semantic reflexivity and contrastive focus in non-reflexive contexts (KÖNIG; SIEMUND, 2000a; SIEMUND, 2002).

reflexives. This means, in effect, that not all expressions that function as reflexive pronouns are governed by RCA (BP emphatic pronouns being a case in point). In the next section, I discuss further problems with Reinhart and Reuland's (1993) RCA which will motivate an even more radical revision of the treatment of reflexives, to be undertaken in Chapter 5.

3.3 ON REFLEXIVITY CONDITION A: THE TWO REFLEXIVES HYPOTHESIS

In this section, I discuss a conjecture that is hardwired into Reinhart and Reuland's (1993) Reflexivity Condition A (RCA): the Two Reflexives Hypothesis (TRH). The TRH was embraced, under various guises, by most of the post-CBT literature as a solution to the problems posed by discursive occurrences of reflexives (LEBEAUX, 1985; MANZINI; WEXLER, 1987; ZRIBI-HERTZ, 1989; CHIERCHIA, 1989; POLLARD; SAG, 1992; BAKER, 1995; STEEDMAN, 1996; GOLDE, 1999; HORNSTEIN, 2001; SAFIR, 2004; BÜRING, 2005; POSTAL, 2006). I start this section by enumerating some basic properties of discursive reflexives that theories need to explain, regardless of whether or not they subscribe to the TRH. I then argue that the TRH (in particular RCT's version) fails to do this because it does not adequately identify the conditions under which the interpretation of reflexives is influenced by discourse properties like empathy or perspective (KUNO; KABURAKI, 1977; OSHIMA, 2007; CHARNAVEL, 2019).

In Chapter 1, I discussed two theories which take the distribution of English reflexives to be fully governed by a single grammatical condition. This UNIFIED approach, as I called it, presumes that a single syntactic generalization can express the contexts in which reflexive anaphors are interpretable. As I argued, this view runs into serious trouble upon confrontation with data like (70) (repeated from Chapter 1).

(70)   a.   That description of *himself* annoyed *DeGaulle* more than you know.

b.   *Albert* was never hostile to laymen who couldn't understand what physicists like *himself* were trying to prove.

c.   *Max* boasted that the queen invited Lucie and *himself* for a drink.

d.   *John*'s favorite topic of conversation is *himself*.

Starting with Lebeaux (1985), many researchers recognized the challenge posed by these examples and sought to provide a kind of damage control to the syntactic theory of anaphora. The general idea was to isolate discourse-governed instances of reflexives like (70) (which I called DISCURSIVE or LOGOPHORIC REFLEXIVES in Chapter 1) from more grammatically

obedient ones – e.g. standard reflexives that are bound by higher coarguments.

This particular way of carving the data suggests that occurrences of English reflexives fall into two fundamentally disjoint categories. There are, on the one hand, PLAIN OCCURRENCES, which obey strict grammatical requirements (e.g. locality); and, on the other, DISCURSIVE OCCURRENCES, which are not conditioned by syntactic rules, but by contextual factors like empathy or point of view. As Charnavel (2019, pg. 19) notes, this way of framing the distinction between the two classes of reflexives entails that plain and discursive occurrences are in complementary distribution: i.e. a reflexive is either plain or discursive, but it cannot be both. Typical examples of each of these categories are given in (71)-(72).[17]

(71) PLAIN OCCURRENCES:

    a. *George* hates *himself*.

    b. *George* seems to hate *himself*.

    c. *George* appears to *himself* to be in love with Dot.

    d. *George* wants to believe in *himself*

    e. *George* expects *himself* to be in the park on Sunday.

(72) DISCURSIVE OCCURRENCES:

    a. *John*'s campaign required that pictures of *himself* be placed all over town.

    b. While looking at the still lake, *John* distinctly heard a voice whispering within *himself*: 'life is wonderful'.

    c. It angered *him* that she tried to attract a man like *himself*.

    d. *Max* expected the queen to invite Mary and *himself* for a drink.

    e. *John* had worked hard to make sure that the twins would be well taken care of. As for *himself*, it was relatively unlikely that anyone would be interested in hiring an ex-convict who had little in the way of professional skills.

As soon as we make the distinction between PLAIN and DISCURSIVE occurrences, many complications introduced into the CBT (especially in its ever-changing definition of domain) can potentially be stripped away. We can focus on defining the syntactic conditions governing plain occurrences, leaving discursive occurrences out of the grammar.

---

[17] There are other terminologies around: e.g. long-distance *vs.* local reflexives (LEBEAUX, 1985; HORNSTEIN, 2001), plain *vs.* exempt reflexives (CHARNAVEL; SPORTICHE, 2016), syntactic *vs.* non-syntactic reflexives (POLLARD; XUE, 1998). The examples in (72) are drawn from Lebeaux (1985, pg. 346), Zribi-Hertz (1995, pg. 339), Zribi-Hertz (1989, pg. 718), Reuland (2011, pg. 28) and Pollard and Sag (1992, pg. 264).

There are different ways one can incorporate this into one's theory. A naive approach could take the plain/discursive distinction to reflect a homonymy between a genuine reflexive form and complex NP that consists in a pronominal attached to an intensifier. The most well-known version of this approach is found in Baker (1995), but the basic insight underlying it comes from early work by Partee (1965), Helke (1970) and Postal (1971). The apparent contradiction between the locality requirement in (71) and the absence of such requirement in (70)/(72) is resolved by claiming that, in the latter cases, what we see is, in fact, the *him*+intensifier form, which patterns distributionally with regular pronominals.

A problem with this alternative, which was already mentioned in Chapter 1, is that it cannot deliver a principled account of why the allegedly pronominal-like homonym of reflexives is excluded from environments like (73) (ZRIBI-HERTZ, 1995; CHARNAVEL, 2019).[18]

(73) a. *It angered *him* that she tried to attract *himself*. (cf. (72c))

b. **Max* expected the queen to invite *himself* for a drink. (cf. (72d))

The contrast between (72) and (73) tells us that we need some systematic criteria for delimiting the contexts in which discursive occurrences flourish. A natural conjecture is that *grammar itself* can determine the conditions in which discursive occurrences emerge on the basis of some syntactic property (i.e. one which distinguishes (72) from (71)/(73)).

This idea is what I call the Two Reflexives Hypothesis (TRH). The hallmark of the TRH, as I see it, is not the notional distinction between plain and discursive occurrences *per se*, but the suggestion that this distinction is definable in syntactic terms. This leads to the conclusion that it is the job of the grammar to decide when reflexives need to obey strict syntactic constraints (e.g. RCA) and when they are free to be interpreted logophorically.

This view is clearly embraced by the RCT. Reuland (2011, pg. 92), for instance, claims that whether or not a reflexive is free from grammatical conditions "is purely determined by its structural position". The classical statement of Reflexivity Condition A (RCA), repeated below, is carefully constructed in order to exclude from its domain discursive occurrences like (72).

---

[18] We cannot attribute the unacceptability of (73) to the failure to meet discourse constraints on discursive occurrences (e.g. focus or perspective), otherwise the parallel cases in (72c) and (72d) would be excluded as well. As Charnavel (2019) shows, the presence of an adequate discourse context is never a sufficient condition to license discursive occurrences – it is, at best (in the case of perspective), a necessary condition. Focus in particular, contrary to what Baker (1995) claims, is not even a necessary condition for discursive occurrences, as examples like (72a) and (72b) show (ZRIBI-HERTZ, 1995, pg. 338-339). Though a contrastive reading of *himself* as a focalized element is certainly possible for these cases given a suitable context, this is not the most natural interpretation. The discourse factor responsible for motivating the use of a reflexive in (72a) and (72b) is clearly the speaker's intention to empathize with the point of view of the person they are talking about.

(74)   REFLEXIVITY CONDITION A (RCA)

A reflexive-marked syntactic predicate is reflexive.

a.   A predicate P is REFLEXIVE-MARKED iff either P is lexically reflexive or one of P's arguments is a reflexive pronoun.

b.   The SYNTACTIC PREDICATE formed of P is P, all of the projections that realize grammatical functions associated with P (i.e. P's SYNTACTIC ARGUMENTS), and the subject of P.

The RCA only cares about reflexive pronouns that are arguments of syntactic predicates. In cases where they are not, reflexives are claimed to be *exempt* from RCA. The theory defines two syntactic contexts in which this can occur, summarized in (75).

(75)   A reflexive is exempt from RCA (MENUZZI, 1999, pg. 160):

a.   when its predicate does not have a subject (in which case the predicate does not qualify as a syntactic one);

b.   when the reflexive itself is not a syntactic argument of a predicate (in which case it does not reflexive-mark anything).

The contexts characterized by (75a) cover reflexives within picture NPs lacking a specifier (cf. (72a)) and locative PPs (cf. (72b)). (75b) describes reflexives that embedded within NP adjuncts (cf. (72c)) and conjunctions (cf. (72d)), as well as A′-reflexives of different varieties: the *as for* reflexive in (72e), as well as other topic and focus reflexives in (76). The latter category arguably includes focus reflexives which are overtly in argument positions but covertly move to A′-positions at LF, where RCA is assumed to apply, like the example in (77).

(76)   a.   It was HIMSELF that *Felix* claimed ordinary people could never understand.

b.   HIMSELF, *Felix* claimed that ordinary people could never understand.

(POSTAL, 2006, pg. 8-9)

(77)   A: Did Jim tell Sue that Jane would kiss Bill in the moonlight?

B: No, *Jim* told Sue that Jane would kiss HIMSELF in the moonlight.

(SAFIR, 1992, pg. 41):

The reason why Reinhart and Reuland feel the need to offer a theoretical account of discursive occurrences in the first place is the perception that these constitute a natural kind. An inference to a natural kind is based on a cluster of observable properties which co-vary in

predictable ways. The core properties of discursive occurrences that theories need to explain are the absence of locality and c-command requirements, which clearly distinguish the reflexives in (72) from those in (71). (This does not mean, of course, that *all* discursive occurrences have non-local and non-c-commanding antecedents. The absence of locality or c-command are merely interpreted as a sufficient conditions to identify discursive occurrences.)

However, as Charnavel and Sportiche (2016) argue, using locality and c-command alone as a criteria for identifying discursive occurrences is unsatisfactory from a methodological standpoint. This practice implies, in effect, that defining discursive occurrences in terms of (the complement of) the properties that are typically used to define constraints on reflexives. This could lead to circular reasoning: given a potential counterexample to one's favored constraint on reflexives (stated in terms of locality or c-command) one could simply redefine c-command or locality in order to make the example in question come out as discursive. In order to avoid circularity, it is necessary to invoke criteria for identifying discursive occurrences that are independent of the syntactic notions in terms of which constraints on reflexives are stated.

A feature which is often mentioned as an independent trait of discursive occurrences as such is the fact that they must receive a logophoric interpretation – i.e. they must refer to an entity whose perspective is being empathized within the discourse at hand (KUNO; KABU-RAKI, 1977; ZRIBI-HERTZ, 1989; GOLDE, 1999; CHARNAVEL; ZLOGAR, 2015). There are conflicting definitions of what exactly counts as a possible perspective-bearer within a discourse (SELLS, 1987; STIRLING, 1993; MINKOFF, 2004; OSHIMA, 2007; CHARNAVEL, 2019). What all of these definitions minimally share, however, is an animacy requirement: in order to be a bearer of point of view with whom the speaker can empathize, an entity has to be animate, or be construed as animate. This means that logophoric interpretations can be detected by exploiting the animacy property, which holds of necessity of all perspective-bearers.[19]

If receiving a logophoric reading is a necessary condition on discursive reflexives, inanimate reflexives are *never* discursive. From this it follows, conversely, that animate occurrences

---

[19] This is a slight simplification, because, as Charnavel and Sportiche (2016, pg. 39) note, grammatically inanimate NPs can be used as proxies for groups of animate agents with whom the speaker can potentially empathize (e.g. institutions, countries, corporations) as well as for agents whose animate status is dubious (e.g. robots, babies, animals), as the following discursive cases collected from the internet show:

(i) a. "*NYPD* Is Now Bypassing Journalists to Write News Stories About *Itself*."
     b. "*NASA's Mars rover Curiosity* is doing a bit of mechanical navel-gazing on the Red Planet, snapping ultra-clear pictures of *itself*."

I assume that in these anthropomorphized uses inanimate NPs also pick out their semantic values from a set of entities capable of holding a point of view.

of reflexives that cannot be replaced by inanimate ones are *always* discursive. A restriction against inanimate antecedents can, therefore, function as a *test* (i.e. a sufficient condition) for identifying discursive occurrences (MINKOFF, 2004; POSTAL, 2006; CHARNAVEL, 2019).

Reflexives within PPs and NP adjuncts as well as picture NP reflexives pass the test, because they are all generally unacceptable with inanimate antecedents. The data below are adapted from Minkoff (2004), Postal (2006) and Charnavel and Zlogar (2015):

(78)   a.   *Winston Felix* claimed that physicists like *himself* were a godsend.

         b. *?*Syntactic Structures* claimed that books like *itself* were a godsend.

(79)   a.   *Bobby* pushed the boat away from *himself*.

         b. *?*The sea* pushed the boat away from *itself*.

(80)   a.   That picture of *himself* made folks think they should avoid *Bill*.

         b. *?Those pictures of *itself* made folks think they should avoid *the cave*.

Though judgments are less sharp, Postal (2006) also mentions reflexives in focalization (81) and topicalization (82) structures as being subject to the animacy constraint:

(81)   a.   It was HIMSELF that *Felix* claimed ordinary people could never understand.

         b. ??It was ITSELF that *Syntactic Structures* claimed ordinary people could never understand.

(82)   a.   HIMSELF, *Felix* claimed that ordinary people could never understand.

         b. ??ITSELF, *Syntactic Structures* claimed that ordinary people could never understand.

In addition to having logophoric interpretations (diagnosed by the animacy test) and not requiring local c-commanding antecedents, discursive reflexives are often claimed to share other traits which distinguish them from plain occurrences. I will go over three of these below.

The first is the presence of strict readings in ellipsis constructions (LEBEAUX, 1985; REINHART; REULAND, 1993; HESTVIK, 1995; HORNSTEIN, 2001). The discursive occurrence examples in (83) both have strict readings, as opposed to the plain occurrence examples in (84): e.g. the second conjunct in (83a) can mean that Bill suspected that the pictures of *John* were going to circulate all over the world, but in (84a) can only mean that Stephen hates Stephen. Sloppy interpretations are available in (83a)/(84a) but in (83b)/(84b) these readings are disfavored due to a gender mismatch.

(83)   a.   *John* suspected that those pictures of *himself* were going to circulate all over the world, and so did Bill.   (sloppy/strict)

       b.   *Horace* was mad at Dolly for trying to lure a sober man like *himself* into marriage and Irene was too.   (*sloppy/strict)

(84)   a.   *George* hates *himself*, and so does Stephen.   (sloppy/*strict)

       b.   *George* wants to believe in *himself* and Marie does too.   (?sloppy/*strict)

This distribution is expected in the RCT, since nothing in the theory entails that anaphors as such need to be interpreted as bound variables. This conclusion only follows for anaphors which are subject to RCA, which requires reflexives to be A-bound by (or co-bound with) a coargument – lest their syntactic predicates would not be reflexive. If, however, an anaphor happens to be exempt from the A-binding requirement imposed by RCA, its relation to its antecedent can be of coreference. So Reinhart and Reuland's hypothesis that discursive occurrences are exempt actually predicts that strict readings are only available in these cases.[20]

The second trait that distinguishes discursive occurrences from plain ones is the allowance of split antecedents. Plural reflexives can be coconstrued with two separate NPs in (85), but not in (86) (HELKE, 1979; LEBEAUX, 1985; OKADA, 1998; CHARNAVEL, 2019):

(85)   a.   *John* protected *Mary* from mobsters like *themselves*.

       b.   *John* told *Mary* that, as for *themselves*, they should leave immediately.

(86)   a.   **John* protected *Mary* from *themselves*.   (KEENAN, 1988, pg. 220)

       b.   **The boys* introduced *the girls* to *themselves*.   (LEBEAUX, 1985, pg. 346)

The third (and last) peculiar property of discursive occurrences I will mention here is the absence of complementary distribution with pronominals, which is illustrated in the contrast between (87) and (88) (BOUCHARD, 1984; LEBEAUX, 1985; BÜRING, 2005).

(87)   a.   *Max* boasted that the queen invited Lucie and {*himself* / *him*} for a drink.

       b.   *John* distinctly heard a voice whispering within {*himself* / *him*}.

       c.   *John* told *Mary* that, as for {*themselves* / *them*}, they should leave immediately.

---

[20] Hestvik (1995) argues that this is too strong. He reports that, in subordination contexts, plain reflexives can also bear strict readings: e.g. **George admires himself more than Stephen does** can be interpreted as saying that Stephen admires *George*, despite the fact that *himself* is subject to RCA. Reinhart and Reuland (1993, pg. 675) offer a tentative explanation for similar counterexamples in Sag (1976) by claiming that reflexives which allow strict interpretations are always foci, and, thus, exempt from RCA. Be that as it may, it does seem to be true that discursive occurrences are more prone to accept strict readings *regardless of the context* than plain ones.

(88)   a.   *George* seems to hate {*himself* / **him*}.

   b.   *George* appears to {*himself* / **him*} to be in love with Dot.

   c.   *George* expects {*himself* / **him*} to be in the park on Sunday.

Since plain reflexives have local antecedents and pronominals have *non*-local ones, whenever a reflexive can be replaced by a pronominal, we can infer that it is not a plain occurrence.

However, we cannot safely claim that a reflexive *is* a plain occurrence solely on the basis of the fact that it cannot be replaced by a pronominal, since pronominals may be excluded from some discursive contexts for independent reasons. For example, all of the properties mentioned above warrant the conclusion that reflexives within conjoined NPs are discursive occurrences. This is correctly predicted by RCT's criteria of exemption in (75) as well. However, in structures like (89), replacing the reflexive with a pronominal violates RCB:

(89)   *Bobby* praised both Marta and {*himself* / **him*}.

Given the properties mentioned above, I still want to say that (89) is a discursive occurrence. But it is one which exceptionally is in complementary distribution with a pronominal (since the latter would violate RCB). In sum, non-complementarity with pronominals entails that a reflexive is discursive, but not vice-versa – cf. Büring (2005, pg. 228) for a similar point.

I do not wish to commit myself too seriously to the wholesale validity of the correlations stated above. It may well be that not all discursive occurrences of reflexives have all of the three properties I mentioned, or that not all plain occurrences lack them. The only properties which I take to be really sufficient conditions to identify discursive reflexives in the present context are: (i) the possibility of non-local binding; (ii) the possibility of coconstrual with non-c-commanding antecedents, (iii) the animacy requirement and (iv) the lack of complementarity with pronouns. The other properties (strict readings and split antecedents) are mentioned here for phenomenological purposes: the goal is simply to identify a natural cluster of attributes associated with discurisve occurrences which call for some sort of explanation. The explanation proposed by Reinhart and Reuland (1993) is that these occurrences are *exempt* from RCA – where the criteria for exemption is syntactically determined (cf. (75)). In the next section, I will show that this explanation, which is a version of the TRH, is not in fact adequate.

3.3.1 Discursive properties in plain reflexives?

A first indication that there may be something amiss in RCT's exemption-based account of discursive reflexives can be seen in connection to Reinhart and Reuland's (1993, pg. 670) analysis of reflexives within NP adjuncts, such as (72c), repeated as (90):

(90)   It angered *him* that she tried to attract a man like *himself*.

As I mentioned above, Reinhart and Reuland (1993) argue that the reflexive in (90) is exempt from the RCA because it is not an argument of a syntactic predicate. However, even though this reflexive is indeed *not* a syntactic argument of a verb, it *is* a syntactic argument of the comparative predicate headed by the adjective *like*. This predicate may or may not count as a syntactic predicate (this depends on whether one thinks predication always implies the presence of a null PRO subject), but note that a structurally parallel case like (91) is not an acceptable discursive occurrence (MENUZZI, 1999, pg. 161):

(91)   *It angered *him* that she tried to attract a man proud of *himself*.

The RCT does not tell us what is going on here. We will see that this is a symptom of a larger problem: since the definition of exemption does not take into account specific semantic properties of predicates that could distinguish (90) from (91), it cannot incorporate the fact that discursive occurrences may be structurally indistinguishable from plain occurrences.

We will now see many examples where Reinhart and Reuland's (1993) theoretical characterization of discursive occurrences as exempt from RCA does does not adequately circumscribe the class of reflexives which exhibit the discursive properties mentioned in the previous subsection. In particular, I show, following Menuzzi (1999), that there are instances of reflexives which are ARGUMENTS OF SYNTACTIC PREDICATES but which, nonetheless, exhibit non-local antecedents, logophoric interpretations and some of other typical traits of discursive occurrences. Despite conforming to the phenomenological characterization of discursive reflexives presented above, these cases do not count as exempt, according to the criteria in (75).

The counterexamples to Reinhart and Reuland's theory of discursive reflexives fall into two categories, classified according to the type of predicates they involve: predicates with expletive subjects (cf. (92)) and comparison predicates (cf. (93)).[21]

---

[21]   The examples in (92) are taken from Kuno (1987, pg. 99), Menuzzi (1999, pg. 164), Culicover (1997, pg. 298) and Postal (2006, pg. 12). The examples in (93) are taken from Safir (1992, pg. 26), Menuzzi (1999, pg. 165), Postal (2006, pg. 11), and the internet. The division between these two classes of examples is done for

(92)   PREDICATES WITH EXPLETIVE SUBJECTS (KUNO, 1987; POLLARD; SAG, 1994):

    a.   *They* made sure [that it was clear to *themselves* that this needed to be done].

    b.   *Paul* claimed that [it would be good for *himself* if Mary left].

    c.   *Mary* said that [it would be very difficult for *herself* to accept your theory].

    d.   *Mary* says that [it was incumbent upon *herself* not to abandon the car after the breakdown].

    e.   *That author* claimed that [there would always be *himself* for you to count on].

(93)   COMPARISON PREDICATES (KOSTER, 1978; NAPOLI, 1989; SAFIR, 1992):

    a.   *The men* said that [the new administrators would be very much like *themselves*].

    b.   *The veterans* think that [the new recruits are more qualified than *themselves*].

    c.   *Claudine* treated [you as inferior to *herself*].

    d.   "*Pears*, a senior project manager in the U.K. for Sony PlayStation, believes [some customers would be similar to *herself*]".

    e.   "[D]oes *she* love her hubby because [he resembles *herself*]?"

The reflexives in (92) and (93) are all blatant violations of RCA: they occupy argument positions, thereby reflexive-marking their syntactic predicates, which are nonetheless not reflexive. Note that these reflexives really are arguments and their predicates really do satisfy RCT's definition of syntactic predicate (i.e. they all contain subjects). So there is no way these cases can be treated as exempt by Reinhart and Reuland's standards, repeated in (94).[22]

---

convenience. In Chapter 5, I will show that these all of these cases actually fall under a single generalization, which also encompasses other cases of discursive reflexives in argument position

Koster (1978, pg. 168), despite being the first to recognize examples like (93), reports a case like (14c) as bad. However, reflexive complements of *resemble* with non-local antecedents are well attested on the internet:

(i)   a.   "In this comedic short, *a psychoanalyst* encounters a patient [who eerily resembles *himself*]."
    b.   "*People* trust people [that resemble *themselves*]."
    c.   "*[T]he child*'s libido flows to an outside person [who in some way resembles *himself*]."

Note that example (ic) not only violates locality, but also the c-command requirement.

[22] For my argument to work, it must be shown that these examples are *not reducible to focus*, which is the only way Reinhart and Reuland can analyze locally free reflexives in argument positions as exempt (cf. (82) above). No focus seems to be necessary in (92)-(93), though more research on the topic is required. Consider the variants of (92b) and (93a) in (i), where focal stress (signaled by caps) does not fall on the reflexive:

(i)   a.   Though he knew that her departure would be for the best, *Paul* never claimed that it would be EASY for *himself* if Mary left.
    b.   A: Did *Claudine* treat HER ENTIRE FAMILY as inferior to *herself*?
       B: No, *she* treated HER HUSBAND as inferior to *herself*.

The judgments here are particularly murky, but a similar argument built on similar examples is sketched in Safir (1992, pg. 41-42). Additional support comes from contrasts such as the one below:

(94)   A reflexive is exempt from RCA (MENUZZI, 1999, pg. 160):

      a.   when its predicate does not have a subject (in which case the predicate does not qualify as a syntactic one);

      b.   when the reflexive itself is not a syntactic argument of a predicate (in which case it does not reflexive-mark anything).

Even though these reflexives do not fall under Reinhart and Reuland's *theoretical* account of discursive occurrences in (94), they do conform to their *descriptive* characterization of this category: they are locally free while not producing strong unacceptability like most cases of RCA violations. In addition to their non-local character, they also share with discursive occurrences most of the other properties we mentioned in Section 3.3.

As is characteristic of expressions which receive logophoric interpretations, the reflexives in (92)-(93) are generally unacceptable with inanimate antecedents, as the examples below illustrate (cf. POSTAL, 2006, pg. 12):

(95)   a.   *That book* claimed that [there would always be *itself* for people to count on].

      b.   *Syntactic Structures* considered [all other books worse than *itself*].

These examples also allow strict readings under ellipsis. The second conjunct in (96a) can mean that John thinks Mary's leaving would be good for *Paul* and, in (96b), it can mean that Pears' boss thinks that the customers will be similar to *Pears*.

(96)   a.   *Paul* thinks that [it would be good for *himself* if Mary left], and John does too.

      b.   *Pears* thinks that [the customers will be similar to *herself*], and so does *her* boss.

Moreover, like other cases of discursive occurrences, the long-distance argument reflexives in (92)-(93) also tolerate split antecedents, as in (97):

(97)   a.   *Mary* told *John* that [it was incumbent upon *themselves* not to abandon the car after the breakdown].

      b.   *The men* told *the company executives* that [the new administrators would be very much like *themselves*].

(ii)   a.   ?*Mary* always thought her brother, but not her father, was similar to *herself*.

      b.   *Mary* always thought her brother, but not her father, was mad at *herself*.

What (ii) suggests is that there is a real difference with respect to non-local binding between non-focused reflexives within comparative predicates (iia) and other kinds of predicates (iib).

Further, all the occurrences of reflexives in (92)-(93) are completely interchangeable with pronominals. This lack of complementarity is the most commonly used test to identify discursive occurrences of reflexives:

(98) a. *Paul* claimed that [it would be good for {*himself / him*} if Mary left].

  b. *Mary* said that [it would be very difficult for {*herself / her*} to accept your theory].

  c. *Claudine* treated [you as inferior to {*herself / her*}].

  d. Does *she* love her hubby because [he resembles {*herself / her*}].

Lastly, many of the contexts in (92) and (93) also allow the reflexives to be coconstrued with non-commanding antecedents – a property which is usually thought of as exclusive to pronominals (BÜRING, 2005). The first two examples below were suggested by Peter Culicover (p.c.) and (99c) and (99d) were collected from the internet:

(99) a. *Bobby*'s ideas are clear to *himself*.

  b. ?*Bobby*'s objectives are going to be difficult for *himself* to achieve.

  c. "If *this young man*'s friends are like *himself*, the evening will be an amusing one."

  d. "*[T]he respondent*'s friends are similar to *himself* on a variety of characteristics."

The argument against Reinhart and Reuland's (1993) account of discursive occurrences as exempt from RCA applies more generally to *any* version of the TRH. There is simply no *syntactic* generalization about the structures of (92) and (93) which can distinguish them from structures in which plain occurrences of reflexives appear. Consider the contrasts below:

(100) a. *Paul* claimed that [it would be good for *himself* if Mary left].

  b. *?*Paul* claimed that [it would help *himself* if Mary left].

(101) a. *She* loves her hubby because [he resembles *herself*].

  b. **She* loves her hubby because [he remembers *herself*].

In all of the examples we've been examining, the reflexives are arguments of predicates which have subjects (i.e. syntactic predicates, in RCT's terms). In this respect, the pairs within (100) and (101) are structurally indistinguishable. Nonetheless, we see a clear contrast in acceptability between the a-cases and the b-cases. This suggests that whatever licenses the reflexives these kinds of contexts must have something to do with the semantic difference between

predicates like *good/resemble* and predicates like *help/remember*. This contradicts the very essence of the TRH, which says that discursive occurrences can be identified purely on the basis of the syntactic properties of the environments in which reflexives appear.

If the RCT wants to account for why particular occurrences of reflexives pattern together with respect to properties like non-locality, logophoric readings, split antecedents, etc., (92) and (93) should fall under the same rubric as discursive reflexives. However, since the reflexives in (92) and (93) are, in relevant respects, syntactically indistinguishable from RCA abiding ones (i.e. they are all arguments of syntactic predicates), the idea that syntax determines when discursive reflexives are acceptable must be rejected.[23] As Dalrymple (1993, pg. 57-58) puts it,

> [t]here is no neat, syntactically-definable division of positions into those in which an anaphor must obey syntactic constraints and those in which an anaphor must obey discourse constraints. [...] [N]o syntactic definition of availability for discourse binding is available.

The TRH is precisely an attempt to provide a syntactic definition of availability for discourse binding. Reinhart and Reuland (1993) built a version of the TRH into their formulation of RCA by conditioning the compliance to RCA to reflexives which reflexive-mark syntactic predicates. The alternative I pursue in Chapters 5-6 strips the TRH away from the theory of anaphora. It treats discursive reflexives as *tolerable violations* of (a simplified version of) RCA which are rendered acceptable under particular pragmatic conditions. The basic idea is that violations of grammatical constraints contribute to a form of markedness. Certain instances of marked violations of grammar may be "saved" from unacceptability by becoming become pragmatically associated with marked interpretations (GRICE, 1989; HORN, 1984; RETT, 2020).

3.4 CONCLUSION

In spite of its advantages, the RCT is not immune to many of the problems that also affected the earlier approaches to anaphora within MGG surveyed in Chapter 1. As we saw in the present chapter, the Chain Theory part of the RCT recreates many of the same problems that

---

[23] It is also not clear whether the TRH (including RCT's version) is immune to the accusations I laid out, in Section 3.3, against homonymy approaches to the plain/discursive distinction. If words are individuated not only by their phonology, but also in terms of their syntactic and semantic properties, plain reflexives and logophorically interpreted exempt reflexives will necessarily count as distinct (albeit related) lexical items. As Charnavel (2019) argues, this seems particularly problematic in light of the fact that the formal coincidence between markers of reflexivity and point-of-view is not parochial to English, but is found in typologically unrelated languages: e.g. Icelandic (MALING, 1984), Japanese (KUNO, 1987; OSHIMA, 2004), Turkish (MAJOR; OZKAN, 2018) and French (CHARNAVEL; SPORTICHE, 2016). Proponents of the TRH would have to stipulate the same unexplained homophony for each of these languages separately.

haunted Condition A and Condition B of the CBT. The simplest way to circumvent these, I argued, is to simply abandon the Chain Condition and incorporate its residues into the Reflexivity Condition A (RCA) and its syntactic analogue – a principle I called Syntactic Condition B.

We also found that a definition of reflexive-marking embodying the dubious concept of *lexical reflexivity* is problematic. We saw that by replacing this notion for the notion of prototypical reflexivity we are able to arrive at a revised version of Reflexivity Condition B (RCB) which is sensitive to compositional and pragmatic properties of predicates and is able to derive typologically well-attested universals of reflexive-marking.

At the end of this exercise, we were left with a theory of anaphora where all of the constraints are stated as correspondence principles between narrow syntax and A-binding in semantics, with no syntactic encoding of semantic dependencies (via indices or Agree):

  (i) RCA requires reflexive arguments to be A-bound within their syntactic predicates;

 (ii) RCB requires prototypically non-reflexive semantic predicates (i.e. predicates whose variables are A-bound by the same antecedent) to have reflexive pronouns arguments;

(iii) Syntactic Condition B requires pronominal arguments (i.e. +R elements) to not be A-bound within their syntactic predicates.

However, even the improved version of the RCT we provisionally arrived at still cannot deal with many of the problems noted throughout this chapter. For instance, Syntactic Condition B is still incompatible with languages like Brazilian Portuguese (BP) which allow pronominals to be locally A-bound in certain contexts. The problem of providing an adequate account for discursive occurrences of reflexives in English also remains, even after we incorporate Chain Visibility into the RCA, as I proposed above. The theoretical characterization of discursive occurrences as exempt from RCA simply does not match the descriptive properties of discursive occurrences mentioned in Section 3.3 (MENUZZI, 1999). The problems that arise in this connection are not specific to any specific statement of RCT: they affect any attempt to define the licensing contexts for discursive occurrences of reflexives in strictly syntactic terms.

In the following chapters, I propose an alternative theory of anaphora which attempts to overcome these problems. I will propose an account where anti-locality effects involving pronominals are relativized to individual languages and where discursive occurrences are not licensed by means of exemption from RCA. Before doing that, however, I need lay out the basic theoretical framework within which my proposal will be stated: the theory of Simpler Syntax.

Part II

SEEKING SIMPLICITY

4 THE THEORY OF SIMPLER SYNTAX

The goal of this chapter is to motivate a grammatical framework which is suitable for stating hypotheses about anaphora that overcome some of the difficulties that upset the previous approaches we have encountered so far. The theory presented here will be version of the SIMPLER SYNTAX (SiSx) approach of Culicover and Jackendoff (2005) and, especially, Culicover (2009, 2013b, 2021). The particular implementation I discuss is more formal than the usual statements of SiSx. The formal underpinnings employed here draw heavily from work on Lexical-Functional Grammar (LFG) (BRESNAN; KAPLAN, 1982a; KAPLAN, 1995; DALRYMPLE, 2001; ASUDEH; DALRYMPLE; TOIVONEN, 2013).[1]

SiSx was born out of a dissatisfaction with the conceptual and formal assumptions of MAINSTREAM GENERATIVE GRAMMAR (MGG) (CHOMSKY, 1957, 1965, 1981, 1995). It differs from MGG in adopting a model-theoretic formalism, an autonomous level of representation for grammatical functions and a richly structured lexicon. Unlike most work in MGG, SiSx is also open to exploring extra-grammatical or functional explanations for grammatical phenomena. Each of these features will be explained in what follows.

In spite of diverging from MGG with respect to core empirical and architectural assumptions, SiSx is firmly dedicated to pursuing what Jackendoff (2007a, chap. 2) identifies as the two founding themes of Generative Grammar: MENTALISM and COMBINATORIALITY.

MENTALISM is the view that language is a product of the mind/brain of individual speakers. SiSx is committed to a particularly strong version of this, which Bresnan and Kaplan (1982b), following Chomsky (1965, pg. 9), dub the COMPETENCE HYPOTHESIS. This is the suggestion that the *same* body of knowledge (i.e. LINGUISTIC COMPETENCE) underlies *every* type of language-related behavior (e.g. speaking, reading, learning). In this approach, the linguist's theoretical constructs are not only psychologically real in an abstract sense, but must be integrated into an account of how language is actually processed and acquired by real speakers.

The second founding theme that MGG shares with MGG is COMBINATORIALITY: i.e. the view that knowledge of language is instantiated as a finite system of *rules* that define (or "generate") an unbounded array of structured expressions. The linguist's explicit formulation of these rules (i.e. the grammar), must, ideally, explain the structures speakers assign to all linguistic utterances – making no principled distinction between pure manifestations of "core

---

[1] Large portions of the present chapter come from Varaschin (to appear), where I provide an explicit comparison between LFG and SiSx.

grammar" (as defined by a system of paramaters) and "peripheral data" (CULICOVER, 1999).

In line with these commitments, SiSx seeks to characterize the human language capacity in a way that is: (i) PSYCHOLOGICALLY PLAUSIBLE, seeking a graceful integration of linguistic theory with what is known about the structure and function of mind/brain (JACKENDOFF, 2007b, 2011b; CULICOVER, 2015); and (ii) FORMALLY AND DESCRIPTIVELY ADEQUATE, representing generalizations of varying granularities with sufficient precision (JACKENDOFF, 2008; CULICOVER, 1999, 2014, 2021).

The remainder of this chapter summarizes some of the conclusions SiSx reaches in the pursuit of implementing these goals. Section 4.1 provides an overview of the SiSx framework and describes some of its general properties. Section 4.2 focuses on what is perhaps the major difference between SiSx and MGG: the fact that MGG adopts a PROOF-THEORETIC FORMALISM while SiSx opts for a MODEL-THEORETIC one. In Section 4.3 I use the basic model-theoretic toolkit to provide a formal characterization of the different structures SiSx attributes to linguistic expressions, as well as a definition of what counts as a WELL-FORMED EXPRESSION. Section 4.4 lays out the structure of the SiSx lexicon, which includes not only words but also constructions and abstract schemas that are customarily thought of as RULES OF GRAMMAR. Lastly, Section 4.5 explores the idea that EXTRA-GRAMMATICAL constraints may also play a role in explaining some of speakers' acceptability judments.

4.1 OVERVIEW

Like other syntactic theories, SiSx is an attempt to describe and explain the language user's ability to establish a correspondence between meaning and sound or gesture. What defines SiSx is the claim that this correspondence should be as minimal as possible – i.e. that syntax should *only* be invoked when other factors (e.g. semantics, prosody, processing) are insufficient to explain the phenomena at hand. This claim is embodied in the Simpler Syntax Hypothesis (CULICOVER; JACKENDOFF, 2005, pg. 5):

(1)   THE SIMPLER SYNTAX HYPOTHESIS (SSH)
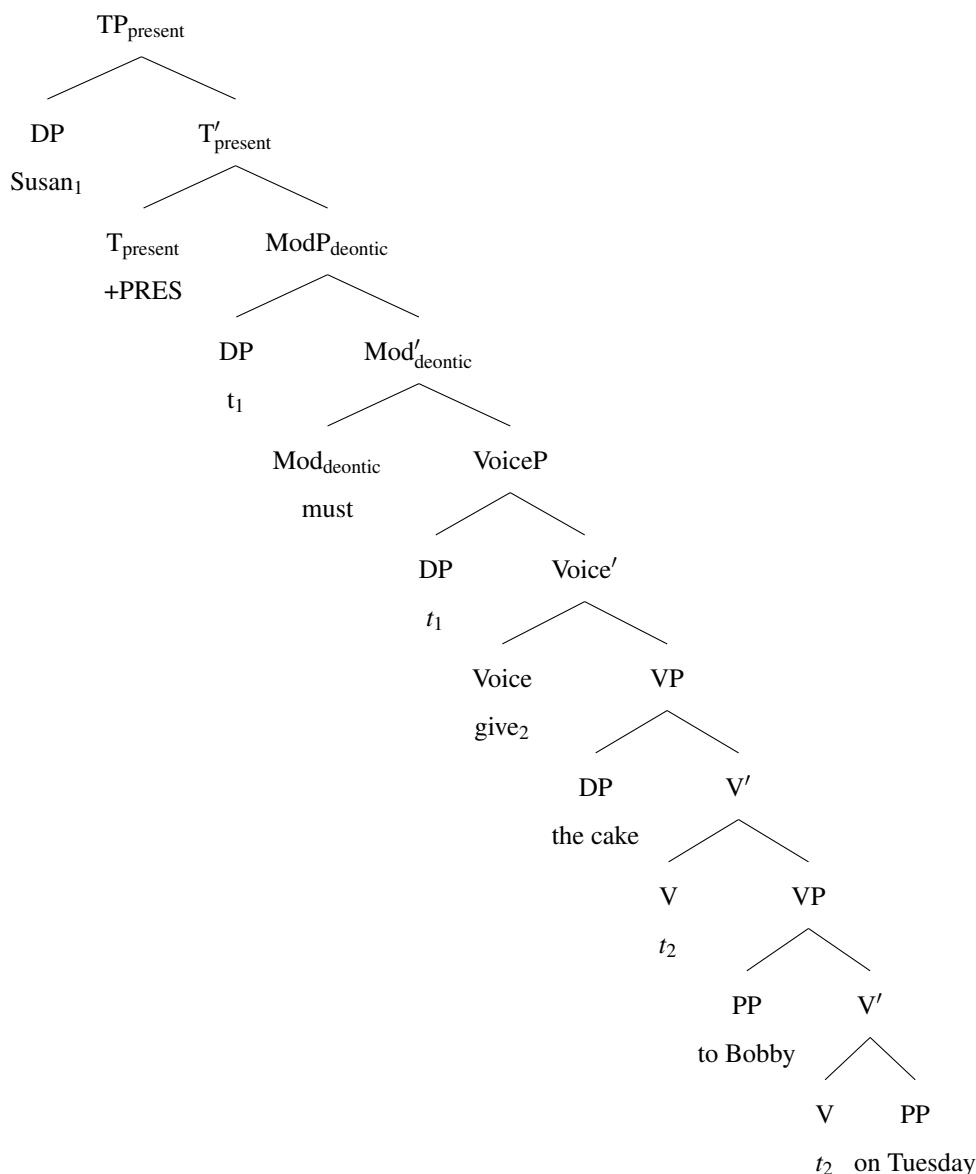
        The most explanatory syntactic theory is one that imputes the minimum structure necessary to mediate between phonology and meaning.
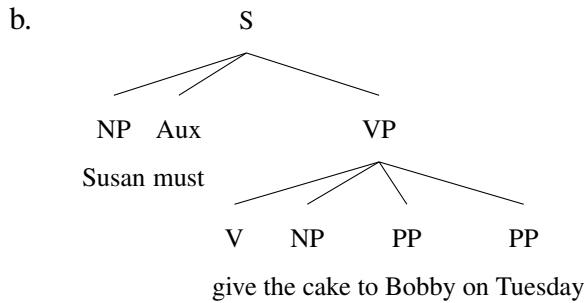
Assuming Chomsky's (1965) notions of descriptive and explanatory adequacy, what the SSH says is that, given a set of descriptively adequate grammars of a language *L*, the one

the theorist should choose (i.e. the more explanatory one) is the one that assigns less syntactic structure to the expressions of *L*. The SSH favors, thus, representational economy (CHOMSKY, 1991; TROTZKE; ZWART, 2014) over other notions of simplicity, such as minimizing the class of possible grammars or the number of principles in particular grammars. The latter two goals have been the main driving forces of MGG since the advent of the Principles and Parameters framework (CHOMSKY, 1973, 1981, 1995).

As an example, contrast the relatively flat constituent structure SiSx assigns to the English sentence *Susan must give the cake to Bobby on Tuesday* in (2b) with the MGG variant in (2a), which is based on Cinque's (1999) decomposition of IP and on the widely adopted VP-shell analysis (LARSON, 1988; KRATZER, 1996; HALE; KEYSER, 1993; CHOMSKY, 1995):

(2) a.

$TP_{present}$
- DP: Susan$_1$
- $T'_{present}$
  - $T_{present}$: +PRES
  - $ModP_{deontic}$
    - DP: t_1
    - $Mod'_{deontic}$
      - $Mod_{deontic}$: must
      - VoiceP
        - DP: t_1
        - Voice$'$
          - Voice: give$_2$
          - VP
            - DP: the cake
            - V$'$
              - V: t_2
              - VP
                - PP: to Bobby
                - V$'$
                  - V: t_2
                  - PP: on Tuesday

b.

```
                    S
         ┌──────┬───────────┐
        NP    Aux          VP
      Susan   must    ┌────┬────┬────┐
                      V    NP   PP   PP
                   give the cake to Bobby on Tuesday
```

MGG opts for structures like (2a) because the grammar that generates them involves *fewer* principles (and is allegedly *more restrictive*) than the one that yields (2b).[2] The idea is that (2a) follows a universal blueprint for structure-building that is virtually *invariant* across languages – one that imposes strict binary branching, endocentricity and a rigid order among heads. Moreover, the hierarchical organization of phrases in (2a) is semantically transparent, reflecting a universal THEMATIC HIERARCHY, in which AGENTS are higher than THEMES, THEMES are higher than GOALS and GOALS are higher than MODIFIERS (see Baker (1997)).

The structure itself, however, is clearly much simpler in (2b): (2b) has fewer degrees of embedding (just two), no empty functional projections (e.g. VoiceP) and no phonetically null elements (traces or deleted copies). Given a suitably flexible interface, (2b) can also be placed in correspondence with a level of SEMANTIC STRUCTURE (JACKENDOFF, 1990b). The semantic properties that (2a) purports to reflect can be more naturally represented in this level, which is independently required to explain inferences that go well beyond what narrow syntax can express.[3] Thus, between representations (2a) and (2b) – the former illustrating simplicity of principles and the latter simplicity of structure – SSH recommends (2b).

A theoretical reason for pursuing the SSH (as opposed to other measures of simplicity) is that it approximates syntactic structures to what is directly inferable from input, thereby reducing the task of the language learner (CULICOVER, 1998, 1999, JACKENDOFF, 2011a). The child has no direct evidence for the traces and empty elements assumed in (2a). As Chomsky (1982, pg. 19) notes, this raises poverty-of-stimulus issues, which call for the invocation of a richer UNIVERSAL GRAMMAR (UG). Insofar as SiSx posits more concrete structures, it

---

[2] The suggestion that (2a) implies a more restrictive grammatical formalism is highly misleading. As Kornai and Pullum (1990) show, as soon as empty elements are introduced, an X′-theory with obligatory binary branching becomes equivalent to an arbitrary context-free grammar that can generate structures like (2b). Similar considerations apply to minimalist descendants of X′-theory (e.g. Chomsky's (1995) Bare Phrase Structure).

[3] For instance, even the rich structure in (2a) fails to encode the inference that Hector is the Source of the cake (in addition to the Agent of *give*), or that cakes are artifacts typically used for eating. The latter influences the interpretation of evaluative adjectives: e.g. a *good cake* is a cake that is good to *eat* (PUSTEJOVSKY, 1995). The phrase-structure formalism has no natural way to represent this.

contributes to the minimalist project of a leaner UG (CHOMSKY, 2005; HORNSTEIN, 2009).

Aside from being more *explanatory*, the option for simpler structures is often more *descriptively adequate* than accounts based on rich uniform representations like (2a). Classic constituency tests, for example, only provide motivation for the the major constituent divisions shown in (2b): VPs, PPs, NPs, etc. The empirical virtues of the SSH also manifest in accounts of specific linguistic phenomena (see Culicover and Jackendoff (2005) and Culicover (2021) for numerous examples). Most arguments for SiSx analyses have the following form:

> Given some phenomenon that has provided putative evidence for elaborate syntactic structure, there nevertheless exist numerous examples which demonstrably involve semantic or pragmatic factors, and in which such factors are [...] impossible to code uniformly into a reasonable syntactic level [...]. Generality thus suggests that, given a suitable account of the syntax–semantics interface, all cases of the phenomenon in question are accounted for in terms of the relevant properties of semantics/pragmatics; hence no complications are necessary in syntax. (CULICOVER; JACKENDOFF, 2005, 5)

As this makes clear, the SSH eschews any kind of covert structure that is motivated exclusively in order to provide a uniform mapping onto semantics. This means that SiSx rejects the SYNTACTOCENTRIC architecture of MGG – i.e. the view that syntax is solely responsible for the combinatorial richness of language (CULICOVER; JACKENDOFF, 2005, pg. 17) –, as well as the assumption of INTERFACE UNIFORMITY – i.e. the view that the interface between syntax and semantics is perfectly transparent (CULICOVER; JACKENDOFF, 2005, pg, 47).

As an alternative, SiSx adopts the PARALLEL ARCHITECTURE of Jackendoff (2002), according to which linguistic structure is determined by (at least) three independent formal systems: phonology, syntax and semantics. In addition, SiSx borrows from LFG the idea of a separate syntactic layer for representing grammatical functions: the GF-tier (CULICOVER; JACKENDOFF, 2005). Each one of these systems is defined by its own characteristic primitives and relations and is connected to the others by means of interfaces of varying complexity:
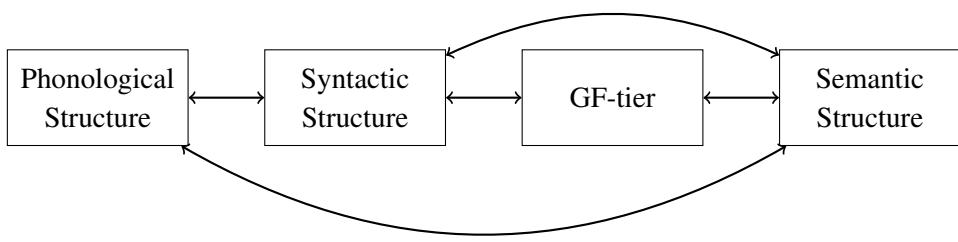

Figure 2 – The Parallel Architecture of SiSx

A well-formed expression must be well-formed in each level, in addition to having well-formed links among the interfaces. A toy example is shown in (3), where natural numbers indicate interface links between the components:[4]

(3)
$$
\begin{bmatrix}
\text{PHON} & /\text{m}\varepsilon\text{ri}_1 \; \text{kis}_2\text{d}_3 \; \text{dʒɒn}_4/_5 \\
\text{SYN} & [_S \text{NP}_1 \; [_{VP} \text{V}_2 \text{ - past}_3 \; \text{NP}_4]]_5 \\
\text{GF} & \text{PRED}_{2,5}: \langle \text{GF}_1: [\textit{3rd,sing,fem}], \text{GF}_4: [\textit{3rd,sing,masc}] \rangle \\
\text{SEM} & \textbf{kiss}'_2(\text{AGENT:mary}_1, \text{PATIENT:john}_4, \text{TIME:past}_3)_5
\end{bmatrix}
$$

The structure in (3) represents the sentence *Mary kissed John*. The most opaque aspect of the formalism is likely the GF-tier, which plays an analogous role to the level of ARG-ST in frameworks like HPSG (MANNING; SAG, 1998; DAVIS; KOENIG; WECHSLER, 2021). The basic units of this level are SYNTACTIC PREDICATES (PREDs) and GRAMMATICAL FUNCTIONS (GFs), which stand for syntactic arguments. Each PRED is an ordered list of GFs. Each GF in this list is represented by the set of its grammatically relevant features (i.e. specifications for NUMBER, PERSON and GENDER). The positions in the list of syntactic arguments are ordered with respect to each other, but they are not explicitly labeled with grammatical function names like SUBJECT or OBJECT. For reasons that will become clear in Section 4.4, these notions are relationally defined as *first GF*, *second GF*, etc., in accordance with the RELATIONAL HIERARCHY in (4) (PERLMUTTER; POSTAL 1977, 1983; KEENAN; COMRIE, 1977; POLLARD; SAG, 1994; BRESNAN, 1995):

(4)   **SUBJECT**   > **OBJECT**   > **OBLIQUE**
       PRED:$\langle$**GF**, ... $\rangle$ PRED:$\langle$GF, **GF**, ... $\rangle$ PRED:$\langle$GF, GF, **GF**$\rangle$

Note, furthermore, that there is nothing in SYN that signals that NP_1 in (3) is pronounced as the string /mɛri/ – this information is phonological, and, as such, it is only represented in PHON. The terminal strings in a tree like (2b) are, thus, not strictly speaking part of the syntactic structure. A similar division between phonological, syntactic and semantic forms is anticipated in Distributed Morphology (HALLE; MARANTZ, 1994; MARANTZ, 1997) as well as in variants of Categorial Grammar that build on Curry's (1963) PHENOGRAMMAR *vs.*

---

[4]   An interface link is well-formed *iff* it instantiates some constraint in the grammar: e.g. the links indicated by subscript 1 in (3) conform to what is stipulated by the lexical entry of *Mary*. The way SiSx represents different kinds of constraints is discussed in Section 4.4. Throughout this work, I will use the attribute-value matrix (AVM) notation adopted in Culicover (2021) for representing both linguistic objects as well as the constraints such objects must satisfy in order to be fully well-formed.

TECTOGRAMMAR distinction (OEHRLE, 1994; KUBOTA; LEVINE, 2020).

Like HPSG and LFG, and SiSx is SURFACE-ORIENTED. A model of grammar is SURFACE-ORIENTED if it posits syntactic structures that are directly associated with observable word strings, with a minimum of empty elements and degrees of embedding. In SiSx, this WHAT-YOU-SEE-IS-WHAT-YOU-GET flavor is suggested by the SSH – which enforces representational economy on phrase-structure configurations – along with the Parallel Architecture – which provides *other* levels for encoding GFs and semantic relations beyond narrow syntax.

Surface-orientation is driven by matters of psychological plausibility. Empty elements are not easily detectable from linguistic input. This raises the question of how they come to be learned (as discussed above in connection to the SSH) and inferred in real-time language processing (SAG; WASOW, 2011). The common conclusion is that they are *not learned*, but constitute part of UG. Though this move does solve the learnability problem (albeit by raising the more difficult question of how these elements evolved in humans), it hardly addresses the concern over language processing. So the *desideratum* of minimizing invisible structure winds up being an important one, at least for linguists committed to MENTALISM.

However, learnability and processing issues do not arise if empty elements *can* be inferred on the basis of language-internal evidence. This is arguably the case in situations where invisible structure systematically alternates with visible material, such as gaps in unbounded dependency structures like (5) (KLUENDER; KUTAS, 1993; CLARK; LAPPIN, 2011).[5]

(5)   Who$_1$ did Joanne say that Bobby should marry t_1?

In these cases SiSx, like other monostratal frameworks, *does* posit phonologically empty elements (i.e. traces) as a kind of "last resort" to maintain the generality of the mapping between form and meaning (CULICOVER; JACKENDOFF, 2005, pg. 304).

However, the status of empty elements in SiSx is very different from their status in MGG: they are not leftovers of transformations, but directly licensed by CONSTRAINTS. This distinction reflects the contrast between the PROOF-THEORETIC design of MGG and the MODEL-THEORETIC flavor of SiSx (POLLARD, 1996; PULLUM; SCHOLZ, 2001; PULLUM, 2013; MÜLLER, 2020). This is the topic of the next section.

---

[5]   For most of these scenarios, it can also be shown that grammars with empty elements are extensionally equivalent to grammars without them. This effectively reduces empty elements to notational devices for stating generalizations more directly and reducing the overall complexity of the grammar (MÜLLER, 2020, chap. 19). If one assumes a simplicity-based evaluation metric like the one in Chomsky (1951), this notational choice actually has empirical consequences for language acquisition (see Chomsky (1965, pg. 45) for a similar point).

4.2 PROOF-THEORETIC AND MODEL-THEORETIC FORMALISMS

A PROOF-THEORETIC FORMALISM (PTF) for grammatical theories relies on the technology of stepwise algorithmic derivations to constructively assign structures to the expressions of a language. As the name suggests, PTFs borrow their tools from the mathematical study of logical proofs. The most notable influence on early PTF work in natural languages (e.g. CHOMSKY, 1959) was Post (1943), who defined rule systems for deriving logical inferences in terms of purely syntactic string manipulation operations called PRODUCTIONS. As an example, consider the production rule expressing MODUS PONENS (where $\vdash$ is the assertion sign and W_n is a variable over strings) (PULLUM, 2013, pg. 493):

(6)   Given a string that matches $\vdash (W_1) \rightarrow (W_2)$ and a string that matches $\vdash (W_1)$, add a new string that matches $\vdash (W_2)$.
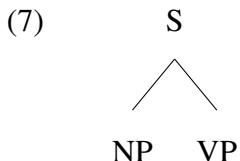
Systems of this sort are general enough to characterize any recursively enumerable set. PTFs assume that all natural languages are recursively enumerable sets of sentences. A sentence, in this framework, is simply an uninterpreted string of symbols. A particular PTF-based grammar of a language $\mathscr{L}$ thus consists of a pair $\langle \Sigma, F \rangle$, where Σ is a finite set of primitive strings and F is a finite set of operations for constructing new strings from previously given ones. The grammar must be be such that an infinite random application of F to Σ yields precisely the sentences that belong to $\mathscr{L}$. The finite sequence of strings obtained in the course of constructing any sentence $\Psi \in \mathscr{L}$ is a DERIVATION of Ψ.

In early work on MGG, the only primitive element assumed was the initial symbol S, which was successively decomposed and manipulated by rewrite rules and different kinds of transformations (CHOMSKY, 1957). Contemporary work views lexical items as primitives and Merge as the sole recursive operation (CHOMSKY; GALLEGO; OTT, 2019).

Despite their popularity, PTFs are not the only way of stating grammars for natural languages. There are also MODEL-THEORETIC FORMALISMS (MTFs), which draw from the semantic side of mathematical logic (TARSKI, 1944,1956). Instead of resorting to stepwise algorithmic derivations to generate sets of strings, MTF-based grammars formulate their basic statements as declarative CONSTRAINTS on the structure of individual expressions. These constraints can be either true or false of any given linguistic structure. The objects that satisfy the constraints (their MODELS, in the logician's sense) are the expressions licensed by the grammar.

The hallmark of MTFs is a sharp distinction between the mathematical structures that constitute linguistic expressions and the propositions of a grammar which are modeled by these expressions (i.e. the CONSTRAINTS). There is no sense in which the constraints of the grammar are responsible for *constructing* the objects they describe. A constraint is merely a partial condition (given in plain prose or in a formal language) on what counts as a well-formed structure.

To take a concrete example, consider the structure in (7). A PTF-based grammar would say that say that (7) is well-formed because it is the yield of the procedural rule in (8) – or, to be more precise, because (7) is a member of the set characterized by a nondeterministic random generation procedure which includes (8). A MTF-based grammar, on the other hand, would say that (7) is a licit linguistic object because it *is a model of* the constraint in (9):

(7)   S
           ╱   ╲
       NP   VP

(8)   S → NP VP

(9)   A node labeled S dominates a node labeled NP followed by a node labeled VP.

The differences between PTFs and MTFs may seem subtle, but they have important implications. The manner of characterizing expressions in PTFs naturally invites the dynamic and procedural metaphors that are routinely employed in the MGG literature. One often reads that certain operations must apply "before" others, that structures are "sent" to some component, or that grammatical roles are defined by the "timing of insertion" of constituents (see Boeckx (2008, pg. 103) for an example). The problem with such locutions is that it is unclear what they are supposed to mean in terms of real-time processing (JACKENDOFF, 2011b).

The practical consequence of this uncertainty has been an increasing difficulty in integrating theories of linguistic competence with empirically sound models of performance, starting from the early 1970s (FODOR; BEVER; GARRETT, 1974). This, in turn, led some linguists to stiffen the notion of competence into an idealization which "denies the need to go beyond itself", thereby isolating linguistics from psychology at large (JACKENDOFF, 2002, pg. 33).

MTFs avoid all such problems, lending themselves to a much more direct relation to processing models (SAG; WASOW, 2011; JACKENDOFF, 2007b; JACKENDOFF, 2011b). Since constraints have no inherent directionality, they can be invoked in any order. They allow the processor to start with fragment of phonology and pass through its mappings to syntax and

semantics as well as go the other way around. The processor can also use constraints to proceed bottom-up or top-down, depending on the context. This accounts for the fact that the processor is "opportunistic" and deploys diverse types of information as soon as they become available (ACUÑA-FARIÑA, 2016). It also makes MTFs neutral with respect to production (which goes from semantics to phonology) and comprehension (which goes from phonology to semantics) – as a proper theory of competence should indeed be.

Moreover, in most MTFs (including the one I assume for SiSx), constraints yield a monotonic mapping from form to meaning – i.e. there are no destructive operations that throw out information inferable from parts of a structure. Information about an expression's phonological, syntactic and semantic representations "is locally distributed across the expression structure in such a way that can be inferred from the parts of the expression" (BRESNAN, 2001, pg. 46). This makes MTFs suitable to deal with the grammaticality of linguistic fragments and with the incremental nature of parsing – another desirable property in light of psychological adequacy.

Pullum and Scholz (2001) and Postal (2004) point out many other advantages of MTFs that are independent of the commitment to psychological adequacy. For example, unlike PTFs, MTFs have a natural way of expressing the phenomena of gradient ungrammaticality:

(10)   a.   *Him is the chair of this department.

         b. **Him is chair the of this departments.

While both of the sentences in (10) are deviant, most speakers have a clear intuition that (10a) is *less* deviant than (10b). This difficult to capture within a pure PTF setting because a PTF-based grammar either generates a linguistic object (in which case it is perfectly grammatical) or does not generate it (in which case it completely ungrammatical). There is no such thing as *partial ungrammaticality* or a *quasi-well-formed sentence*.[6]

MTF-based grammars have no difficulty in this respect. Since linguistic expressions are kept separate from the constraints that they instantiate, it is perfectly coherent to claim that some structures may have more non-licensed "parts" than other structures and that the number of non-

---

[6]   The common answer to this conundrum within MGG is to consign gradience in judgments entirely to matters of performance (SCHÜTZE, 1996). Rather than offering any genuine insight, I take this solution to be simply the expression of a limitation which is forced by the PTF-based design of MGG. A notable exception to this trend is Chomsky (1961), who actually attempts to make gradient ungrammaticality compatible with PTFs. His account involves a function which maps members of the complement of the set generated by a PTF-based grammar (i.e. $\mathscr{L}^C$, for any language $\mathscr{L}$) to numbers representing their degree of ungrammaticality. This number is attained by comparing the sequence of strings in $\mathscr{L}^C$ to similar sequences of strings which are elements of $\mathscr{L}$. See Pullum and Scholz (2001) for a detailed critique of Chomsky's proposal.

licensed parts determines their degree of (un)grammaticality. We can characterize the contrast in (10) by saying that (10a) has only one "part" that fails to instantiate the grammar (namely, the accusative subject), while (10b) has three ill-formed "parts": the accusative subject, the order between *chair* and *the* and the number mismatch between *this* and *departments*.

This issue is related to another major difference between PTF and MTF-based grammars: the fact that the latter do not presuppose that well-formed linguistic expressions form a closed set with a fixed cardinality (say, $\aleph^0$), nor that they must necessarily be finite in length (LAN-GENDOEN; POSTAL, 1984; PULLUM; SCHOLZ, 2001; PULLUM, 2013). A set of syntactic constraints concerning what counts as a licit linguistic structure does not entail anything about how many objects instantiating that structure exist or what size they can have. This is in fact a desirable property of a formalism for stating grammars, as Pullum (2020, pg. 18) notes:

> No syntactician should have to pretend there are serious answers to questions about how many sentences exist, or how long they can get. On the assumption that there are infinitely many sentences, all of the extremely long ones are surely unacceptable to every speaker (one can hardly call a sentence acceptable if it is not even humanly utterable, e.g. if uttering it or listening to it would take longer than any human can remain awake). But the extremely long sentences are the overwhelming majority. In other words, most of any infinite set of grammatical sentences will be totally unacceptable. It would be preferable if we did not have to make (or deny) that assumption about the language. An MTS grammar provides exactly that kind of neutrality, because MTS constraints make statements about what expressions are like, structurally, with no entailments about how many there are, or how big they are allowed to be.

SiSx can be entirely stated as an MTF.[7] The core assumption in MTFs is that linguistic expressions actually *have structure* and that the job of the linguist is to *describe* these structures by making general statements about their properties and relations in the form of constraints. Therefore, if SiSx is to be characterized as a full-blown MTF, it is of utmost importance to be very clear about what the theory takes to be the structure of linguistic expressions, how these

---

[7] There is, however, one slight deviation. In describing the SiSx view, Jackendoff (2011b) claims that, in addition to a set of grammatical constraints, speaker's linguistic competence must also include UNIFICATION. Unification is a procedure for solving a set of constraints by combining the information they encode. This allows us to discover the structures that a set of constraints describes (SHIEBER, 1986). Jackendoff's claim preserves, thus, the PTF notion that grammars include some sort of constructive algorithm that defines a set of expressions.

However, if one views the set of constraints as equivalent to the grammar, unification can be eliminated along with the residues of PTFs. Instead of defining well-formedness in terms of membership in the set of solutions unification provides to all of the constraints in the grammar, we can define it in purely model-theoretic terms: a structure is well-formed *iff* all of its parts satisfy some constraint in the grammar (a full definition is given in Section 4.3.4). To be sure, constraints *can* be coupled with a solution algorithm to construct members of a particular recursively enumerable set. This may be useful for the purposes of grammar engineering or for proving theorems about the equivalence of certain MTF and PTF-based grammars. But it is by no means a necessity. The way speakers actually *construct* their utterances should be the object of empirically grounded models of performance (POLLARD, 1996; POSTAL, 2004).

objects can be described and what it means for some expression to satisfy a set of constraints. These are the issues I turn to next.

4.3   STRUCTURES, DESCRIPTIONS AND WELL-FORMEDNESS

As we saw in Figure 2, SiSx posits four levels of representation: PHONOLOGICAL STRUC-TURE, SYNTACTIC STRUCTURE, the GF-TIER and SEMANTIC STRUCTURE. Each of these levels is designed to model different aspects of linguistic expressions. The structures SiSx ascribes to linguistic expressions thus abide by what Jackendoff (1997, pg. 41) calls REPRESENTA-TIONAL MODULARITY: "The overall idea is that the mind/brain encodes information in some finite number of distinct representational formats or 'languages of the mind". What individuates a module, in this approach, is not its functional role as an input system or its *informational encapsulation*, as in Fodor (1983), but the formal properties of the representations it deals with.

A fundamental assumption of this kind of framework is that "there are different kinds of informational dependencies among the parts of a sentence, and that these are best expressed using different formal structures" (KAPLAN, 1995, pg. 10). This avoids having to overload a single representation format (e.g. hierarchical phrase-markers) with the burden of modeling types of linguistic information as diverse as linear order, syntactic constituency, grammatical functions and inference. This is precisely the same intuition that led to the projection architecture of LFG (BRESNAN; KAPLAN, 1982a; BRESNAN, 2001; DALRYMPLE, 2001), which is one of the major inspirations for the formalization of SiSx I present here.

The representations pertaining to each level of linguistic organization can be mathematically defined in terms of RELATIONAL STRUCTURES: i.e. as finite sets of primitives and relations defined over these primitives. Structures in each level are, in turn, in systematic correspondence to structures in other levels. Given well-defined structure types and correspondences between them, the role of the grammatical analysis is to ensure that only appropriate structures and correspondences are assigned to linguistic expressions. This is accomplished by the grammar, which I take to be a set of constraints. Constraints are simply descriptions of structures and their correspondences (often *partial* descriptions) in a suitable formal language. In SiSx these are implemented in the most obvious way: by listing the defining relations pertaining to each level and the appropriate correspondences between the primitive units of each level.

In what follows, I provide definitions for the structures in phonology, syntax, the GF-tier and semantics as well as characterization of the correspondences between these structures. I

also discuss what it means for a structure in the domain of linguistic objects to be well-formed in relation to the set of constraints in the grammar.

4.3.1 Phonological, Syntactic and Semantic Structures

In this section, I define PHONOLOGICAL STRUCTURES (PHON), SYNTACTIC STRUCTURES (SYN) and SEMANTIC STRUCTURES (SEM). I also give examples of constraints over these structures. Discussion of the GF-tier is reserved for a separate subsection, as this is the least familiar aspect of the formalism and also the one that most deviates from the implementation assumed in previous SiSx work (CULICOVER; JACKENDOFF, 2005).

The primitives units in PHON are strings of sounds (i.e. segments or lists thereof). To enhance readability, I henceforth represent strings in plain orthography and variables over strings as φ_n. I will not have much to say about PHON here – a more thorough discussion would require probing into the internal structure of segments and suprasegmental properties such as tone, syllable and rhythm (GOLDSMITH, 1976; NESPOR; VOGEL, 1986; BOSCH, 2011). Setting such issues aside, I assume that the sole relations native to PHON are INCLUSION ($\subseteq$) and PRECEDENCE ($\prec$), defined as follows (CULICOVER, 2021):

(11) INCLUSION ($\subseteq$)

   For any two strings φ_1 and φ_2, $\varphi_1 \subseteq \varphi_2$ *iff* pronouncing φ_2 involves pronouncing φ_1.

(12) PRECEDENCE ($\prec$)

   For any three strings φ_1, φ_2 and φ_3, $\varphi_3 = \varphi_1 \prec \varphi_2$ *iff*

   (i) $\varphi_1 \subseteq \varphi_3$

   (ii) $\varphi_2 \subseteq \varphi_3$

   (iii) When φ_3 is pronounced, φ_1 is pronounced and φ_2 is pronounced later.

   (iv) $\neg\exists\varphi_4[\varphi_4 \subseteq \varphi_1 \wedge \varphi_4 \subseteq \varphi_2]$ (i.e. φ_1 and φ_2 do not overlap).

Precedence defines a strict partial order on the set of strings: it is a *transitive* (if $\varphi_1 \prec \varphi_2$ and $\varphi_2 \prec \varphi_3$, then $\varphi_1 \prec \varphi_3$), *irreflexive* (for any φ_n, $\varphi_n \nprec \varphi_n$) and *asymmetric* (if $\varphi_1 \prec \varphi_2$, $\varphi_2 \nprec \varphi_1$) relation. I typically represent inclusion by placing the substrings of a given string in between two forward slashes. Assuming this convention, $\varphi_1 \subseteq \varphi_3$ is equivalent to $/\varphi_1/_3$.

The phonological structure of the expression *Bobby, come on over for dinner* is, thus, the string /Bobby$\prec$come$\prec$on$\prec$over$\prec$for$\prec$dinner/. We can use this simple precedence relation to define the following constraint on the phonology of the idiom *by and large*. This allows us to

capture the fact that the sequence *large and by* is not licensed as the PHON of this idiom:

(13) $\left[\text{PHON} \quad \text{/by} \prec \text{and} \prec \text{large/} \right]$

In particular examples of PHON objects and their descriptions, I often omit the "$\prec$" symbol and represent precedence simply by left-right organization of characters on the page.

The primitive units of SYN, in turn, are NODES and LABELS. There are two fundamental relations in this level: the MOTHER function (*M*), which maps nodes onto nodes, and the LABEL function (*L*), which maps nodes onto labels. In more explicit terms, each representation in SYN is an ROOTED UNORDERED TREE, which consists in a quadruple $\langle N, Q, M, L \rangle$, where:[8]

(14)   *N*: set of nodes (n_1, n_2, n_3, ..., n_n)

Q: set of category labels (e.g. V, NP, Aux, S, PP)

M: a partial function from *N* into *N* (the MOTHER function)[9]

*L*: a function from *N* into *Q* (the LABEL function)

In order to understand why syntactic trees are ROOTED, we need to define the relations DOMINANCE (*D*) and IMMEDIATE DOMINANCE (*ID*). We say that a node *x* immediately dominates a node *y* *iff* *x* is the value of *y* under *M*. More formally, for any $x, y \in N$, $\langle x, y \rangle \in ID$ *iff* $x \in M^{-1}[\{y\}]$. The dominance relation *D* is the transitive closure of immediate dominance: $\langle x, y \rangle \in D$ *iff* there is a sequence of nodes $x_0, x_1, \ldots x_n$ such that $x_0 = x$ and $x_n = y$ and for every x_i, $0 \leq i < n$, the pair $\langle x_i, x_{i+1} \rangle \in ID$. The rootedness of SYN objects simply means that in every well-formed tree there is exactly one node that dominates every node. This is expressed by the following axiom (PARTEE; MEULEN; WALL, 1990, pg. 443):
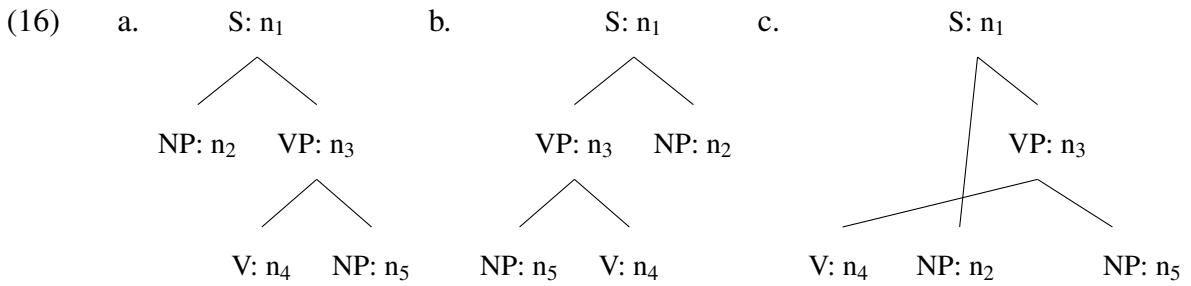
(15)   THE SINGLE ROOT AXIOM

$\exists x \in N[\forall y \in N[\langle x, y \rangle \in D]]$

The trees in SYN are UNORDERED because the linear left-right arrangement of nodes does not make a difference to the syntactic representation. The only relations that matter to SYN are the MOTHER and LABEL. This entails that the diagrams in (16) all represent exactly the same syntactic object (category label symbols are written to the left of every node):

---

[8] The graph-theoretic definition of trees as elementary syntactic objects is largely borrowed from Kaplan (1995). See McCawley (1968, pg. 245), and Partee, Meulen and Wall (1990, pg. 443) for similar characterizations.

[9] The fact that *M* is defined as a *function* excludes the possibility multi-dominance (i.e. structures where a node has more than one mother) from SYN representations. *M* has to be a *partial* function because there is always one node in each tree which does not have a mother of its own: the root node at the top of the tree.

(16)   a.   S: n_1 b.   S: n_1 c.   S: n_1

NP: n_2 VP: n_3 VP: n_3 NP: n_2 VP: n_3

V: n_4 NP: n_5 NP: n_5 V: n_4 V: n_4 NP: n_2 NP: n_5

The fact that order is represented only in PHON *via* the PRECEDENCE relation does not mean that syntax is *unrelated* to linear order. In languages like English, there is a fairly rigid correspondence between order and syntactic structure: e.g. strings that correspond to syntactic heads tend to precede strings that correspond to syntactic complements. However, instead of encoding these facts directly in SYN (e.g. by defining a PRECEDENCE as a relation between nodes), we can use the Parallel Architecture of SiSx (see Figure 2 above) to state them as CORRESPONDENCE CONSTRAINTS between SYN and PHON (see 4.3.3 below).

We describe a structure in SYN simply by listing the defining relations *M* and *L* that hold among its primitive elements (nodes and labels). The tree in multiply diagrammed in (16) is the smallest structure that *satisfies* (or *is a model of*) the description in (17):[10]

(17) $M(n_2) = n_1 \land M(n_3) = n_1 \land M(n_4) = n_3 \land M(n_5) = n_3 \land L(n_1) = S \land L(n_2) = NP \land$
         $L(n_3) = VP \land L(n_4) = V \land L(n_5) = NP$

If (17) were a constraint in a particular grammar, we could say that (17) licenses the structure in (16). In addition to (names of) primitive elements and relations that define each level of representation, some constraints also employ variables and predicates which do not correspond directly to structures in the domain of described objects. A particularly clear example of this can be given for SYN objects, though these devices will be used for describing PHON, SEM and GF-tier structures as well. The following constraint is the analogue to the traditional phrase-structure rule for licensing an English PP (i.e. PP → P NP) in the present framework:

(18) $M(x_2) = x_1 \land M(x_3) = x_1 \land L(x_1) = PP \land L(x_2) = P \land L(x_3) = NP$

Note that (18) does not use names of nodes (n_1, n_2, … n_n) but variables over nodes (x_1, x_2, … x_n). For perspicuity, I adopt the usual labeled bracketing notation for representing both syntactic structures and the constraints they must satisfy. So (18) is equivalent to (19) (the fact

---

[10]   (17) is also true of larger structures that include (16) as a subtree. Kaplan (1995, pg. 14) notes that "structures that satisfy descriptions form a semi-lattice that is partially ordered by the amount of information they contain".

that P and NP stand for labels assigned to node *variables* is signaled by italics):

(19) $\left[\text{SYN} \quad [_{PP} \; P \; NP] \right]$

Some constraints also employ variables for labels. As an example, consider the abstract schema requiring every constituent of type XP to dominate a constituent of type X – i.e. the SiSx analogue to HPSG's HEAD FEATURE PRINCIPLE (POLLARD; SAG, 1994, pg. 34):

(20) $\left[\text{SYN} \quad [_{XP} \ldots X \ldots] \right]$

Since the *X* symbol is a variable over categories, it can be instantiated by any label (e.g. V, N, P, A, Adv). This is also indicated by placing *X* in italics. The "..." symbols around X are simply a finite specification of an unbounded disjunction of descriptions of labeled nodes.

Next, consider the representations in SEM. I assume that the basic units in SEM are the MEANINGFUL EXPRESSIONS (MEs) of Montague's Intensional Logic (MONTAGUE, 1974), which receive a model-theoretic interpretation along familiar lines. As is customary, each ME in SEM is assigned to a semantic type which determines the kind of denotation it has. The notion of type is defined as follows (where *e* is short for *entity*, and *t* is short for *truth-value*):

(21) a. *e* and *t* are types.

  b. If *a* and *b* are types, then $\langle a, b \rangle$ is a type.

Types defined by (21a) denote primitive objects: entities and truth-values. Types defined by the recursive clause in (21b) are called FUNCTIONAL TYPES because they are interpreted as functions from things of type *a* to things of type *b*; for example, a ME of type $\langle e, t \rangle$ corresponds to a function from entities to truth-values. The notion of a Meaningful Expression of type *a* (ME_a) is defined as follows:

(22) a. Every variable and constant of type *a* is in ME_a.

  b. If $\alpha \in \text{ME}_a$ and *u* is a variable of type *b*, then $\lambda u[\alpha] \in \text{ME}_{\langle b, a \rangle}$.

  c. If $\alpha \in \text{ME}_{\langle a, b \rangle}$ and $\beta \in \text{ME}_a$, then $\alpha(\beta) \in \text{ME}_b$.

  d. If $\alpha, \beta \in \text{ME}_\alpha$, then $\alpha = \beta \in \text{ME}_t$.

  e. If $\phi, \psi \in \text{ME}_t$, then $\neg\phi, [\phi \wedge \psi], [\phi \vee \psi], [\phi \rightarrow \psi], [\phi \leftrightarrow \psi] \in \text{ME}_t$.

  f. If *u* is a variable and $\phi \in \text{ME}_t$, then $\forall u[\phi], \exists u[\phi] \in \text{ME}_t$.

  g. If *u* is a variable of type *a* and $\phi \in \text{ME}_t$, then $\iota u[\phi] \in \text{ME}_a$.

For the purposes of representing thematic roles, I assume a neo-Davidsonian overlay to the standard Montagovian system outlined above. Thematic predicates (AGENT, PATIENT, EX-PERIENCER, etc.) denote relations between individuals and the events they partake in, as in Parsons (1990). However, instead of representing quantification over events directly, I use the abbreviated notation adopted by Culicover (2021), where labels for thematic roles are indexed to argument positions of predicates which introduce an implicit quantification over events. In this set up, the SEM representation for *Joanne broke the glass yesterday* is (23a), which is equivalent to the standard representation in (23b):

(23) a. $\mathbf{break}'(\text{AGENT} : \text{joanne}, \text{PATIENT} : \iota x[\mathbf{glass}'(x)], \text{TIME:yest})$

b. $\exists e[\mathbf{break}'(e) \wedge \text{AGENT}(e, \text{joanne}) \wedge \text{PATIENT}(e, \iota x[\mathbf{glass}'(x)]) \wedge \text{TIME}(e, \text{yest})]$
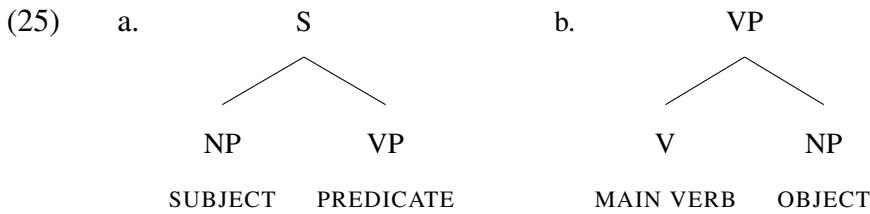
An example of a SEM constraint can be found on the lexical entry for (the transitive variant of) *eat*, which is given in (24). Numerical indices indicate correspondences between the different structures (the precise meaning of coindexing links is clarified in Section 4.3.3):

(24)
$$
\begin{bmatrix}
\text{PHON} & \text{eat}_1 \\
\text{SYN} & \text{V}_1 \\
\text{SEM} & \lambda x[\lambda y[\mathbf{eat}'_1(\text{AGENT} : y, \text{PATIENT} : x)]]
\end{bmatrix}
$$

The parts of the SiSx architecture presented so far (PHON, SYN and SEM) are fairly standard in that there are close analogues to each of them in virtually all grammatical frameworks. The only component of Figure 2 left to discuss is the GF-tier. Since this level of representation is somewhat more contentious, the next subsection attempts not only to define its structures, but also to motivate its existence.

4.3.2 The GF-tier

In any theory, grammatical functions (GFs) such as SUBJECT, OBJECT and OBLIQUE serve as abstract relators between a class of surface syntactic properties (e.g. linear order, agreement and case marking) and semantic roles. MGG assumes that these abstract GFs can be represented in the same format as syntactic groupings. This implies that GFs are epiphenomena of constituent structure configurations. An early statement of the MGG view is found in Chomsky (1965, pg. 68-74), who claims that notions like SUBJECT and OBJECT are definable in terms of the structural positions in (25) (see also Chomsky (1955, 254-255)):

(25) a.

$$S$$
NP   VP
SUBJECT   PREDICATE

b.

$$VP$$
V   NP
MAIN VERB   OBJECT

On this approach, syntactic theory does not need to make explicit reference to GFs *per se* because whatever rules and operations involve notions like SUBJECT and OBJECT can be stated in terms of the configurations in (25).

SiSx, like Relational Grammar and LFG before it, rejects this view, which Bresnan (2001) dubs the CONFIGURATIONAL DESIGN OF UG. Consider what it implies for the sentence in (26):

(26)   Brad seems to like Janet.

In (26), *Brad* behaves like the SUBJECT of two predicates: the one headed by *seem* (where it establishes agreement) and the one headed by *like* (where it gets interpreted semantically). The configurational design requires that each of these GFs be realized in different positions, which *Brad* has to occupy simultaneously. This, however, is technically impossible in a typical phrase-structure system, since it entails multi-dominance. The alternative is to posit a sequence of phrase-markers in which these positions are occupied at separate stages, as in (27):

(27)   seems [$_S$ Brad [$_{VP}$ to like Janet]] $\Rightarrow$ [$_S$ Brad$_i$ [$_{VP}$ seems [$_S$ t_i to like Janet]]]

The configurational design thus calls for operations that map phrase-markers onto phrase-markers – i.e. syntactic transformations. Note, however, that these mappings are simply a way to encode the effects of multi-dominance in a system that does not naturally allow for it.

Though this might seem plausible for English (where SUBJECTS typically correspond to the configuration in (25)), it is less appealing for languages like Russian, where word order is freer and GFs are signaled mainly by case endings on nouns. A derivation for the Russian OVS sentence (28) would have to look like (29) (KALLESTINOVA, 2007, pg. 30):

(28)   Vaz-u   razbila Olj-a
        vase-ACC broke   Olya-NOM
        'Olya broke the vase'

(29)   [$_S$ Olja [$_{VP}$ razbila vazu]] $\Rightarrow$ [$_{S'}$[$_{VP}$ razbila vazu]$_i$ [$_S$ Olja t_i]]

        $\Rightarrow$ [$_{S''}$ vazu$_k$ [$_{S'}$ [$_{VP}$ razbila t_k]$_i$ [$_S$ Olja t_i]]]

The SUBJECT and OBJECT in (29) are base-generated in the positions signaled in (25) and then scrambled to where they are actually pronounced via roll-up movements (BAILYN, 2003). The resulting structure represents "several types of information that seem quite dissimilar in nature" (KAPLAN; ZAENEN, 1995, pg. 137): GFs like SUBJECT and OBJECT, linear order, dominance relations and syntactic categories. The option for these hybrid representations thus leads to "the attenuation of the classical properties of constituency [...] and to the concomitant adoption of increasingly abstract surface structures" (BRESNAN, 2001, pg. 46).

SiSx rejects this on the grounds of REPRESENTATIONAL MODULARITY. Dominance and syntactic categories are naturally represented in a phrase-structure system. Order is delegated to PHON. The organization of GFs, however, has different formal properties (e.g. multi-dominance) that justify positing a separate component. This is the GF-tier in SiSx.
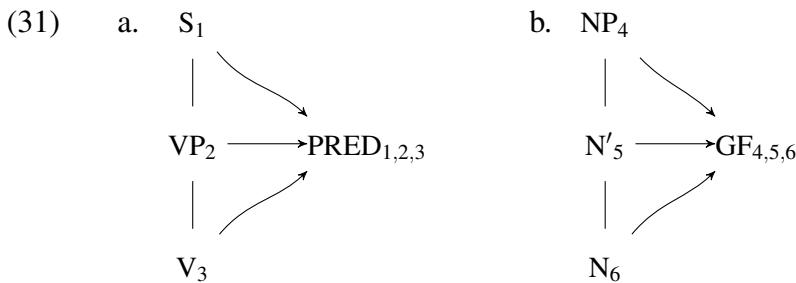
There are many ways GFs can be formalized as part of an autonomous representation. I follow the guidelines of Kaplan's (1995) formalization of f-structure in LFG, treating GF-tier objects (GFOs) as functions in the mathematical sense. There are two kinds of GFOs: (i) PREDs, which represent syntactic predicates and (ii) GFs, which represent syntactic arguments and their ϕ-features. GFs map from elements in a set of primitive attribute symbols (A1) to primitive value symbols (V) while PREDs maps from the set of list attributes (A2) to the set of GFs. The general notion of a GFO is the union of the set of GFs and the set of PREDs.

(30)   a.   A1= {PERS, NUM, GEND}

       b.   A2= {FIRST, REST}

       c.   V= {*1st*, *2nd*, *3rd*, *sing*, *plur fem*, *masc*, *neut*}

       d.   GF: A1 $\rightarrow$ V

       e.   PRED: A2 $\rightarrow$ GF

       f.   GFO = GF $\cup$ PRED
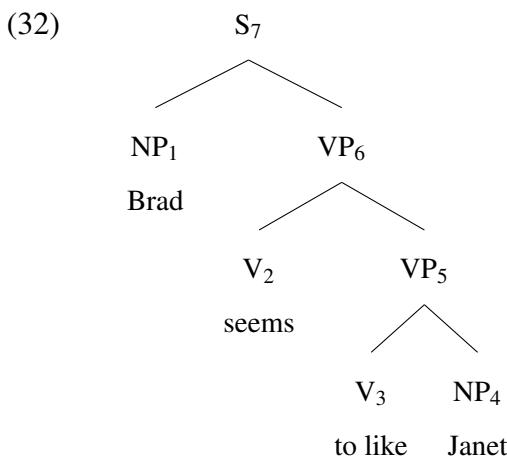
The fact that agreement is fundamentally dependent on GFs instead of syntactic configuration (DALRYMPLE, 2001) is modeled by including the ϕ-features PERS, NUM and GEND as primitive attributes of GFs. The attributes FIRST and REST are the standard way of encoding ordered lists in feature structure terms (SHIEBER, 1986, pg. 29). These allow us to dispense with labels for particular GFs. Notions like SUBJECT and OBJECT are, therefore, not primitives of SiSx: they are defined RELATIONALLY in terms of a hierarchy of arguments within a predicate, as in HPSG (MANNING; SAG, 1999). So, PRED(FIRST) represents the SUBJECT of

PRED, PRED(REST) the OBJECT or COMPLEMENT of PRED, etc.[11]

Each instance of PRED and GF is written with subscripted indices that serve to distinguish it from other GFOs of the same type and to mark its correspondences with other linguistic structures. As in LFG, each node representing a lexical head in SYN is typically linked to the same GFOs as all of its projections (in fact, these links are the way headedness is represented in SiSx). For example, a finite V is linked to the same PRED as the VP and S that dominate it and an N is linked to the same GF as its N$'$ and NP mothers, as schematized below:

(31) a. S_1 b. NP_4

$$VP_2 \longrightarrow PRED_{1,2,3} \qquad N'_5 \longrightarrow GF_{4,5,6}$$

$$V_3 \qquad\qquad\qquad N_6$$

Accordingly, the GF-tier structure SiSx associates with (32) (repeated from (26)) is (33). Each AVM in (33) lists the pairs that are members of the GFO named on its left.

(32)
$$S_7$$

NP_1 VP_6

Brad

V_2 VP_5

seems

V_3 NP_4

to like   Janet

(33)

$$\text{PRED}_{2,6,7}: \begin{bmatrix} \text{FIRST} & \text{GF}_1: \begin{bmatrix} \text{PERS} & \textit{3rd} \\ \text{NUM} & \textit{sing} \\ \text{GEND} & \textit{masc} \end{bmatrix} \end{bmatrix} \text{PRED}_{3,5}: \begin{bmatrix} \text{FIRST} & \text{GF}_1 \\ \\ \text{REST} & \text{GF}_4: \begin{bmatrix} \text{PERS} & \textit{3rd} \\ \text{NUM} & \textit{sing} \\ \text{GEND} & \textit{fem} \end{bmatrix} \end{bmatrix}$$

---

[11] This is the main formal difference between the GF-tier units in SiSx and f-structures in LFG. Patejuk and Przepiórkowski (2016), however, argue that LFG should incorporate a list of unlabeled arguments as well. Following Alsina (1996), they show that most GF labels redundantly represent information already available in morphosyntax and s(emantic)-structure. Borrowing ideas from HPSG, they propose to replace GF attributes by a single ordered DEPS list which looks like SiSx's GF-tier. This also allows a direct encoding of the functional hierarchy, which is used in LFG analyses of binding (FALK, 2001) and control (BRESNAN, 1982a).

For compactness, I write only the values for the PERS, NUM, GEND attributes and adopt the usual list abbreviation for FIRST/REST feature structures. So (33) is equivalent to (34):

(34) $\text{PRED}_{2,6,7}$:$\langle$GF$_1$:[*3rd,sing,masc*]$\rangle$ $\text{PRED}_{3,5}$:$\langle$GF$_1$, GF$_4$:[*3rd,sing,fem*]$\rangle$

A peculiarity of GF-tier representations is that they lack the unlimited embedding we find in phrase-structure. Each PRED in the GF-tier is represented as a self-contained unit. There is no sense in which the PRED that corresponds to *like* in (33)/(34) ($\text{PRED}_{3,5}$) is embedded under the one that corresponds to *seem* ($\text{PRED}_{2,6,7}$). Both are completely detached from one another.

For our purposes, the most important part of the GF-tier structures in (33)/(34) is the fact that the GF linked to *Brad* (GF$_1$) occupies the first position in two distinct PRED lists: the one linked to the projections of *seem* ($\text{PRED}_{2,6,7}$) and the one linked to the projections of *like* ($\text{PRED}_{3,5}$), thereby capturing the observation that *Brad* functions as the subject of two different clauses. This direct encoding of multi-dominance – which is also central to LFG's analysis of raising (BRESNAN, 1982a) – renders transformations like (27) unnecessary.

A similar analysis is possible for raising-to-object structures like *Bobby expects Amy to sing*. The only difference in comparison to (33)/(34) is that, in raising-to-object, the GF which is shared between the the two PREDs is the second GF in the PRED list that corresponds to the main verb – i.e. GF$_3$ in the full structure depicted in (35):

(35)
$$
\begin{bmatrix}
\text{PHON} & \text{Bobby}_1 \text{ expects}_2 \text{ April}_3 \text{ to} \prec \text{sing}_4 \\
\text{SYN} & [_S \text{ NP}_1 \text{ }[_{VP} \text{ V}_2 \text{ NP}_3 \text{ VP[INF]}_4]_5]_6 \\
\text{GF} & \text{PRED}_{2,5,6}\text{:}\langle\text{GF}_1\text{:[\textit{3rd,sing,masc}], GF}_3\text{:[\textit{3rd,sing,fem}]}\rangle \text{ PRED}_4\text{:}\langle\text{GF}_3\rangle \\
\text{SEM} & \textbf{expect}'_2(\text{EXPERIENCER:bobby}_1, \text{THEME:}\textbf{sing}'(\text{AGENT:april}_3)_4)_6
\end{bmatrix}
$$

The autonomy of GFs in SiSx also makes it possible to state mappings between GFs and SYN without specifying syntactic configuration or linear order. So, for dependent-marking languages like Russian, GFs can be linked directly to Ns with the appropriate case morphology. This is captured by the correspondence constraint in (36) (CULICOVER, 2009, pg. 154).[12]

---

[12] The AVMs in (36), *qua* constraints, are able to specify *partial* information about the linguistic objects they describe. The constraints in (36) only specify the *position* of structurally cased-marked Ns within a PRED, but not the particular featural content of the GFs which correspond to these Ns. The *PRED* and *GF* symbols in (36) are, thus, VARIABLES over PREDs and GFs in the model. This is indicated by placing *PRED* and *GF* in italics. So the constraints in (36) are compatible with GFs bearing any possible combination of person, number and gender specifications. The idea that word parts such as case endings can carry information about GFs bypassing syntax is shared with LFG (BRESNAN, 2001). The proposal sketched in (36)-(37) bears a particularly close resemblance to the CONSTRUCTIVE CASE theory proposed by Nordlinger (1998).

(36) a. $\begin{bmatrix} \text{SYN} & [_S \ldots N[\text{NOM}]_1 \ldots]_2 \\ \text{GF} & PRED_2:\langle GF_1, \ldots \rangle \end{bmatrix}$ b. $\begin{bmatrix} \text{SYN} & [_S \ldots N[\text{ACC}]_2 \ldots]_3 \\ \text{GF} & PRED_3:\langle GF_1, GF_2 \ldots \rangle \end{bmatrix}$

This proposal avoids abstract *ad hoc* MGG derivations like (29), opening the possibility of licensing flat structures. A SiSx analysis for (28) in this spirit could be something like (37). Note that configuration does not play a role in determining GFs in this case. (This does not mean that it cannot play a role in defining information structure properties, which are not being represented in (37).)

(37) $\begin{bmatrix} \text{PHON} & \text{Vaz-u}_3 \text{ razbila}_2 \text{ Olj-a}_1 \\ \text{SYN} & [_S \, N[\text{ACC}]_3 \, V_2 \, N[\text{NOM}]_1]_4 \\ \text{GF} & PRED_{2,4}:\langle GF_1:[\textit{3rd,sing,fem}], GF_3:[\textit{3rd,sing,neut}]\rangle \\ \text{SEM} & \mathbf{break}'_2(\text{AGENT:olya}_1, \text{PATIENT:}\iota x[\mathbf{vase}'(x)]_3)_4 \end{bmatrix}$

Since SiSx is not committed to an exhaustive mapping from SYN nodes to GFs, the class of GFs that have to be represented within the PRED of any given syntactic projection can be smaller than in GF-centered theories like LFG.[13] Only the elements subcategorized by a head – direct NP or CP arguments and oblique objects of non-thematic prepositions – actually need to appear on the PREDs that correspond to the projections of these heads. There are usually no more than three of these GF-bearing arguments per PRED.

What is peculiar about subcategorized elements is that, in general, their morphosyntactic forms are unrevealing about their semantic roles (CULICOVER; JACKENDOFF, 2005, pg. 182). Adjuncts, by contrast, wear their semantic roles on their morphosyntactic sleeves: e.g. *via* inherent case specification or choice of preposition. In English, for instance, PPs headed by

---

[13] Most of the richness that is present in SYN and SEM is absent from the GF-tier, which makes the GF-tier a much *simpler* level than LFG's f-structure. This derives from the fact that SiSx builds upon a more radical version of representational economy than the one LFG assumes – one that applies not only to phrase structure, but to ALL LEVELS OF GRAMMAR. If some correspondences *can* be stated as direct relations between SYN and SEM, SiSx can do this without invoking an intermediate mapping through the GF-tier. This, however, is only possible because SiSx also abandons the assumption of INTERFACE UNIFORMITY, which is pervasive in MGG and survives – albeit in a much lighter fashion – in LFG's architecture (KAPLAN, 1987; ASUDEH, 2006; FINDLAY, 2017). It is the idea that the mapping to semantics is established uniformly on the basis of GFs that forces LFG to populate f-structure with semantically relevant SYN information.

SiSx's more sparing use of GFs is partly motivated by the commitment to what Jackendoff (2011a) calls the EVOLUTIONARY CONSTRAINT – namely, the idea that the architecture of grammar should be compatible with a plausible evolutionary scenario. With this in mind, Jackendoff (2002, pg. 261) speculates that the GF-tier is probably "the latest developing part of the architecture", since its properties are asymmetrically dependent upon the existence of articulated systems of constituent structure and semantics – i.e. the latter two components can exist without the GF-tier, but not vice-versa. It is hard to reconcile the LFG architecture – where f-structures are essential to the mapping between c-structure and semantics – with these considerations.

near and *under* are always LOCATIONS, while those headed by *during* and *after* are invariably interpreted as TIMES. As a consequence, correspondence rules for these elements can be stated directly as relations between SYN and SEM, circumventing the GF-tier (a possibility anticipated in Figure 2). Subcategorized arguments that occur *within* adjuncts (e.g. complements of predicative Ps) receive GFs like any other argument of a syntactic predicate.

The general idea, then, is that every *n*-level lexical projection endowed with subcategorization properties corresponds to a list of GFs (a PRED), where each member represents a subcategorized syntactic argument along with its ϕ-features. Henceforth, I will often omit the multiple links to the same GFO associated with the different projections of a head and represent the correspondences to the GF-tier either at the X^0 or XP levels alone. The structure for the sentence *Bobby ate the cake with the fork during the party* is, therefore, (38).

$$(38) \quad \begin{bmatrix} \text{PHON} & \text{Bobby}_1 \text{ ate}_2 \text{ the} \prec \text{cake}_3 \text{ with}_4 \text{ the} \prec \text{fork}_5 \text{ during}_6 \text{ the} \prec \text{party}_7 \\ \text{SYN} & [_S \text{ NP}_1 [_{VP} \text{ V}_2 \text{ NP}_3 [_{PP} \text{ P}_4 \text{ NP}_5] [_{PP} \text{ P}_6 \text{ NP}_7]]] \\ \text{GF} & \text{PRED}_2{:}\langle \text{GF}_1{:}[\textit{3rd,sing,masc}], \text{GF}_3{:}[\textit{3rd,sing,neut}]\rangle \\ & \text{PRED}_4{:}\langle \text{GF}_5{:}[\textit{3rd,sing,neut}]\rangle \text{ PRED}_6{:}\langle \text{GF}_7{:}[\textit{3rd,sing,neut}]\rangle \\ \text{SEM} & \mathbf{eat}'_2(\text{AGENT}{:}\text{bobby}_1, \text{PATIENT}{:}\iota x[\mathbf{cake}'(x)]_3, \\ & \text{INSTRUMENT}{:}\iota y[\mathbf{fork}'(y)]_5, \text{TIME}{:}\mathbf{during}'_6(\iota z[\mathbf{party}'(z)]_7)) \end{bmatrix}$$

Note that the PREDs that correspond to *with the fork* (PRED$_4$) and *during the party* (PRED$_6$) are detached from the PRED that corresponds to the main clause (PRED$_2$). Each lexical projection is linked to a separate PRED, as we saw above. In this way, the role of PREDs is analogous to that of KERNEL SENTENCES in Harris (1957) and Chomsky (1957), GOVERNING CATEGORIES in Chomsky (1981), ARG-ST in HPSG (MANNING; SAG, 1998) and, most importantly, SYNTACTIC PREDICATES in Reinhart and Reuland (1993).

Each PRED contains a maximum of three GF members: the SiSx equivalents of what LFG calls CORE GFs (BRESNAN, 2001, pg. 96) (SUBJ, OBJ and OBJ2) and Relational Grammar calls relations 1, 2 and 3. These are the GFs that most strongly justify an autonomous tier for GFs in the first place, because, in addition to being subcategorized, they are the typical targets for phenomena like agreement, raising, passive and structural case-marking.[14] None of these

---

[14] The elements that correspond to OBJ2 (or relation 3 in Relational Grammar) are oblique objects of governed prepositions like *to* and *about* in *talked about April to Joanne* or *on* in *relies/counts on Bobby*. These positions are assigned to GFs within their predicates because they are subject to passivization (i) and can host raised expletives from embedded VPs (ii) (POSTAL; PULLUM, 1988):

can be stated in terms of direct correspondences between SEM and SYN. Passive and raising in English, for instance, are able to target elements that are not semantic arguments of their verbs, as in (39), as well as semantically empty expletives like *it* and *there*, as in (40).

(39)   a.   Bobby was expected to win.

       b.   Bobby seems to be expected to win.

(40)   a.   There was expected to be a party for Bobby.

       b.   It seems to be possible to leave early.

What we need to handle these facts is a way of manipulating "the status of *syntactic* arguments, irrespective of their semantic status and their syntactic position" (CULICOVER; JACKENDOFF, 2005, pg. 189). This is precisely what the GF-tier provides. For (39), the GF-tier is able to express the fact that a SYN node can realize a GF in a clause where it is not a semantic argument. For (40), the GF-tier is able to represent a syntactic dummy as bearing a GF without requiring this GF to correspond to anything in SEM. What happens in (39)-(40) is that it is precisely these *autonomous* (i.e. non-semantically linked) GFs which end up being targeted by passive and raising. Note that the subjects in both (39) and (40) also trigger agreement with the main-clause verbs, which do not assign them a θ-role. Hence, licensing conditions for passive, raising and agreement all find a comfortable home on the GF-tier.

However, if the GF-tier is to provide a coherent formulation of these phenomena, it cannot be limited to GFs assigned within *verbal* predicates, contrary to the practice of Culicover and Jackendoff (2005) and the suggestions in Smirnova and Jackendoff (2017). At least APs and PPs must map to PREDs on the GF-tier, given the possibility of raising in (41) (BRESNAN, 2001, chap. 12) and the fact that, in the Portuguese example in (42), agreement occurs both with the embedded adjective *bonitos* ('pretty') and the matrix raising verb *parecem* ('seem'):[15]

(i)   a.   Bobby was being talked about all night.
       b.   Bobby is relied on by all of his friends.

(ii)   a.   Amy relies on it to be raining.   (POSTAL, 2004, pg. 92)
       b.   "He's counting on there to be a soft place to fall on the other side."   (example from the Web)

Prepositional passives involving predicative PPs (e.g. *This bed has been slept in*) call for a separate treatment, since they seem to be subject to peculiar discourse related constraints that don't apply to passivization in general. A particular analysis of these cases based on Davison (1980) and Menuzzi (2004) is sketched Chapter 6, where it will play an important role in motivating my own analysis of discursive reflexives in English.
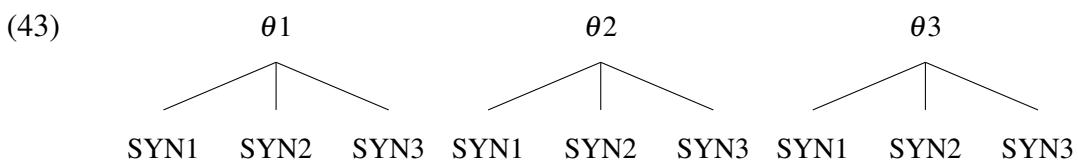
[15] Consider (42) in more detail. If we want to uphold that agreement is a local phenomenon, confined roughly to the domain of subcategorization (i.e. the governing category of Chomsky (1981)), we are forced to say either that there is an unpronounced constituent which is local to the adjective or that the adjective projects a PRED of its own, where the GF that corresponds to *os garotos* ('the boys') is the first argument. I assume the latter.

(41) a. Susan seems happy.

   b. Susan seems in a bad mood.

(42) Os   garotos parecem bonitos.
   DET.3.PLUR boy.3.PLUR seem.3.PLUR pretty.3.PLUR.MASC
   'The boys seem pretty.'

I will assume that every head capable of subcategorizing an argument maps to a PRED and that every subcategorized argument (including arguments of V, P, A and N) corresponds to a member of the PRED that corresponds to head that selects it. The latter requirement is the SiSx analogue to the Case Filter in the theory of Government and Binding (CHOMSKY, 1981).

The general point of positing an autonomous levels for GFs is to simplify the correspondence between form and meaning. Instead of having to state separately all of the possible syntactic realizations for each individual θ-role (AGENTS can be omitted in imperatives, realized as overt sisters of VPs in declaratives, as complements of *by* in passives; PATIENTS can be objects in transitives, subjects in passives, etc.), we can set up the GF-tier as an intermediate level between SYN and SEM and retain a simple uniform mapping between θ-roles and GF configurations. On the basis of a constant GF arrangement, we can then specify a correspondence to a proprietary set of SYN realizations, which may be shared by different θ-roles (e.g. AGENT and PATIENT both share the possibility of realized as sisters of VP in different contexts).

By so doing, we avoid the inconvenience of having to state the same set of mappings to SYN over and over for different θ-roles. So instead of having something like (43) (where $\theta 1$, $\theta 2$, $\theta 3$ are θ-roles and SYN1, SYN2, SYN3 are unrelated syntactic expressions), GFs allow us to have something like (44) (where $\langle \dots \text{GF} \dots \rangle$ represents any given arrangement of GFOs):

(43) $\theta 1$ $\theta 2$ $\theta 3$

    SYN1 SYN2 SYN3 SYN1 SYN2 SYN3 SYN1 SYN2 SYN3

(44) $\theta 1$ $\theta 2$ $\theta 3$

     $\langle \dots \text{GF} \dots \rangle$

   SYN1 SYN2 SYN3

Each position of a GF within a PRED, is, thus, an equivalence class of SYN realizations

that share a characteristic mapping to θ-roles in SEM (BRESNAN, 2001, pg. 95). This allows us to capture commonalities between different constructions, both within a language as well as across languages, without resorting to derivational devices like transformations.

The following subsection discusses what it means for a particular unit in one level of representation to be in CORRESPONDENCE to a unit in another level of representation. This will clarify the precise nature of the numerical indices used in the AVMs throughout this chapter. We will also see some further examples of correspondence constraints involving the GF-tier.

4.3.3 Correspondences

Throughout this chapter, I have been presenting the Parallel Architecture of SiSx (see Figure 3, repeated below), which contains four different modules for representing linguistic information: PHONOLOGICAL STRUCTURE (PHON), SYNTACTIC STRUCTURE (SYN), the GF-TIER (GF) and SEMANTIC STRUCTURE (SEM). Repeatedly, I have been alluding to the fact that that these different modules are connected to one another via systematic correspondences. I now clarify how these correspondences are modeled within the present framework.
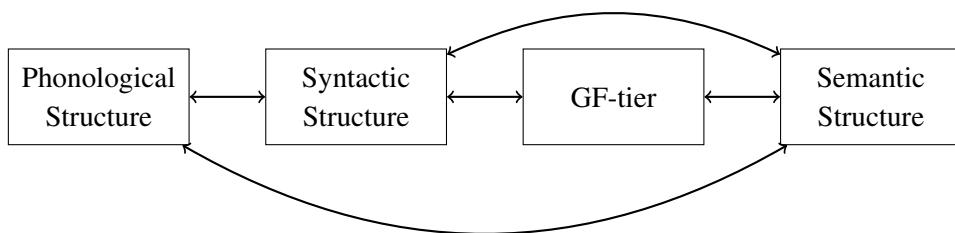
Figure 3 – The Parallel Architecture of SiSx

The theory of SiSx posits five types correspondence among the different structures which comprise linguistic expressions. These are depicted by the double arrows in Figure 3. In formal terms, a correspondence will be a binary symmetric relation between the minimal units in each level of representation. The five correspondences posited by SiSx are defined in (45).

(45)   **Correspondences:**

    a.   A symmetric relation $C_{PHON\text{-}SYN}$ that holds between strings and nodes.

    b.   A symmetric relation $C_{SYN\text{-}GF}$ that holds between nodes and GFOs.

    c.   A symmetric relation $C_{GF\text{-}SEM}$ that holds between GFOs and MEs.

    d.   A symmetric relation $C_{PHON\text{-}SEM}$ that holds between strings and MEs.

    e.   A symmetric relation $C_{SYN\text{-}SEM}$ that holds between nodes and MEs.

As with all other defining properties of linguistic objects in SiSx, correspondences can also be used in DESCRIPTIONS of modeled structures in order to state well-formedness constraints. An example of a PHON-SYN correspondence constraint requiring heads to precede their complements is the following (where $\prec$ is the PRECEDENCE relation over strings, φ_4 and φ_5 are variables over strings, x_1, x_2 and x_3 are variables over nodes, M is the MOTHER function and L is the LABEL function):

(46)　$M(x_2) = x_1 \land M(x_3) = x_1 \land L(x_1) = XP \land L(x_2) = X \land L(x_3) = YP \land \varphi_4 \prec \varphi_5$
　　　　$\land \, C_{\text{PHON-SYN}}(\varphi_4, x_2) \land C_{\text{PHON-SYN}}(\varphi_5, x_3)$

As with other constraints, logical formulae like (46) are abbreviated by AVMs. Each correspondence relation is depicted by coindexing of structures in different levels. The constraint in (46) is, therefore, equivalent to (47):

(47)　$\begin{bmatrix} \text{PHON} & \varphi_1 \prec \varphi_2 \\ \text{SYN} & [_{XP} \, X_1 \, YP_2] \end{bmatrix}$

An example of linguistic object which is licensed by (46)/(47) is given in (48):

(48)　$\begin{bmatrix} \text{PHON} & \text{into}_1 \text{ the} \prec \text{woods}_2 \\ \text{SYN} & [_{PP} \, P_1 \, NP_2] \end{bmatrix}$

Though structures in individual levels of representation may be of theoretical interest to the linguist, most of the action in the grammar is carried out by correspondences between levels. The kind of linguistic object that will most often concern us is what Culicover (2021) calls a CONSTRUCT: i.e. an object consisting (minimally) of a string in PHON, a node in SYN, a Meaningful Expression in SEM and a correspondence between each of the units in these structures. In order to license constructs, the grammar needs to include CORRESPONDENCE CONSTRAINTS not only between PHON and SYN (like (47)), but also between the other levels. Following Culicover (2021), I sometimes refer to correspondence constraints as CONSTRUCTIONS.

A more abstract example of a construction (now involving the GF-tier) is a VALENCE FRAME like the one for transitive Vs. The demand that a verb V subcategorize for an object can be thought of as correspondence constraint requiring the presence of a at least two members on the PRED that corresponds to V. This is captured by (49) below. This constraint is satisfied by

the particular transitive variant of *eat* that appears in the construct in (38) above.

(49)
$$\begin{bmatrix} \text{SYN} & V_1 \\ \text{GF} & PRED_1{:}\langle GF_2, GF_3 \dots \rangle \end{bmatrix}$$

Similarly, a valence template for raising to subject verbs like *seem*, which is instantiated by (34) above, is a correspondence constraint like the following:

(50)
$$\begin{bmatrix} \text{SYN} & [_{VP}\ V_2\ VP[\text{INF}]_3] \\ \text{GF} & PRED_2{:}\langle GF_1 \rangle\ PRED_3{:}\langle GF_1 \dots \rangle \end{bmatrix}$$

Subject-verb agreement in languages like English can also be thought of as a correspondence constraint – this time requiring marking in SYN and PHON and an imposition that the first argument on the PRED that corresponds to the verb bear certain features. A suitable template for a 3rd person verb like is something like (51). When this construction is used in conjunction with the lexical entries for particular verbs, it licenses forms like *eats*, *wants*, *believes*, *runs*, etc.

(51)   ENGLISH SUBJECT-VERB AGREEMENT CONSTRUCTION
$$\begin{bmatrix} \text{PHON} & \varphi_1 \prec s_3 \\ \text{SYN} & [_V\ V_1\ \text{Affix}_3] \\ \text{GF} & PRED_1{:}\langle GF_2{:}[3rd, sing], \dots \rangle \end{bmatrix}$$

A slightly more complex example of a correspondence constraint also involving the GF-tier is the BRAZILIAN PORTUGUESE OBJECT CLITIC CONSTRUCTION in (52) (see Culicover and Jackendoff (2005, pg. 193) for a simplified version). Parentheses indicate optionality and ⊕ stands for the APPEND relation often employed in HPSG (POLLARD, 1997; MÜLLER, 2020). APPEND combines two lists into one, preserving the order of the elements in each list:

(52)   BRAZILIAN PORTUGUESE OBJECT CLITIC CONSTRUCTION
$$\begin{bmatrix} \text{PHON} & \varphi_2 \prec \varphi_3 \\ \text{SYN} & [_{VP}\ Clitic_2\ V_3] \\ \text{GF} & PRED_3{:}\langle GF_1(, GF_4) \rangle \oplus \langle GF_2 \rangle \oplus (\langle \dots \rangle) \end{bmatrix}$$

This construction has the effect of realizing the second or third GF in a syntactic predicate (i.e. a DIRECT OBJECT or an OBLIQUE) as a clitic which precedes the string that corresponds

to the verb in PHON. In order to capture this, I separate the PRED list that corresponds to the verb ($PRED_3$) in three parts: one for the arguments that precede the clitic (one argument, if the clitic stands for the direct object, two arguments, if the clitic is oblique) one for the clitic itself, and another (optional one) for any arguments that may follow the clitic. This last list will have one member only if the clitic corresponds to a direct object of a ditransitive verb.

The complex correspondence constraint in (52) licenses the Portuguese VPs in (53):

(53) a. me   viu
          CL.1SG saw
          'saw me'

      b. me   mostrou pra   Marta
          CL.1SG showed to-DET.FEM Marta
          'showed me to Marta'

      c. me   mostrou a   Marta
          CL.1SG showed DET.FEM Marta
          'showed Marta to me'

For (53a), (52) licenses the structure in (54a), where the clitic stands for direct object. For (53b), (52) licenses in (54b), where the clitic is also a direct object, but one which is followed by an oblique. For (53c), on the other hand, (52) licenses (54c), where the clitic corresponds to an oblique followed by a direct object:

(54) a. $\left[\begin{array}{ll} \text{PHON} & \text{me}_2 \text{ viu}_3 \\ \text{SYN} & [_{VP} \text{ Clitic}_2 \text{ V}_3] \\ \text{GF} & PRED_3:\langle GF_1 \rangle \oplus \langle GF_2:[\textit{1st, sing}] \rangle \\ & = PRED_3:\langle GF_1, GF_2:[\textit{1st, sing}] \rangle \end{array}\right]$

      b. $\left[\begin{array}{ll} \text{PHON} & \text{me}_2 \text{ mostrou}_3 \text{ pra} \prec \text{Marta}_4 \\ \text{SYN} & [_{VP} \text{ Clitic}_2 \text{ V}_3 \text{ PP}_4] \\ \text{GF} & PRED_3:\langle GF_1 \rangle \oplus \langle GF_2:[\textit{1st, sing}] \rangle \oplus \langle GF_4 : [\textit{3rd, sing, fem}] \rangle \\ & = PRED_3:\langle GF_1, GF_2:[\textit{1st, sing}], GF_4:[\textit{3rd, sing, fem}] \rangle \end{array}\right]$

      c. $\left[\begin{array}{ll} \text{PHON} & \text{me}_2 \text{ mostrou}_3 \text{ a} \prec \text{Marta}_4 \\ \text{SYN} & [_{VP} \text{ Clitic}_2 \text{ V}_3 \text{ NP}_4] \\ \text{GF} & PRED_3:\langle GF_1, GF_4:[\textit{3rd, sing, fem}] \rangle \oplus \langle GF_2:[\textit{1st, sing}] \rangle \\ & = PRED_3:\langle GF_1, GF_4:[\textit{3rd, sing, fem}], GF_2:[\textit{1st, sing}] \rangle \end{array}\right]$

Due to the pervasiveness of correspondences and their importance in the framework of SiSx, the architecture in Figure 2 can also be called a CORRESPONDENCE ARCHITECTURE – a term often used in LFG circles (ASUDEH, 2006; FINDLAY, 2017). This kind of architecture sets SiSx and LFG apart from sign-based theories like HPSG and SBCG (POLLARD; SAG, 1994; SAG, 2012). The latter use the same kind of data structure to model all aspects of linguistic objects: i.e. typed feature-structures. Different types of information are not related by means of modular correspondences, but in virtue of being values assigned to different attributes of the same sign, with each attribute representing a different type of linguistic information. The design of HPSG/SBCG does not make it clear that phonology, syntax and semantics are autonomous combinatorial systems. Combinatoriality only exists at the level of signs as a whole (e.g. in features like DTRS, which take lists of *signs* as values, instead of syntactic nodes).

Even though SiSx follows HPSG/SBCG in using AVMs to represent all aspects of linguistic objects, its basic ontology is much closer to LFG's: each linguistic level is conceptualized as an autonomous formal system in its own right. Just as in LFG, this requires positing correspondence principles to link the objects independently defined by each of these systems.

There are great benefits to adopting a correspondence architecture beyond the elegant factorization they provide for linguistic analyses. One example of the utility of these kinds of architectures which goes beyond the domain of grammatical theory concerns the relationship between linguistics and studies in the evolution of language.

Most evolutionary psychologists assume that the emergence of human language was gradual, involving a series of incremental steps (protolanguages), each of which offered some adaptive advantage over the previous one (PINKER; BLOOM, 1990; CORBALLIS, 2017; DENNETT, 2017; FITCH, 2017; BOECKX, 2017; BOER et al., 2020). Given the absence of a fossil record, one of the main ways to investigate the particular stages of this incremental process is reverse-engineering: i.e. asking what components of language are advantageous without the whole system in place (JACKENDOFF, 2005, PROGOVAC, 2016, 2019).

Saying that language is composed by a set of separable modules makes it easier to imagine how these stages could have come about: e.g. we can think of a first stage in the evolution of language where linguistic expressions consist only of PHON and SEM (a possible model of primate calls), a subsequent stage where SYN is added, a third stage where GFs are introduced, increasing complexity, and so on. Each of these stages is logically possible and may leave residues (so-called LINGUISTIC FOSSILS) in constructions of modern-day languages (JACKENDOFF,

2002; PROGOVAC, 2015). Correspondence architectures give us these possible stages for the evolution of language essentially for free (JACKENDOFF, 1999, 2010). Theories that posit a single representational format which is responsible for the full range of information encoded by linguistic utterances are typically forced to adopt a highly implausible 'saltationist' scenario, whereby language with all its architectural complexity emerges in a single non-incremental step (HORNSTEIN, 2009; MIYAGAWA et al., 2014; BERWICK; CHOMSKY, 2016).

Correspondence architectures are also better suited for integration with theories of other cognitive faculties than syntactocentric models (see Jackendoff (2007b, 2011b) for full versions of this argument). It is a given that the mind includes relations between non-linguistic representations. For instance, visual and haptic information relate to a modality-independent understanding of the spatial structure of objects (MARR, 1982). This spatial structure, in turn, relates to language in a way that allows us to talk about what we perceive (JACKENDOFF, 1987; LANDAU; JACKENDOFF, 1993). Actions are also spatially guided, requiring an interface between spatial structure and schemas encoding action patterns. It does not make any sense to think of any of these representations as being *algorithmically derived* from any other – they are, rather, related in virtue of modular correspondences.

In this sense, the correspondence architecture of SiSx sees the internal components of language as "connected to each other in the same way as language is connected with the rest of the mind, and in the same way as other faculties of mind are connected to each other" (JACKENDOFF; AUDRING, 2020, pg. 8). Though many details about how such connections work remain unknown, SiSx seems better suited for fruitful cross-disciplinary dialogue with cognitive science than MGG, which opts for a syntactocentric derivational design.

4.3.4 Well-formedness

Now that we have an adequate understanding of what linguistic expressions and the substructures that comprise them are like, we can consider how these expressions are rendered WELL-FORMED by that constraints that comprise the grammar. Roughly speaking, an expression will be deemed well-formed if all of its parts fully instantiate some constraint in the grammar. The discussion here will follow closely the exposition in Culicover (2021, chap. 2).

There are two relations that hold between parts of linguistic expressions and grammatical constraints: instantiation by identity and instantiation by substitution. Both of these are included under the definition of TERM INSTANTIATION in (55) (CULICOVER, 2021):

(55) TERM INSTANTIATION:

A term τ in a linguistic expression instantiates a term T in a constraint *iff*:

a. τ is identical to T (instantiation by identity);

b. τ is a value of type α and T is a λ-bound or free variable of type α (instantiation by substitution).

The notion of 'term' is used here to refer to the basic units and relations that define structures in all levels of representation of a linguistic expression (PHON, SYN, GF, SEM). As an illustration, consider the construct that represents the sentence *Bobby eats the cake*:

(56)
$$\begin{bmatrix} \text{PHON} & \text{Bobby}_1 \text{ eats}_2 \text{ the} \prec \text{cake}_3 \\ \text{SYN} & [_S \text{ NP}_1 [_{VP} \text{ V}_2 \text{ NP}_3]]_5 \\ \text{GF} & \text{PRED}_{2,5}{:}\langle \text{GF}_1{:}[\textit{3rd,sing,masc}], \text{GF}_3{:}[\textit{3rd,sing,neut}]\rangle \\ \text{SEM} & \textbf{eat}'_2(\text{AGENT:bobby}_1, \text{PATIENT:}\iota x[\textbf{cake}'(x)]_3)_5 \end{bmatrix}$$

The instantiation of the SEM structure in (56) involves both identity and substitution. The meaning and the argument structure of *eat* in (56) instantiate the constraints imposed by the lexical entry in (57) by identity. In both (56) and (57), the meaning of *eat* is $\textbf{eat}'$ and its argument structure involves the AGENT and PATIENT relations. The terms 'bobby' and '$\iota x[\textbf{cake}'(x)]$' in (56), however, instantiate terms in (57) by substitution. This is possible because x and y in (57) are variables of type e and 'bobby' and '$\iota x[\textbf{cake}'(x)]$' are values of type e.[16]

(57)
$$\begin{bmatrix} \text{PHON} & \text{eat}_1 \\ \text{SYN} & \text{V}_1 \\ \text{SEM} & \lambda x[\lambda y[\textbf{eat}'_1(\text{AGENT}:y, \text{PATIENT}:x)]] \end{bmatrix}$$

Any given term in a linguistic expression may instantiate terms in multiple constructions simultaneously and in different ways. The instantiation of 'bobby' in the SEM tier of (56) instantiates by substitution the λ-bound variable in (57) and by identity of the term present in

---

[16] Instantiation by substitution is also illustrated by the way in which (56) instantiates the HEAD-INITIAL CON-
STRUCTION in (47) repeated below:

(i)
$$\begin{bmatrix} \text{PHON} & \varphi_1 \prec \varphi_2 \\ \text{SYN} & [_{XP} X_1 \, YP_2] \end{bmatrix}$$

To be more precise: the V and NP in (56) instantiate *X* and *YP* (respectively) in (i) by substitution. The relations
of precedence in PHON and MOTHER in SYN in (56) are instantiated by identity with the relations in (i).

the SEM tier of the lexical constraint associated with *Bobby*, which is given in (58):

(58)
$$\begin{bmatrix} \text{PHON} & \text{Bobby}_1 \\ \text{SYN} & \text{NP}_1 \\ \text{SEM} & \text{bobby}_1 \end{bmatrix}$$

On the basis of the notion of TERM INSTANTIATION, we can define what it means for a linguistic expression as a whole to instantiate a particular constraint:

(59) CONSTRAINT INSTANTIATION

An expression ε fully instantiates a constraint C *iff*

a. for all terms T in C, there is some term τ in ε such that τ instantiates T; and

b. all descriptions of relations between levels in C are true of ε (i.e. the correspondences in ε are models of those in C).

Given (59) we can finally define the notion of WELL-FORMED EXPRESSION:

(60) An expression ε is WELL-FORMED *iff*

a. for every term τ in ε, τ instantiates a term T in some constraint C and

b. for every such C, ε fully instantiates C.

The definition in (60) imposes two conditions on well-formed expressions. The first one says that every element in the expression must instantiate some element in some grammatical constraint. This entails that, for every well-formed expression ε, there is no part of ε which is not adequately licensed by the grammar. The second condition requires that any constraint involved in making an expression well-formed must be fully instantiated in that expression. This means, in effect, that expressions cannot pick and choose parts of constraints to instantiate: they either instantiate an entire constraint, or not instantiate it at all (CULICOVER, 2021). This happens because a grammatical constraint is, formally speaking, a logical conjunction. In order for the conjunction to be *true* of an individual expression ε, all of its conjuncts must be true of ε. What the two conditions in (60) imply, in other words, is that a well-formed expression ε must be a MODEL of some subset of the constraints in the grammar. This definition adequately captures the model-theoretic flavor of of SiSx.

According to (60), a structure may be ill-formed because it fails to instantiate constraints

on the form of expressions (PHON, SYN and their correspondences):

(61) $*\begin{bmatrix} \text{PHON} & \text{Curly}_1 \text{ girl}_4 \text{ saw}_2 \text{ the}_3 \\ \text{SYN} & [_S \text{ NP}_1 [_{VP} \text{ N}_4 \text{ V}_2 [_{NP} \text{ Det}_3]_5]] \\ \text{SEM} & \mathbf{see}'_2(\text{EXPERIENCER:curly}_1, \text{THEME:}\iota y[\mathbf{girl}'_4(y)]) \end{bmatrix}$

However, in addition, (60) also predicts ill-formedness for structures like (62), which fail to instantiate constraints that specify correspondences between form and meaning:

(62) $*\begin{bmatrix} \text{PHON} & \text{Curly}_1 \text{ saw}_2 \text{ the}_3 \text{ girl}_4 \\ \text{SYN} & [_S \text{ NP}_1 [_{VP} \text{ V}_2 [_{NP} \text{ Det}_3 \text{ N}_4]_5]] \\ \text{SEM} & \mathbf{fall}'_2(\text{THEME:}\iota x[\mathbf{horse}'(x)]_1) \end{bmatrix}$

There is nothing wrong with the linguistic forms employed in (62): the dominance and labeling relations in SYN and the precedence relations in PHON all instantiate constraints that are part of the grammar of English. The problem lies in the mapping between PHON/SYN and SEM. The constraints stored in the English lexicon do not license a correspondence between the string *Curly* and the meaning $\iota x[\mathbf{horse}'(x)]$ or between *saw* and $\mathbf{fall}'$; in addition, they require *the girl* to correspond to a semantic term $\iota y[\mathbf{girl}'(y)]$, which is absent from (62). These violations show that the notion of well-formedness in (60) is broader than the purely syntactic notion of string validity typically assumed within PTFs (CHOMSKY, 1957).

Another interesting property of the definition of well-formedness in (60) is that it entails a kind of CONFIRMATIONAL HOLISM – a trait which Quine (1951) ascribes to scientific theories in general. Just as no scientific hypothesis, taken in isolation, can be refuted or confirmed by any piece of data (because ancillary hypotheses can always be invoked in a way that preserves the original hypothesis), no individual constraint a linguist proposes as part of a grammar can be refuted or confirmed by the well-formedness of a particular linguistic expression (see also Pullum (2020)). Only the grammar as a whole (i.e. the complete set of constraints) can be confirmed or falsified. This happens because well-formedness in SiSx is not determined by satisfaction of ALL of the constraints in the grammar. For an expression ε to be considered well-formed, it is only required that each of the terms in ε the expression satisfy SOME constraint. Even though one constraint may be in conflict with some properties of ε, there can always be ANOTHER constraint in the grammar which makes those properties of ε well-formed.

As an exemple, consider the general constraint which requires heads, in English, to pre-

cede their complements in PHON. This SiSx analogue to the HEAD-INITIAL PARAMETER was captured by (47) above and is repeated below:

(63)
$$\begin{bmatrix} \text{PHON} & \varphi_1 \prec \varphi_2 \\ \text{SYN} & [_{XP} X_1 \ YP_2] \end{bmatrix}$$

A putative counterexample to (63) is the English adposition *ago*, which must follow its complement (**ago a long time*, *a long time ago*) (CULICOVER, 1999; HUDDLESTON; PULLUM, 2002). If this expression is indeed a head (and it does behave like normal adpositional heads in other respects), its properties are in contradiction with (63), which describes the overwhelming majority of English expressions.

If well-formedness was defined in terms of satisfaction of ALL of the constraints in the grammar, the behavior of *ago* would oblige us to give up on (63).[17] But this is not necessary in SiSx. The definition of well-formedness in (60) allows us to keep (63) within our grammar as a kind of *default* and introduce item-specific constraints to license the order of head-final adpositions. The following is a lexical entry for *ago* which accounts for its odd properties:

(64)
$$\begin{bmatrix} \text{PHON} & \varphi_1 \prec \text{ago}_2 \\ \text{SYN} & [_{PP} \text{P}_2 \ NP_1] \end{bmatrix}$$

A SiSx grammar can, therefore, include "heads precede their complements", "*ago* is an head" and "*ago* follows its complement" among its statements and still be satisfiable. An expression like *a long time ago* would be a model of such grammar. The set of constraints in the grammar is allowed to be INCONSISTENT in a way that is useful to capture both general patterns as well as idiosyncrasies which deviate from what is otherwise the norm. As a result, the theory does not have to downplay the importance of broad 'macro-parametric' properties like (63) nor of quirky distributional facts which are often linked to a small subset of lexical items.

---

[17] As Pullum (p.c.) notes, the only way around this, within a non-holistic setting, would be to stipulate a higher-level category for syntactic heads which includes, as a subsort, a category for head-initial heads (including head-initial Ps like *on*) and head-final heads (including head-final Ps like *ago*). One could then state separate precedence constraints for each of these classes. (63) would only apply to the former. In addition, we would need to posit separate entries for the adposition *notwithstanding*, which can both precede as well as follow its complement (*notwithstanding Bobby*, *Bobby notwithstanding*). These options introduce a kind of complexity which the violable constraints of holistic grammars manage to avoid. In this way, allowing constraints to be violated makes grammars "more compact" (LASCARIDES et al., 1996). More compact grammars can be extensionally equivalent to less compact ones (and this may well be the case here), but compactness makes an empirical difference under the assumption that language acquisition abides by a kind of evaluation metric that favors simpler (e.g. shorter) grammars (CHOMSKY, 1965; CULICOVER, 2013b).

4.4 THE LEXICON AND ITS STRUCTURES

Up to now, I have talked mostly about how SiSx represents the structure of linguistic expressions and how these expressions are licensed. This section focuses more on the kinds of CONSTRAINTS that are responsible for licensing these expressions. A widespread assumption is that these constraints fall into two radically different classes, depending on whether they apply to WORDS and their internal parts or to larger PHRASAL UNITS. This view is famously expressed in Bresnan & Mchombo's (1995, pg. 181) LEXICAL INTEGRITY PRINCIPLE:

(65)   THE LEXICAL INTEGRITY PRINCIPLE

Words are built out of different structural elements and by different principles of composition than syntactic phrases.

Most theories (MGG, LFG and most variants of HPSG) enforce (65) by separating the LEXICON from the RULES OF (PHRASAL) GRAMMAR. The latter are responsible for the organization of novel phrases while the former is supposed to register idiosyncrasies as well as capture some partial regularities among stored items (in the form of LEXICAL RULES).[18]

SiSx argues that there is much to be gained by abandoning this distinction. The first step of the argument involves asking what the lexicon is. Due to its mentalist commitment, SiSx frames this issue in essentially psycholinguistic terms, taking the lexicon to be whatever the language user has to learn and store in long-term memory. The argument then goes on to show that a lexicon thus conceived must contain entries of such variety and complexity that a sharp distinction between lexical items and grammatical rules becomes artificial (JACKENDOFF, 1997; CULICOVER; JACKENDOFF; AUDRING, 2017; JACKENDOFF; AUDRING, 2020). The slippery slope from words to rules of grammar prompts SiSx to view the latter as *part* of the lexicon, as in Construction Grammar (GOLDBERG, 1995). This looks natural under a model-theoretic design, where lexicon and grammar are equally stated as constraints.

---

[18] In its contemporary form, this distinction dates back to Chomsky's (1970) LEXICALIST HYPOTHESIS. In that framework, however, the divide between LEXICAL RULES and RULES OF GRAMMAR overlapped with the distinction between CONSTRAINTS and ALGORITHMS. In a model-theoretic formalism – where ALL rules are stated as constraints – these two kinds of rules can only be distinguished by the types of variables they contain: variables on lexical constraints range over word-like elements and the ones on grammatical constraints range over phrases. Lexical Integrity is, then, a requirement that constraints containing different types of variables involve fundamentally different relations (i.e. "different principles of combination"): e.g. constraints on word formation should not mention long-distance relationships between items, like the ones found in phrasal grammar. Though this is requirement is formulable in a model-theoretic setting, it is not clear whether it can be empirically justified. See Bruening (2018) for some relevant discussion.

A typical instance of a lexical item is an individual word like *cow*. SiSx, following the Parallel Architecture in Figure 3, treats this as a correspondence constriant, linking a small piece of phonology, a syntactic category and a meaning, as in (66):

(66)
$$\begin{bmatrix} \text{PHON} & \text{cow}_1 \\ \text{SYN} & \text{N}_1 \\ \text{SEM} & \lambda x[\mathbf{cow}'_1(x)] \end{bmatrix}$$

As mentioned above, the same format can be used to represent valence. For items whose argument structure properties do not follow from general linking rules, valence information has to be specified on a case by case basis (i.e. as a part of individual lexical entries). If we assume that syntactic arguments are optional by default, verbs with obligatory OBJECTS must include this information in their lexical entries, as in (67) (JACKENDOFF, 2002):

(67)
$$\begin{bmatrix} \text{PHON} & \text{devour}_1 \\ \text{SYN} & \text{V}_1 \\ \text{GF} & PRED_1{:}\langle GF_2, GF_3 \rangle \\ \text{SEM} & \lambda y.\lambda x[\mathbf{devour}'_1(\text{AGENT}{:}x, \text{PATIENT}{:}y)] \end{bmatrix}$$

Valence in information, in this case, is stated in terms of the GF-tier and not SYN because the OBJECT of *devour* (i.e. GF_3) can be passivized (e.g. *The salad was devoured*), in which case it does not correspond to a verb-adjacent NP (more on passive below). Note, however, that the GFOs in (67) are not names for a concrete GFOs – they function as variables that impose contextual restrictions on structures where the verb might appear.

This formalism can also encode the properties of lexical heads which subcategorize for complements with an idiosyncratic categorial status in SYN. The verb *depend*, for example, subcategorizes for an NP within a PP headed by *on*, as in (68):

(68)
$$\begin{bmatrix} \text{PHON} & \text{depend}_1 \text{ on}_2 \; \varphi_3 \\ \text{SYN} & [_{\text{VP}} \text{V}_1 \; [_{\text{PP}} \text{P}_2 \; NP_3]]_4 \\ \text{SEM} & \lambda y[\lambda x[\mathbf{depend}'_1(\text{EXPERIENCER}{:}x, \text{THEME}{:}y)]](\sigma_3)_4 \end{bmatrix}$$

Italicized elements and Greek letters represent typed variables that must be contextually instantiated in order for the item to be licensed (CULICOVER, 2021). These are what give

lexical items their combinatoric potential.[19]

Productive morphology receives a similar treatment. Since regular forms *can* be computed online – and *must* be so computed in agglutinative languages like Turkish (HANKAMER, 1989) – we cannot require every one of them to be stored in the lexicon (JACKENDOFF, 1997; JACKENDOFF, 2002). Therefore, regular affixes must have their own lexical entries with variables specifying the phonology, category and semantics of their putative roots – as was also assumed in American Structuralist models of immediate constituent analysis (BLOOMFIELD, 1933). (69) is a possible entry for the English past suffix.

$$
(69) \quad \begin{bmatrix} \text{PHON} & \varphi_2\text{-ed}_1 \\ \text{SYN} & [_V \; V_2\text{-Affix}_1]_3 \\ \text{SEM} & \lambda P_{\langle event,t \rangle}[\lambda e_{event}[P(e) \wedge \text{TIME}(e, \text{past})]]_1(\sigma_2)_3 \end{bmatrix}
$$

Note that, as far as SiSx is concerned, there is no deep formal distinction between the *syntactic* combinatoriality of the verb in (68) and the *morphological* combinatoriality of the affix in (69). The only difference has to do with the nature of the variable in SYN: NP_3 in (68) is the label assigned to a phrasal node variable while V_2 in (69) is a label assigned to a word-like node variable. SiSx has no separate MORPHOLOGICAL level of representation beyond PHON, SYN, the GF-tier and SEM (JACKENDOFF; AUDRING, 2020).

A lexicon conceived in these terms should also contain a variety of multiword entries (CULICOVER; JACKENDOFF; AUDRING, 2017). Among these are idioms with fully specified material on all tiers, such as *kick the bucket*. In SiSx, these expressions can be stored as whole phonological/syntactic units, linked to noncompositional semantics, as in (70). We know that this particular idiom instantiates the canonical syntactic structure of an English VP because *kick* inflects just like an ordinary verb (e.g. *John kicked the bucket*, *if John kicks the bucket, then...*, *John seems to be kicking the bucket*, etc.).

$$
(70) \quad \begin{bmatrix} \text{PHON} & \text{kick}_1 \; \text{the}_2 \; \text{bucket}_3 \\ \text{SYN} & [_{VP} \; V_1 \; [_{NP} \; \text{Det}_2 \; N_3]]_4 \\ \text{SEM} & \lambda x[\textbf{die}'(\text{PATIENT}:x)]_4 \end{bmatrix}
$$

Like the verb in (68) and the affix in (69), some idioms have variables that grant them

---

[19] In particular, φ_1, φ_2, φ_n are variables over strings in PHON and σ_1, σ_2, σ_n are variables over Meaningful Expressions in SEM. Italicized elements in SYN can stand either for variables over node labels (in which case the letters X, Y and Z are often used) or for variables over nodes themselves. The latter is the case in (68).

combinatorial potential of their own. These are cases like *stab NP in the back*, *put NP on ice* and *catch NP's eye*. Here is a lexical entry for this last one:

(71)
$$\begin{bmatrix} \text{PHON} & [\text{catch}_1\ \varphi_2\text{'s}_3\ \text{eye}_4]_5 \\ \text{SYN} & [_{\text{VP}}\ V_1\ [_{\text{NP}}\ NP_2\text{-GENITIVE}_3\ N_4]]_5 \\ \text{SEM} & \lambda x[\textbf{notice}'(\text{EXPERIENCER}:\sigma_2, \text{THEME}:x)]_5 \end{bmatrix}$$

The entries in (70) and (71) pose a kind of ordering paradox for theories that assume a radical separation between grammar and lexicon, as prescribed by the Lexical Integrity. The information that *kick the bucket* and *catch NP's eye* are VPs has to be stated *in* the lexicon, because their semantics is idiosyncratic. However, the phrase-structure rule that generates VPs can only apply *outside* the lexicon.

In addition to these cases, the lexicon also has to include a class of CONSTRUCTIONAL IDIOMS that use normal syntax to unusual (i.e. noncompositional) semantic ends (JACKEND-OFF, 1997; JACKENDOFF, 2002). An example is the SOUND+MOTION CONSTRUCTION in (72)(LEVIN; HOVAV, 1995; GOLDBERG; JACKENDOFF, 2004):

(72)   The car [$_{\text{VP}}$ rumbled around Bobby].

$\textbf{go}'(\text{THEME}:\iota z[\textbf{car}'(z)], \text{PATH}:\textbf{around}'(\text{bobby}), \text{EFFECT}:\textbf{rumble}'(\iota z[\textbf{car}'(z)]))$

Syntactically, the VP in (72) is merely a sequence of a verb followed by a PP. Its semantics is unusual because the verb is not interpreted as a functor over the PP, but as specifying the EFFECT of a motion that is not codified by any of the words in the sentence. The effect of the motion, is, moreover, predicated of whatever is interpreted as the THEME (i.e. the entity undergoing the motion). A lexical entry with these properties is sketched in (73).

(73)   SOUND+MOTION CONSTRUCTION (adapted from Culicover (2013b, pg. 42)):
$$\begin{bmatrix} \text{SYN} & [_{VP}\ V_1\ PP_2]_3 \\ \text{SEM} & \lambda x[\textbf{go}'(\text{THEME}:x, \text{PATH}:\sigma_2, \text{EFFECT}:\sigma_1(x))]_3 \end{bmatrix}$$

What is peculiar about constructional idioms is that the SYN tier in their lexical entries consists *entirely* of variables that are completely unlinked to phonology. This makes them much more rule-like than word-like. However, since their interpretation does not follow from general principles, they have to be explicitly learned and stored just like words are (see Jackendoff

(1997) and Culicover (1999)).[20]

Two further examples of constructional idioms along with their relevant (simplified) lexical entries proposed in the SiSx literature are show below. (74) represents the RESULTATIVE CONSTRUCTION and (75) the DITRANSITIVE CONSTRUCTION (cf. JACKENDOFF, 1990b, GOLDBERG, 1995, 2006; ASUDEH; GIORGOLO; TOIVONEN, 2014):

(74)   RESULTATIVE CONSTRUCTION (adapted from Goldberg and Jackendoff (2004))

    a.   Annie [kissed Will unconscious].

    b.
$$\begin{bmatrix} \text{PHON} & [\varphi_1\ \varphi_2\ \varphi_3]_4 \\ \text{SYN} & [_{VP}\ V_1\ NP_2\ AP_3]_4 \\ \text{SEM} & \lambda x[\textbf{cause}'(\text{AGENT}:x, \text{THEME}:\sigma_3(\sigma_2), \text{MEANS}:\sigma_1)]_4 \end{bmatrix}$$

(75)   DITRANSITIVE CONSTRUCTION (adapted from Culicover (2021)):

    a.   Laurie [kicked Curly the ball].

    b.
$$\begin{bmatrix} \text{PHON} & [\varphi_1\ \varphi_2\ \varphi_3]_4 \\ \text{SYN} & [_{VP}\ V_1\ NP_2\ NP_3]_4 \\ \text{SEM} & \lambda x[\textbf{transfer}'(\text{SOURCE}:x, \text{GOAL}:\sigma_2, \text{THEME}:\sigma_3, \text{MEANS}:\sigma_1)]_4 \end{bmatrix}$$

Language turns out to be full of constructional idioms like these (GOLDBERG, 1995; JACKENDOFF, 2008; CULICOVER, 1999). However, insofar as recognizing their existence commits us to syntactically complex lexical items without phonology, nothing stops us from seeing general syntactic and correspondence constraints – usually thought of as part of the GRAMMAR – in the same way. The rule for licensing a transitive VP can be construed as a declarative schema for licensing a particular configuration of labeled nodes, as in (76):

(76)   TRANSITIVE VP SCHEMA (adapted from Jackendoff (2002, pg. 180)):
$$\begin{bmatrix} \text{SYN} & [_{VP}\ V\ NP] \end{bmatrix}$$

As far as SiSx is concerned, this is simply one of the possibilities allowed by the system: a lexical item with no idiosyncratic phonology or semantics, just syntactic category variables

---

[20] The existence of defective lexical items lacking structures in some level is not surprising in a correspondence architecture. Jackendoff (1997, pg. 94) notes that there are words with phonology, syntax and no meaning (e.g. expletives), others with meaning, phonology and no syntax (*hello, ouch, yes*) and even sequences with nothing but phonology (*e-i-e-i-o, inka-dinka-doo, tra-la-la*). All of these are clearly stored in long-term memory and recognized in the same way typical words are. Moreover, they fit into the phonotactic and stress patterns of English. This indicates that, though some of them have no syntax, they are still part of language. The only reason for excluding them from the lexicon is syntactocentrism – which is abandoned in SiSx.

arranged in a particular configuration. In this respect, SiSx deviates from variants of Construction Grammar which require every syntactic configuration to be paired with a meaning (GOLDBERG, 1995, 2006).
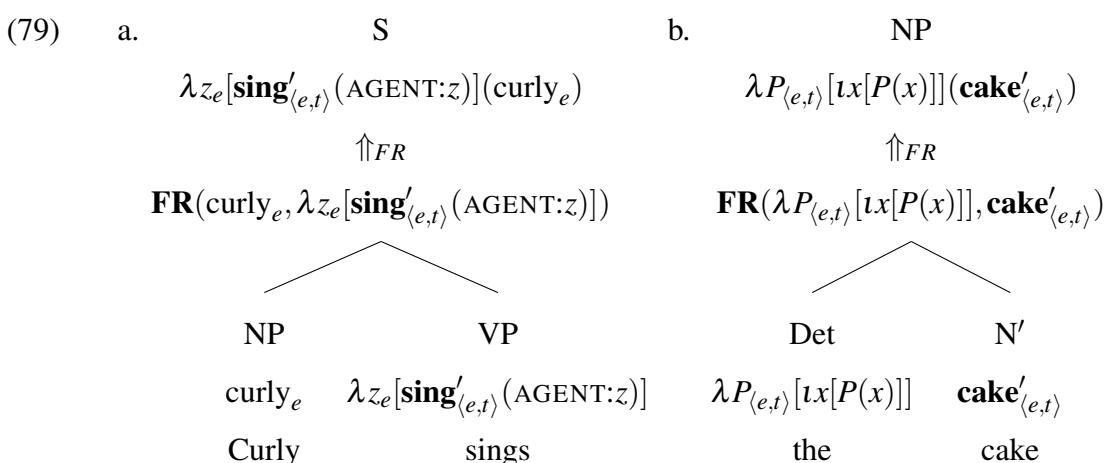
The default principle of compositional type-driven interpretation can also be represented as an abstract lexical item which licenses a maximally general correspondence between syntactic variables and meaning variables of the appropriate type. This is captured by the construction in (77). The symbol '**FR**' stands for Klein & Sag's (1985) FUNCTIONAL REALIZATION operator, whose (somewhat simplified) definition is given in (78).

(77)   COMPOSITIONALITY CONSTRUCTION

$$\begin{bmatrix} \text{SYN} & [_X X_1 \ Y_2 \ \ldots \ Z_n]_3 \\ \text{SEM} & \mathbf{FR}(\sigma_1, \sigma_2, \ldots, \sigma_n)_3 \end{bmatrix}$$

(78)   If τ is a logical type and Σ is a multiset consisting of typed Meaningful Expressions σ_1, ..., σ_n then $\mathbf{FR}_\tau(\sigma_1, ..., \sigma_n)$ denotes a set of logical expressions of type τ that are derived by exhaustively applying some σ_i to some σ_k until each member of Σ has been consumed exactly once.

Using **FR** in the SEM of (77) allows us to say that the semantics of a mother node is the result of exhaustively applying the semantics of its daughters to each other in a manner fully driven by their semantic types. Consider the schematic illustrations below, where the SEM terms corresponding to each node are written below the node's label:

(79)   a.

$$\begin{array}{c} \text{S} \\ \lambda z_e[\mathbf{sing}'_{\langle e,t\rangle}(\text{AGENT}{:}z)](\text{curly}_e) \\ \Uparrow_{FR} \\ \mathbf{FR}(\text{curly}_e, \lambda z_e[\mathbf{sing}'_{\langle e,t\rangle}(\text{AGENT}{:}z)]) \end{array}$$

NP	VP
curly_e	$\lambda z_e[\mathbf{sing}'_{\langle e,t\rangle}(\text{AGENT}{:}z)]$
Curly	sings

b.

$$\begin{array}{c} \text{NP} \\ \lambda P_{\langle e,t\rangle}[\iota x[P(x)]](\mathbf{cake}'_{\langle e,t\rangle}) \\ \Uparrow_{FR} \\ \mathbf{FR}(\lambda P_{\langle e,t\rangle}[\iota x[P(x)]], \mathbf{cake}'_{\langle e,t\rangle}) \end{array}$$

Det	N′
$\lambda P_{\langle e,t\rangle}[\iota x[P(x)]]$	$\mathbf{cake}'_{\langle e,t\rangle}$
the	cake

As we see in (79), if we have two semantic terms σ_1 and σ_2, both $\sigma_1(\sigma_2)$ and $\sigma_2(\sigma_1)$ may be instances of the more general semantic constraint $\mathbf{FR}(\sigma_1, \sigma_2)$, depending on the semantic

types of σ_1 and σ_2. A simple abstract construction like (77) thus covers all of the possible scenarios of Heim & Kratzer's (1998) Functional Application rule.[21]

In a similar vein, the principles guiding linking hierarchies – such as the one in LFG's Lexical Mapping Theory (BRESNAN; KANERVA, 1989) – can be formalized, within SiSx, as constructions that establish a correspondence between GFO and SEM variables. (80) represents the rule that says that the highest thematic argument maps to the first GF in a PRED.

(80)   LINKING CONSTRUCTION (adapted from Culicover and Jackendoff (2005, pg. 185))

$$\begin{bmatrix} \text{GF} & PRED_2{:}\langle GF_1, \dots \rangle \\ \text{SEM} & \sigma_2(\theta{:}\sigma_1, ...) \end{bmatrix}$$

Correspondences between GF positions and SYN can be stated as abstract lexical items as well. The canonical correspondence for SUBJECTS and (transitive) OBJECTS in English are (81a) and (81b), respectively:

(81)   ARGUMENT STRUCTURE CONSTRUCTIONS

a. $$\begin{bmatrix} \text{SYN} & [_S\ NP_1\ VP_2] \\ \text{GF} & PRED_2{:}\langle GF_1, \dots \rangle \end{bmatrix}$$

b. $$\begin{bmatrix} \text{SYN} & [_{VP}\ V_2\ NP_1] \\ \text{GF} & PRED_2{:}\langle GF_3, GF_1 \dots \rangle \end{bmatrix}$$

In this set-up, passivization can be seen as a more complex strategy for linking the GF-tier to SYN, as in (82) below. The same applies to relation-changing constructions in other languages (e.g. applicatives, anti-passives) (CULICOVER, 2009).

(82)   PASSIVE CONSTRUCTION (adapted from Culicover and Jackendoff (2005, pg. 203))

$$\begin{bmatrix} \text{PHON} & \varphi_1\ (by_2\ \varphi_3) \\ \text{SYN} & [_{VP}\ V[\text{PASSIVE}]_1\ \dots\ ([_{PP}\ P_2\ NP_3])] \\ \text{GF} & PRED_1{:}\langle GF_3, PRED*_1{:}\langle GF_4, \dots \rangle \rangle \end{bmatrix}$$

---

[21] This general constraint on form and interpretation does not need to be instantiated by ALL well-formed expressions in a language. Idioms, for instance, are a notable exception to compositionality, as prescribed by (77). This does not pose any theoretical difficulty for SiSx because the definition of grammaticality we adopted in Section 4.3.4 says that, for a linguistic object to be licensed, it suffices that each of its terms fully instantiate SOME constraint. This entails that a linguistic object can fail to satisfy a given constraint and still be grammatical *as long as* there is some other constraint in the grammar which it satisfies. For instance, the idiom in (72) fails to meet the compositional construction in (77). Since there is another (more specific) construction which it satisfies (the SOUND+MOTION CONSTRUCTION in (73)), SiSx predicts that (72) is well-formed.

The construction in (82) looks very much like a non-derivational version of the Relational Grammar account of passivization (PERLMUTTER; POSTAL, 1977). It expresses two fundamental intuitions: (i) that the first GF argument of the clause (i.e. the "logical subject", GF_3) is demoted to an optional *by*-phrase (without disrupting the link between this GF and its θ-role, as defined by (80)); and (ii) that the second GF argument (namely, GF_4) gets mapped to SYN like a typical SUBJECT would in virtue of (81a). This last result is accomplished by placing GF_4 (i.e. the "logical object") within an embedded PRED list, also linked to the verb in SYN – a similar device is also employed in the HPSG account of passivization proposed by Manning and Sag (1998).[22] The fact that this second PRED is distinct from the PRED which contains it (albeit both are linked to *V*-PASSIVE) is signaled by placing an asterisk after the latter. A concrete example of a linguistic object which instantiates (82) is given in (83):

(83)
$$
\begin{bmatrix}
\text{PHON} & \text{The}\prec\text{cake}_1 \text{ was}\prec\text{eaten}_2 \text{ by}_3 \text{ Hector}_4 \\
\text{SYN} & [_S \text{ NP}_1 \, [_{VP} \text{ V[PASSIVE]}_2 \, [_{PP} \text{ P}_3 \text{ NP}_4]]] \\
\text{GF} & \text{PRED}_2{:}\langle GF_4{:}[\textit{1st, sing, masc}], \text{PRED*}_2{:}\langle GF_1{:}[\textit{1st,sing,neut}]\rangle\rangle \\
\text{SEM} & \mathbf{eat}'_2(\text{AGENT:hector}_4, \text{THEME:}\iota x[\mathbf{cake}'(x)]_1)
\end{bmatrix}
$$

The PASSIVE CONSTRUCTION applies in the same way if the passivized NP is an oblique object (e.g. Bobby was relied on by his friends). It can also apply to raising to object predicates like *expect*. In these cases, the logical object which gets promoted to subject position is also the subject of an infinitival embedded predicate. An example is given in (84) (cf. (35)):

(84)
$$
\begin{bmatrix}
\text{PHON} & \text{Bobby}_1 \text{ was}\prec\text{expected}_2 \text{ to}\prec\text{sing}_3 \text{ by}_4 \text{ Joanne}_5 \\
\text{SYN} & [_S \text{ NP}_1 \, [_{VP} \text{ V[PASSIVE]}_2 \text{ VP[INF]}_3 \, [_{PP} \text{ P}_4 \text{ NP}_5]]] \\
\text{GF} & \text{PRED}_2{:}\langle GF_5{:}[\textit{1st, sing, fem}], \text{PRED*}_2{:}\langle GF_1{:}[\textit{1st,sing,masc}]\rangle\rangle \\
 & \text{PRED}_3{:}\langle GF_1\rangle \\
\text{SEM} & \mathbf{expect}'_2(\text{EXPERIENCER:joanne}_5, \text{THEME:}\mathbf{sing}'(\text{AGENT:bobby}_1)_3)
\end{bmatrix}
$$

The present treatment requires passive verbs to be linked to simultaneously two distinct

---

[22] This also happens to be the main technical reason why GFs in SiSx are unlabeled. If GFs were defined in terms of substantive roles (e.g. SUBJ, OBJ), as in LFG, a constructional account of relation-changing rules like PASSIVE would involve replacing one function name by another, in violation of monotonicity (BRESNAN, 2001). LFG and HPSG avoid this problem by stating PASSIVE as a LEXICAL RULE (BRESNAN, 1982b; POLLARD; SAG, 1987). For evidence that lexical accounts of argument structure (like the one found in LFG and HPSG) are superior to the SiSx constructional account sketched here, see Müller (2020). For a lexical account of PASSIVE in SiSx (which resembles the LFG/HPSG one), see Culicover (2021).

PRED lists, one of which embedded within the other. Consider what this implies for the analysis of the interaction between passive and raising in an example like *The cake seems to have been eaten by Hector*, whose simplified structure is depicted in (85):

$$
(85) \quad
\begin{bmatrix}
\text{PHON} & \text{The} \prec \text{cake}_1 \text{ seems}_2 \text{ to} \prec \text{have} \prec \text{been} \prec \text{eaten}_3 \text{ by}_4 \text{ Hector}_5 \\
\text{SYN} & [_S \text{ NP}_1 \ [_{VP} \text{ V}_2 \ [_{VP} \text{ V[PASSIVE]}_3 \ [_{PP} \text{ P}_4 \text{ NP}_5]]]] \\
\text{GF} & \text{PRED}_2{:}\langle \text{GF}_1{:}[\textit{1st,sing,neut}] \rangle \\
& \text{PRED}_3{:}\langle \text{GF}_5{:}[\textit{1st, sing, masc}], \text{PRED*}_3{:}\langle \text{GF}_1 \rangle \rangle \\
\text{SEM} & \mathbf{seem}'_2(\mathbf{eat}'_3(\text{AGENT:hector}_5, \text{THEME:}\iota x[\mathbf{cake}'(x)]_1))
\end{bmatrix}
$$

In principle, raising could target the first element in either one of PREDs linked to the passive variant of *eat* (i.e. GF_5 of PRED_3 or GF_1 of PRED*_3). The reason why only GF_1 can in fact be chosen to be realized as the matrix subject is because the position of GF_5 in SYN is independently stipulated to be adjacent to *by* by the PASSIVE CONSTRUCTION itself. This is why a structure like **Hector seems the cake to have been eaten by* is ungrammatical.

SiSx's rule-like lexical entries can play two roles in the grammar: a GENERATIVE ROLE, where they are used to license novel structures like (83)-(85); and a RELATIONAL ROLE, where they function in way reminiscent of nodes in an inheritance hierarchy, "lending" their structure to other independently stored items (JACKENDOFF; AUDRING, 2020).

The generative role of lexical entries can be defined in terms of the MODELING relation that holds between a linguistic expression and a lexically stored constraint (or a set thereof): e.g. the construct in (83) is a MODEL of (i.e. SATISFIES) the phrasal constraint on passive structures in (82). The relational role of lexical entries can be defined in terms of ENTAILMENT between separate constraints stored in the lexicon. A lexical entry α entails an entry β *iff* every linguistic object which is a model of α is a model of β. When a specific lexical entry α entails a more general entry β we can say that α includes β. In this sense, the *kick the bucket* idiom in (70) includes the more general VP construction in (76), which, in turn, includes a more abstract X-bar schema $[_{XP} \text{ X YP}]$ (JACKENDOFF, 1977).

Likewise, if particular passive or past tense verbs happen to be overtly stored due to high frequency, they will include the past tense and passive schemas in (69) and (82). These relational links can be represented in a relational hierarchy, where the more dominated nodes entail the less dominated ones. SiSx assumes that, other things being equal, a lexical item with relational links should be easier to store and learn than one without such links (JACKENDOFF, 1975). A

fundamental assumption in SiSx is that all lexical entries that play a generative role can be can also play a relational role, because any expression at all can, in principle, be stored as a lexeme if it is used frequently enough (JACKENDOFF; AUDRING, 2020).[23]

There is an obvious connection between this relational function of lexical entries and the hierarchy of signs in HPSG (SAG, 1997, 2012). These devices all do the work of lexical rules in earlier approaches going back to Chomsky (1970). But there is a difference: since many of SiSx's abstract entries can *also* be used generatively, unmarked lexical properties (e.g. regular morphology, valence) can, *in principle*, be kept out of individual lexemes. There is no need to list separately the active, passive and regular past tense forms for *all* verbs. These forms can be rendered well-formed by the interaction of abstract items like (81b), (82) and (69) (respectively) (CULICOVER; JACKENDOFF, 2005, pg. 188). In HPSG terms, it is as if schemas like (81b), (82) and (69) were, at once, information that can be inherited by particular lexically stored types of signs and rules to license novel signs that are not in the lexicon.

The SiSx view, is, in sum, that rules of grammar are lexical items. There is a *continuum* from stereotypical words, which specify fully linked phonology, syntax, and semantics (cf. (66)), through idioms with a few variables (cf. (71)), constructional idioms with *nothing but* variables (cf. (73)-(75)) to fully general rules (cf. (76)-(82)), from which many constructions can inherit structure. All of these things are stated in the same format: as declarative schemas, either licensing structures at a single level (e.g. (76)) or establishing correspondences between various levels (e.g. (66)). Theories which adopt a rigid lexicon/grammar distinction, must draw an artificial line somewhere in this *continuum*.

4.5   CONSTRAINTS OUTSIDE THE GRAMMAR

If language is indeed integrated into the larger ecology of the mind, it is expected that the grammatical constraints which reflect a speaker's knowledge of language are not all there is to explain the (un)acceptability of sentences. Since Miller and Chomsky (1963), the influence of EXTRA-GRAMMATICAL factors on linguistic judgments has been a major topic of investigation – one that is very much relevant within SiSx. An important consequence of this is that the notion

---

[23] Since the modeling objects in SiSx are not typed-feature structures, SiSx relational hierarchies are not literally type inheritance hierarchies, as we see in HPSG (POLLARD; SAG, 1994). They are , rather, a kind of template hierarchy of the kind employed withn LFG (DALRYMPLE; KAPLAN; KING, 2004; ASUDEH; DALRYM-PLE; TOIVONEN, 2013). Type hierarchies represent inheritance relations between linguistic objects in the model, while the template hierarchies represent entailment (or inclusion) relations between *descriptions* of linguistic objects. A template in this sense is simply a summary or an abbreviation of a set of linguistic constraints.

of GRAMMATICALITY (i.e. well-formedness in the sense of Section 4.3.4)) is kept separate from that of ACCEPTABILITY, which is understood to be a matter of *language use* (e.g. pragmatics, processing, etc.). This distinction was famously expressed by Chomsky (1965, pg. 11):

> The notion "acceptable is not to be confused with "grammatical". Acceptability is a concept that belongs to the study of performance, whereas grammaticalness belongs to the study of competence. [...] Grammaticalness is only one of many factors that interact to determine acceptability.

There are two logically possible scenarios where acceptability and grammaticality may be at odds with one another: (i) those where an expression is unacceptable but grammatical and (ii) those where an expression is acceptable but ungrammatical. The former possibility has often been explored in connection to the phenomena of constraints on unbounded dependencies. I review some of this debate in 4.5.1. The scenario of acceptable ungrammaticality is less often discussed, but it will play an important role in my account of discursive reflexives in Chapter 5-6. I give a brief introduction to this topic in 4.5.2.

4.5.1 Unacceptable grammaticality

The most hotly debated instance of unacceptable grammaticality concerns certain instances of UNBOUNDED DEPENDENCIES. The hallmark of unbounded dependencies is the presence of a GAP, by means of which a constituent in a non-canonical position (i.e. a FILLER) acquires its semantic role. I will follow HPSG (POLLARD; SAG, 1994, pg. 161) in treating gaps as simple lexical items like (86):

(86) GAP LEXEME

$$\begin{bmatrix} \text{PHON} & \varepsilon_1 \\ \text{SYN} & \text{XP[gap]}_1 \\ \text{SEM} & X_1^{gap} \end{bmatrix}$$

The PHON of a gap is the empty string (ε), its SYN can be any phrasal category marked with a [GAP] feature and its SEM is simply a gap variable of the appropriate semantic type (X_1^{gap}).[24] The effect of logically binding a gap can be reproduced by an abstract construction that establishes a correspondence between an arbitrarily long S containing an XP[GAP] – a

---

[24] A gap variable only receives a meaning if it is logically bound. In contrast, when non-gap variables are unbound, the interpretive procedure can assign to them an arbitrary denotation in accordance to their semantic type (COOPER, 1979). For instance, while the non-gap variable in **run′**(AGENT: x) can be assigned to any entity in the model, which allows its formula to receive a truth-value, the one in **run′**(AGENT: x^{gap}) cannot.

slash category S|XP, interpreted as an S which is missing an XP – and a property which results from λ-abstraction over the gap variable. This is signaled by the SEM in (87), where "/" is a term replacement function s.t., for MEs α, β of the same type, α/β is the result of replacing occurrences of free variables in α by β (see Muskens (2003) for a similar proposal):

(87)   GAP-BINDING CONSTRUCTION (adapted from Culicover (2021, chap.7))

$$\begin{bmatrix} \text{SYN} & [_{S|XP} \dots XP[\text{GAP}]_1 \dots]_2 \\ \text{SEM} & \lambda Z[\sigma(\dots X_1^{gap}/Z\dots)]_2 \end{bmatrix}$$

SiSx also needs a phrase-structure construction akin to (76) in order to license fillers in the left-periphery of clauses. (88) accomplishes this effect:

(88) $\begin{bmatrix} \text{SYN} & [_S \ XP \ S|XP] \end{bmatrix}$

Consider how this works in the simple case of topicalization in (89) (I ignore the GF-tier and the information structure status of topics). The construction in (86) licenses an empty NP as the complement of *Annie kissed*. This sequence, in turn, gets interpreted as a property (i.e. $\lambda z[\textbf{kiss}'(\text{AGENT:annie}, \text{PATIENT:}z)]$) in virtue of (87). (88) licenses a filler (i.e. *Ali*) in sentence-initial position. Due to the COMPOSITIONALITY CONSTRUCTION in (77), the property attained by (87) is applied to the semantics of the filler, yielding the right interpretation.

(89) $\begin{bmatrix} \text{PHON} & \text{Ali}_1, /\text{Annie}_2 \text{ kissed}_3 \ \varepsilon_5/_4 \\ \text{SYN} & [_S \ NP_1 \ [_{S|NP} \ NP_2 \ [_{VP}V_3 \ NP[\text{GAP}]_5]]_4] \\ \text{SEM} & \lambda z[\textbf{kiss}'(\text{AGENT:annie}_2, \text{PATIENT:}x_5^{gap}/z)]_4(\text{ali}_1) \\ & = \lambda z[\textbf{kiss}'(\text{AGENT:annie}_2, \text{PATIENT:}z)]_4(\text{ali}_1) \\ & = \textbf{kiss}'(\text{AGENT:annie}_2, \text{PATIENT:ali}_1) \end{bmatrix}$

A similar structure is ascribed to the *wh*-question in (90). I follow Culicover (2021) in positing a quantifier-like entry for the *wh*-word, as in (91).

(90)   [What [did [Laurey say Curly bought ε]]]?

(91) $\begin{bmatrix} \text{PHON} & \text{what}_1 \\ \text{SYN} & NP_1 \\ \text{SEM} & \lambda P[\textbf{WH}x(P(x))]_1 \end{bmatrix}$

The GAP-BINDING CONSTRUCTION licenses a property interpretation for the portion of (90) which excludes the *wh*-phrase (*Laurey say Curly bought ε*). This property, in turn, is fed as an argument to the **WH** quantifier (licensed in initial position by (88)), which ends up binding a variable corresponding to the gap. (92) illustrates the β-reductions in the SEM tier of (90):

(92) $\quad \lambda P[\mathbf{WH}x(P(x))](\lambda z[\mathbf{say}'(\text{AGENT:laurey}, \text{THEME:}\mathbf{buy}'(\text{AGENT:curly}, \text{THEME:}z))])$

$\quad = \mathbf{WH}x(\lambda z[\mathbf{say}'(\text{AGENT:laurey}, \text{THEME:}\mathbf{buy}'(\text{AGENT:curly}, \text{THEME:}z))](x))$

$\quad = \mathbf{WH}x(\mathbf{say}'(\text{AGENT:laurey}, \text{THEME:}\mathbf{buy}'(\text{AGENT:curly}, \text{THEME:}x)))$

The only additional machinery we need to for (intersective) relative clauses is a construction which interprets attachment of a gappy S missing an NP to an N head as property which results from the intersection between the denotations of N and S:

(93) $\quad \begin{bmatrix} \text{SYN} & [_{N'}\ N_1\ S|NP_2]_3 \\ \text{SEM} & \lambda x[\sigma_1(x) \wedge \sigma_2(x)]_3 \end{bmatrix}$

Given the property reading that (87) and (86) ascribes for the string *Annie kissed ε*, (93) is able to license the relevant part of the structure of *the man (who) Annie kissed ε* as in (94):

(94) $\quad \begin{bmatrix} \text{PHON} & /\text{man}_1\ /\text{Annie} \prec \text{kissed} \prec \varepsilon/_2/_3 \\ \text{SYN} & [_{N'}\ N_1\ S|NP_2]_3 \\ \text{SEM} & \lambda x[\lambda y[\mathbf{man}'(y)]_1(x) \wedge \lambda z[\mathbf{kiss}'(\text{AGENT:annie}, \text{PATIENT:}(z)]_2(x)]_3 \\ & = \lambda x[\mathbf{man}'(x) \wedge \mathbf{kiss}'(\text{AGENT:annie}, \text{PATIENT:}(x)] \end{bmatrix}$

The constraints in (86)-(88) and the standard principle of compositional interpretation are, therefore, all SiSx needs to model the syntactic and semantic effects of unbounded dependencies.[25] The dependency between the filler and the gap is represented as variable-binding, while a phonologically null *XP[GAP]* in SYN guarantees that the subcategorization requirements of the head that licenses the filler are locally satisfied.

However, since this mechanism assumes that gaps can be freely introduced into representations, it does not explain why sentences like (95) are bad:

---

[25] Note incidentally that the type-driven compositional rule in (77) make the presence of subject gaps in sentences like *Who sang?* unnecessary. In those cases, the **WH** quantifier can combine directly with the bare property semantics of the VP, with no need to invoke the GAP-BINDING CONSTRUCTION. There is aso no need to posit a null complementizer equivalent to *who* or *that* for zero-relatives like *the man Annie kissed* (cf. (94)): the GAP-BINDING CONSTRUCTION is sufficient to license the property interpretation for *Annie kissed ε*.

(95)   a.   *Who does that Laurie baked a pie for ε irritate Jud?

        b.   *Curly is the kind of guy who going to the party with ε would please Laurie.

It is entirely possible to derive a perfectly well-formed structure for (95) given the principles laid out so far. Most approaches to unbounded dependencies take this "overgeneration" to be a flaw and attempt to encode into the grammar restrictions that prevent gaps from occurring in contexts like (95). Ross (1967) famously took these cases to be violations of the Sentential Subject Constraint, which prohibited gaps from occurring within sentential subjects. Sentential subjects were, thus, considered to be ISLANDS with respect to extraction. Chomsky (1977, 1986a) later subsumed these and many other island effects under the principle of Subjacency.

From the point of view of SiSx, it is not clear whether it is the grammar's responsibility to embody substantive restrictions to account for unacceptable unbounded dependencies like (95). Upon closer examination, there does not seem to be a purely grammatical characterization of precisely the contexts in which filler-gap patterns are judged to be 'bad' by speakers. The explanation for most (if not all) island effects must, therefore, lie outside of the grammar, in pragmatics, discourse structure or in processing complexity. A growing body of literature points to this conclusion as well (KLUENDER, 1991,1992, 2004; KLUENDER; KUTAS, 1993; SAG; HOFMEISTER; SNIDER, 2007; HOFMEISTER et al., 2007, HOFMEISTER; SAG, 2010; HOFMEISTER; CASASANTO; SAG, 2013; , CHAVES, 2013; CHAVES; DERY, 2014,2019; CULICOVER, 2013a, 2013b). I briefly discuss a small piece of the empirical evidence against grammatical theories of islands. Space limitations prevent me from getting into the details of particular performance-based alternatives.

The suspicion that something is amiss in purely grammatical accounts of island phenomena comes from the observation that concrete proposals tend to be both too weak and too strong. Any reasonable set of constraints we can imagine will be too weak because it is likely to fall short of an explanation for data like (96)-(97), originally due to Erteschik-Shir (1973, pg. 84).

(96)   a.   What did Harry claim that marriage did ε for him?

        b.   ?What did Harry comment that marriage did ε for him?

        c.   *What did Harry transcribe that marriage did ε for him?

(97)   a.   What did Will say that he saw ε in Kansas City?

        b.   ?What did Will shout that he saw ε in Kansas City?

        c.   *What did Will eulogize that he saw ε in Kansas City?

If we posit a grammatical condition to disallow gaps within *that*-complements in general, then (96a) and (97a) would be incorrectly excluded. A grammatical constraint which freely allows gaps, on the other hand, licenses all of the structures above. This calls for an account of the gradient decrease in acceptability within each example.

The latter is hard to provide in purely syntactic terms. It is, of course, possible to assign different syntactic categories to the complements of *claim* and *say*, on the one hand, and *transcribe* and *eulogize*, on the other. This can be done by positing a syntactic feature – call it [BRIDGE] – which has a negative value for complements of *transcribe* and *eulogize* ([-BRIDGE]) and a positive one for *claim* and *say* ([+BRIDGE]). We could then state a principle that impedes gaps from occurring within [-BRIDGE] complements, ruling out (96c) and (97c). But this move would be nothing more than a stipulation that invokes a suspicious syntactic feature, not used anywhere else in the grammar.[26] Moreover, it is not clear how it could extend to (96b) and (97b), which are slightly better than (96c) and (97c) but worse than (96a) and (97a).

The ultimate explanation for the facts in (96)-(97) is plausibly related to a complex interaction between semantic, pragmatic and frequency factors (RICHTER; CHAVES, 2020). For instance, simple verbs of saying (e.g. *say* and *claim*) impose less restrictions on extraction than verbs that specify MANNER of speaking (e.g. *shout* and *eulogize*) (ERTESCHIK-SHIR, 2006). More frequent verbs (e.g. *claim* and *comment*) also impose less restrictions on extraction than less frequent ones (e.g. *transcribe*). Wherever the proper balance between these factors lie, the reason why acceptability decreases from a to the c-cases in (96)-(97) is simply not statable in terms of any reasonable grammatical generalization.

There are also cases in which grammatical principles that purport to account for island phenomena are too strong – i.e. they exclude sentences that are actually acceptable. I observed above that Ross (1967) and Chomsky (1977, 1986a) posited principles to prevent gaps from occurring within sentential subjects in order to explain (95). There are, however, examples analogous to these which are, under suitable conditions, judged to be reasonably acceptable by most speakers (KLUENDER, 2004; SAG; HOFMEISTER; SNIDER, 2007; CHAVES, 2013; CULICOVER; WINKLER, 2018). The examples below repeat (95) along with their more acceptable counterparts:

---

[26] As Culicover (p.c.) points out, in order to implement this syntactic solution, one would have to assume that [BRIDGE] is an innate feature which is initially specified for a negative value. Positive evidence, such as (96a) and (97a), would allow some complements to be specified as [+BRIDGE]. The idea that [BRIDGE] is innate is highly implausible in light of recent concerns about evolvability (CHOMSKY, 2005; JACKENDOFF, 2011a).

(98)   a.   *Who does [that Laurie baked a pie for ε] irritate Jud?

          b.   Who would [baking a pie for ε] irritate you?

(99)   a.   *Curly is the kind of guy who [going to the party with ε] would please Laurie.

          b.   Curly is the kind of guy who [going to the party with ε] would be a treat.

The contrasts above are plausibly due to a difference in the amount of discourse referential processing (GIBSON, 2000; KLUENDER, 2004; CULICOVER; WINKLER, 2018). Kluender (2004) argues that since the subject in (98a) is a finite clause, there is an implicit reference to a temporal event. This reference is absent for the non-finite form in (98b), which makes the sentence in question less complex in processing terms.

This kind of temporal reference is absent from both of the examples in (99). However, Culicover and Winkler (2018) propose that (99a) is more complex, in referential terms, than (99b) due to the presence of the additional referring expression *Laurie*. This extra-element – which Culicover and Winkler (2018) call the 'Uninvited Guest' – increases complexity in referential processing, thereby reducing acceptability. This observation also explains why the presence of another gap in the sentence (i.e. a PARASITIC GAP) makes many Subject Island violations more acceptable:

(100)   a.   *Who would [us talking to ε] bother Jud?

            b.   Who would [us talking to ε] bother ε?

(101)   a.   *This is a list of professors who [inviting ε to a party] would offend the dean.

            b.   This is a list of professors who [inviting ε to a party] would offend ε.

Sentential Subjects are not the only kinds of island environments for which this kind of pragmatic amelioration is found. Another famous example are NPs which contain an S as a subconstituent. As Ross (1967) famously observed, gaps occurring within such COMPLEX NPs are generally bad, as (102) illustrates:

(102)   *Who did Amy hear [the claim that Bobby is dating ε]?

However, a grammatical condition that simply excludes these configurations in principle (e.g. Ross' COMPLEX NP CONSTRAINT) is likely to rule out some acceptable filler-gap configurations as well (CULICOVER; JACKENDOFF, 2005, pg. 336):

(103)   Who did Amy make [the claim that Bobby is dating ε]?

The example in (103) is syntactically indistinguishable from (102). Nonetheless, it is perfectly acceptable. In order to explain this contrast, Culicover and Jackendoff (2005) suggest an explanation along the lines of Kroch (1998): (102) presupposes the existence of *the claim* while (103) doesn't. The unacceptability of (102) follows from a general principle of discourse which says that a gap cannot be interpretively dependent on an operator if its reference is part of a presupposition in the common ground. This principle extends to contrasts like (104), which are also hard to account for in purely syntactic terms.

(104)   a.   *Who did he buy that picture of ε?   (presupposes there is a picture)

       b.   Who did he buy a picture of ε?   (no presupposition)

The debate on whether all island constraints reduce to extra-grammatical factors is still very much ongoing (see Newmeyer (2016) for a survey). What this section meant to illustrate is that the SiSx view – which might seem too unconstrained at first glance – could turn out to be just what the data requires. If there is no grammatically coherent characterization of when unbounded dependencies are unacceptable, then island constraints should not be built into the rules that license these dependencies (in SiSx terms, they should not be registered as conditions on the GAP-BINDING CONSTRUCTION). On this view, sentences that incur in island violations are not technically ungrammatical, but merely unacceptable for performance-related reasons.[27]

4.5.2 Acceptable ungrammaticality and paragrammaticality

The other logically possible case of mismatch between grammaticality and acceptability occurs when a structure which is not licensed by the grammar is, nonetheless, judged acceptable by speakers in virtue of performance factors or paragrammatical conventions concerning language use (SEARLE, 1975; MORGAN, 1978). I use the general label ACCEPTABLE UN-GRAMMATICALITY for such cases, following previous work (EMONDS, 1970; OTERO, 1972; BACH; HARNISH, 1979; SOBIN, 1994; GIBSON; THOMAS, 1999; CASASANTO; SAG,

---

[27] Extra-grammatical accounts of island constraints have a long history in SiSx. They go as far back as Jackendoff and Culicover (1972). In this early paper, the authors propose that "perceptual strategy constraints on acceptability" explain otherwise puzzling contrasts like (i):

(i)   a.   Who did John give a book to ε?
      b.   *Who did John give ε a book?

Jackendoff and Culicover (1972) argue that this contrast is also difficult to capture in grammatical terms. Their idea is that the identification of gaps must be appropriately triggered by something in the linear string encountered by the processor. Since there nothing in the string in (ib) that signals that there is a gap adjacent to the verb (because *give a book* is a plausible string without a gap), the structure is judged to be unacceptable.

2008; LEIVADA; WESTERGAARD, 2020). A famous example is the following line from the 1973 Wings song "Live and Let Die", where the preposition *in* appears one too many times:

(105)   "But if this ever-changing world **in** which we live **in** makes you give in and cry . . . "

The question of when to posit acceptable ungrammaticality is delicate, as the notion can easily be interpreted as an *ad hoc* attempt to save a failed grammatical hypothesis: given that a grammar does not predict a structure α, but α nonetheless occurs, it is all too easy to attribute α's existence to some extraneous factor, keeping one's grammar immune to criticism. That such accounts carry a heavy burden is recognized by Pullum (1985, pg. 239):

> To postulate ungrammaticality for a class of examples that are naturally and immediately interpreted in a consistent way by all speakers of the language [. . .] is to make a highly controversial claim, one whose support requires detailed empirical and theoretical reasoning to show that no possible grammatical account of the phenomenon in question could be formulated within known grammatical theories.

However, if one assumes that speakers' linguistic competence can be described by a grammar and that this competence interacts with other cognitive systems to yield language-related behaviors, one can never know *a priori* whether the fact that some sentence is judged as acceptable is to be explained in terms of grammar or extra-grammatical factors.[28] Aside from being a corollary of the competence/performance distinction, the existence of ungrammatical acceptability can also be empirically established without much controversy for data like (105).

Other clear examples are discussed under the rubric of GRAMMATICALITY ILLUSIONS in the psycholinguistic literature (FRAZIER, 1985; PHILLIPS; WAGERS; LAU, 2011; CHRISTENSEN, 2016). The most famous of these is perhaps the COMPARATIVE ILLUSION illustrated in (106) (MONTALBETTI, 1984; WELLWOOD et al., 2018):

(106)   More people have been to Russia than I have.

Though sentences like (106) are (initially) judged to be acceptable by most speakers, it is generally agreed that they have no well-formed meaning. The ungrammaticality stems from the fact that the matrix subject *more people* presumes a comparison between two sets of entities,

---

[28] On conceptual grounds, instances of acceptable ungrammaticality should be no more controversial than those of unacceptable grammaticality. I suspect that this is not how most of the literature sees the issue: the idea of unacceptable grammaticality seems to be much more widespread. The reason for this, I believe, is mostly psychological. Saying that a structure which is widely used is ungrammatical evokes an unpleasant reminiscence of the stubborn school grammarian dictating rules of "proper language".

but the subordinate comparative clause provides no corresponding comparison set (PHILLIPS; WAGERS; LAU, 2011). Contrast (106) with the well-formed comparative in (107):

(107)   More people have been to Russia than elephants have.

Another plausible set of acceptable ungrammatical expressions are subjectless declarative clauses. It is widely known that English generally requires subjects in declarative finite clauses. However, as Valian (1991) notes, in specific contexts, subjects can tolerably be omitted:

(108)   a.   Seems to be interesting.

          b.   Hate it.

          c.   Want to go to lunch now?

The problem is that it is hard to characterize in purely grammatical terms (i.e. in terms of the different kinds of linguistic structures and their correspondences) the contexts in which declarative finite clauses in English can appear without subjects. Simply changing the verbs in (108), for instance, substantially reduces acceptability, as (109) shows:

(109)   a. ??Appears to be interesting.

          b. ??Loathe it.

          c. ??Willing to go to lunch now?

Since we cannot distinguish (108) from (109) in terms of sensible generalizations over PHON, SYN, GF, SEM and their correspondences, it is better to regard the both (108) and (109) as ungrammatical and explain the acceptability of (108) by some other means.

There are different accounts of what exactly makes (105), (106) and (108) acceptable despite the fact that they are not licensed by reasonable formulations of the grammar of English. Extra-grammatical factors such as frequency, 'good-enough' processing heuristics and coherence are likely to play a role (FERREIRA; PATSON, 2007; CASASANTO; SAG, 2008; LEIVADA; WESTERGAARD, 2020). What is common to all of these cases, however, is that the ungrammaticality of each example is itself unrelated to the reasons that make it acceptable. Acceptability comes from an external source, which has nothing to do with ungrammaticality.

In Chapter 6, I examine two scenarios where, somewhat paradoxically, ungrammaticality actually plays a role in explaining *why* a structure is acceptable. One of these involves passivization out of predicative PPs (e.g. *This bed has been slept in*, *This spoon has been eaten*

with, etc.). The other involves discursive occurrences of English reflexives (e.g. *That picture of himself made Bobby cry*, *Playwrights like himself always annoy Charlie*). I will argue that the failure to be licensed by the grammar is a crucial part of what makes these kinds of structures acceptable (albeit marked). The basic idea is that ungrammaticality, in these situations, is systematically interpreted as a form of MARKEDNESS, thereby driving hearers to infer a MANNER IMPLICATURE for these structures (GRICE, 1989; LEVINSON, 2000; RETT, 2020).

Another phenomenon where acceptability and ungrammaticality are intimately connected concerns the occurrence of pre-verbal *please* in questions that signal requests (SADOCK, 1974):

(110)   Can you please pass the salt?

Without *please*, the sentence in (110) could be used as a literal question about the hearer's physical abilities. With *please*, however, (110) can only be used as a request. Note that *please* is generally unacceptable in literal questions, but acceptable when used in imperatives:

(111)   a.   *Why do you please pass the salt?
          b.   Please pass the salt!

Stating the conditions under which *please* can be used in questions seems, therefore, to require crucial reference to speakers' intentions. We need to say that *please* is only acceptable in the pre-verbal position of a question when the speaker intends the question to be interpreted as an indirect request (i.e. like an imperative is typically interpreted). This, however, is not something that can be naturally represented as a grammatical constraint in SiSx.[29]

The simplest account of these facts, which is the one proposed by Bach and Harnish (1979, pg. 198-202), involves saying that (110) and (111a) are both ungrammatical, but that (110) becomes acceptable for pragmatic reasons. The idea is that the grammatical constraints simply *do not license* a pre-verbal *please* in questions under any circumstance. However, the fact that *please* is used in (110) can signal that the speaker intends to be performing a non-literal request and is not merely asking a question about the hearer's abilities. In other words, the very

---

[29]   An alternative, which was adopted by Sadock (1974), is to claim that (110) is ambiguous between an interrogative and an imperative deep structure and that *please* can only occur within the latter. A SiSx variant of this approach would be to claim that questions can optionally be interpreted as containing an imperative operator in SEM and that *please* is constructionally licensed whenever this operator is present. A problem with this view, which is a version of Ross' (1970) Performative Hypothesis, is that it misses the fact that (110) is merely an *indirect* request, and not *literally* imperative. According to this kind of reasoning, even clear declarative clauses like *I'd like you to please pass the salt* would need have an underlying imperative structure (in SYN or SEM), given the presence of *please* (BACH; HARNISH, 1979; LEVINSON, 1983).

fact that the speaker violates a grammatical rule (i.e. the rule restricting the pre-verbal *please* to imperatives) contributes to making the request interpretation for (110) acceptable.

Since the violations of grammar implied by such structures instantiate standing linguistic patterns, it is convenient to refer to them separately under the rubric of PARAGRAMMATICAL-ITY, borrowing the term from Bach and Harnish (1979). The characteristic feature of paragrammatical uses of language is the existence of a stable pattern of interpretation which is not (yet) incorporated into the grammar: e.g. speakers systematically employ pre-verbal *please* in questions to convey indirect requests, despite the fact that this usage is not grammatically licensed. As we will see in Chapter 6, something similar applies to some prepositional passives and to the use of non-local reflexives to communicate point-of-view.[30]

The explanation for these patterns is not given in terms of the grammar (i.e. the patterns are not licensed by a construction) – this is what makes them UNGRAMMATICAL, in the technical sense defined in Section 4.3.4. Rather, paragrammatical patterns typically emerge due to pragmatic pressures that motivate an inference to some kind of non-truth-conditional meaning – usually a generalized implicature, in the sense of Levinson (1983, 2000).

This does not mean that speakers have to perform an inferential process on-line whenever they encounter a paragrammatical usage: the existence of a pragmatically motivated inference merely explains *why* the pattern exists and not *how* it arises for any given example. On many occasions, these pragmatic inferences may be SHORT-CIRCUITED by usage conventions (MORGAN, 1978). This is plausibly the case in the association between questions and indirect requests that accounts for the acceptability of pre-verbal *please* in (110), as well as in the prepositional passive and non-local reflexive cases we examine in Chapter 6.

The presence of a convention does not change the fact that the form-meaning correspondences in paragrammatical uses are established pragmatically and are not grammatically licensed. As Searle (1975) and Morgan (1978) argue, not all conventions that govern linguistic behavior need to be factored in as part of a grammar. Knowledge of language may include more than what we can state in terms of a context-independent system of constraints.

As with the unacceptable grammatical data discussed in 4.5.1, it is always possible, in principle, to provide a grammar-internal explanation for these cases, thereby eliminating the mismatch between grammaticality and acceptability. The problem with such explanations is that

---

[30] Violations of grammar implied by paragrammatical structures are, therefore, typically not caused by the linguistic forms themselves, but by the failure to instantiate the canonical correspondences between form and meaning. This is why expressions that are phonologically or syntactically ill-formed (e.g. *btkph llsdtkea* or *Book the the yes run*) are never associated with acceptable paragrammatical uses.

they need to resort to *ad hoc* extensions of the vocabulary for stating grammatical constraints: one needs the grammar to refer to frequency or familiarity to explain the acceptability of (105) and to speaker intentions to account for the distribution of *please* in (110)/(111a). I assume that such extensions are illegitimate, insofar as they place within the descriptive language with which we state constraints notions that do not pertain to PHON, SYN, GFs or SEM.[31]

For any of the cases examined in this section, it is also possible to make the grammar more permissive, making all of the structures in question *grammatical*. We could, for instance, regard *both* (108) and (109) well-formed and explain the anomaly of (109) by some other means. The decision between this alternative and what I suggested above hinges on whether we take (108) or (109) to be the marked exception to the norm. Treating (109) as marked, as opposed to (108), implies rejecting the widely accepted view that English is a not null subject language.

Likewise for the other examples. If we stated a grammar that licenses (110), we would have to devise an extra-grammatical explanation for why speakers judge negatively most other instances of pre-verbal *please* in questions (cf. (111a)). This misses the fact that (110) is indeed the *marked* case – i.e. unlike questions in general, it conveys a non-literal request meaning.

The notion of acceptable ungrammaticality provides us, therefore, with an interesting way of approaching markedness. Rather than allowing more structures to be grammatical and explaining markedness extra-grammatically (as I proposed for island violating structures), we can restrict the range of grammatical structures, treating markedness as a kind of ungrammatical-

---

[31] Pragmatic inferences that manage to creep their way into the truth-conditional content of utterances may also be seen as instances of acceptable ungrammaticality of the paragrammatical sort. Consider the much discussed case of scalar implicatures (LEVINSON, 1983):

(i)   Some students failed the exam.

Although the literal meaning of *some* is associated with the existential quantifier $\exists$, (i) would, under normal circumstances, implicate the stronger proposition that *not every* ($\neg\forall$) student failed the exam. This implicature derives from Grice's Maxim of Quantity: if a cooperative speaker intended to convey that *every student* failed the exam, they would have used the stronger form *every*, rather than the weaker *some*. Given that they did not do so, we can infer that, as far as the speaker is concerned, is not the case that every student failed the exam. It is by now widely recognized that this inference from $\exists$ to $\neg\forall$ is truth-conditionally relevant. We can see this by noting that the upper-bounded $\neg\forall$ reading can be embedded under the scope of propositional operators (LEVINSON, 2000; CHIERCHIA, 2004; SAUERLAND, 2004; SIMONS, 2010; RECANATI, 2010):

(ii)   Joanne believes that some students failed the exam.

A sentence like (ii), when uttered in a typical context, attributes to Joanne the belief that some but not all of the students the student failed the exam. This implies that the $\neg\forall$ inference is actually computed as part of the embedded proposition and is, therefore, in some sense, part of the meaning of (ii). Rather than saying that *some* is lexically ambiguous between $\exists$ and $\neg\forall$, we can say that $\neg\forall$, though not literally licensed by any of the constructions instantiated in (i)-(ii), is accepted within the meaning of these sentences for pragmatic reasons. In other words, the meaning the grammar licenses for sentences containing *some* does not exhaust the full meaning speakers typically accept. In this sense, the scalar $\neg\forall$ readings are ungrammatical, but acceptable.

ity which is "saved" from unacceptability by extra-grammatical factors. Both of these options are theoretically available in SiSx and there is no algorithm for deciding between them – as is typical of alternative hypotheses in science in general.

This approach also gives us a novel way to understand the process of constructionalization and constructional change (TRAUGOTT; TROUSDALE, 2013; SMIRNOVA, 2015; FLACH, 2020). Whenever speakers come up with new forms for new communicative purposes or attach new interpretations to old forms, they are creating a new pairing between form and meaning which is, at its inception, not yet licensed by their grammars. However, such pairings are only able to persist in language use if they are PARAGRAMMATICAL – i.e. if they are consistently judged as acceptable by a relevant group of speakers, at least in a restricted set of circumstances.

With repeated use, these newly established paragrammatical pairings become regularized and less dependent on grammatically unstatable pragmatic factors. This is when they become genuine constructions, fully integrated into the grammar. The process of constructionalization thus implies that new form-meaning pairings undergo at least the following stages:[32]

(112)   unacceptable and ungrammatical ⇒ paragrammatical ⇒ grammatical

In the next chapters, we will see that this overall picture of constructional change, which built upon the idea that ungrammatical structures may be judged acceptable for pragmatic reasons, is particularly useful to explain the patterns we observe in connection to discursive occurrences of reflexives in English and other languages.

4.6 CONCLUSION

In this chapter, I introduced and provided motivation for SIMPLER SYNTAX (SiSx). SiSx is a mentalistic, model-theoretic theory of grammar which posits a richly structured lexicon that includes both traditional lexical items as well as more regular and productive patterns. This is the framework within which I will state my hypotheses in subsequent chapters.

A striking characteristic of SiSx from the point of view of mainstream syntactic theory

---

[32] As I said above, the shift from unacceptable and ungrammatical to paragrammatical is generally mediated by pragmatics. To take an example from the history of English, in the early stages of the transition from volitional *willan* to future tense *will*, the purely temporal meaning was not yet grammatically licensed as part of the interpretation (BYBEE; DAHL, 1989; BYBEE; PERKINS; PAGLIUCA, 1994). Nonetheless, since desiring that an event E occurs implies that E is in the future, speakers were able to arrive at the temporal interpretation by pragmatically loosening the volitional meaning. This loosening operation, whereby the conditions of application of a predicate are relaxed, is another often discussed example of pragmatic influence on the truth-conditional content (SPERBER; WILSON, 1985; CARSTON, 1997; RECANATI, 2004, 2010).

is the absence of movement – i.e. of sequences of phrase markers related by transformations. Once we liberate ourselves from the assumptions Mainstream Generative Grammar (MGG), the notion of movement can be seen simply as *one* way to encode the fact that a token of information can be non-locally distributed (i.e. "shared") within different parts of a string without violating restriction against multi-dominance imposed by phrase-structure based systems.

As we saw throughout this chapter, there are different ways of modeling the *effect* of movement without adhering to a transformational design. In the place of A-movement, SiSx resorts to structure-sharing in the GF tier, following the spirit of LFG (BRESNAN; KAPLAN, 1982a). In the place of A′ movement, SiSx appeals to directly licensed gaps and to a set of constraints which guarantee a semantic relation between fillers and an empty position – a move that resembles the HPSG account of unbounded dependencies (POLLARD; SAG, 1994).

The general view of grammar SiSx ends up with is this: Explanations about our intuitions regarding which structures are possible divide between grammatical constraints (as recorded in the lexicon) and extra-grammatical factors (pragmatics, processing, etc.). The former tend to correlate with sharp judgments, while the latter tend to show more variability and dependence on contextual factors (CULICOVER, 2013c).

Sources of universals are mostly confined to extra-grammatical factors and to the pressure to reduce the complexity of the set of grammatical constraints (CULICOVER, 2013b). The latter implies a kind of simplicity-based evaluation metric, in the sense of Chomsky (1965). Even though anything that is statable in terms of PHON, SYN, GFs and SEM can in principle be a grammatical constraint, some kinds of constraints are simply too complex for human learners to grasp. Such constraints, therefore, tend not to appear in the world's languages. In conjunction with the extra-grammatical sources of universals, the simplicity metric ranking the complexity of constraints makes up part of what Chomsky (2005) calls THIRD FACTOR properties.

Overall, SiSx arrives at very minimalist conception of Universal Grammar (UG) – as it happens, one that conforms (in an unorthodox way) to what Baker (2008) calls the BORER-CHOMSKY CONJECTURE: the hypothesis according to which all parameters of variation among languages are attributed to individual properties of lexical items. In this respect, SiSx agrees with MGG. But the difference between SiSx and MGG is that, as discussed in Section 4.4, lexical items are highly structured and include what we usually think of as rules of grammar. The result is that most aspects of speakers' knowledge of language end up being potentially subject to variation, respecting only the limits imposed by general cognition and complexity.

5 A SIMPLER SYNTAX OF REFLEXIVES: A CLOSER LOOK AT ENGLISH

The central goal of this chapter is to argue for a grammatical theory of English reflexives that is both simple and unified. What this means, effectively, is that the interpretive properties of reflexives in English will be subsumed under a single grammatical constraint – one that, roughly, requires reflexives to be bound by higher arguments of their syntactic predicates.

Given that the literature on the topic is teeming with contentious debate, more needs to be said in favor of this simple theory. Unified approaches like the one advanced here became unfashionable sometime after Chomsky (1986b) – the last installment of the Classical Binding Theory (CBT). This happened largely because, as we saw in Chapter 1, the CBT underestimated the full diversity of reflexives in English (not to mention other languages).

The popular antidote for diversity is division. Accordingly, syntacticians sought to state rules which explicitly guaranteed a schism between the well-behaved law-abiding plain reflexives and the fancy-free anything-goes discursive occurrences. The former have to meet grammatically identifiable requirements while the latter are explicitly liberated from any such obligations (i.e. they are EXEMPT), and are, instead, used to signal empathy towards a perspective-bearer (or a set thereof, in the case of quantified NPs) in the discourse.

This idea that conditions on reflexives are essentially disjunctive – in the sense of implying reflexives are *either* subject to grammatical requirements *or* to discursive ones – is part of what I call the Two Reflexives Hypothesis (TRH). In addition to that, the TRH also claims that *grammar itself* regulates which kinds of conditions apply to which kinds of reflexives, according to their syntactic contexts. Chapters 2-3 showed that the TRH found a particularly clear and insightful mode of expression in the reflexivity-and-chains theory (RCT). Nevertheless, I argued that it has unsurmountable problems which call for a new kind of approach.

My own proposal seeks to develop and refine the account of English discursive reflexives outlined in Menuzzi (1999), deploying the framework of Simpler Syntax (SiSx) laid out in Chapter 4. The basic idea is that there is only *one* type of grammatically licensed reflexive in English. These reflexives are licensed by a construction which establishes a correspondence between syntactic predicates (PREDs) containing reflexives and reflexive λ-terms in semantics. Discursive reflexives emerge, not as a result of a specific construction, but whenever the grammatical constraint governing reflexives has no way of being felicitously fulfilled: e.g. when the PRED that contains the GF that corresponds to the reflexive is *anti-reflexive*. This concept in-

cludes the typical examples that motivate the TRH (e.g. reflexives within conjunctions, picture NPs), as well as all the cases the TRH fails to take account of.

The fact that reflexives are nonetheless used in these structures, despite not having their correspondence to semantics licensed by the grammar, triggers a non-literal interpretive strategy which is responsible for their marked logophoric readings, along the lines suggested in Menuzzi (2004). Discursive reflexives are thus treated as cases of PARAGRAMMATICALITY, insofar as their interpretation consistently fails to comply with the *only* constraint the grammar imposes on reflexives. If this approach is on the right track, we can have our cake and eat it too, as far as the problems noted with respect to the CBT and the TRH are concerned: our theory of reflexives can be both simple (as proponents of the TRH strive for) *and* unified (as the CBT was).

In section 5.1, I outline a basic framework for addressing binding constraints within SiSx. In section 5.2, I formulate the constraint which is responsible for licensing the core instances of English reflexives. In section 5.3, I give a preliminary characterization of discursive reflexives which serves as a suitable alternative to the TRH.

5.1 PRONOUNS AND BOUND VARIABLES IN SIMPLER SYNTAX

Before moving on to my treatment of reflexives, it useful to lay out what the basic Reinhartian picture regarding bound anaphora sketched in Chapter 2 looks like within Simpler Syntax (SiSx). The main innovation I propose here is that English reflexives, insofar as they include pronominals as part of their internal structure, are subject to licensing constraints on pronominals, in addition to the extra constraints associated with reflexive constructions *tout court*. As we saw in Chapter 1, this is an old idea, whose roots goes back to Chomsky (1955).

In keeping with previous assumptions, I maintain the view that pronouns of all sorts correspond to variables and that what distinguishes pronominals from anaphors is the property of referential independence. In the present context, I interpret this as meaning that only pronominals lexically correspond to *free variables* in SEM. This can modeled by the construction in (1). For the reasons mentioned in Chapter 3, variables do not come equipped with syntactic indices. Rather, I assume an interpretive procedure where pronominal variables that remain free are assigned an arbitrary value in the model, as in Cooper (1979) and Szabolcsi (1987).

(1)   PRONOMINAL CONSTRUCTION
$$\begin{bmatrix} \text{SYN} & \text{NP[PRON]}_1 \\ \text{SEM} & x_1 \end{bmatrix}$$

In terms of the underlying logic used in linguistic descriptions, I assume that complex category symbols like NP[PRON] (or for that matter, NP, N′, N^0) correspond to conjunctions of labels on nodes. So the SYN constraint in (1) is equivalent to the following logical formula: $L(y) = N \land L(y) = \text{PHRASAL} \land L(y) = \text{PRON}$, where y is a variable over nodes.

Certain types of pronominals may correspond to variables that carry presuppositions regarding the gender and the discourse roles of their putative values (COOPER, 1983; HEIM; KRATZER, 1998). These take the form of semantic definedness conditions, which I represent in curly brackets, following Schlenker (2003b).

(2)   a.

$$\begin{bmatrix} \text{PHON} & \text{I}_1 \\ \text{SYN} & \text{NP[PRON]}_1 \\ \text{GF} & \textit{PRED}: \langle \text{GF}_1: \textit{[1st, sing]} \rangle \\ \text{SEM} & z\{+\text{SPEAKER}(x) \land -\text{ADDRESSEE}(x)\}_1 \end{bmatrix}$$

   b.

$$\begin{bmatrix} \text{PHON} & \text{her}_1 \\ \text{SYN} & \text{NP[PRON]}_1 \\ \text{GF} & \textit{PRED}: \langle \textit{GF}, \ldots \text{ GF}_1: \textit{[3rd, sing, fem]} \rangle \\ \text{SEM} & x\{+\text{FEM}(x) \land -\text{SPEAKER}(x) \land -\text{ADDRESSEE}(x)\}_1 \end{bmatrix}$$
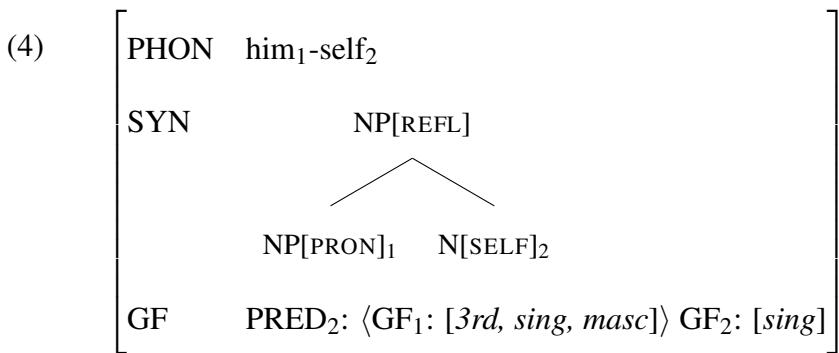
The complete SiSx structure assigned to the sentence *I kiss her* is something like the following:

(3)

$$\begin{bmatrix} \text{PHON} & \text{I}_1 \text{ kissed}_2 \text{ her}_3 \\ \text{SYN} & [_\text{S} \text{ NP[PRON]}_1 [_\text{VP} \text{ V}_2 \text{ NP[PRON]}_3]] \\ \text{GF} & \text{PRED}_2: \langle \text{GF}_1: \textit{[1st, sing]}, \text{GF}_3: \textit{[3rd, sing, fem]} \rangle \\ \text{SEM} & \textbf{kiss}'_2(\text{AGENT}: z\{+\text{SPEAKER}(x) \land -\text{ADDRESSEE}(x)\}_1, \\ & \quad \text{PATIENT}: x\{+\text{FEM}(x) \land -\text{SPEAKER}(x) \land -\text{ADDRESSEE}(x)\}_3) \end{bmatrix}$$

The assertion associated with the semantics of (3) is simply that an arbitrary individual z kissed an arbitrary individual x. However, (3) also encodes the presupposition that z is the speaker but not the addressee and that x is a female who is neither the speaker nor the addressee. If the possible values for the variables z and x do not meet these definedness conditions, the semantics of (3) will be simply undefined. For the most part, presuppositions associated with variables will be omitted in the representations to come.

I mentioned above that only pronominals can correspond to free variables. Part of what

allows some anaphors like the English reflexive *himself* to be used as free variables as well (e.g. in discursive occurrences such as *Stories about **himself** annoy **Curly***) is the fact that they contain a pronominal as a subconstituent. The partial description proposed in (4) is directly inspired by the decompositional analysis of reflexives sketched in Chapter 2. Note that the reflexive NP there is mapped to two distinct GF objects: $PRED_2$, which includes the GF of the pronoun, and GF_2, which represents the functional features of the reflexive itself.

$$(4) \quad \begin{bmatrix} \text{PHON} & \text{him}_1\text{-self}_2 \\ \\ \text{SYN} & \begin{array}{c} \text{NP[REFL]} \\ \diagdown \\ \text{NP[PRON]}_1 \quad \text{N[SELF]}_2 \end{array} \\ \\ \text{GF} & \text{PRED}_2\text{: } \langle \text{GF}_1\text{: } [\textit{3rd, sing, masc}] \rangle \text{ GF}_2\text{: } [\textit{sing}] \end{bmatrix}$$

For the sake of readability, I annotate the correspondences between SYN and GF solely on the nodes that correspond to syntactic terminals (either X^0s or non-branching XPs). However, as clarified in the previous chapter, I assume that every lexical X^0 has the same correspondences to GF-tier as all of its projections (X', X'', XP, etc.). In the case of (4), this means that both NP[REFL] and N[SELF] map to the same GF-tier objects $PRED_2$ and GF_2. This kind of double linking to the GF-tier, where a head and its projections function both as a potential arguments and as a syntactic predicate is characteristic of NPs in general.

In virtue of including a pronominal as a subconstituent, the construction for English reflexives *entails* the construction for English pronominals: whatever objects satisfy the former also satisfy the latter. A consequence of this relational link is that any constraint that mentions pronominals will also potentially apply to structures containing reflexives.

This grants us a new insight into a fundamental idea of Reinhart's (1983) seminal work. As we saw in Chapter 2, Reinhart argued that there is a general principle governing the bound-variable interpretation of pronouns as such, which was neutral to the distinction between pronominals and reflexives. In her system, this BOUND ANAPHORA CONDITION, as she called it, was encoded as part of the translation procedure in (5), which was one among the many rules specifying the correspondence between narrow syntax and semantic structures (where, for any string Φ and any NP β which occupies an A-position in Φ, Φ^β/x is the result of replacing β and all occurrences of pronouns coindexed with and c-commanded by β in Φ by x):

(5) $[_{S'} \Phi] \Rightarrow [_{S'} \beta(\lambda x.\Phi^{\beta}/x)]$

I assume that something like the bound anaphora condition in (5) is indeed operative in the grammar of English. However, instead of stating it a constraint on the bound-variable reading of pronouns of all types, I propose to restrict it to pronominals. Insofar as reflexives include pronominals as parts of their structure, they will wind up being subsumed by it. I am, therefore, treating the non-local binding characteristic of pronominals as the basic binding pattern, at least for English. The formulation below employs the term replacement function (/) in SEM (where, for any SEM objects α, β, α/β is the result of replacing free variables in α by β):

(6)   ENGLISH BOUND ANAPHORA CONSTRUCTION

$$\begin{bmatrix} \text{SYN} & [Y_3 \ldots [X_2 \ldots NP[\text{PRON}]_1]] \\ \text{GF} & PRED_3{:}\langle \ldots GF_4 \ldots \rangle \; PRED_2{:}\langle \ldots GF_1 \ldots \rangle \\ \text{SEM} & \sigma_4'(\lambda x[\sigma_3'(\ldots \sigma_4'/x \ldots \sigma_1'/x \ldots)]) \end{bmatrix}$$

The constraint in (6) says that a pronominal may be bound by a GF antecedent if the latter occupies a PRED whose corresponding SYN node c-commands the SYN node of PRED containing the pronominal. This merges the dominance-sensitivity of Reinhart's bound anaphora condition with the predicate-based anti-locality encoded by the Syntactic Condition B of Chapter 3. If we posit something like (6), we do not need separate principles governing bound anaphora (e.g. (5)) and syntactic anti-locality involving pronominals.

Since (6) comes with a built-in anti-locality, it does correctly rule out binding in (7):

(7)   a. $$\begin{bmatrix} \text{PHON} & \text{Every actress}_1 \text{ admires}_2 \text{ her}_3 \\ \text{SYN} & [_S NP_1 \, [_{VP} V_2 \, NP[\text{PRON}]_3]] \\ \text{GF} & PRED_2{:} \; \langle GF_1{:} [\textit{3rd, sing, fem}], GF_3{:} [\textit{3rd, sing, fem}]\rangle \end{bmatrix}$$

       b. $$\begin{bmatrix} \text{PHON} & \text{Every actress}_1 \text{ visualized}_2 \text{ her}_3 \text{ winning}_4 \\ \text{SYN} & [_S NP_1 \, [_{VP} V_2 \, NP[\text{PRON}]_3 \, VP_4]] \\ \text{GF} & PRED_2{:} \; \langle GF_1{:} [\textit{3rd, sing, fem}], GF_3{:} [\textit{3rd, sing, fem}]\rangle \\ & PRED_4{:} \; \langle GF_3{:} [\textit{3rd, sing, fem}]\rangle \end{bmatrix}$$

The problem with (7) is that the GFs that correspond to the pronominal and its antecedent occupy the same local PRED. The construction in (6) only licenses bound variable readings for pronominals whose antecedents are in a distinct (and hierarchically superior) PRED. This is

precisely what we see in a structure like (8). For compactness, I omit the thematic role labels on arguments and abbreviate the generalized quantifier $\lambda Q[\forall x[\mathbf{actress}'(x) \rightarrow Q(x)]]$ as $\forall(\mathbf{actress}')$:

(8)

$$
\begin{bmatrix}
\text{PHON} & \text{Every actress}_1 \text{ thinks}_2 \text{ Bobby}_3 \text{ hates}_4 \text{ her}_5 \\
\text{SYN} & [_S \text{ NP}_1 \; [_{VP} \text{ V}_2 \; [_S \text{ NP}_3 \; [_{VP} \text{ V}_4 \text{ NP}_5]]]] \\
\text{GF} & \text{PRED}_2\text{: } \langle \text{GF}_1\text{: } [\textit{3rd, sing, fem}] \rangle \\
& \text{PRED}_4\text{: } \langle \text{GF}_3\text{: } [\textit{3rd, sing, masc}], \text{GF}_5\text{: } [\textit{3rd, sing, fem}] \rangle \\
\text{SEM} & \forall(\mathbf{actress}')_1(\lambda z[\mathbf{think}'_2(\forall(\mathbf{actress}')_1/z, \mathbf{hate}'_4(\text{bobby}_3, y_5/z))]) \\
& = \forall(\mathbf{actress}')(\lambda z[\mathbf{think}'(z, \mathbf{hate}'(\text{bobby}, z))])
\end{bmatrix}
$$

The variable associated with *her* (y_5) is replaced by a variable which is bound by the antecedent *every actress*. This structure is a model of (6) because the PRED that includes GF_1 (PRED_2) corresponds to a node that c-commands the PRED that includes GF_5 (PRED_4).

The pronominals contained within English reflexives are also subject to the general bound anaphora constraint in (6). However, since the *self* N that heads the reflexive NP introduces a separate PRED that takes the GF corresponding to the pronominal as an argument (cf. (4)), (6) is trivially satisfied even in apparently local contexts like (9).

(9)

$$
\begin{bmatrix}
\text{PHON} & \text{Every actress}_1 \text{ admires}_2 \text{ her}_3\text{-self}_4 \\
\text{SYN} & \\
\end{bmatrix}
$$

$$
\begin{bmatrix}
\text{GF} & \text{PRED}_2\text{: } \langle \text{GF}_1\text{: } [\textit{3rd, sing, fem}], \text{GF}_4\text{: } [\textit{sing}] \rangle \\
& \text{PRED}_4\text{: } \langle \text{GF}_3\text{: } [\textit{3rd, sing, fem}] \rangle \\
\text{SEM} & \forall(\mathbf{actress}')_1(\lambda z[\mathbf{admire}'(\forall(\mathbf{actress}')_1/z, y_3/z)]_2) \\
& = \forall(\mathbf{actress}')(\lambda z[\mathbf{admire}'(z, z)])
\end{bmatrix}
$$

What happens here is that the PRED introduced by N[SELF] and its projections (PRED_4) allows the pronominal contained within it to satisfy the anti-locality presumption encoded in

(6). In other words, the reflexive itself creates a kind of barrier that automatically guarantees that the pronominal variable it contains will not be locally bound. The structure in (9) is, in this sense, not much different from the structure our grammar would assign to other sentences containing pronominals embedded within NPs, such as the following:

(10)   *Every actress* admires *her* mother.

In (10), it is the PRED introduced by (the projections of) *mother* that allows the binding of the pronoun within the NP to avoid violating the anti-locality imposed by (6). By positing a common structure to reflexives and possessive NPs like the one in (10), we are able to express the common diachronic tendency to develop reflexive pronouns out of pronominals attached to body-part and other kinds of inalienable possession nouns (SAFIR, 1996; DÉCHAINE; WILTSCHKO, 2017). The idea that special nominal morphemes like *self* and body-part nouns protect pronominal variables from being locally bound is also an essential part of more recent RCT accounts of how reflexivity is licensed (REULAND, 2011).

However, the typical function of the *self* morpheme in English is not only to *license* reflexivity/local binding, but also to *enforce* it. This is not captured by anything I have said above. In and of itself, (6) does not rule out non-local binding of reflexives in contexts analogous to (8). In this sense, reflexives and other kinds of possessive NPs clearly diverge:

(11)   *Every actress* thinks Bobby hates {*herself / *her* mother}.

In the next section, I propose a constraint on English reflexives that rules out non-local binding in structures like (11), while, at the same time, providing a novel way to understand how non-locally bound discursive reflexives emerge in language use.

5.2   INTRODUCING THE ENGLISH REFLEXIVE CONSTRUCTION

The central paradox theories of reflexives need to deal with is how to manage the conflict between the rigid locality enforced upon plain occurrences of reflexives and the looseness which is characteristic of discursive occurrences. In Chapter 3, I argued that theories which define the distinction between plain and discursive occurrences in terms of syntactic criteria fail to account for the full range of phenomena they purport to explain. Reinhart and Reuland (1993), for example, characterize discursive reflexives as being EXEMPT from their Reflexivity Condition A (RCA), where exemption is determined by the syntactic contexts in (12):

(12)   A reflexive is exempt from RCA (MENUZZI, 1999, pg. 160):[1]

     a.   when its predicate does not have a subject (in which case the predicate does not qualify as a syntactic one);

     b.   when the reflexive itself is not a syntactic argument of a predicate.

As we saw in Chapter 3, this syntactic characterization of exempt reflexives incorrectly excludes cases like (13)-(14), which exhibit the same crucial properties that justify the postulation of exempt reflexives in the first place (e.g. non-locality, animacy restrictions, sloppy readings in ellipsis contexts, etc.).

(13)   a.   *They* made sure that [it was clear to *themselves* that this needed to be done].

     b.   *Mary* said that [it would be very difficult for *herself* to accept your theory].

     c.   *Mary* says that [it was incumbent upon *herself* not to abandon the car].

(14)   a.   *The men* said that [the new administrators would be very much like *themselves*].

     b.   *The veterans* think that [the new recruits are more qualified than *themselves*].

     c.   "[D]oes *she* love her hubby because [he resembles *herself*]?"

The fact of the matter is that there is no obvious syntactic generalization that encompasses all and only discursive occurrences of reflexives. Any purely syntactic account is either going to be too weak, allowing too many reflexives to be discursive, or, as in the case of (12), too strong – excluding perfectly good occurrences of discursive reflexives such as (13)-(14).

In light of this conundrum, it proves convenient to let some non-structural factor determine when reflexives are governed by discursive constraints. I discuss below what this non-structural factor is likely to be, but now I want to consider two alternative ways to implement this insight. We can: (i) enrich the binding constraint governing reflexives so that it specifies non-syntactic conditions for exemption (as in Koster (1978) and Safir (1992)), or (ii) treat discursive occurrences as external to the grammar proper – as violations of the grammatical constraint on reflexives which are rendered acceptable under particular non-syntactic conditions.

The former alternative presupposes that, in addition to a grammatical constraint governing plain reflexives (e.g. something like RCA), there must be a separate non-syntactic constraint which is responsible for licensing discursive occurrences. This entails that discursive reflexives

---

[1]   All of the notions invoked by (12) have direct counterparts in terms of the GF-tier: the notion of syntactic predicate corresponds to a PRED, the notion of syntactic argument corresponds to a GF within a PRED and the notion of subject corresponds to the first GF of a PRED that is associated with a verbal or adjectival projection.

are, in principle, unmarked structures because they are grammatically licensed like plain re-flexives are. The latter alternative, in turn, implies that discursive reflexives are *always* marked with respect to plain occurrences because they are not grammatically licensed on all linguistic levels. In particular, their correspondence to SEM is, albeit stable, pragmatically rather than constructionally mediated. This makes them paragrammatical, in the sense of Chapter 4.

I opt for a version of the latter approach. Motivation for it comes from the fact that the acceptability of discursive reflexives in English is subject to contextual effects that are difficult to encode in purely grammatical terms. This situation contrasts with what we find languages where discursive anaphors are genuinely grammaticalized (e.g. Japanese and Icelandic).

As we have seen, a pervasive trait of discursive reflexives in English is their logophoric interpretation – i.e. the fact that they must refer to a perspective-bearer whom the speaker em-pathizes with (KUNO, 1987; OSHIMA, 2007; CHARNAVEL, 2019). However, the mere pres-ence of a suitable logophoric antecedent seems not to be enough to make a discursive occurrence acceptable. Zribi-Hertz (1989, pg. 771) observes that, in addition, discursive reflexives need to find their antecedents within the domain of the MINIMAL SUBJECT OF CONSCIOUSNESS, which is, roughly, the category that includes the nearest logophoric antecedent available.

(15) a. *?*Bill* said that Tom saw a picture of *himself* in the post office.

  b. *?What bothered *Bill* was that Tom had seen a picture of *himself* in the post office.

  c. *?Why is *Bobby* letting the boy spit on *himself*?

What makes cases like these bad, on Zribi Hertz's account, is the fact that the discursive reflexive is bound across a potential logophoric antecedent, outside of the domain of its minimal subject of consciousness. All of these examples can be improved if an NP denoting entity that cannot be construed as a logophoric antecedent (e.g. an inanimate) replaces the intervening NP, as in (16), adapted from Pollard and Sag (1992) and Chomsky (1973).

(16) a. *Bill* said that the Times had printed [a picture of *himself*] in the Sunday edition.

  b. What bothered *Bill* was that the Times had printed [a picture of *himself*] in the Sunday edition

  c. Why is *Bobby* letting the honey drip on *himself*?

Pollard and Sag (1992, pg. 273) note further that using a definite picture NP enhances the acceptability of discursive reflexives, even when there are intervening logophoric antecedents:

(17)   a.   *Donald* was upset. What *he* finally realized was that Jill was going to publish that picture of *himself* with Putin in the Sunday edition.

    b.   *Donald* finally realized that if Jill decides to publish that picture of *himself* with Putin in the Sunday edition of the Times, there might be some backlash.

What the contrast between (15) and (16)/(17) shows is that discursive reflexives in English are subject to a rather malleable locality constraint – one which is neither permissive to the point of always allowing long-distance binding (cf. (15)) nor rigid to the point of always forbidding it (cf. (16)-(17)). Even if we guarantee the presence of a logophoric antecedent (as I did above), the acceptability discursive reflexives in English depends on heterogeneous set of factors such as linear distance, definiteness and potentially other properties of the discourse. It is implausible one could come up with a unified grammatical characterization of these contexts that could serve as a statement of the licensing condition for discursive reflexives.

Something different happens in languages where discursive anaphors are genuinely grammaticalized. In such cases, logophoricity and a handful of purely structural constraints (often related to tense and mood) suffice to determine when discursive readings are acceptable.

Consider the case of *zibun* in Japanese and *sig* in Icelandic. Both of these anaphors have been analyzed as being logophoric (KUNO, 1987; SELLS, 1987; OSHIMA, 2004; CHARNAVEL, 2019). However, unlike what we see in English, they do not show resistance to finding their logophoric antecedents across other potential logophoric binders, as the examples in (19), taken from Anderson (1986, pg. 66-67) and Yashima (2015, pg. 137), show:

(18)   *Jón* segir [að María viti   [að Ólafur vilji   [að Billi meiði   *sig*]]].
       John says that Mary knows.SBJV that Olaf   wants.SBJV that Bill hurts.SBJV SE
       '*John* says that Mary knows Olaf wants Bill to hurt *him*.'

(19)   *Taroo*-wa [Hanako-ga [*zibun*-ga wairo-o   moratta koto]-o   Ziroo-ni tugeta
       Taro-TOP Hanako-NOM self-NOM bribe-ACC receive COMP-ACC Jiro-DAT told
       to]   omotteiru.
       COMP think
       '*Taro* thinks that Hanako told Jiro that *he* had received a bribe.'

Any individual which bears the relevant logophoric role can be an antecedent for *sig* and *zibun*, as long as stable grammatical licensing conditions are in place: Japanese *zibun* is acceptable pretty much across the board and Icelandic *sig* is acceptable as long as the path between *sig* and its antecedent does not cross an indicative clause boundary.

The acceptability of logophoric anaphors in languages where logophoricity is grammat-

icalized is, therefore, much less restricted than it is in English – i.e. it is not subject to vague factors like "minimal subject of consciousness". If English logophoric anaphors were similarly grammaticalized, we would expect them to behave in a similar way. However, examples (15)-(17) show that the acceptability English logophoric reflexives is still very much dependent on contextual factors which are hard to formalize in purely grammatical terms.

To make things more vivid, consider how the four logically possible alternatives below fare in explaining the contrast between between the widespread acceptability of logophors in Icelandic/Japanese and the more restricted acceptability of logophoric *himself* in English:

(i) Logophoric anaphors are not grammaticalized in Icelandic/Japanese but are in English.

(ii) Logophoric anaphors are not grammaticalized in either Icelandic/Japanese and English.

(iii) Logophoric anaphors are grammaticalized both in Icelandic/Japanese and English.

(iv) Logophoric anaphors are grammaticalized in Icelandic/Japanese but not in English.

We can safely discard (i) as an explanation for the contrast because it gets the data exactly backwards: what we find is that logophoric anaphors are generally *less* acceptable in English than in Icelandic and Japanese. Furthermore, it is hard to define the environments where logophoric *himself* is acceptable in terms of grammar alone. This is not the case for logophors in Japanese or Icelandic. This is why there is no reason to accept alternative (ii).

Hypothesis (iii) is what most people would be initially inclined to accept. However it faces an important challenge, which also affects hypothesis (ii): if there is no difference in how the grammar treats logophoric anaphors between the two groups of languages, we need to find some extra-grammatical explanation for why English speakers are more reluctant to accept logophoric anaphors than speakers of Icelandic and Japanese. In concrete terms, there must be some extra-grammatical story that accounts for why, among the two (allegedly grammatical) analogous structures in (20)-(21), only the former is unacceptable.

(20)   *?*John* says that Mary knows that Olaf wants Bill to destroy pictures of *himself*.

(21)   *Jón* segir [að María viti   [að Ólafur vilji   [að Billi meiði *sig*]]].
        John says that Mary knows that Olaf   wants that Bill hurts SE

It is very hard to imagine what such an account could be. It is implausible, for instance, that speakers of Icelandic and English differ in their referential processing – e.g. that English speakers have less working memory capacity and, therefore, do not tolerate binding of

logophoric *himself* across potential logophoric antecedents in the same way speakers of Icelandic do. The most plausible alternative is, therefore, some variant of (iv).

Saying that English are not fully grammaticalized amounts to saying that is no DISCURSIVE REFLEXIVE CONSTRUCTION that *licenses* English discursive reflexives in a set of environments that could be individuated in terms of the vocabulary of PHON, SYN, GFs and SEM. The only grammatical construction which is relevant for English reflexives is the one that covers the core cases of coargument reflexives.

Given the model-theoretic foundations laid out in Chapter 4, claiming that a construct *C* is not licensed by a grammar *G* (i.e. that *C* is not a model of any of the constraints included in *G*) is equivalent to saying that *C* is UNGRAMMATICAL according *G*. This, of course, does not entail that *C* is never present in language use: it may well be the case that extra-grammatical factors make *C* consistently acceptable to some or even all speakers who internalize *G*, despite *C* not being licensed by *G*. It may even be that there is an extra-grammatical usage preference or usage convention favoring *C* (SEARLE, 1975; MORGAN, 1978; BACH; HARNISH, 1979; MILLER, 2013). This is what I propose for discursive reflexives in English.

In addition to its empirical virtues, the idea that discursive reflexives are not grammatically licensed is also preferable on the grounds of theoretical parsimony. To see how, let us return, for the moment, to the vocabulary of the reflexivity-and-chains theory (RCT). It is clear that keeping discursive reflexives out of the grammar allows us to strip away some of the complications that need to be loaded into the RCA with the purpose of excluding discursive reflexives from its domain of application – not only the problematic cases in (13)-(14), but also the more usual discursive occurrences like picture NP reflexives. In other words, if we no longer need to build into RCA the adequate conditions for exemption, RCA itself could be much simplified.

This is, in essence, the argument Menuzzi (1999, pg. 173) makes for treating discursive reflexives as *tolerable violations* of the RCA. If we merge his version of RCA with Chain Visibility, along the lines I argued for in Chapter 3, we arrive at the following:

(22)   SIMPLIFIED REFLEXIVITY CONDITION A

        If R is a grammaticalized reflexive anaphor, then R is bound by a higher argument of a syntactic predicate P where R serves as a syntactic argument.

(23)   The SYNTACTIC PREDICATE formed of (a head) P is P and all of the projections that realize P's grammatical functions (i.e. P's syntactic arguments).
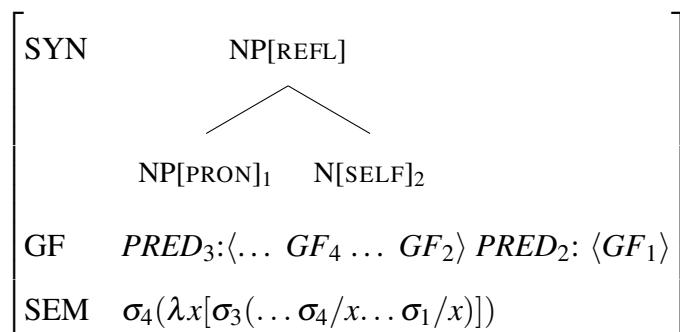
There are two main innovations here. The first is the elimination of the subject require-ment on the definition of syntactic predicates, which was present on all earlier statements of the RCA. This was previously motivated by the need to exempt reflexives within picture NP and locative PPs from the effects of RCA. If we no longer want to view discursive occurrences as exempt, the subject requirement becomes entirely unnecessary.

The second innovation is that, instead of conditioning the application of RCA to struc-tures where a reflexive is an argument of a predicate (i.e. reflexive-marked structures), (22) actually *requires* reflexives to occupy argument positions. This, along with the first innovation mentioned above, has the effect of turning the RCA into a unified constraint that potentially ap-plies to *every instance of reflexive* in a language, regardless of its context – much like Chomsky (1981) did in his formulation o Condition A of the CBT (MENUZZI, 1999, pg. 173).

Unlike the original RCA, (22) is not a condition on a restricted class of predicates (reflex-ive marked syntactic ones), but on the class of reflexive anaphors as a whole. If we assume the indexless theory of variable binding argued for in Chapter 3, (22) can be interpreted as a state-ment about the conditions in which reflexives can be translated as bound variables in semantic structure. In other words, (22) turns RCA essentially into a CONSTRUCTION.

This is where the framework of SiSx becomes particularly handy. The constructional nature of (22) can be made even clearer by phrasing (22) as an ENGLISH REFLEXIVE CON-STRUCTION, which translates the RCT parlance into SiSx analogues:

(24)   ENGLISH REFLEXIVE CONSTRUCTION (ERC)

$$
\begin{bmatrix}
\text{SYN} & \text{NP[REFL]} \\
& \diagup \diagdown \\
& \text{NP[PRON]}_1 \quad \text{N[SELF]}_2 \\
\text{GF} & PRED_3{:}\langle \dots\ GF_4\ \dots\ GF_2 \rangle\ PRED_2{:}\ \langle GF_1 \rangle \\
\text{SEM} & \sigma_4(\lambda x[\sigma_3(\dots \sigma_4/x \dots \sigma_1/x)])
\end{bmatrix}
$$

The syntactic part of ERC says that a reflexive (NP[REFL]) is composed of a pronominal (NP[PRON]) and a *self* noun (N[SELF]), which serves as its head. The GF tier encodes two pieces of information: (i) that N[SELF] and its projections introduce their own PRED, which contains the GF that corresponds to NP[PRON] as an argument (as in (4)), and (ii) that the GF that corresponds to projections of N[SELF] is the argument of a PRED that has a higher GF

(GF_4). This alone entails that there are no reflexive subjects in English (e.g. *Heself arrived). The mapping to the SEM requires, further, that the variable that corresponds to the pronominal (σ_1) be bound by the semantics of a GF that outranks the GF of the projections of N[SELF]. This makes ERC a special case of the general English Bound Anaphora Construction in (6).

We can think of N[SELF] in (24) as an operator that forces the variable that corresponds to the pronominal to be bound by a higher GF coargument of its projections (N[SELF] and NP[REFL]), as in Sauerland (2013), Patel-Grosz (2013), McKillen (2016) and Déchaine and Wiltschko (2017). Simplifying a bit, we can characterize ERC as a principle that requires reflexives to be bound by higher syntactic coarguments. This entails that the syntactic predicate where NP[REFL] is an argument should be reflexive, as in the RCA. This is clear in (24): the semantic counterpart of $PRED_3$ (σ_3) has two arguments that are bound by the same λ-operator.

ERC turns the principle governing reflexives into a language-specific lexical constraint. Since this constraint is, basically, a statement of the semantic contribution of reflexives, there is no need to regard it as part of a dedicated Binding Theory or Reflexivity module of grammar – it is simply one grammatical construction in the English lexicon, among others.

Moreover, for this same reason, the acceptability of discursive occurrences, in spite of their failure to be licensed by ERC, is no more suspicious than coercion and other kinds of standardized non-literality (BACH; HARNISH, 1979; LEVINSON, 2000). All of these phenomena can be seen as tolerable violations of lexically encoded constraints on the meanings of expressions – i.e. as special cases of acceptable ungrammaticality.

The ungrammaticality in question is not due to syntax (like the ungrammaticality of *Won boy the*), but to a failure to adequately instantiate the grammatically stipulated correspondences between SYN/GF and SEM. In this broad sense, a structure where *Can you pass the salt?* is linked to the meaning of a request (rather than to that of a question) is also ungrammatical (SEARLE, 1975; MORGAN, 1978; BACH; HARNISH, 1979). Since these ungrammatical form-meaning pairings are systematically interpreted as acceptable in certain contexts, they qualify as instances of paragrammaticality, in the sense of Chapter 4.

Before getting into the details of how ERC enters into my account of discursive reflexives, we need to make sure that the ERC has the same effect as Reinhart and Reuland's (1993) RCA for some baseline cases of coargument reflexives. Consider first the simplest possible structure, where an object reflexive is bound by the subject of an ordinary transitive verb (for simplicity, I omit the internal ϕ-feature specification of GFs in most of the following examples):

(25)
$$
\begin{array}{ll}
\text{PHON} & \text{Susan}_4 \text{ admires}_3 \text{ her}_1\text{-self}_2 \\
\text{SYN} & [_S \text{ NP}_4 \text{ } [_{VP} \text{ V}_3 \text{ } [_{NP[REFL]} \text{ NP[PRON]}_1 \text{ N[SELF]}_2]]] \\
\text{GF} & \text{PRED}_3{:}\langle \text{GF}_4, \text{GF}_2 \rangle \text{ PRED}_2{:}\langle \text{GF}_1 \rangle \\
\text{SEM} & \lambda P[P(\text{susan})]_4(\lambda x[\mathbf{admire}'_3(\lambda P[P(\text{susan})]_4/x, y_1/x)]) \\
& = \lambda P[P(\text{susan})](\lambda x[\mathbf{admire}'(x,x)])
\end{array}
$$

The construct above is correctly licensed by ERC because all of the terms and correspondences in (24) are fully instantiated in (25). In SYN, the NP[REFL] dominates NP[PRON]and N[SELF]. The GF of NP[PRON] corresponds to an argument of the PRED linked to N[SELF] and the GF linked to the latter is, in turn, an argument of the PRED linked to the verb. In the SEM level, the verbal PRED corresponds to a reflexive λ-term where the variable that corresponds to NP[PRON] is bound by (the semantics of) a higher GF coargument of the latter. In virtue of the interpretive rule for NPs, this higher GF argument corresponds, in SEM, to a generalized quantifier ($\lambda P[P(\text{susan})]$) which combines with the reflexive predicate to yield (after the β-reduction steps) $\mathbf{admire}'(\text{susan}, \text{susan})$. This gives us the correct semantics: the sentence is true *iff* the pair $\langle \text{susan, susan} \rangle$ is a member of the set defined by $\mathbf{admire}'$.

ERC also licenses binding in cases involving verb selecting two objects, as shown below:

(26)
$$
\begin{array}{ll}
\text{PHON} & \text{Susan}_4 \text{ protected}_3 \text{ Bobby}_5 \text{ from her}_1\text{-self}_2 \\
\text{SYN} & [_S \text{ NP}_4 \text{ } [_{VP} \text{ V}_3 \text{ NP}_5 \text{ } [_{NP[REFL]} \text{ NP[PRON]}_1 \text{ N[SELF]}_2]]] \\
\text{GF} & \text{PRED}_3{:}\langle \text{GF}_4, \text{GF}_5, \text{GF}_2 \rangle \text{ PRED}_2{:}\langle \text{GF}_1 \rangle \\
\text{SEM} & \lambda P[P(\text{susan})]_4(\lambda x[\mathbf{protect}'_3(\lambda P[P(\text{susan})]_4/x, \text{bobby}_5, y_1/x)]) \\
& = \lambda P[P(\text{susan})](\lambda x[\mathbf{protect}'(x, \text{bobby}, x)])
\end{array}
$$

However, the structure underlying an unacceptable binding example like ***Susan's father admires herself*** is not licensed by ERC – which is exactly what we want:

(27)
$$
\begin{array}{ll}
\text{PHON} & \text{Susan's}_4 \text{ father}_5 \text{ admires}_3 \text{ her}_1\text{-self}_2 \\
\text{SYN} & [_S \text{ } [_{NP} \text{ NP[GEN]}_4 \text{ N}_5] \text{ } [_{VP} \text{ V}_3 \text{ } [_{NP[REFL]} \text{ NP[PRON]}_1 \text{ N[SELF]}_2]]] \\
\text{GF} & \text{PRED}_3{:}\langle \text{GF}_5, \text{GF}_2 \rangle \text{ PRED}_5{:}\langle \text{GF}_4 \rangle \text{ PRED}_2{:}\langle \text{GF}_1 \rangle \\
\text{SEM} & \lambda P[P(\text{susan})]_4(\lambda x[\mathbf{admire}'_3(\iota z[\mathbf{father}'_5(z, \lambda P[P(\text{susan})]_4/x)], y_1/x)]) \\
& = \lambda P[P(\text{susan})](\lambda x[\mathbf{admire}'(\iota z[\mathbf{father}'(z,x)], x)])
\end{array}
$$

The problem with (27) is that the GF whose semantics binds the variable that corresponds to the pronominal within NP[REFL] is not a GF that appears on the same PRED as NP[REFL] itself: to put matters more simply, the reflexive and its antecedent are not coarguments of the same syntactic predicate. As a result, the syntactic predicate that ensues ($PRED_3$) is not actually reflexive – i.e. its GF arguments do not translate as variables bound by the same λ-operator.[2]

Similarly, instances of non-local binding, such as ***Susan** thinks Bobby hates **herself*** are not licensed by ERC. To see why, consider the simplified structure below:

(28)

PHON $Susan_6$ $thinks_5$ $Bobby_4$ $hates_3$ her_1-$self_2$

SYN

$$
\begin{array}{c}
S \\
\diagup \diagdown \\
NP_6 \quad VP \\
\diagup \diagdown \\
V_5 \quad S \\
\diagup \diagdown \\
NP_4 \quad VP \\
\diagup \diagdown \\
V_3 \quad NP[\text{REFL}] \\
\diagup \diagdown \\
NP[\text{PRON}]_1 \quad N[\text{SELF}]_2
\end{array}
$$

GF $PRED_5$: $\langle GF_6 \rangle$ $PRED_3$: $\langle GF_4, GF_2 \rangle$ $PRED_2$:$\langle GF_1 \rangle$

SEM $\lambda P[P(\text{susan})]_6(\lambda z[\textbf{think}'_5(\lambda P[P(\text{susan})]_6/z, \textbf{hate}'(\text{bobby}_4, y_1/z)_4)])$

$= \lambda P[P(\text{susan})](\lambda z[\textbf{think}'(z, \textbf{hate}'(\text{bobby}, z))])$

The problem with (28) is that the pronominal within the reflexive is not bound by a higher coargument of the reflexive. As a result, no reflexive predicate is formed. Note, however, that the binding of the pronominal in (28) itself does conform to the general constraint in (29). In fact, (28) is identical in structure to the well-formed pronominal binding example in (8).

---

[2]   On quite independent grounds, this example also fails to meet the compositional principle that ensures that the SEM of any branching SYN node γ is the result of applying the SEM of one of γ's daughter nodes to the others – i.e. the type-driven COMPOSITIONALITY CONSTRUCTION of Chapter 4. This means that, in (27), the SEM of *Susan* cannot reach out of the NP headed by *father* to combine with the SEM of the predicate headed by *admire*. A compositional interpretation would dictate that the SEM of *Mary* must combine with the SEM of *father* before it combines with **admire**$'$ (see Reuland (2011, chap. 2)).

(29) ENGLISH BOUND ANAPHORA CONSTRUCTION

$$\begin{bmatrix} \text{SYN} & [Y_3 \dots [X_2 \dots NP[\text{PRON}]_1]] \\ \text{GF} & PRED_3:\langle\dots GF_4\dots\rangle \; PRED_2:\langle\dots GF_1\dots\rangle \\ \text{SEM} & \sigma_4'(\lambda x[\sigma_3'(\dots \sigma_4'/x\dots \sigma_1'/x\dots)]) \end{bmatrix}$$

The specific part of (28) that violates the grammar is not the pronominal per se, but the *self* morpheme, which, in order to be licensed by ERC, enforces local coargument binding. This point will be important later on in my account of discursive occurrences.

More complex examples of constructs licensed by the ERC are given below. In each of these, structures licensed by ERC interact with those licensed by other constructions: the imperative in (30), raising to subject in (31), raising to object in (32) and subject control in (33).

(30)
$$\begin{bmatrix} \text{PHON} & \text{Defend}_3 \; \text{your}_1\text{-self}_2 \\ \text{SYN} & [_S \; V[\text{IMP}]_3 \; [_{NP[\text{REFL}]} \; NP[\text{PRON}]_1 \; N[\text{SELF}]_2]] \\ \text{GF} & PRED_3:\langle GF_4, GF_2\rangle \; PRED_2:\langle GF_1\rangle \\ \text{SEM} & \lambda P[P(\text{you})]_4(\lambda x[\mathbf{defend}_3'(\lambda P[P(\text{you})]_4/x, y_1/x)]) \\ & = \lambda P[P(\text{you})](\lambda x[\mathbf{defend}'(x,x)]) \end{bmatrix}$$

(31)
$$\begin{bmatrix} \text{PHON} & \text{Susan}_5 \; \text{seems}_4 \; \text{to}{\prec}\text{admire}_3 \; \text{her}_1\text{-self}_2 \\ \text{SYN} & [_S \; NP_5 \; V_4 \; [_{VP} \; V[\text{INF}]_3 \; [_{NP[\text{REFL}]} \; NP[\text{PRON}]_1 \; N[\text{SELF}]_2]]] \\ \text{GF} & PRED_4:\langle GF_5\rangle \; PRED_3:\langle GF_5, GF_2\rangle \; PRED_2:\langle GF_1\rangle \\ \text{SEM} & \lambda P[P(\text{susan})]_5(\lambda z[\mathbf{seem}_4'(\lambda x[\mathbf{admire}_3'(\lambda P[P(\text{susan})]_5/x, y_1/x)](z))]) \\ & = \lambda P[P(\text{susan})](\lambda z[\mathbf{seem}'(\lambda x[\mathbf{admire}'(x,x)](z))]) \\ & = \lambda P[P(\text{susan})](\lambda z[\mathbf{seem}'(\mathbf{admire}'(z,z))]) \end{bmatrix}$$

(32)
$$\begin{bmatrix} \text{PHON} & \text{Susan}_5 \; \text{expects}_4 \; \text{her}_1\text{-self}_2 \; \text{to}{\prec}\text{win}_3 \\ \text{SYN} & [_S \; NP_5 \; [_{VP} \; V_4 \; [_{NP[\text{REFL}]} \; NP[\text{PRON}]_1 \; N[\text{SELF}]_2] \; [_{VP} \; V[\text{INF}]_3]]] \\ \text{GF} & PRED_4:\langle GF_5, GF_2\rangle \; PRED_3:\langle GF_2\rangle \; PRED_2:\langle GF_1\rangle \\ \text{SEM} & \lambda P[P(\text{susan})]_5(\lambda x[\mathbf{expect}_4'(\lambda P[P(\text{susan})]_5/x, \lambda z[\mathbf{win}_3'(z)](y_1/x))]) \\ & = \lambda P[P(\text{susan})](\lambda x[\mathbf{expect}'(x, \lambda z[\mathbf{win}'(z)](x))]) \\ & = \lambda P[P(\text{susan})](\lambda x[\mathbf{expect}'(x, \mathbf{win}'(x))]) \end{bmatrix}$$

(33)

$$
\begin{bmatrix}
\text{PHON} & \text{Susan}_5 \text{ tries}_4 \text{ to}\prec\text{admire}_3 \text{ her}_1\text{-self}_2 \\[4pt]
\text{SYN} & [_S \text{ NP}_5 \ [_{VP} \text{ V}_4 \ [_{VP} \text{ V[INF]}_3 \ [_{NP[REFL]} \text{ NP[PRON]}_1 \text{ N[SELF]}_2]]]] \\[4pt]
\text{GF} & \text{PRED}_4{:}\langle\text{GF}_5\rangle \ \text{PRED}_3{:}\langle\text{GF}_6, \text{GF}_2\rangle \ \text{PRED}_2{:}\langle\text{GF}_1\rangle \\[4pt]
\text{SEM} & \lambda P[P(\text{susan})]_5(\lambda z[\textbf{try}'_4(\lambda P[P(\text{susan})]_5/z, \\[2pt]
& \quad \lambda Q[Q(z)]_6(\lambda x[\textbf{admire}'_3(\lambda Q[Q(z)]_6/x, y_1/x)]))]) \\[4pt]
& = \lambda P[P(\text{susan})](\lambda z[\textbf{try}'(z, \lambda Q[Q(z)](\lambda x[\textbf{admire}'(x,x)]))]) \\[4pt]
& = \lambda P[P(\text{susan})](\lambda z[\textbf{try}'(z, \lambda x[\textbf{admire}'(x,x)](z))]) \\[4pt]
& = \lambda P[P(\text{susan})](\lambda z[\textbf{try}'(z, \textbf{admire}'(z,z))])
\end{bmatrix}
$$

In each of these examples, ERC licenses a reflexive without requiring devices such as empty categories or movement operations in SYN. Most of this work is turned over to the GF tier. For instance, the fact that the highest ranked GF in (30) does not correspond to anything in SYN is part of the imperative construction, whose general statement is shown below:

(34)   IMPERATIVE CONSTRUCTION

$$
\begin{bmatrix}
\text{SYN} & [_S \text{ V[IMP]}_3 \dots] \\[4pt]
\text{GF} & \textit{PRED}_3{:} \ \langle \text{GF: } [2nd]_4 \dots \rangle \\[4pt]
\text{SEM} & \lambda P[P(\text{you})]_4(\sigma_3)
\end{bmatrix}
$$

Likewise, the fact that *Susan* in (31) is, at once, the highest ranked syntactic argument of *seems* and of *like* is determined by the raising construction, which is formalized in terms of structure-sharing on the GF tier, as we saw in Chapter 4. Structure sharing also occurs between the second GF of the predicate *expect* and the first GF of *win* in the raising to object structure (32).[3] Sample lexical entries of subject and object raising verbs are shown below:

---

[3]   The fact that ERC is stated as a constraint on reflexives, rather than on predicates, is what enables it to avoid some of the artificial stipulations Reinhart and Reuland (1993) had to appeal to in order to account for reflexives in raising to object structures. Since the reflexive in (32) is a syntactic argument of both *expect* and *win* (it is the object of the former and the subject of the latter), *both* of these predicates are reflexive marked (i.e. both of them contain a reflexive as an argument). The classical RCA thus requires both of them to be reflexive – a condition which is only met for the predicate headed by *expect*. In order to avoid this conclusion, Reinhart and Reuland had to reclaim an analysis of Chomsky's (1955) according to which the lower verbs in raising-to-object structures move up to form a complex syntactic predicate with the main verb, as in (i).

(i)   Susan₃ [[expects [to win]₁] t_1 herself₃].

In (i), the complex predicate is both reflexive-marked and reflexive, as required by RCA. However, under ERC (or Menuzzi's Simplified Condition A in (22), for that matter), the complex syntactic structure in (32) is not necessary. The reflexive in (32) is licensed because the pronominal within it is bound by a higher argument of one of the reflexive's syntactic predicates (the one headed by *expect*). The fact that it is not also bound by a higher argument of the other of the reflexive's syntactic predicates (the PRED headed by *win*) is irrelevant.

(35) $$\begin{bmatrix} \text{PHON} & \text{seem}_2 \\ \text{SYN} & [_{VP}\ V_2\ [_{VP}\ V[\text{INF}]_3\ \ldots]] \\ \text{GF} & \text{PRED}_2{:}\langle GF_1 \rangle\ \text{PRED}_3{:}\langle GF_1 \ldots \rangle \\ \text{SEM} & \lambda z[\mathbf{seem}'_2(\sigma_3(z))] \end{bmatrix}$$

(36) $$\begin{bmatrix} \text{PHON} & \text{expect}_2 \\ \text{SYN} & [_{VP}\ V_2\ NP_4\ [_{VP}\ V[\text{INF}]_3\ \ldots]] \\ \text{GF} & \text{PRED}_2{:}\langle GF_1, GF_4 \rangle\ \text{PRED}_3{:}\langle GF_4 \ldots \rangle \\ \text{SEM} & \lambda y[\lambda z[\mathbf{expect}'_2(z, \sigma_3(y))]] \end{bmatrix}$$

In (33), there is no structure sharing between the GFs of the controller and the controllee, but the reflexive, nonetheless, imposes a reflexive interpretation on the PRED that corresponds to *admire*. The fact that the identical variables of the predicate $\mathbf{admire}'$ ultimately (after the β-reductions) get bound by the generalized quantifier $\lambda P[P(\text{susan})]$ is determined by the particular SEM associated with subject control (BACH, 1979; DOWTY, 1985). The following is a sample lexical entry for the subject control verb *try* that accounts for the interaction with ERC:

(37) $$\begin{bmatrix} \text{PHON} & \text{try}_1 \\ \text{SYN} & [_{VP}\ V_1\ [_{VP}\ V[\text{INF}]_2\ \ldots]\ ] \\ \text{GF} & PRED_1{:}\ \langle GF_3 \rangle\ PRED_2{:}\ \langle \text{GF}_4 \ldots \rangle \\ \text{SEM} & \sigma_3(\lambda z[\mathbf{try}'_1(\sigma_3/z, \lambda Q[Q(z)]_4(\lambda x[\sigma_2(\lambda Q[Q(z)]_4/x)]))]) \\ & = \sigma_3(\lambda z[\mathbf{try}'_1(z, \lambda Q[Q(z)]_4(\lambda x[\sigma_2(x)]))]) \\ & = \sigma_3(\lambda z[\mathbf{try}'_1(z, \lambda x[\sigma_2(x)](z))]) \\ & = \sigma_3(\lambda z[\mathbf{try}'_1(z, \sigma_2(z))]) \end{bmatrix}$$

What is important is that, in all of the constructs in (30)-(33), reflexive NPs are arguments of a syntactic predicate and the variable of the pronoun they contain is bound by one of their higher coarguments. This entails that these syntactic predicates that contain NP[REFL] are also reflexive – i.e. they correspond to SEM terms where two variables are bound by the same λ-operator. All of these structures are, thus, models of a grammar which includes the ERC. This is crucially not the case for discursive reflexives, as we will see in the next section.

5.3 THE ENGLISH REFLEXIVE CONSTRUCTION AND DISCURSIVE OCCURRENCES

So far, my constructional account has the same effect as a version of Reinhart and Reuland's (1993) RCA incorporating Chain Visibility. The main empirical differences between my account and RCA only come to light when we consider discursive occurrences of reflexives. In my system, these cases all wind up being acceptable violations of ERC.

Within a theory like SiSx, the claim that discursive reflexives violate ERC is simply a short-hand way of saying that their correspondences fail to be licensed by ERC and that there is no alternative construction in the grammar that licenses them – i.e. no DISCURSIVE REFLEXIVE CONSTRUCTION. Reasons for this latter claim were provided in the previous section. What this means is that discursive occurrences are, strictly speaking, ungrammatical in English.

In order to avoid confusion, I want to reiterate that this concept of ungrammaticality is distinct from the narrower concept which is usually employed within theoretical linguistics. In a narrower sense, ungrammaticality refers solely to syntactic ill-formedness. However, in the sense I am using the term, I mean to include structures that are syntactically well-formed, but whose correspondences to SEM are not licensed by any construction in the grammar. For instance, a structure where *Rain is wet* is linked to meaning **won**$'$(AGENT: bobby) is ungrammatical in this sense (CULICOVER, 2021). In this case, however, unlike what we find for discursive occurrences, ungrammaticality correlates with unacceptability.

As we will see throughout this section, what distinguishes ungrammatical but acceptable violations of ERC from genuinely unacceptable violations (e.g. cases like ***Bobby's mother likes himself***) is the fact that they occur in contexts where ERC has no way of being felicitously fulfilled: either when the reflexive is not a GF argument of a PRED or when it's PRED is ANTI-REFLEXIVE. This generalization covers the usual cases of discursive reflexives discussed by Reinhart and Reuland (1993) as well as the outlier cases mentioned in Chapter 3.

5.3.1 Discursive reflexives as violations of ERC
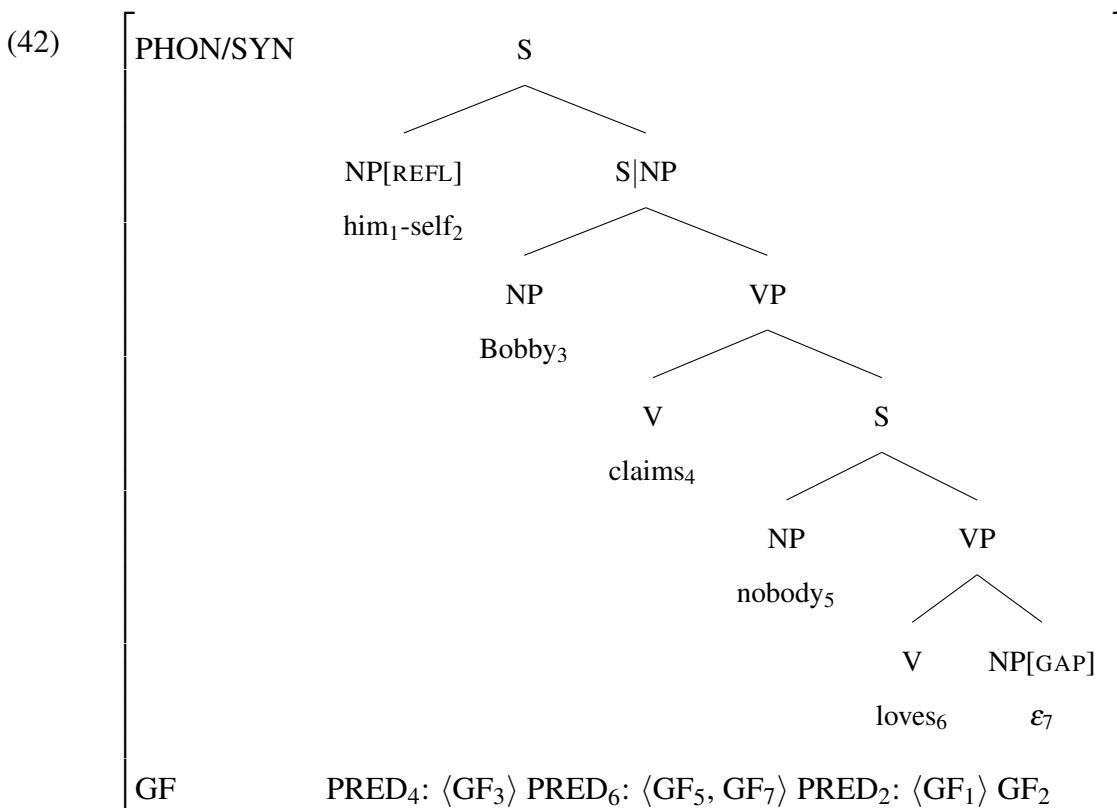
In this section, I show that all of the discursive occurrences of reflexives (i.e. those that instantiate the typical discursive properties mentioned in Chapter 3) constitute violations of ERC. Particularly, the interpretations of discursive reflexives do not match those that are dictated by ERC: their failure to be licensed lies precisely in the correspondence between their SYN/GF and SEM structures. In the next subsection, I discuss how these kinds of violations of ERC are distinguished from unacceptable violations. Some representative examples are shown below:

(38)   a.   It was *himself* that *Bobby* claimed that Susan should have married.

          b.   *Himself*, *Bobby* claims nobody loves.

(39)   a.   *Bobby* knew that Marta would never fall in love with a dandy like *himself*.

          b.   *Bobby* said that Joanne invited Susan and *himself* for a drink.

(40)   a.   *Bobby* felt that nobody appreciated that picture of *himself*.

          b.   *Bobby* heard this voice whispering near *himself*.

The reflexives in (38)-(40) are all subsumed by Reinhart and Reuland's (1993) exemption-based account of discursive occurrences: the ones in (40) are arguments of subjectless predicates (cf. (41a)) and the ones in (38)-(39) are not considered arguments at all (cf. (41b)).

(41)   A reflexive is exempt from RCA (MENUZZI, 1999, pg. 160):

          a.   when its predicate does not have a subject;

          b.   when the reflexive itself is not a syntactic argument of a predicate.

As desired, on my account, all of the examples in (38)-(40) come out as violations of ERC. The reflexives in (39) violate ERC simply because their GFs are not members of any PRED – i.e. they are not syntactic arguments. This is the same reason why they were exempt from the RCA. To see how, consider the partial simplified representation for (38b) below:

(42)

$$
\begin{bmatrix}
\text{PHON/SYN} & \\
& \\
\text{GF} & \quad \text{PRED}_4\text{: } \langle \text{GF}_3 \rangle \; \text{PRED}_6\text{: } \langle \text{GF}_5, \text{GF}_7 \rangle \; \text{PRED}_2\text{: } \langle \text{GF}_1 \rangle \; \text{GF}_2
\end{bmatrix}
$$

PHON/SYN tree:

- S
  - NP[REFL] — him$_1$-self$_2$
  - S|NP
    - NP — Bobby$_3$
    - VP
      - V — claims$_4$
      - S
        - NP — nobody$_5$
        - VP
          - V — loves$_6$
          - NP[GAP] — ε_7

It is the GF that corresponds to the gap (GF_7) that appears as the second syntactic argument of the PRED headed by *loves* ($PRED_6$). The GF that corresponds to the reflexive (GF_2) is left dangling out: i.e. it is not a member of any PRED. This type of configuration is not licensed by ERC, which requires the GFs that correspond to reflexives to be members of a PRED.

The reflexives in (39) are also violations of ERC – but for a reason that differs from Reinhart and Reuland's (41b). Though it is certainly true that the reflexives in (39) are not arguments of a *verb*, I want to argue that they are arguments of other kinds of syntactic predicates.[4]

In the case of (39a), *himself* is an argument of the predicate headed by *like*, which I take to be the same one that appears in sentences such as *Charlie is like Mary*. The reason why (39a) counts as a violation of ERC is because, even though the GF of the reflexive is a member of the PRED linked to *like*, it is not bound by the SEM of a higher coargument of this PRED. In fact, the PRED that contains the NP[REFL] as an argument is not even reflexive:

(43)
$$\begin{bmatrix} \text{PHON} & \text{dandy}_3 \text{ like}_4 \text{ him}_1\text{-self}_2 \\ \text{SYN} & [_{N'} \text{ N}_3 \text{ } [_{AP} \text{ A}_4 \text{ } [_{NP[REFL]} \text{ NP[PRON]}_1 \text{ N[SELF]}_2]]] \\ \text{GF} & GF_3 \text{ PRED}_4\text{: } \langle GF_3, GF_2 \rangle \text{ PRED}_2\text{: } \langle GF_1 \rangle \\ \text{SEM} & \lambda x[\mathbf{dandy}'_3(x) \wedge \mathbf{like}'_4(x, y_1)] \end{bmatrix}$$

In the partial representation above, the variable that corresponds to the pronominal (y_1) is free in the context of the predicate that corresponds to *like*. The pronominal within the reflexive is only bound (if at all) at the level of the PRED that corresponds to *know*,

In (39b), the reflexive is an argument of a PRED which is linked to the conjunction itself. For my narrow purposes, I assume that NP conjunctions are licensed by (44), which employs Partee and Rooth's (1983) generalized conjunction operator ⊓ defined in (45).

(44)
$$\begin{bmatrix} \text{SYN} & [_{NP} \text{ NP}_1 \text{ Conj}_3 \text{ NP}_2] \\ \text{GF} & GF_3\text{:[NUM=}plur\text{] PRED}_3\text{: } \langle GF_1, GF_2 \rangle \\ \text{SEM} & \sigma_1 \sqcap_3 \sigma_2 \end{bmatrix}$$

(45)   a.   For *p* and *q* of type *t*, $p \sqcap q =_{def} p \wedge q$

   b.   For *P* and *Q* of any conjoinable type other than *t* (where *P* and *Q* are of the same type), $P \sqcap Q =_{def} \lambda R[P(R) \sqcap Q(R)]$

---

[4]   There something odd about saying, as Reinhart and Reuland do, that the reflexives in (39) are not syntactic arguments of any predicate. In addition to empirical difficulties raised in Chapter 3, this claim seems to fly in the face of the intuition that motivates the Case Filter (CHOMSKY, 1981), which, in SiSx parlance, amounts to the requirement that all the GFs of all NPs in non-A$'$ positions be members of PRED lists.

According to (44), in addition to having their own GFs, NP conjunctions also introduce a PRED list, whose members correspond to each of the NP conjuncts. Since SiSx models agreement in terms of the GF tier, (44) also includes the requirement that the GF that corresponds to the conjunction itself is carries a plural value for its number feature.[5]

The relevant part of the structure underlying the discursive occurrence example in (39b) is shown in (46). The terms $\lambda z[\lambda Q[Q(z, \text{susan})]]_4$ and $\lambda y[\lambda S[S(y, x_1)]]$ are the type-lifted meanings that contain the semantics of *Susan* and the pronominal within the reflexive (respectively):

(46)

$$
\begin{array}{ll}
\text{PHON} & \text{Susan}_4 \text{ and}_3 \text{ him}_1\text{-self}_2 \\[4pt]
\text{SYN} & [_{NP} \text{ NP}_4 \text{ Conj}_3 \ [_{NP[\text{REFL}]} \text{ NP[PRON]}_1 \text{ N[SELF]}_2]] \\[4pt]
\text{GF} & \text{GF}_3:[plur] \text{ PRED}_3:\langle \text{GF}_4, \text{GF}_2 \rangle \\[4pt]
\text{SEM} & \lambda z[\lambda Q[Q(z, \text{susan})]]_4 \sqcap_3 \lambda y[\lambda S[S(y, x_1)]] \\[4pt]
& = \lambda w[\lambda z[\lambda Q[Q(z, \text{susan})]](w) \sqcap \lambda y[\lambda S[S(y, x)]](w)] \\[4pt]
& = \lambda w[\lambda Q[Q(w, \text{susan})] \sqcap \lambda S[S(w, x)]] \\[4pt]
& = \lambda R[\lambda w[\lambda Q[Q(w, \text{susan})](R) \sqcap \lambda S[S(w, x)](R)]] \\[4pt]
& = \lambda R[\lambda w[R(w, \text{susan}) \wedge R(w, x)]]
\end{array}
$$

Even after all of the β-reductions that ensue from applying clause (45b) of the definition of the generalized conjunction, it is still not the case that the pronominal within the reflexive (the variable x) is bound by a higher coargument of the PRED of which the reflexive is an argument (namely, PRED$_3$). The only higher coargument of PRED$_3$ is *Susan*, whose semantics does not bind the semantics of *him*. If *him* winds up being bound, this only occurs at the level of the PRED headed by *say*, which does not contain the GF that corresponds to the reflexive.

To summarize, all of the occurrences of reflexives in (38)-(39), which Reinhart and Reuland classified as exempt in virtue of not being arguments of syntactic predicates (cf. (41b)), come out as violations of ERC: none of them contain reflexives whose pronominal subparts are bound by GF higher coarguments of the reflexives themselves. This is exactly what we want.

---

[5] A positive argument for assigning GFs to NP conjuncts comes from languages that have grammatical gender agreement, such as Brazilian Portuguese (BP). In cases where the two conjuncts are grammatically feminine, the conjunction itself also triggers feminine agreement:

(i) [A   calça   e   a   bermuda] são bonit {-as   / *-os}.
     the.FEM pants.FEM and the.FEM shorts.FEM are pretty {-FEM.PLUR / -MASC.PLUR}
     'The pants and the shorts are pretty.'

If the conjunction inherits feminine features from its conjuncts, then it is necessary to represent agreement features on the conjuncts. Since SiSx represents agreement features on GFs, NP conjuncts must receive GFs.

The cases in (40) are more interesting. These examples were considered exempt in virtue of Reinhart and Reuland's (1993) definition of syntactic predicate, which, as we saw, requires the presence of subjects. Since the ERC eliminates this otherwise unmotivated stipulation the predicates in these structures now count as garden-variety syntactic predicates (i.e. as PREDs). Since the reflexives therein do realize GFs associated with the heads that select them, they count as syntactic arguments as well. However, these reflexives violate ERC because they (or, rather, the pronominals they contain) are not bound by higher GF coarguments.

In (40a), the reflexive is an argument of the syntactic predicate headed by *picture*. For the moment, we can think of this noun as having a meaning of type $\langle e, \langle e, t \rangle \rangle$ (i.e. something like $\lambda x.\lambda y.\textbf{picture-of}'(y, x)$) (PARTEE; BORSCHEV, 1998, 2003) – i.e. it denotes a relation between a representation and a depicted entity (this proposal will be refined shortly below). So the relevant part of the structure containing the reflexive is the following:

(47)

$$
\begin{bmatrix}
\text{PHON} & \text{that}_4 \text{ picture} \prec \text{of}_3 \text{ him}_1\text{-self}_2 \\
\text{SYN} & [_{\text{NP}} \text{ Det}_4 \text{ N}_3 \text{ } [_{\text{NP[REFL]}} \text{ NP[PRON]}_1 \text{ N[SELF]}_2]] \\
\text{GF} & \text{GF}_3 \text{ PRED}_3\text{: } \langle \text{GF}_2 \rangle \text{ PRED}_2\text{: } \langle \text{GF}_1 \rangle \\
\text{SEM} & = \lambda P[\iota z[P(z)]]_4(\lambda x[\textbf{picture-of}'_3(x, y_1)]) \\
& = \iota z[\textbf{picture-of}'(z, y)]]
\end{bmatrix}
$$

Note that the SEM predicate that corresponds to PRED$_3$ is not one where one coargument binds another. In fact, this predicate is not reflexive at all. This entails that the structure SiSx ascribes to (40a) does not instantiate the semantic constraint imposed by ERC.

The remaining example covered by Reinhart and Reuland's exemption account is (40b). In this case, the reflexive is an argument of the syntactic predicate headed by the locative preposition *near*, which denotes a spatial relation. This predicate also fails to meet the requirements imposed by ERC because neither of its two arguments is in a binding relation with respect to the other. Following Reinhart and Reuland's (1993, pg. 688) analysis for a similar example, I propose that, in fact, the semantics of (40b) is something like (48), where the semantic term that corresponds to the PRED linked to *near* predicates over the variable that corresponds to the pronoun (x) and an existentially bound event argument:

(48) $\lambda P[P(\text{bobby})](\lambda x[\exists e[\textbf{hear}'(x, \iota z[\textbf{voice}'(z) \wedge \textbf{whisper}'(e, z) \wedge \textbf{near}'(e, x_1)]]]])$

Now consider the counterexamples to Reinhart and Reuland's exemption account, which

involve expletive subjects and comparison predicates. These are examples like the following:

(49)   a.   *They* made sure that [it was clear to *themselves* that this needed to be done].

       b.   *Mary* said that [it would be very difficult for *herself* to accept your theory].

       c.   *Mary* says that [it was incumbent upon *herself* not to abandon the car].

(50)   a.   *The men* said that [the new administrators would be very much like *themselves*].

       b.   *The veterans* think that [the new recruits are more qualified than *themselves*].

       c.   *Joanne* loves Bobby because [Bobby resembles *herself*].

Just like the cases in (39)-(40), the reflexive NPs in (49)-(50) correspond to GFs that are members of PRED lists (i.e. they are arguments of syntactic predicates). Nonetheless, none of these PREDs meet the constraints imposed by ERC: the GF that corresponds to the reflexive is not bound by the semantics of a higher GF within the same PRED. The PREDs in question are not even reflexive: the SEM terms they are linked to do not contain variables that are bound by the same λ-operator. Therefore, all of these examples violate of ERC in much the same way as (40). (In fact, the adjective *like* is the head of a syntactic predicate in both (50a) and in (39a).) As an example, consider the simplified partial structure for (50c):

(51)
$$
\left[
\begin{array}{ll}
\text{PHON} & \text{Bobby}_4 \text{ resembles}_3 \text{ her}_1\text{-self}_2 \\
\text{SYN} & [_S \text{ NP}_4 \text{ }[_{VP} \text{ V}_3 \text{ }[_{NP[REFL]} \text{ NP[PRON]}_1 \text{ N[SELF]}_2]]] \\
\text{GF} & \text{PRED}_3{:}\langle \text{GF}_4, \text{GF}_2 \rangle \text{ PRED}_2{:}\langle \text{GF}_1 \rangle \\
\text{SEM} & \lambda P[P(\text{bobby})]_4(\lambda x[\textbf{resemble}'_3(x, y_1)])
\end{array}
\right]
$$

The projections of the reflexive correspond to a member of PRED$_3$ (as they should), but the correspondence between PRED$_3$ and SEM is not the one licensed by ERC. In fact, if (51) were licensed by ERC, its semantics would have to be (52), which involves a reflexive predicate:

(52) $\lambda P[P(\text{bobby})]_4(\lambda x[\textbf{resemble}'_3(\lambda P[P(\text{bobby})]_4/x, y_1/x)])$
      $= \lambda P[P(\text{bobby})](\lambda x[\textbf{resemble}'(x, x)])$

What we find, in sum, is that the simple statement of ERC given in section 5.2 only licenses prototypical cases where a reflexive is bound by a higher coargument, making its syntactic predicate reflexive. The different varieties of discursive occurrences all come out as violations of ERC. This explains why these occurrences are often perceived as marked by most speakers and need to be motivated by discourse considerations (MENUZZI, 2004).

5.3.2 Discursive occurrences and anti-reflexivity

As it stands, the account here states that VIOLATING ERC (rather than exemption) is a necessary condition for a reflexive to be a discursive occurrence. In order to make this work, however, we must determine what distinguishes acceptable violations of ERC in discursive occurrences (including the problematic cases in (49)–(50)) from genuinely ungrammatical violations like (53):

(53) a. *Susan*'s father admires *herself*.   (cf. the structure in (27))

  b. *Susan* thinks Bobby hates *herself*.   (cf. the structure in (28))

  c. *The actress defended himself.

This is easier for the discursive occurrences in A′ positions, which fall under Reinhart and Reuland's generalization in (12b): those cases, exemplified in (38) above, are distinguished from ungrammatical violations of ERC in virtue of the fact that the GFs of the reflexives therein are not members of any PRED list (i.e. they are not arguments of a syntactic predicate).

It is for discursive occurrences that do correspond to GF arguments of PREDs – i.e. those that fell under Reinhart and Reuland's (12a) ((39)–(40)) and the outliers (49)–(50) – that the problem becomes more stringent. I propose that, in all of these cases, the reflexive NPs are members of PREDs that are ANTI-REFLEXIVE in the sense defined in (54). This is what makes these ERC violations acceptable (for reasons I'll elaborate on below).

(54) A PRED π is ANTI-REFLEXIVE *iff* the reflexive interpretation of π is semantically uninformative: i.e. either tautologically true, contradictory, idempotent or undefined.

  (i) The INTERPRETATION of π is the semantic term π' that corresponds to π.

  (ii) A semantic term π' is REFLEXIVE *iff* π' contains at least two argument terms that are bound by the same λ-operator.

Like the similar notion of *implied non-coreference* in Safir (1992), anti-reflexivity is an interpretive property that reflects the inherent semantics of certain predicates. If a reflexive interpretation for a predicate is tautological, contradictory, idempotent or undefined, it will be either either true, false or indeterminate in all circumstances. This makes the interpretation uninformative or trivial in a sense that is solely determinable on semantic grounds, independently of factual or contextual knowledge. Whether a PRED is anti-reflexive is, therefore, a purely an-

alytic property, in the standard logical sense.[6] (Though, as we will see below, pragmatic factors enter into how anti-reflexivity gives rise to logophoric interpretations for discursive reflexives.)

Most predicates do not have the property defined in (54). It does not apply, for instance, to the ungrammatical cases in (53), as there is nothing that disfavors reflexive interpretations for the PREDs linked to *admire*, *hate* and *defend*. All of these correspond to semantic terms for which reflexivity is perfectly informative. There is absolutely nothing odd about the sentences in (55), where these PREDs receive run-of-the-mill reflexive readings:

(55)   a.   *Susan's father* admires *himself.*

          b.   *Susan* hates *herself.*

          c.   *The actress* defended *herself.*

Since the reflexives in (54) are GF arguments of the PREDs linked to *admire*, *hate* and *defend*, and since, moreover, these PREDs are *not* anti-reflexive, as (55) shows, the violations of ERC in (53) do not make for acceptable discursive occurrences.

Given that anti-reflexivity is a strictly semantic notion, the explanation given for (55) applies in exactly the same way to (56a) – despite the fact that *married* is slightly anomalous under a reflexive interpretation, as we can see in (56b).

(56)   a.   **Bobby* said that Susan never married *himself.*

          b.   ?*Bobby* married *himself*

What makes (56b) odd is not the semantics of *married* but pragmatic infelicity that stems from the understanding that, in most legal systems, marriage to oneself is not permitted. This kind of world knowledge does not enter into the definition of anti-reflexivity, which pertains to

---

6   The semantic nature of (54) is a conceptual problem for the Menuzzi's (1999) theory, which is, in other respects, similar to what I propose here. Since he defines reflexivity in terms of coindexation in narrow syntax (i.e. a predicate is reflexive *iff* at least two of its arguments are coindexed), he has to resort to an obscure notion of "incompatibility with syntactic reflexivity" to play the same role as (54). So, since (i) contains an anti-reflexive predicate, coindexation between subject and object should be banned on purely syntactic grounds:

(i)   # $Bobby_3$ is similar to $himself_3$.

However, there is nothing *syntactically* ill-formed about (i). As a matter of fact, (i) is not even *semantically* anomalous: it just happens to be tautologically true. This would make uttering (i) *pragmatically infelicitous* – at least in contexts where speakers are expected to be informative. However, as far a modular syntax is concerned, there is simply no way to filter out uninformative or infelicitous indexings.

Curiously, Menuzzi (2004, pg. 139) rejects an account which includes a property like anti-reflexivity as part of the definition of which reflexives are subject to RCA on the grounds that "we would have to refer to a semantic property of the predicate, which would add suspicion to the claim that [RCA] applies to syntactic predicates". However, this is precisely what he has to do in order to count cases like (i) as *syntactically* anti-reflexive.

semantics alone. So the predicates in both (55) and (56a) fail to qualify as anti-reflexive. The same cannot be said for the genuinely discursive reflexives in (49) and (50) or for the standard cases of discursive reflexives within N and P-headed predicates, as I discuss below.

Anti-reflexivity and predicates with expletive subjects

Let us first look at the expletive subject cases in (49). Reflexive interpretations are undefined, and, therefore, uninformative for these examples because the semantic terms that correspond to PREDs therein (**clear$'$**, **difficult$'$** and **incumbent$'$**) have, as arguments, an experiencer and a proposition corresponding to the extraposed clause. The relevant examples are repeated below, along with their corresponding simplified semantic structures:

(57)   a.   *They* made sure that [it was clear to *themselves* that this needed to be done].

       a$'$.   **clear$'$**(EXPERIENCER: y, THEME: **this needed to be done$'$**)

       b.   *Mary* said that [it would be very difficult for *herself* to accept your theory].

       b$'$.   **difficult$'$**(EXPERIENCER: y, THEME: **accept$'$**(y, your-theory))

       c.   *Mary* says that [it was incumbent upon *herself* not to abandon the car].

       c$'$.   **incumbent$'$**(EXPERIENCER: y, THEME: **abandon$'$**(AGENT:y, THEME: the-car))

A reflexive interpretation would require identifying the experiencer and theme arguments of each of the semantic terms above. However, these arguments could not, in principle, be identified due to a clash in semantic types – in Montagovian terms, the experiencer is of type e and the propositional theme is of type $\langle s,t \rangle$ (a function of worlds to truth-values).

Since reflexive interpretations for the predicates in (49) would be semantically uninformative (i.e. indeterminate in all circumstances), sentences like (58) are expected to be anomalous (at least in contexts where informativity is at stake). This is in fact what happens:

(58)   a.   #*Bobby* is clear to *himself*.

       b.   #*Bobby* is difficult for *himself*.

       c.   #*Bobby* is incumbent upon *himself*.

What makes the examples in (49) anti-reflexive is, therefore, not the presence of the expletive by itself, but the semantic properties of the predicates involved. Therefore, a similar effect is predicted – uniquely by my approach, as far as I'm aware – for the non-extraposed variants of these structures. Example for (49b) and (49c) are given below:

(59)   a.   *Mary* says that [not abandoning the car was incumbent upon *herself*].

          b.   ?*Mary* said that [Charlie's leaving would be very difficult for *herself*].

Though the judgments I gathered for cases like these were not particularly robust, this result seems to be correct. Since the semantic type of the arguments **incumbent′** in (49c) is preserved (59a), the predicate counts as anti-reflexive for both (49c) and (59a). Discursive reflexives are, thus, expected in both cases alike.

If the crucial factor that contributes to the acceptability of (49) is anti-reflexivity, rather than the presence of the expletive itself, we might expect to find some instances of reflexives within sentences with expletive subjects that are not acceptable discursive occurrences. Indeed, we do find such cases, as the contrast in (60) shows:

(60)   a.   *Paul* claimed that [it would be helpful to *himself* if Amy left].

          b. ?**Paul* claimed that [it would help *himself* if Amy left].

(61)   a.   *Bobby* thinks that [it would be possible for *himself* to win].

          b. ?**Bobby* thinks that [it would scare *himself* if he lost].

Though neither the judgements nor the underlying semantic factors are as clear here as in (49), the contrasts above might plausibly be related to anti-reflexivity as well. The data below suggest that only *helpful* and *possible* are linked to genuinely anti-reflexive PREDs:

(62)   a.   #*Paul* is helpful to *himself*.

          b.   *Paul* helps *himself*.

(63)   a.   #*Bobby* is possible for *himself*.

          b.   *Bobby* scares *himself*.

Anti-reflexivity and comparison predicates

Anti-reflexivity also adequately identifies reflexives in comparison predicates like (50), repeated below, as discursive, but for different reasons. Reflexive interpretations are disfavored there because they would result either in tautology (cf. (64b)-(65b)) or in contradiction (cf. (66b)) – both of which are semantically uninformative and infelicitous in usual contexts:

(64)   a.   *The men* said that [the new administrators would be very much like *themselves*].

          b.   #*The new administrators* are like *themselves*.

(65)   a.   *Joanne* loves Bobby because [Bobby resembles *herself*].

           b.   #*Joanne* resembles *herself*.[7]

(66)   a.   *The veterans* think that [the new recruits are more qualified than *themselves*].

           b.   #*The recruits* are more qualified than *themselves*

When the predicates in (50) are replaced by predicates which are not anti-reflexive in this sense, the sentences become bad, as expected (MENUZZI, 2004, pg. 138):

(67)   a.   *The men said that [the new administrators would be very proud of *themselves*].

           b.   *The veterans thought that [the new recruits would be nice to *themselves*].

           c.   *Joanne loves Bobby because [Bobby remembers *herself*].

Other contrasts involving non-local binding out of adjectival phrases are amenable to a similar explanation. The following examples are adapted from Safir (1992, pg. 5):

(68)   a.   *Bobby* considers Joanne similar to *himself*.

           b.   *Bobby considers Joanne hostile to *himself*.

(69)   a.   *Bobby* is attracted to women unlike *himself*.

           b.   *Bobby is attracted to women angry at *himself*.

Anti-reflexivity and classic examples of discursive reflexives

Anti-reflexivity unifies the discursive occurrences that are problematic for Reinhart and Reuland's exemption-based account (cf. (49)-(50) and (68)-(69)) to the more usual examples of discursive reflexives in argument positions such as (39)-(40), repeated below as (70)-(71).

(70)   a.   *Bobby* knew that Marta would never fall in love with a dandy like *himself*.

           b.   *Bobby* said that Joanne invited Susan and *himself* for a drink.

---

[7]   These kinds of cases make it clear why anti-reflexivity is not reducible to the more restrictive notions of L-analiticity and G-triviality proposed by Gajewski (2002) and Chierchia (2013). L-analytic/G-trivial sentences are those that are true or false in virtue of their logical structure alone – i.e. under all significant rewritings of their non-logical constants. A way to make reflexive interpretations for the predicates *resemble*, *similar* and *like* qualify as L-analytic/G-trivial is to argue for lexical decompositions that include the logical constant =. However, this is not a view I want to commit to, since it would be incompatible with the theories espoused by Gajewski (2002) and Chierchia (2013), who claim that L-analytic/G-trivial sentences are always ungrammatical. When normal concerns about informativeness are not at stake, a predicate like *resemble* is completely acceptable with reflexive interpretations: *The logician said that he developed a formal system where **every entity** resembles **itself***. If reflexive uses of *resemble* were, in fact, L-analytically/G-trivially true, such cases would have to be treated as literally ungrammatical. The same would apply to all predicates where reflexivity is either tautological or contradictory. I find this conclusion highly implausible.

(71)   a.   *Bobby* felt that nobody appreciated that picture of *himself*.

     b.   *Bobby* heard this voice whispering near *himself*.

The reflexive in (70a) is a GF argument of the PRED linked to *like*, which is the same one that appears in comparative constructions such as (64a) above. As we saw, this predicate is anti-reflexive because it is tautologically true under a reflexive interpretation (cf. (64b)).[8]

In (70b), the GF of the reflexive is a member of the PRED that corresponds to the conjunction. There are good reasons for viewing conjunctions as anti-reflexive predicates as well – a point raised, in slightly different terms, by Safir (1992, pg. 40). Regardless of the semantic type of its conjuncts, a reflexive interpretation for a conjunction is always uninformative. For the examples below, I assume that *do so* and *that* are VP and sentential pro-forms, respectively:

(72)   a.   #*Bobby* and *himself* went to visit Joanne.

     b.   #Bobby *kissed April* and *did so*. (did so=kissed April)

     c.   #Bobby kissed April and *that* happened. (that=Bobby kissed April)

In contrast to the other cases mentioned so far, reflexive interpretations associated with the conjunction are not analytically true, false or undefined. What makes them anti-reflexive is simply the fact that the function that results from a reflexive interpretation of the PRED linked to the conjunction is idempotent: i.e. whenever $\lambda R[R \sqcap R]$ is applied to a term Q, yielding $Q \sqcap Q$, the outcome is always equivalent to Q, in virtue of the generalized idempotent law $P \sqcap P \equiv P$. In other words, the conjunction itself does not add any information in a reflexive setting beyond the information already contained in its (repeated) conjuncts.

Let us now consider the cases of reflexives within P and N predicates in (71). The reflexive in (71a) is an argument of the syntactic predicate linked to *picture*. As mentioned above, this noun has a dyadic meaning which establishes a representational relation between a depiction and a depicted entity. I assume that anti-reflexivity is a semantic property of representational relations *in general*: i.e. if x is representationally related to y, x and y are necessarily distinct. So, for these kinds of predicates, reflexive interpretations are always *undefined*.

We can formally enforce this by treating depicted objects as INDIVIDUAL CONCEPTS – i.e. partial functions from possible worlds to individuals – *as per* the analysis suggested in

---

[8]   As I mentioned in Chapter 3 we see a contrast between reflexive embedded within NP adjuncts headed by *like* and those whose heads do not correspond to anti-reflexive PREDs in the same way *like* does:

(i)   a.   It angered *him* that she tried to attract a man like *himself*.
     b.   *It angered *him* that she tried to attract a man proud of *himself*.

Partee (1986). The noun *picture*, in this set up, instead of having a meaning of type $\langle e, \langle e, t \rangle \rangle$ (as is usually assumed), has a meaning of type $\langle \langle s, e \rangle, \langle e, t \rangle \rangle$: i.e. in set-theoretic terms, *picture* denotes a relation between entities and individual concepts. Imposing a reflexive interpretation on predicates of this type amounts to a type error (i.e. an identification between elements of non-overlapping types) in the same way as the attempt to interpret the predicates in (58) reflexively.

This means that, in a phrase like *the picture of Bobby*, *Bobby* is type-shifted to denote the individual concept of Bobby instead of Bobby himself, as in the representation below:[9]

(73)

$$
\begin{bmatrix}
\text{PHON} & \text{the}_4 \text{ picture} \prec \text{of}_3 \text{ Bobby}_2 \\
\text{SYN} & [_{NP} \text{ Det}_4 \text{ N}_3 \text{ NP}_2] \\
\text{GF} & \text{GF}_3 \text{ PRED}_3: \langle \text{GF}_2 \rangle \\
\text{SEM} & \lambda P[\iota z[P(z)]]_4 (\lambda f_{\langle s,e \rangle} [\lambda x[\textbf{picture-of}'_3(x,f)]](\lambda w[\text{bobby}^w]_2)) \\
& = \lambda P[\iota z[P(z)]](\lambda x[\textbf{picture-of}'(x, \lambda w[\text{bobby}^w])]) \\
& = \iota z[\textbf{picture-of}'(z, \lambda w[\text{bobby}^w])]
\end{bmatrix}
$$

There is positive evidence in favor of this. Since individual concepts are defined as partial functions, they need not have values specified in all possible worlds. Note, for example, that an NP like *the picture of Santa Claus* does not typically entail that Santa Claus exists in the actual world. Likewise, one cannot do existential generalization over the indefinite complement in *the picture of a man who has two heads*: e.g. if someone says *Bobby bought the picture of a man who has two heads*, we cannot infer that a man who has two heads exists in the actual world.

This is not a property of N-complements in general: the phrases *the brother of Santa Claus* and *the neighbor of a man who has two heads* do entail that Santa Claus and a two-headed man exist. The explanation for the contrast lies on of the semantic type of each N-complement: in the case of *picture*, the complement is an individual concept of type $\langle s, e \rangle$, rather than an *e*-typed argument, as is the case for the complement of *brother* and *neighbor*. A similar analysis applies to other N predicates that constitute environments for discursive reflexives (e.g. *story*, *rumor*, *statue*). See Abbott (2011) and Varaschin (2020) for more details.

The empirical consequences of anti-reflexivity also differ from those of most predicate-based theories with regards to the behavior of reflexives within NPs. Approaches like Pollard

---

[9] I want to remain neutral about how semantic composition works for these cases. It could be that *picture* is lexically relational and that the preposition *of* is simply an identity function on its individual concept argument (i.e. it maps individual concepts to individual concepts), as (73) implies, or, alternatively, that *picture* is lexically monadic (i.e. of type $\langle e, t \rangle$), but is type-shifted to a relational denotation in order to combine with its individual concept argument. See Partee and Borschev (1998, 2003) for more discussion of these alternatives.

and Sag's (1992) and Reinhart and Reuland's (1993) entail that only anaphors lacking a higher syntactic coargument count as exempt. They predict, therefore, that reflexives embedded within NPs containing possessives are not exempt and, thus, cannot be bound outside of the NP. This prediction was shown to be incorrect by the experiments in Asudeh and Keller (2001) and Runner and Kaiser (2005). Speakers do overwhelmingly accept structures like (74):

(74)   a.   *Bobby* found Amy's pictures of *himself*.

         b.   *Bobby* heard Joanne's stories about *himself*.

Since the property of anti-reflexivity is inherent to the semantics of *picture* and *stories*, it does not make a difference for my account whether these Ns are linked to PREDs where the reflexive has a higher coargument. If anti-reflexivity is all that is required to spring discursive occurrences, the presence or absence of a possessive is irrelevant.

Reflexive complements of most predicative prepositions (prepositions that correspond to PREDs) also exhibit properties of discursive occurrences – e.g. non-complementarity with pronouns – which can be accounted for in terms of anti-reflexivity. This applies to (71b) above as well as to the following examples from Reinhart and Reuland (1993, pg. 661; 686):[10]

(75)   a.   *Max* saw a gun near {*himself* / *him*}.

         b.   *Lucie* said that Max saw a ghost next to {*herself* / *her*}.

         c.   *Lucie* counted five tourists in the room apart from {*herself* / *her*} .

When we attempt to impose reflexive interpretations upon the PREDs linked to the predicative Ps in (75), we get anomalous results akin to the ones we saw in (58) and (64)-(66).

(76)   a.   #*Bobby* is near *himself*.

         b.   #*Lucie* is next to *herself*.

         c.   #*Lucie* is apart from *herself*.

This suggests that the PREDs linked to *near*, *next to* and *apart from* are anti-reflexive.

---

[10] This is not the case for *all* predicative/spatial prepositions. The following examples are not anti-reflexive:

(i)   a.   Bobby rolled *the carpet* over *itself*.

      b.   *Julie* stepped on *herself*.

Since I do not take the presence of subjects to be necessary for the constitution of syntactic predicates (i.e. PREDs), the reflexives in (i) are no longer analyzed as discursive occurrences, as they were in Reinhart and Reuland (1993, pg. 687). For me, they are simply plain reflexives are licensed in accordance to ERC.
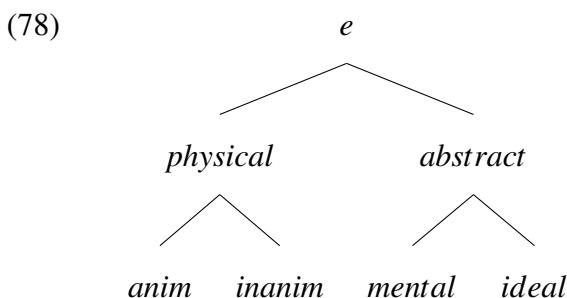
The reason is related to the fact that these PREDs are all interpreted as functions specifying locations. If we build into our semantic model theory the folk physics assumption that the same body cannot occupy two regions of space at once, reflexive interpretations for locational predicates will be tautologically false or undefined, pending on whether one takes locational disjointness to be a presupposition or part of the at-issue content of spatial predicates.

Other possible applications of anti-reflexivity

The hypothesis that anti-reflexivity contributes to the acceptability of discursive reflexives despite the fact that the latter violate the correspondences to SEM specified by the grammar makes further predictions beyond the ones mentioned above. A full examination of these is beyond the scope of this work. But consider the following examples involving verbs of creation:

(77)   a.   #*Bobby* founded *himself.*

       b.   #*Bobby* designed *himself.*

       c.   #*Bobby* wrote *himself.*

       d.   #*Bobby* composed *himself.*

The PREDs linked to these verbs are arguably anti-reflexive: no interpretation is defined for an identification between their necessarily animate AGENTS and inanimate or abstract THEMES. We can formalize these selectional features as lexically specified typing restrictions on arguments – a proposal reminiscent of McCawley (1968), who treated selectional restrictions as presuppositions. This requires a lattice of expanded types like (78), which is based on proposals by Copestake and Briscoe (1991) and Pustejovsky (2012):

(78)
$$e$$

physical   *abstract*

anim   *inanim*   *mental*   *ideal*

If we assume the richer system of types in (78), the semantic type of the passive verbs in (77) will be $\langle anim, \langle inanim, t \rangle \rangle$. In this set up, imposing a reflexive interpretation to these predicates amounts to a type error in much the same way as the attempt to interpret the predicates in (58) reflexively. This explains the following contrasts involving binding across a subject from

an adjectival or sentential passive *by*-phrase:

(79)   a.   *Bobby* quit the company founded by *himself*.

           b. *?*Bobby* quit the room destroyed by *himself*.

(80)   a.   *Bobby* told me to see a building designed by *himself*

           b. *?*Bobby* told me to see a barber admired by *himself*.

(81)   a.   *Bobby* said that this book was written by *himself*.

           b. *?*Bobby* said that this book was burned by *himself*.

(82)   a.   *Bobby* doesn't listen to songs composed by *himself*.

           b. *?*Bobby* doesn't listen to songs despised by *himself*.

Since the reflexives in the a-cases in (79)-(82) are arguments of syntactic predicates which are anti-reflexive, they are correctly predicted to be discursive on my approach.

Note that the concept of anti-reflexivity, as I stated it, is a property of PREDs, rather than of particular pairs of arguments. This issue becomes relevant if we turn our attention to three-place predicates like *distinguish*, *compare* and *prefer*, whose two internal arguments are presupposed to be disjoint in a way that might suggest anti-reflexivity is at stake:

(83)   a. #Bobby distinguished *Joanne* from *herself*.

        b. #Bobby compared *Joanne* to *herself*.

        c. #Bobby preferred *Joanne* to *herself*.

The reason why I don't want to include these cases under the definition of anti-reflexivity is because reflexives serving as non-subject arguments of these PREDs do not seem to trigger acceptable discursive readings like the other cases mentioned in this section.

(84)   a. ?*Bobby* says Amy can't distinguish {Marta from *himself* / *himself* from Marta}.

        b. ?*Bobby* told me that Amy compared {Marta to *himself* / *himself* to Marta.}

        c. ?*Bobby* thinks Amy prefers {Marta to *himself* / *himself* to Marta.}

This is a problem for Safir's (1992) theory, where the condition for discursive occurrences is stated in terms of the concept of *implied non-coreference*, which is defined relative to pairs of arguments, rather than to syntactic predicates as a whole. My account is more restrictive than his. Since the predicates in (83)/(84) *can* be reflexive (with respect to their subject and either

one of the two object arguments), they are not anti-reflexive. Therefore, the reflexives in (84) are not predicted to be acceptable discursive occurrences.

Safir (1992, pg. 34) also mentions as potentially problematic predicates associated with verbs like *outeat*, *kiss* and *collide*. Despite appearing to be anti-reflexive on an intuitive level, these predicates do not trigger the non-local binding characteristic of discursive occurrences:

(85)   **Bobby* said that Amy had {outeaten / kissed / collided with} *himself*.

I do not want to claim that the PREDs linked to these verbs are anti-reflexive. As Safir himself points out, it is perfectly informative to say something like *Bobby outate himself this time*, taking into consideration the fact that Bobby's identity remains constant across different eating events. The case of *kiss* and *collide* are similar to like the example of *marry* mentioned above in that reflexive interpretations are merely unusual for reasons related entirely to world-knowledge. With some imagination, it is possible to construe examples where reflexive interpretations seem more or less plausible for these PREDs (e.g. *Bobby kissed himself on the foot*, *The giant worm collided with itself.*)

Why the Two Reflexives Hypothesis failed

We are now in a position to understand more clearly the failure of the Two Reflexives Hypothesis (TRH) – i.e. the thesis according to which discursive occurrences of reflexives can be distinguished from plain occurrences on the basis of purely syntactic properties. Reinhart and Reuland's (1993) specific version of the TRH is repeated below:
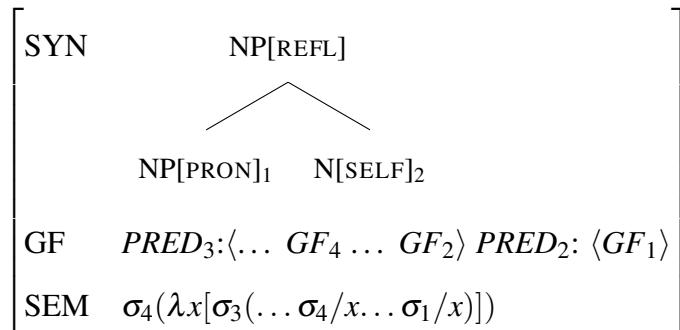
(86)   A reflexive is exempt from RCA (MENUZZI, 1999, pg. 160):

    a.   when its predicate does not have a subject;

    b.   when the reflexive itself is not a syntactic argument of a predicate.

The TRH tries to characterize discursive occurrences in argument positions according to a syntactic property of their predicates: namely, the absence of subjects (cf. (86a)). This was shown to be impossible in Chapter 3. We now know why. The crucial trait of discursive occurrences of reflexives in argument positions is, in reality, a semantic property: the property of anti-reflexivity defined in (54). Only discursive occurrences that fall under (86b) (e.g. reflexives in A′ positions) can be described in terms of a syntactic property: the property of not being linked to a GF argument of a PRED.

To summarize, the hypothesis I am proposing as an alternative to the TRH is the following: Only the violations of ERC which occur with reflexives that are not GF arguments of a PRED (i.e. non-argument reflexives) or within anti-reflexive PREDs result in acceptable discursive occurrences that give rise to logophoric readings. What both of these scenarios have in common is that they are contexts where ERC (repeated below) has no way of being fulfilled.

(87) ENGLISH REFLEXIVE CONSTRUCTION (ERC)

$$\begin{bmatrix} \text{SYN} & \text{NP[REFL]} \\ & \overwidetilde{\text{NP[PRON]}_1 \quad \text{N[SELF]}_2} \\ \text{GF} & PRED_3{:}\langle \ldots GF_4 \ldots GF_2 \rangle \; PRED_2{:} \langle GF_1 \rangle \\ \text{SEM} & \sigma_4(\lambda x[\sigma_3(\ldots \sigma_4/x \ldots \sigma_1/x)]) \end{bmatrix}$$

What the ERC says, to put matters simply, is that a reflexive is licensed if it is bound by a higher GF argument of its PRED. As we saw, this implies that the PRED in question will be reflexive – i.e. at least two of its arguments (the ones corresponding to the pronoun and its antecedent) will be bound by the same λ-operator. If the reflexive does not correspond to a GF argument of a PRED or if its PRED cannot be reflexive in principle, violating ERC is, in some sense, inevitable. This inevitability is what makes violations of ERC in discursive occurrences give rise to logophoric readings. The idea then is that since ERC is the only construction specifying the semantic contribution of reflexives in English (and this is what makes my account unified), when ERC cannot apply, reflexives must be interpreted non-literally, by means of a pragmatic procedure. This is captured in the form of the LOGOPHORIC STRATEGY:

(88)   LOGOPHORIC STRATEGY

When the GF corresponding to a reflexive cannot be the argument of a reflexive PRED, the pronominal within the reflexive is interpreted as a variable whose value is picked up from the set of entities whom the speaker empathizes with in the discourse context.

ERC states that the lexical duty of reflexives is to be bound by higher coarguments, thereby signaling syntactic reflexivity. When they *cannot* do so (either because they are not GF arguments of a PRED or because of anti-reflexivity), they must be interpreted in accordance to some extra-grammatical heuristic. This is where logophoric readings kick in.

One certainly needs to be more specific about *why* discursive occurrences thrive precisely in anti-reflexive and non-argument contexts (as opposed to any other context where ERC is violated). Furthermore, one can also wonder why violations of ERC are made acceptable by *logophoric* readings, rather than other kinds of readings (e.g. a reading where the referent of the pronoun is interpreted as denoting an exceptionally large entity). The next chapter addresses both of these issues, proposing a specific hypothesis concerning why logophoric readings emerge precisely in contexts where ERC has no way of being fulfilled. In other words, it provides a rationale for the Logophoric Strategy in (88).

I argue there, following Menuzzi (2004), that violations of constraints on reflexives are only acceptable in the contexts specified by (88) because these contexts are appropriate for the derivation of logophoric interpretations as a kind of Manner implicature. In so doing, hearers are guided by the accessibility status of the particular kinds of pronouns in their language (ARIEL, 1990, 2001). My argument there will draw heavily on Davison's (1980) similar approach to certain kinds of prepositional passives. Though I believe this hypothesis is on the right track, it is logically independent and much more tentative than the analysis developed thus far.[11]

5.4 CONCLUSION

Most of the early approaches in Mainstream Generative Grammar – with the notable exception of Partee (1965) and Postal (1968, 1971) – presupposed that constraints on reflexives applied uniformly to all instances of reflexives in a language. These theories provided, therefore, *unified* constraints, which did not distinguish, *inter alia*, different types of grammatical licensing conditions for reflexives according to their syntactic contexts. In the Classical Bind-

---

[11] Even though (88) accounts for more cases of logophors in argument positions than Reinhart and Reuland's exemption-based account summarized in (86) above, it is not clear whether data like the following, from Zribi-Hertz (1989, pg. 709) and Baker (1995, pg. 68), can be analyzed in the terms I propose here:

(i) a. But *Rupert* was not unduly worried about Peter's opinion of *himself*.
     b. And that was exactly it, *he* thought, *he* really did not care too much what happened to *himself*.

(ii) a. If *Cassandra* has filled my bed with fleas, I am sure they must bite *herself*.
     b. [...] and let *his* net that *he* hath hid catch *himself*.

If we construe the PREDs linked to *opinion* and *happen* in (i) as anti-reflexive – the former being tantamount to a representational relation and the latter taking a proposition and an experiencer as arguments –, the discursive nature of the reflexives therein would follow from the Logophoric Strategy. The reflexives in (ii) are more problematic. Even though it is *technically* feasible to assign the status of *foci* to these reflexives (for which the the LS would then predict discursive readings), this does not seem to be the most insightful analysis . I take (ii) to reflect a literary or stylized variant of English – one in which ERC does not hold in the same way as in Standard American English. See Golde (1999) for an account along these lines.

ing Theory (CBT), for instance, all anaphors in all languages (including English reflexives) were subject to the same c-command and locality requirements embodied in Condition A.

Since then, however, the idea that a unified grammatical constraint applies to all anaphors (or even to all *reflexive anaphors*) has become a minority position for very good reasons. Prime among these is the challenge posed by typological diversity. As we saw throughout this Thesis, there are languages that contain expressions that meet the notional characterization of anaphors but which, nonetheless, do not pattern distributionally like English reflexives. Moreover, even if we restrict ourselves to an overstudied language like English the sheer diversity of reflexives makes unified accounts too complicated. For instance, there is not a single requirement implicit in CBT's statement of Condition A which is not shaken by discursive occurrences of reflexives, which need not be c-commanded nor local with respect to their antecedents.

In light of many exceptions to general principles, syntacticians of various persuasions have been forced to partition the domain of reflexives roughly between those to which rigid grammatical constraints apply and those which are subject to looser (often vaguely stated) discourse conditions. The idea here is that, in order to account for plain cases of reflexivization without having to compromise or complicate one's favored syntactically-defined constraint on reflexives, one can restrict the range of cases to which this constraint applies and provide non-syntactic (e.g. discourse-related) principles of interpretation for the remaining cases.

After the original insights in Partee (1965) and Postal (1968), this Two Reflexives Hypothesis (TRH), as I called it, was popularized in a modern context by Lebeaux (1985). It is assumed (under various guises) in approaches as diverse as Manzini and Wexler (1987), Sells (1987), Zribi-Hertz (1989), Chierchia (1989), Pollard and Sag (1992), Reinhart and Reuland (1993), Baker (1995), Steedman (1996), Hornstein (2001), Büring (2005) and Postal (2006). Crucially, in order to make the distinction between grammatical and grammatically exempt reflexives seem principled, the conditions for exemption are built into in the grammatical conditions themselves. It is the claim that the distinction between two kinds of reflexives is encoded *in the grammar* which I take to be the essential feature of the TRH.

The TRH is a radical departure from the CBT because it explicitly acquits an entire class of reflexives from grammatical rules of construal and przoposes to account for their interpretation by means of qualitatively different constraints – e.g. discourse constraints having to do with empathy and point of view (POLLARD; SAG, 1992; GOLDE, 1999; CHARNAVEL, 2019).

For reasons which became clear in Chapter 3, this kind of approach is problematic – at

least for languages like English (see Safir (1992), Dalrymple (1993), Pollard and Xue (1998), Menuzzi (1999), and Charnavel and Sportiche (2016) for similar arguments). Upon closer inspection, it turns out to be surprisingly hard to isolate the syntactic conditions under which a reflexive can be exempt from syntactic constraints. The problem with the TRH, therefore, is not with the suggestion that reflexives can be subject to both grammatical and non-grammatical conditions, but the claim that *grammar* (or, more narrowly, *syntax*) determines which of these conditions applies for any given case.

It appears, therefore, that neither the CBT nor the TRH which was proposed as an alternative to it provide a fully satisfactory account for reflexives. The central goal of this chapter was to offer a general theoretical framework that overcomes the main obstacles that hindered these previous approaches and to suggest new insights to lesser known problems of anaphora.

The main point I want to argue for is that having a violable constraint giving rise to logophoric interpretations pragmatically provides the flexibility to reconcile the distributional variability of English reflexive forms with the idea that there is a unified principle underlying all uses of reflexives. By keeping the condition on reflexives simple and allowing it to be violated in certain contexts, we can avoid the main problems which jeopardized previous unified approaches like the CBT, as well as the empirical shortcomings of the TRH, which formed an integral part of Reinhart and Reuland's (1993) reflexivity-and-chains theory.

The particular constraint I proposed – the ERC – captures the fact that there are core cases of reflexives, whose licensing conditions are defined by the grammar, and discursive cases, whose acceptability is explained by functional or pragmatic factors. Core reflexives are always bound by a local coargument of the same syntactic predicate. Every other case of reflexivization is, in principle, marked. This is compatible with the experimental results in Burkhardt (2002, pg. 25), which attest that "the interpretation of logophoric reflexives poses a higher burden on the processor than that of coargument reflexives". As Menuzzi (2004) notes, this approach also explains why discursive occurrences are often perceived as odd outside of context: they are violations of a grammatical constraint whose unacceptability can only be mitigated under particular pragmatic conditions. These will be further investigated in the next chapter. Another corollary of this approach, which seems to empirically correct, is that the acceptability of discursive reflexives is more subject to inter-speaker variation than that of reflexives which straightforwardly conform to ERC. In the next chapter, I discuss extensions of this general framework to other languages, such as Brazilian Portuguese and Japanese.

6 A PRAGMATIC ACCOUNT OF LOGOPHORIC READINGS

In Chapter 1, I presented a historical survey of unified approaches to reflexives within Mainstream Generative Grammar. I focused particularly in two prominent versions of this approach: a transformational account of reflexivization, represented mainly by Lees and Klima (1963), and the Classical Binding Theory, whose most prominent expression is found in Chomsky (1981). Each of these, as we saw, is representative of a particular strategy in handling potential objections to unified principles: the transformational theory theory opts for enriching structures and derivations while the Classical Binding Theory almost always opts for making the principles responsible for the interpretation of reflexives more complex.

Chapter 2 examined an alternative to unified approaches to reflexives: Reinhart and Reuland's (1993) theory, which incorporates a version the Two Reflexives Hypothesis (TRH). The TRH is motivated by the fact that previous unified approaches tend to leave out of the picture an entire class of uses of reflexives, which I called DISCURSIVE OCCURRENCES. The most salient trait of discursive occurrences – in addition to their apparent obliviousness to grammatical strictures – is their logophoric interpretation: their semantic value is picked up from the set of entities with whose point of view the speaker emphasizes in the context (i.e. the set of EMPATHY LOCI, in the sense of Charnavel (2019)). What defines the TRH is the claim that the anaphoric behavior of reflexives in English is governed by an essentially disjunctive constraint, which specifies syntactic licensing conditions for both plain coargument reflexives and discursive occurrences alike. Chapter 3 argued that the TRH fails because there is no way to distinguish plain and discursive reflexives solely on the basis of syntactic properties.

In Chapter 5, I suggested a fourth strategy, which preserves unified approaches from the counterexamples raised by proponents of the TRH, while at the same time avoiding the problems faced by the TRH: instead of enriching structure or introducing complexity into the principles (which usually entails introducing complexity into Universal Grammar), we can keep both the principles and the structures simple if we posit a single grammatical constraint on reflexives and allow it to be violated in certain contexts. These violations can, in turn, explain how logophoric readings are pragmatically derived by means of a non-literal interpretive strategy.

I proposed that the principle governing English reflexives is a construction, which I called the ENGLISH REFLEXIVE CONSTRUCTION (ERC), which is a special case of the general construction governing pronominals. The ERC basically establishes a correspondence between

morphosyntactic reflexivity (i.e. PREDs where a *self*-marked pronominal occupies a non-subject position) and semantic reflexivity (a SEM term where at least two variables corresponding to two arguments of the reflexive's PRED are bound by the same λ-operator). Precisely because English reflexives include pronominals as part of their structure, is possible to use reflexives in English in ways that do not meet the licensing conditions for reflexives but are nevertheless rendered acceptable under particular circumstances.

In this chapter, I discuss how the violations of ERC we find in discursive occurrences give rise to logophoric readings. I argue, following Menuzzi (2004), that the ungrammaticality of discursive reflexives is systematically interpreted as a form of MARKEDNESS, which gives rise to logophoric readings as Manner implicatures (GRICE, 1989; RETT, 2020). Menuzzi's insight was inspired by Davison's (1980) approach to peculiar passives. I elaborate my own analysis of logophoric readings on in terms of an interaction between Horn's (1984, 2005) neo-Gricean pragmatic principles and Ariel's (1990) accessibility theory.

In section 6.1, I lay out Horn's basic neo-Gricean framework, building on subsequent developments by Katzir (2007) and Rett (2015, 2020). In section 6.2, I explore how these principles can be applied to the analysis formulate a version of Davison's (1980) analysis of peculiar passives. What interests me about this analysis is the fact that it derives an acceptable marked reading from what is, essentially, an ungrammatical structure. The idea that ungrammaticality may result in (marked) acceptability is a crucial ingredient in my account of discursive reflexives. In section 6.3, I discuss how the constraint on English reflexives (ERC) interacts with the neo-Gricean pragmatic principles to yield the logophoric readings. Lastly, in section 6.4 I discuss how the constructional approach developed for English reflexives in the previous and the present chapter might be extended to other languages.

6.1 A NEO-GRICEAN THEORY OF PRAGMATIC INFERENCE

Horn (1984) argues that Grice's (1989) Conversational Maxims can be reduced to two opposing forces: one reflecting the hearer's communicative interests, and the other the speaker's.[1] The first is vaguely related to Grice's (first) Maxim of Quantity (*Make your contribution as informative as required*) and the second to his Maxim of Relation (*Be relevant*):

---

[1] This idea is based on Zipf (1949), who proposes two opposing economy principles: Auditor's Economy (a drive towards eliminating ambiguity) and Speaker's Economy (a drive towards simplifying linguistic forms). Zipf (1949, pg. 21) says that the first force, if left untampered, would "tend to increase the size of a vocabulary to a point where there will be a distinctly different word for each different meaning", whereas the second would "tend to reduce the size of the vocabulary to a single word by unifying all meanings behind a single word".

(1)   Q PRINCIPLE (Hearer-based)

Make your contribution sufficient

Say as much as you can (given R)

(2)   R PRINCIPLE (Speaker-based)

Make your contribution necessary;

Say no more than you must (given Q)

The Q Principle sets a lower-bound on how much information a speaker has to give (i.e. "the speaker must communicate at least *p*"). As a result, it can be exploited to derive upper-bounded implicatures, which allow the hearer to infer that anything more informative than what was said does not hold (i.e. "the speaker must communicate no more than *p*"). For Horn, the Q Principle is what is derives scalar implicatures. If a speaker says *some students passed the exam*, we can infer the the upper-bounded reading that *not all students passed* because we assume that the speaker is following the Q principle and is, therefore, saying as much as they can.

The R Principle, in turn, is an upper-bounding condition, placing a maximum limit on how much a speaker has to say. It embodies Zipf's (1949) idea that speakers are guided by a principle of least effort, which orients them to use the minimal forms capable of conveying their intended messages. As consequence, the R Principle produces lower-bounded implicatures that enrich the information conveyed by the utterances beyond what is strictly encoded by the sentence (i.e. "the speaker must communicate at least *p* and possibly more"). Horn (1984, pg. 14) cites indirect speech acts as typical examples of R-based implicatures (SEARLE, 1975):

> [. . .] if I ask you whether you can pass me the salt, in a context where your abilities to do so are not in doubt, I license you to infer that I am doing something more than asking you whether you can pass the salt – I am in fact asking you to do it. (If I know for a fact that you can pass me the salt, the yes-no question is pointless; the assumption that I am obeying the [R Principle] allows you to infer that I mean something more than what I say.)

The Q and R Principles outlined above are often in direct collision. As Horn (1984, pg. 15) puts it "[a] speaker obeying only Q would tend to say everything she knows on the off-chance that it might prove informative, while a speaker obeying only R would probably, to be on the safe side, not open her mouth". Out of the conflict between the antithetical requirements of Q and R emerges a synthesis Horn (1984, pg. 22) calls the DIVISION OF PRAGMATIC LABOR (i.e. an "equilibrium"). He expresses this in terms of the following principle:

(3)   DIVISION OF PRAGMATIC LABOR

The use of a marked (relatively complex and/or prolix) expression when a corresponding unmarked (simpler, less 'effortful') alternate expression is available tends to be interpreted as conveying a marked message (one which the unmarked alternative would not or could not have conveyed).

The idea is that, whenever a speaker employs a marked form μ, when a corresponding unmarked form τ is available, the hearer uses the R Principle to infer that whatever makes μ marked with respect to τ was, in fact, not unnecessary, but needed in order to convey the particular message they intend to convey. As a result, we get an association between marked forms and marked meanings and unmarked forms and unmarked meanings:

> The unmarked alternative τ tends to become associated (by use or – through conventionalization – by meaning) with unmarked situation S, representing stereotype or salient member of the extension of μ. The marked alternative μ tends to become associated with the complement of S with respect to the original extension of τ. (HORN, 1984, pg. 22)

As the formulation of Horn's Division of Pragmatic Labor in (3) already anticipates, the actual operation of the Q and R Principles in language use typically involve a comparison between sets of alternative or competing expressions. Quantity implicatures arise when competitors with roughly the same complexity differ in informativity. Manner implicatures arise when competitors with roughly the same informativity differ in complexity/markedness.

Horn's program is somewhat vague and admits different kinds of executions, depending on one's interests. For my purposes, I will assume a slightly more formal version of each of the principles in (1)-(3), based on work by Katzir (2007) and Rett (2015, 2020). A necessary step towards formalizing Horn's neo-Gricean theory characterizing notions like *informativity* and *complexity*. The former can be understood, as is standard, in terms of a (generalized) notion of entailment (HORN, 1972; LEVINSON, 1983; MATSUMOTO, 1995):

(4)   a.   For any constructs μ and τ with corresponding SEM terms μ' and τ', $\mu \supseteq_i \tau$ (μ INFORMATIONALLY INCLUDES τ) *iff* every SEM term M of type t containing μ' entails a semantic term T of type t containing τ', where M is identical to T except for the presence of μ' and τ'.

       b.   If $\mu \supseteq_i \tau$ and $\tau \supseteq_i \mu$ then $\mu =_i \tau$ (μ and τ are equally informative)

       c.   If $\mu \supseteq_i \tau$ and $\mu \neq_i \tau$, then $\mu \supset_i \tau$ (μ is more informative than τ).

The definition in (4) is a fair approximation of the concept of informativity that enters into the definition of Horn scales, which are crucial for the generation of scalar Q implicatures. The definition in (4) clearly accounts for the fact, in (5), each item is more informative than the item that follows it (LEVINSON, 1983, pg. 134):

(5)   ⟨all, most, many, some, few⟩

For example, the quantified determiner *all* counts as more informative than *some* because every SEM term containing the meaning of *all* will entail a corresponding SEM term containing the meaning of *some*. More simply put, every situation that makes a sentence of the form *all XYZ* true will also make a sentence of the form *some XYZ* true.

For the notion of *complexity* (or *markedness*), I assume an effort-based definition, as already suggested in Chapter 3. This similar to the definition of complexity assumed within Relevance Theory (SPERBER; WILSON, 1995) and some approaches to language change and universals (GIVÓN, 1991; WURZEL, 1998; HAWKINS, 2004; CULICOVER, 2013b):

(6)   EFFORT-BASED COMPLEXITY

For any constructs μ and τ, $\mu <_c \tau$ (μ is LESS COMPLEX than τ) *iff* using (e.g. processing, articulating or understanding) μ consistently involves less effort than using τ. If neither $\mu <_c \tau$ nor $\tau <_c \mu$, then $\mu =_c \tau$ (μ and τ are equally complex).

Unlike informativity, the notion of complexity employed here is not based on grammar alone. Grammatical factors like syntactic embedding or segmental length are certainly part of what contributes to the effort speakers have to put into using linguistic forms. But other factors such as frequency and compatibility with non-structural processing heuristics also come into play (SPERBER; WILSON, 1995). Consider the case below:

(7)   a.   Bobby hates birthdays.

        b.   Bobby abhors birthdays.

While (7a) and (7b) are equally complex and equally informative with respect to grammatical structure, I still want to say (7b) is more complex overall than (7a) given the lower frequency of the verb *abhors* in comparison to *hates*. That this contributes to complexity can be seen from the fact that (7b) gives rise to a Manner implicature which suggests that (7b) is more dramatic or intense than (7a). A purely structural notion of complexity, like the one proposed

by Katzir (2007), has to way to capture these kinds of frequency effects.[2]

On the basis of the aforementioned notions of informativity (cf. (4)) and complexity (cf. (6)), we can formalize Horn's neo-Gricean theory as follows. The definitions below are adapted, in a somewhat simplified fashion, from Rett (2020, pg. 73):

(8)   Q ALTERNATIVES

For any construct μ, the set of Q alternatives for μ is written as $A_Q(\mu)$

$A_Q(\mu) := \{\tau : \tau =_c \mu\}$.

(9)   Q PRINCIPLE (REVISED)

Do not use μ if there is another construct $\tau \in A_Q(\mu)$ such that both (i) $\tau \supset_i \mu$, and (ii) τ is assertable (i.e. the speaker believes τ is applicable and relevant in the context).

(10)   R ALTERNATIVES

For any construct μ, the set of R alternatives for μ is $A_R(\mu)$

$A_R(\mu) := \{\tau : \tau =_i \mu\}$

(11)   R PRINCIPLE (REVISED)

Do not use μ if there is another expression $\tau \in A_R(\mu)$ such that both (i) $\tau <_c \mu$, and (ii) τ is assertable (i.e. the speaker believes τ is applicable and relevant in the context).

I do not wish to claim that the revisions above cover all of the ground that falls under Horn's principles in (1)–(2). They are merely meant to capture the essence of Horn's proposal: the idea that communication is governed by two diametrically opposing constraints: one saying that a speaker should not use a less informative form if a more informative and equally complex one is available (cf. (9)) and the other saying that a speaker should not use a more complex form if a less complex and equally informative one is available (cf. (11)).

The revised version of the Q Principle handles scalar inferences in an obvious way. Speakers should not say things like *some students passed the exam* if they are in a position to assert the equally complex but more informative *all of the students passed the exam*. Therefore, if a speaker says *some students . . .*, we assume that they are not in a position to assert that the

---

[2] Unlike most kinds of implicature, the inference (7a) is DETACHABLE: it is possible to find another way of saying the same thing (namely, that Bobby does not like birthdays) in the same context without the implication in question. Grice (1989, pg. 39) himself recognizes, however, that NON-DETACHABILITY is only a property of implicatures for which "the manner of expression plays no role in the calculation". It is widely recognized, therefore, that Manner implicatures (i.e. implicatures that arise in virtue of *how* things are said) and even some Quantity implicatures are detachable (RETT, 2020). Given that Manner implicatures are often tied to the presence of particular forms, they are also harder to cancel – i.e. they have a lower degree of CANCELLABILITY than most other kinds of implicatures. On the basis of these facts, Rett (2020) concludes that non-detachability and cancellability are not necessary conditions for conversational implicatures.

stronger form: i.e. we derive the implicature that *not all students passed the exam.*

The R Principle is what gives rise to Manner implicatures (GRICE, 1989, pg. 37).

(12)   a.   Miss X sang "Home Sweet Home".

      b.   Miss X produced a series of sounds that corresponded closely with the score of "Home Sweet Home".

The two sentences in (12) are truth-conditionally equivalent: i.e. from a purely objective perspective, they characterize the same states of affairs in the world. As a consequence (12a) is among the R Alternatives for (12b) – i.e. (12a) $\in A_R$((12b)). Since (12a) is, by all measures, less complex than (12b) (i.e. (12a) $<_c$ (12b)), the R Principle says that a speaker who is in a position to assert (12a) should not assert (12b). Therefore, if a speaker does happens to use (12b), the only way we can make sense of his behavior is by assuming that they were not, in fact, in a position to assert (12a). We infer, accordingly, that the meaning the speaker is trying to convey is not quite the one we would typically infer from (12a) – i.e. it is atypical or marked in a special sense. The most obvious supposition was that Miss X's performance was terrible.[3]

Rett (2020, pg. 73) summarizes the effect of the R Principle on marked constructions in the form of a MARKED MEANING PRINCIPLE. The (adapted) formulation in (13) captures the part of Horn's Division of Pragmatic Labor relevant to my analysis of logophoric reflexives:

(13)   MARKED MEANING PRINCIPLE

      For any constructs μ and τ such that $\tau \in A_R(\mu)$ and $\tau <_c \mu$, using μ carries the implicature **atypical**$'(\tau')$.

Before going over the details of how logophoric readings are derived from these principles, I discuss a particular interpretation of Davison's (1980) analysis of PECULIAR PASSIVES along similar lines which serves as an inspiration for the implicature-based account of discursive reflexives proposed here (MENUZZI, 2004).

---

3   Something similar happens in cases of periphrastic synonyms discussed by McCawley (1978): e.g. the relationship between *pale red* and *pink*. The latter is a member of the set of R Alternatives for the former, given that both forms arguably share the same denotation. Since *pale red* is more complex than *pink*, the R Principle predicts that *pale red* should not be used in a context where *pink* is assertable. Therefore, if a speaker uses the more marked form *pale red*, we would typically assume that something special is intended that would not be conveyed by *pink* alone. As a result, the meanings conveyed by *pink* and *pale red* are carved up in a way that "the central, canonical part of the denoted property comes to be associated with the unmarked phrase, while the sidelined, atypical part of the property comes to be associated with the marked phrase" (RETT, 2020, pg. 60). This explains why *pale red* implies a shade of red which is not as pale as the one implied by other similar constructs lacking a suitable R Alternative, such as *pale green* or *pale blue* (MCCAWLEY, 1978, pg. 246).

6.2 DAVISON'S (1980) ANALYSIS OF PECULIAR PASSIVES

Davison coins the term PECULIAR PASSIVE to refer to instances of prepositional passives out of predicative PPs. She gives the following examples (DAVISON, 1980, pg. 44-45):

(14) a. This chair has been sat on (by somebody).

   b. This cup was drunk out of by Napoleon.

   c. This spoon has been eaten with.

   d. That knife had been cut with too often without being sharpened.

These examples are problematic for most standard accounts of passivization because they imply the promotion of a non-object argument – i.e. an NP that is not governed by the verb that carries the passive morphology. In (16a), for instance, the argument promoted to subject position is *the chair*, which is not governed by *sit*, but by the preposition *on*, which heads its own syntactic predicate. This is also what distinguishes such cases from instances of prepositional passives out of oblique objects (e.g. *Bobby was relied on by all of his friends*, *Amy was laughed at*), which are typically understood as being governed by the verb (MARANTZ, 1984).

The Passive Construction in (15), from Chapter 4 is not spared from the problems posed by (14). The trouble comes from the same sorts of considerations raised above: (15) only licenses a right bracket around the second GF of a PRED (thereby allowing it to be realized as a subject) if the GF in question is a member of a PRED linked to a verb with a PASSIVE feature. It does not promote a GF argument of a PRED linked to one head (e.g. a preposition) to a position within a PRED linked to a different head (e.g. a verb). In the case of (14), all of the GFs that are realized as the subjects of their verbs are members of PREDs linked to prepositions:

(15) PASSIVE CONSTRUCTION

$$
\begin{bmatrix}
\text{PHON} & \varphi_1 \ (\text{by}_2 \ \varphi_3) \\
\text{SYN} & [_{VP} \ V[\text{PASSIVE}]_1 \ \ldots \ ([_{PP} \ P_2 \ NP_3])] \\
\text{GF} & PRED_1{:}\langle GF_3, PRED^*_1{:}\langle GF_4, \ldots \rangle\rangle
\end{bmatrix}
$$

Given that the structures in (14) contradict the generalization that passives only target (oblique and direct) objects of verbs, an alternative would be to extend the Passive Construction in a way that allows the promotion of non-objects within predicative PPs. This would allow peculiar passives to be licensed in exactly the same way ordinary passives are.

The problem with this proposal, as Menuzzi (2004, pg. 143) points out, is that it misses the fact that peculiar passives are subject to restrictions that do not apply passives in general. For an NP to be passivized out of a predicative PP, it has to denote a specific or definite referent. With indefinite subjects, peculiar passives like those in (14) become anomalous. No definiteness effect is found in simple passives that promote direct objects or obliques (cf. (17)).

(16)   a. ??A chair has been sat on (by somebody).

         b. ??A knife had been cut with too often without being sharpened.

(17)   a.   A window was destroyed by one of the students.

         b.   A leader is relied on to build the organization.

Assuming that definiteness is a condition on discourse topics (cf. (18)), Davison (1980) interprets (16) as evidence that the promoted arguments in peculiar passives are subject to a topicality requirement. For ordinary passivized arguments, topicality is optional, as (17) suggests.

(18)   a.   As for {that chair / *a chair}, Bobby sat on it.

         b.   As for {this cup / *a cup}, it was drunk out of by Napoleon.

Davison (1980) observes that, in addition to a topicality requirement, peculiar passives are also associated with specific rhetorical effects that are not necessarily present in their corresponding active sentences – or, for that matter, in ordinary passives either. These effects are typically associated with the implication that the subject of the sentence has acquired some remarkable property in virtue of the event described by the sentence. In other words, subjects promoted in peculiar passives are generally understood as being AFFECTED (KIM, 2009; BEAVERS, 2011). Consider the pairs below (MENUZZI, 2004, pg. 144-145):

(19)   a.   This chair has been sat on by Fred.

         b.   Fred sat on this chair.

(20)   a.   This spoon has been eaten with by someone.

         b.   Someone ate with this spoon.

(21)   a.   This cup was drunk out of by Napoleon.

         b.   Napoleon drank out of this cup.

The passive and active sentences above all share the same truth-conditional content. How-

ever, the passive variants all convey extra-assumptions that are not necessarily shared by their active counterparts: (19a) implies that the chair was physically affected (in a negative way) by Fred having sat on it; (20a) implies that the spoon shows some visible sign of having been eaten with (e.g. it is dirty), (21a) indicates that the cup has picked up some kind of special status in virtue of the fact that an famous figure like Napoleon has used it – to use Bolinger's (1977) terms, it acquires a kind "aura" due to its connection to a historical figure.

When we are not able to infer that the subject has acquired a significant property, peculiar passives are generally unacceptable (HUDDLESTON; PULLUM, 2002, pg. 1446). We get, thus, contrasts like the following (RIDDLE; SHEINTUCH, 1983, pg. 536):

(22)   a.   This bed has been slept on.

        b. ??This bed has been slept near.

(23)   a.   This house was once lived on by FDR.

        b. ??This continent was once lived on by FDR.

While sleeping on a bed can easily change a bed's properties, sleeping near an object does not typically affect it in any significant way. Similarly, if a famous person has lived on a particular house, it may be relevant to think of the house as acquiring the property of being permanently associated with that person – i.e. the house becomes a kind of "historical monument" (DAVISON, 1980, pg. 54). No such thing occurs if we are talking about something as large as an entire continent. We seem to view continents as locations that are too large to be affected by the simple fact the fact that some specific person has lived there (BOLINGER, 1977, pg. 77).

In light of these peculiar properties, it becomes tempting to state a special separate construction for peculiar passives that incorporates the topicality requirement and the aforementioned rhetorical effect as licensing conditions. This would allow us to say that peculiar passives are distinct from ordinary passives (given that they are licensed by a separate construction), while nonetheless preserving the view that peculiar passives are grammatically licensed in the usual way. An approach along these lines is considered in Culicover (2021).

There are problems with this view. The factors that impinge on the acceptability of peculiar passives are quite varied in nature and far from categorical in their effects. As the contrasts in (22)-(23) reveal, they are largely influenced by subjective judgements about who is important and about what counts as a relevant change regarding the properties of an entity (i.e. what counts as genuine AFFECTEDNESS). These are not the kinds of notions that can be naturally

expressed with the vocabulary for stating grammatical constraints I developed in Chapter 4: e.g. we cannot distinguish (23a) from (23b) in terms of sensible generalizations over PHON, SYN, GF, SEM and their correspondences.

Since these subtle subjective factors cannot be incorporated into the statement of constructions, any attempt to treat peculiar passives as grammatically licensed will inevitably overgenerate. For example, given that peculiar passives with temporal, directional and accompaniment phrases are sometimes acceptable (cf. (24)), any statement of a Peculiar Passive Construction will count the structures (25) as grammatical as well (DAVISON, 1980, pg. 45-46):

(24)   a.   Freddie consented to being tagged along with.

        b.   That marker should be walked toward very careful – there is a minefield.

        c.   This crisis must be lived through.

(25)   a.   ??John doesn't like being driven with to New York.

        b.   ??Mecca should be prayed toward by the faithful.

        c.   ??This crisis must be lived during.

As we saw in connection to island phenomena in Chapter 4, a grammar's failure to predict unacceptability is not necessarily a problem. It could well be that the grammar of English licenses both (24) and (25) and that (25) is rendered unacceptable in virtue of some extragrammatical factor. The problem is that this would not account for the basic intuition that structures like (24), despite being acceptable, are nonetheless *marked*, both in comparison to their active counterparts as well as to regular passives. The greater markedness of peculiar passives is evident in the fact that they are rare and subject to more restrictions than both actives and normal passives (DAVISON, 1980, pg. 63). Simply stating a Peculiar Passive Construction does not explain *why* peculiar passives are uncommon and subject to these extra restrictions.

Rather than positing a specific construction to license the passivization of non-objects (or extending the passive construction to such cases), we can view peculiar passives as tolerable violations of the grammatical constraint on passives – i.e. as instances of ACCEPTABLE UNGRAMMATICALITY, in the sense of Chapter 4. This is the way Menuzzi (2004) interprets Davison's (1980) proposal. I now propose a reformulation of Menuzzi's (2004) somewhat vague rendition of Davison's account into the the neo-Gricean framework sketched above.

Since the promotion of prepositional object to subject position in (24) and (25) fails to be licensed by the Passive Construction, both examples come out as ungrammatical. This ungram-

maticality can be interpreted as a form of complexity, in the sense of (26):

(26) EFFORT-BASED COMPLEXITY

For any constructs μ and τ, $\mu <_c \tau$ (i.e. μ is less complex than τ) *iff* using (e.g. processing, articulating or understanding) μ consistently involves less effort than using τ. If neither $\mu <_c \tau$ nor $\tau <_c \mu$, then $\mu =_c \tau$ (i.e. μ and τ are equally complex).

The reasoning behind this is simple: resorting to fully grammaticalized constructions is a way of reducing effort of use in computing correspondences between form and meaning, insofar as constructions establish stable instructions about how to interpret particular linguistic structures. In other words, a construct which instantiates a constructional link to a truth-conditional interpretation $\mathscr{I}'$ is generally easier to understand and process than one that does not instantiate a constructional link to $\mathscr{I}'$, but requires the speaker to infer $\mathscr{I}'$ on extra-grammatical grounds.

Therefore, if a speaker were to gratuitously use an ungrammatical/unlicensed construct μ to communicate an interpretation $\mathscr{I}'$ while there exists a perfectly grammatical/licensed construct τ that encodes $\mathscr{I}'$, they would be violating the R Principle. If, however, the speaker does use the ungrammatical construct μ in this scenario, the only way to interpret their behavior is to infer, as per the Marked Meaning Principle, that they intend to communicate something beyond $\mathscr{I}'$. I represent this marked interpretive effect as **atypical′**$(\mathscr{I}')$.[4]

(27) R ALTERNATIVES

For any construct μ, the set of R alternatives for μ is $A_R(\mu)$

$$A_R(\mu) := \{\tau : \tau =_i \mu\}$$

(28) R PRINCIPLE (REVISED)

Do not use μ if there is another expression $\tau \in A_R(\mu)$ such that both (i) $\tau <_c \mu$, and (ii) τ is assertable (i.e. the speaker believes τ is applicable and relevant in the context).

(29) MARKED MEANING PRINCIPLE

For any constructs μ, τ such that $\tau \in A_R(\mu)$ and $\tau <_c \mu$, using μ carries the implicature **atypical′**(τ')

---

4 The way **atypical′** is interpreted varies widely across contexts. For example, indirect speech acts also appeal to the hearer's inferential abilities when a more direct mode of expression is available (SEARLE, 1975; BACH; HARNISH, 1979). The question *Can you pass the salt?* indirectly conveys a request interpretation that is not grammatically licensed for that particular form (it is ungrammatical, in the sense of Chapter 4), but which could be directly encoded in terms of the simpler imperative *Pass the salt!*. In this case, the atypical interpretation associated with the marked form is related to politeness (LAKOFF, 1977; BROWN; LEVINSON, 1987).

Consider how this applies to the case of peculiar passives in (30).

(30) a.   This bed has been slept on (by someone).

  b.   Someone slept on this bed.

On the account being proposed here, the promotion of the prepositional object to the derived subject position of a verb in (30a) makes the overall structure ungrammatical, because the complement of the preposition does not correspond to a GF that belongs to the verb's PRED list. The failure of licensing lies, therefore, in the correspondence between the SYN and GF levels. Let us call this unlicensed pattern of linking between SYN and GFs P-PASS, in contrast to the canonical (i.e. licensed) linking between prepositional GF objects and sisters of P in SYN we see in (30b), which I will call N-PASS:

(31) a.   P-PASS$=\begin{bmatrix} \text{SYN} & [_S \text{ NP}_1 \text{ } [_{VP} \text{ V[PASSIVE]}_2 \text{ } [_{PP} \text{ P}_3 \text{ }] \text{ }]] \\ \text{GF} & \text{PRED}_2{:}\langle\text{GF, PRED*}_2{:}\langle\text{GF}_1 \ldots \rangle\rangle \text{ PRED}_3{:}\langle\text{GF}_1\rangle \end{bmatrix}$

  b.   N-PASS$=\begin{bmatrix} \text{SYN} & [_S \ldots \text{ } [_{VP} \text{ V}_2 \text{ } [_{PP} \text{ P}_3 \text{ NP}_1 \text{ }]\ldots \text{ }]] \\ \text{GF} & \text{PRED}_2{:}\langle \ldots \rangle \text{ PRED}_3{:}\langle\text{GF}_1 \ldots \rangle \end{bmatrix}$

I want to maintain that that both of the patterns in (31) are linked to literal interpretations that are true of exactly the same situations in the world – i.e. P-PASS $=_i$ N-PASS. It follows from this that N-PASS $\in$ A$_R$(P-PASS). However, ungrammaticality renders the P-PASS pattern more complex than N-PASS for the reasons mentioned above – i.e. N-PASS $<_c$ P-PASS. The only way to make sense of the fact that a speaker may use P-PASS in these conditions, the Marked Meaning Principle says, is if P-PASS conveys some atypical meaning as an implicature.

This atypical meaning associated with P-PASS is the rhetorical effect noted above: the inference that the derived subject was AFFECTED in some remarkable way by the event described. Whenever this affectedness implicature is compatible with speaker's general assumptions about the world and other features of the context, the peculiar passive is judged to be acceptable.

The reason why speakers infer *this* particular atypical meaning for peculiar passives, rather than other some other random one (e.g. that the derived subject denotes an object with an unusual color or size) has to do with the fact that what makes peculiar passives ungrammatical (and, therefore, complex) is the P-PASS pattern and not some other random violation of grammar (e.g. a failure of agreement) which would only produce unacceptability.

What is special about the P-PASS pattern is that it has N-PASS among its R alternatives. The P-PASS pattern is not, therefore, some random violation of the grammar, but one which is consistently associated with the same truth-conditions as N-PASS. I want to argue that the affectedness implicature emerges for P-PASS precisely because it is the complement (i.e. the negation) of the inference speakers would typically draw from N-PASS.

What one usually infers from a case where an NP is realized as the complement of a locative or an instrumental preposition is an interpretation where the P-object is not affected in any remarkable way by the event described in the sentence. If a speaker says *Bobby slept on this bed*, we would usually not take *the bed* as being significantly affected by Bobby's action. That is, the prototypical interpretation inferable from a locative P-complement instantiating N-PASS (i.e. being realized as a P-adjacent NP) is that of NON-AFFECTEDNESS. When this pattern is explicitly avoided by a speaker who uses P-PASS, the hearer will infer that the prototypical non-affectedness presumption associated with N-PASS is what is being explicitly avoided, and therefore, derive affectedness as an implicature.[5]

To summarize, the fundamental insight I take from Davison (1980) and Menuzzi (2004) is the somewhat paradoxical idea of using ungrammaticality to explain why a certain structures are acceptable. All that we need to accomplish this is a theory of pragmatic inference (the neo-Gricean approach outlined above) plus the assumption that ungrammaticality contributes to complexity. This approach allows the rhetorical and discourse effects associated with peculiar passives to be "factored out from individual constructions so that they are not part of the syntactic or pragmatic conditions on rule application" (DAVISON, 1980, pg. 63). In the terms I have been adopting, we can preserve the view that there is only one unified passive construction in English (namely, (15)), which applies only to direct objects and obliques, while, at the same time, accounting for the acceptability and rhetorical effects associated with peculiar passives.

In the next section, I show that this exact same idea can be extended to explain how logophoric readings come to be associated with the violations of ERC implicit in discursive occurrences of reflexives. This will allow us to maintain a simple unified account of reflexives – one where ERC is the core reflexive construction in English – in light of the apparently overwhelming variety that threatens previous unified approaches like the CBT.

---

[5] Since the particular violation of the grammar involved in P-PASS is consistently and recurrently associated with a stable interpretive effect – namely, the negation of the non-affectedness presumption typically conveyed by N-PASS – it is plausible to regard it as an instance of PARAGRAMMATICALITY in the sense of Chapter 4. As we will see in connection to discursive reflexives, the claim that a pattern of interpretation is paragrammatical is not necessarily inconsistent with the claim that it is conventionalized to some degree. The conventions in question can be conventions that govern language use rather than grammatical meaning conventions (MORGAN, 1978).

6.3   DERIVING LOGOPHORIC READINGS

Most theories of English reflexives merely take note of the phenomenal correlation between discursive occurrences and logophoric readings, without actually attempting to explain *why* this correlation holds. In this section, I show that the logophoric readings are derived from the same kinds of inferential pragmatic processes that gives rise to the affectedness effect associated with peculiar passives: the failure to be licensed by a constraint (i.e. ungrammaticality) is interpreted as markedness, which, in turn, leads to an implicature that consists in a denial of what would typically be conveyed by an unmarked alternative. This atypicality implicature, in the case of discursive reflexives, is the logophoric reading.

In the course of the analysis of discursive reflexives developed in Chapter 5, we found ourselves in the situation of having a set of sentences where reflexives are acceptable, but somewhat marked, and a set of syntactically parallel sentences where reflexives are unambiguously unacceptable. Examples of each of these groups are repeated below:

(32)   a.   *Bobby* said that Joanne invited Susan and *himself* for a drink.

         b.   *The men* said that the new administrators would be very much like *themselves*.

         c.   *Joanne* loves Bobby because Bobby resembles *herself*.

         d.   *Bobby* considers Joanne similar to *himself*.

(33)   a.   *Bobby said that Joanne invited *himself* for a drink.

         b.   *The men said that the new administrators would be very proud of *themselves*.

         c.   *Joanne loves Bobby because Bobby defends *herself*.

         d.   *Bobby considers Joanne hostile to *himself*.

As we saw in previous chapters, there is no relevant syntactic difference between (32) and (33). Attempts to make a distinction between these cases within non-syntactic parts of grammar are also somewhat implausible, given that the acceptability of non-local binding in (32) is subject to non-structural contextual effects, in contrast to what we find languages where discursive anaphors are genuinely grammaticalized. We are thus, faced with two options:

(34)   a.   OPTION 1: state our grammar so that it excludes both (32) and (33) and account for the acceptability of (32) on extra-grammatical grounds, or

         b.   OPTION 2: state our grammar so that it allows both (32) and (33) and explain the unacceptability of (33) on extra-grammatical grounds.

I adopt a version of Option 1. Option 2 constitutes a significant departure from tradition, insofar as it denies that there is a general syntactic constraint against the non-local binding of reflexives in English. As far as I know, the only proposal that approximates it is the tentative reformulation of the Binding Theory proposed in Pollard (2005). The main problem with this alternative is that it does not explain why the reflexives in (32), though acceptable, are still formally *marked* (i.e. restricted in their distribution), not only in comparison to pronominals (cf. (35)) but also with respect to plain locally bound reflexives (cf. (36)).

(35) a. *Bobby* said that Joanne invited Susan and *him* for a drink.

      b. *The men* said that the new administrators would be very much like *them*.

      c. *Joanne* loves Bobby because Bobby resembles *her*.

(36) a. *Bobby* invited *himself* for the party.

      b. *The men* are proud of *themselves*.

      c. *Joanne* defends *herself*.

Since Option 2 implies that the discursive reflexives in (32) are grammatically licensed in the same way as the pronominals and plain reflexives in (35)-(36), it does not explain why only the sentences in (32) are associated with marked interpretations: namely, logophoric readings, where the value of the pronoun is presumed to reflect the perspective taken by the speaker.

In addition to the animacy requirement mentioned in Chapter 3, logophoric readings have further properties that distinguish them from ordinary readings for pronouns and reflexives in contexts like (35)-(36). For instance, logophoricity implies that discursive reflexives will, for the most part, be in complementary distribution with epithets like *the idiot* (see Charnavel and Zlogar (2015) for some caveats). The reason for this is that, while discursive reflexives are logophoric (i.e. they must have an empathy locus as their value), epithets are *anti-logophoric* (i.e. they must *not* have an empathy locus as their value) (DUBINSKY; HAMILTON, 1998).

In the examples below, partly based on Kuno (1987, pg. 121, 131), the acceptability of reflexives *vs.* epithets flips depending on whether the antecedent is construed as an empathy locus or not. There is no such complementarity between epithets and pronominals, which are acceptable in all of these structures (JACKENDOFF, 1972; DUBINSKY; HAMILTON, 1998):

(37) a. According to *Bob*, the paper was written by Amy and {*himself* / **the idiot* / *him*}.

      b. Speaking of *Bob*, the paper was written by Amy and {**himself* / *the idiot* / *him*}.

(38)   a.   *Bob* says that the customers are similar to {*himself* / **the idiot* / *him*}.

        b.   Amy hired *Bob* because the customers are similar to {**himself* / *the idiot* / *him*}.

(39)   a.   *Bob* knew that the new guys would look a lot like {*himself* / **the idiot* / *him*}.

        b.   Amy said about *Bob* that the new guys would look a lot like {**himself* / *the idiot* / *him*}.

(40)   a.   *Bobby* would never say that Joanne resembles {*himself* / **the idiot* / *him*}.

        b.   Joanne would never say to *Bobby* that she resembles {**himself* / *the idiot* / *him*}.

(41)   a.   *John* was going to get even with Mary. That picture of {*himself* / **the idiot* / *him*} in the paper would really annoy her, as would the other stunts he had planned.

        b.   Mary was quite taken aback by the publicity *John* was receiving. That picture of {**himself* / *the idiot*/ *him*} in the paper had really annoyed her.

What we see above is that the distribution of discursive reflexives is more *restrictive* than that of pronominals. I interpret this as further evidence for the hypothesis that discursive reflexives receive *atypical* interpretations, reflecting their marked status *vis-à-vis* pronominals. The atypicality is the implication that the referent of the reflexive is a perspective-bearer.

The marked status of logophoric readings is particularly clear in the case of discursive reflexives within locative PPs (CANTRALL, 1974). Consider the examples in (42), which are meant to describe the scene depicted in Figure 4:

(42)   a.   *The man* is facing away from us, with {the dog / *the fireplace} behind *himself*.

        b.   *The man* is facing away from us, with {the dog / the fireplace} behind *him*.

Figure 4 – Man and dog standing near a fireplace (CHARNAVEL; ZLOGAR, 2015, pg. 11)
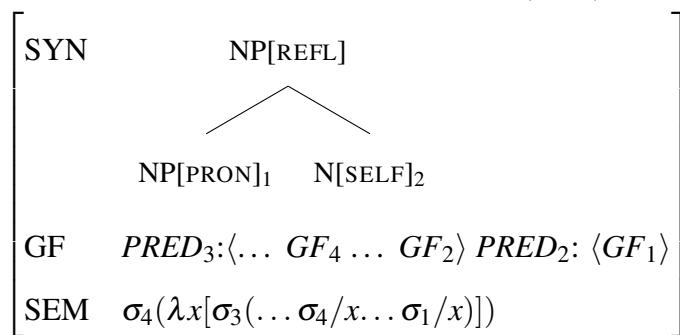
From the point of view of the man in the picture, the fireplace is in front and the dog is behind. If, however, we imagine ourselves from the perspective of a person observing the scene, we would see a man with a dog in front of him and a fireplace behind him. The contrast in (42) shows that, while the bare pronominal can describe the situation from both perspectives (the perspective of the man *and* the perspective of a third-party observer), the reflexive can *only* describe the scene from the man's perspective. What this shows us is that the reflexive forces a logophoric interpretation where its referent *necessarily* counts as a perspective-bearer for the sentence, while the pronoun is merely *compatible* with this interpretation.[6]

The general pattern we find, therefore, is that discursive reflexives are restricted to a set of readings which is properly contained in the set of readings associated with corresponding pronominals. In other words, the interpretation of discursive reflexives is more specific or *atypical* than that of pronominals. The atypicality consists in the implication of a particular perspective or point-of-view, which is not necessarily conveyed by the pronominal. Only an account that explicitly predicts the formal markedness of discursive reflexives can actually *explain* this correlation between marked forms and atypical meanings, rather than stipulate it.

All that we need for this purpose is (i) some explanation of why discursive reflexives are perceived to be marked, and (ii) some principle that associates marked forms with atypical meanings. This is exactly the same situation that we found ourselves in while analyzing peculiar passives. A similar account in terms of ACCEPTABLE UNGRAMMATICALITY is, thus, justified.

As with peculiar passives, the markedness of discursive reflexives follows from the fact that constitute violations of a grammatical construction – in the case of reflexives, the ERC:

(43)   ENGLISH REFLEXIVE CONSTRUCTION (ERC)

$$\left[\begin{array}{ll} \text{SYN} & \text{NP[REFL]} \\ & \overbrace{\text{NP[PRON]}_1 \quad \text{N[SELF]}_2} \\ \text{GF} & PRED_3:\langle \ldots GF_4 \ldots GF_2 \rangle \; PRED_2: \langle GF_1 \rangle \\ \text{SEM} & \sigma_4(\lambda x[\sigma_3(\ldots \sigma_4/x \ldots \sigma_1/x)]) \end{array}\right]$$

---

[6] Cantrall's (1974, pg.148-149) example illustrates the same point (my judgments are slightly different from his):

(i)   a.   *They* placed their guns, as *they* looked at it, in front of {*themselves* / *them*}.
       b.   *They* placed their guns, as I looked at it, in front of {*themselves / *them*}.

The reflexive is only acceptable if its value coincides with the perspective independently established by the parenthetical. The pronominal, in turn, can be used regardless of whose perspective is being empathized with.

However, the conclusion of section 5.3 was that discursive occurrences do not result from *any* random violation of ERC, but only violations that emerge in two specific circumstances: when a reflexive does not correspond to a GF argument of a PRED or when the reflexive's PRED is anti-reflexive. In both of these scenarios, there is simply no way a hearer could assign to the string containing the reflexive a structure (i.e. a pairing of PHON, SYN, GF and SEM) where ERC is fulfilled: violating ERC is inevitable. Therefore, when a speaker uses a reflexive in these contexts, they are, in a sense, OSTENSIBLY VIOLATING ERC. Discursive reflexives are essentially linked to these contexts where ERC is ostensibly violated.

The fact that ERC is ostensibly violated whenever a speaker uses a discursive reflexive structure (D-REFL) makes it clear to the hearer that the semantic effect of ERC – local variable binding – is not what is not what the speaker intends to communicate. Therefore, any instance of D-REFL will always evoke other structures not containing a reflexive that express the same truth-conditional content as D-REFL, but which the speaker is deliberately avoiding. Let us call these alternative structures N-REFL. Since each D-REFL always implies a violation of a grammatical constraint, it will have a good chance of being more complex (in the sense of (44)) than any N-REFL. If this is the case, D-REFL will count a marked structure.

(44)   EFFORT-BASED COMPLEXITY

For any constructs μ and τ, $\mu <_c \tau$ (i.e. μ is less complex than τ) *iff* using (e.g. processing, articulating or understanding) μ consistently involves less effort than using τ. If neither $\mu <_c \tau$ nor $\tau <_c \mu$, then $\mu =_c \tau$ (i.e. μ and τ are equally complex).

In most of the cases, N-REFL can be an equivalent structure containing a pronominal. For the sample D-REFL structures in (45), these N-REFL alternates are (46):

(45)   a.   *Bobby* said that Joanne invited Susan and *himself* for a drink.

b.   *Bobby* felt that nobody appreciated that picture of *himself.*

c.   *The men* said that the new administrators would be very much like *themselves.*

(46)   a.   *Bobby* said that Joanne invited Susan and *him* for a drink.

b.   *Bobby* felt that nobody appreciated that picture of *him.*

c.   *The men* said that the new administrators would be very much like *them.*

Consider what this implies in the context of the neo-Gricean approach outlined above (HORN, 1984; RETT, 2020). The relevant definitions are repeated below:

(47)   ENTAILMENT-BASED INFORMATIVITY

For any constructs μ and τ with corresponding SEM terms μ' and τ', $\mu =_i \tau$ (i.e. μ is informationally equivalent to τ) *iff* every SEM term M of type t containing μ' entails a semantic term T of type t containing τ' and vice versa, where M is identical to T except for the presence of μ' and τ'.

(48)   R ALTERNATIVES

For any construct μ, the set of R alternatives for μ is $A_R(\mu)$

$A_R(\mu) := \{\tau : \tau =_i \mu\}$

(49)   R PRINCIPLE (REVISED)

Do not use μ if there is another expression $\tau \in A_R(\mu)$ such that both (i) $\tau <_c \mu$, and (ii) τ is assertable (i.e. the speaker believes τ is applicable and relevant in the context).

(50)   MARKED MEANING PRINCIPLE

For any constructs μ, τ such that $\tau \in A_R(\mu)$ and $\tau <_c \mu$, using μ carries the implicature **atypical$'(\tau')$**

According to the definition of informativity in (47), D-REFL $=_i$ N-REFL. This makes N-REFL a member of the set of R alternatives of D-REFL – i.e. N-REFL $\in A_R$(D-REFL). Since the only noticeable difference between D-REFL and N-REFL (at least for (45)-(46)) is the presence of the *self* morpheme, which implies a violation of a ERC for D-REFL, it is straightforward to assume that N-REFL $<_c$ D-REFL. The R Principle prescribes, accordingly, that speakers should not use D-REFL, given that the existence of an equally informative and less complex competitor N-REFL. If speakers do, nonetheless, use D-REFL in such circumstances, the Marked Meaning Principle says that they must be implicating some atypical meaning other than the one normally associated with N-REFL. This is where the logophoric reading emerges.[7]

But *what* is the the meaning associated with N-REFL that speakers so ostensibly wish to avoid by using an ungrammatical D-REFL? To put it another way: what is the interpretive effect

---

[7]   I assume that, in the cases where ERC is NON-OSTENSIBLY VIOLATED, such as (i), the set of R alternatives is not sufficiently salient, given that the structures are all ones where ERC could in principle be fulfilled. Since there is no salient R alternative to the reflexive, the Marked Meaning Principle does not generate a logophoricity implicature to "save" the structure from being unacceptable.

(i)   a.   *Bobby* said that Joanne invited *himself* for a drink.
      b.   *Bobby* felt that nobody appreciated *himself*.
      c.   *The men* said that the new administrators would be very proud of *themselves*.

In other words, since an ERC-compliant interpretation is in principle possible for (i), it is simply too difficult for hearers to conjure up N-REFL R alternatives for such cases. The hearer sticks with an assignment of structure which is coherent with the grammar. This blocks the derivation of a logophoricity implicature.

typically communicated by N-REFL whose negation consists in a logophoric reading? I want to argue here, following the insights of Accessibility Theory, that logophoricity is inferred on the basis of a negation of the ACCESSIBILITY MARKING STATUS of the members of A_R(D-REFL) (ARIEL, 1990, 2001; ALMOR, 1999, 2000; ARNOLD; GRIFFIN, 2007, ARNOLD, 2010).

Accessibility Theory is builds on the idea that discourse representations are not flat, but involve a ranking of referents in accordance with their relative degree of accessibility – where the latter is understood as a property of non-linguistic mental representations that determines their ease of retrieval in real-time processing (ARNOLD, 2010). In the case of discourse referents, accessibility is plausibly a consequence of ATTENTION: i.e. a referent is more accessible to the extent that discourse participants devote more attention to it (GUNDEL; HEDBERG; ZACHARSKI, 1993; ARNOLD, 2010; NAPPA; ARNOLD, 2014; ARNOLD; LAO, 2015).

There are many linguistic factors that correlate with this cognitive property of accessibility: other things being equal, a more recently mentioned referent is more accessible than a less recently mentioned one; a referent invoked by a subject is more accessible than one invoked by the object, etc. (GIVÓN, 1983; BRENNAN, 1995; CHAMBERS; SMYTH, 1998; ARNOLD, 2008). As an illustration, consider the contrast in (51):

(51)   a. ??Charlie and Frank finished watching a movie. He didn't like it.
       b.   Charlie and Frank finished watching a movie. Charlie was the one who picked it out. He didn't like it.

The pronominal *he* is inappropriate in (51a) because both Charlie and Frank are equally accessible (i.e. equally attended to) in the context of the second clause. In (51b), on the other hand, *he* can successfully refer to Charlie, because Charlie is the most accessible referent at the point where the pronominal is encountered. In fact, it would be odd for a speaker to use *he* to refer to Frank in that context. The lower accessibility of Frank would justify repeating the name of the antecedent. Repetition of *Charlie* in (51b), in turn, would have been redundant and would have resulted in diminished acceptability (GORDON; HENDRICK, 1998).

This brief discussion indicates that, in order to communicate successfully, speakers must provide hearers with just enough cues to identify which discourse referent (or set thereof) they intend to pick out. This is made possible by the fact that different types of NP function as specialized markers for different degrees of accessibility (ARIEL, 1990, 2001). That is, there is a correlation between particular NP forms and the accessibility status of referents, which

instructs hearers in the task of referential processing. For my purposes, I assume the following Form-Accessibility Correlation (FAC), based on the distributional findings of Ariel (1990):

definite description   proper name   pronominal   reflexive

⟵————————————————————————————————————⟶

LOW ACCESSIBILITY   HIGH ACCESSIBILITY
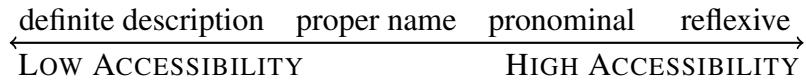
Figure 5 – The Form-Accessibility Correlation (FAC)

I assume that the correlations represented in the FAC are not encoded in the grammar or in lexical entries, but reflect inferential processing routines triggered by NPs of different types. The reason why NPs come to be associated with a processing routine that instructs hearers to infer a referent with a particular degree of accessibility is determined by factors such as informativity and prosodic salience (e.g. stress or length) (ARIEL, 1990).

Given this background, we can now understand *why* logophoric readings come to be associated with discursive reflexives. We saw above that structures containing discursive reflexives (D-REFL) are always more complex than truth-conditionally equivalent and salient R alternatives not containing a reflexive (N-REFL). The Marked Meaning Principle thus prescribes that, by using D-REFL when an N-REFL is available, the speaker must be signaling that an inference typically triggered by a less complex N-REFL is not intended.

I want to argue here that the inference associated with less complex N-REFL alternatives that speakers wish to deny by using D-REFL is the accessibility status of the NP forms involved in N-REFL. Consider the following illustration:[8]

(52)   a.   *Biden* said that a picture of *himself* is on the front page of every paper.

        b.   *Biden* said that a picture of *him* is on the front page of every paper.

When speakers use the marked D-REFL strategy in (52a), they are implying that the accessibility status associated with the less complex grammatically-compliant alternative (52b) is not the accessibility status they want to convey. According to the FAC, this can either mean that the speaker wants to convey that the referent of *himself* (Biden) has a higher accessibility than pronominals or a lower accessibility (akin to the accessibility signaled by names).

The latter is incompatible with the fact that Biden has been recently mentioned and can be unambiguously identified as the referent of the anaphoric expression. Therefore, only the former option applies: by using a D-REFL in (52a), the speaker wishes to convey that Biden

─────────────

8   Note that (52a), but not (52b), necessarily carries a *de se* reading: only (52b) is compatible with a scenario where Biden says "A picture of the president is on the front page of every paper" without knowing that he himself is the president. This is further confirmation of the marked status of discursive reflexives. See Chierchia (1989) and Postal (2006) for more discussion on the connection between logophoricity and *de se* readings.

has a higher accessibility status than what a pronominal would imply: i.e. that Biden is just as accessible as if it were the referent of a local subject. If this assumption is compatible with other features of the discourse (e.g. if other linguistic cues make Biden a plausible empathy locus in the context), the sentence containing D-REFL will be judged acceptable.

Therefore, in general, the atypical meaning a speaker intends to convey by violating ERC with D-REFL is that the accessibility status associated with bare pronominals is being avoided: i.e. that the semantic value of the reflexive is an entity (or a set thereof) that is highly accessible, or to borrow Schlenker's (2005a) terms, SUPER-SALIENT in the context. In addition to the speaker and the hearer, the entities whose perspective the speaker empathizes with (the EMPATHY LOCI) are always highly accessible in this sense (SELLS, 1987; GOLDE, 1999; OSHIMA, 2007; CHARNAVEL, 2019; VARASCHIN; CULICOVER; WINKLER, in press).

A caveat has to be made, however, for structures where discursive reflexives can occur, but a pronominal cannot, given *other* constraints in the grammar. Such cases are rare, but they do occur when reflexives are locally bound by a coargument of their semantic predicate within coordinate NPs and exclusion phrases (REINHART; REULAND, 1993, pg. 707):

(53)   a.   *Benjamin* praised Johanna and {*himself* / ??*him*}.

        b.   *Benjamin* shaved every barber except {*himself* / ??*him*}.

        c.   *Benjamin* will punish every barber apart from {*himself* / ??*him*}.

The first thing to note is that that the reflexives in (53) are indeed discursive: they are violations of ERC within anti-reflexive PREDs. The conjunction PRED is anti-reflexive because its semantics is idempotent. The exclusion PREDs linked to *except* or *apart from*, in turn, are anti-reflexive because its two arguments are explicitly affirmed to be disjoint: the referent of the pronoun in (55b)-(55c) is singled out from the set of barbers (SAFIR, 1992, pg. 40).

In spite of this, reflexives within these PREDs do not alternate with pronominals, since the latter would violate Reinhart and Reuland's (1993) semantically-based Reflexivity Condition B: i.e. all of the sentences in (53) have reflexive predicates in their SEM representations, so overt marking by a reflexive is necessary:

(54)   a. $\lambda x.[\textbf{praised}'(\text{AGENT: } x, \text{THEME: } x)]$

        b. $\lambda x.[\textbf{shave}'(\text{AGENT: } x, \text{PATIENT: } x)]$

        c. $\lambda x.[\textbf{punish}'(\text{AGENT: } x, \text{PATIENT: } x)]$

I do not have a fully developed account for why logophoric reflexives are acceptable in (53). But I would like to suggest the following idea (see Menuzzi (1999, pg. 182) for a different, more detailed account which is compatible with the same general approach advocated here). Reflexives in (53) are interpreted logophorically because they compete with a less complex R alternative containing an accented pronoun in a parenthetical-like phrase, which is separated from the host clause by comma intonation:

(55)　　a.　*Benjamin* praised Joanna – and *him*.

　　　　b.　*Benjamin* shaved every barber – except *him*.

　　　　c.　*Benjamin* will punish every barber – apart from *him*.

If (53) has (55) among its R alternatives and (55) is less complex than (53) (and it plausibly is, given that (53) violates ERC and (55) violates nothing), the structures in (53) will turn out to be subject Marked Meaning Principle. If the Marked Meaning Principle operates, logophoricity implicatures are derived along the lines discussed above.[9]

The essence of my proposal for deriving the logophoric readings of discursive reflexives can be summarized as follows. When speakers employ a discursive reflexive, they are using a structure that fails to be licensed by the core construction pertinent to reflexives in English: the ERC. Given that the violation of ERC that occurs with discursive reflexives is *ostensive* (i.e. there is no way to the speaker *could* be obeying ERC in those contexts), it can be interpreted as a form of markedness, or complexity, insofar as alternative structures for expressing the same meaning are made immediately salient. In other words, in cases where the grammatically licensed correspondences for the reflexive cannot apply in principle, it is always immediately clear that the speaker is trying to communicate something that could have been more simply and directly encoded by another construct – typically, one containing a pronominal.

Simply using a complex structure while a simpler truth-conditionally equivalent structure is salient and assertable would be a violation of the R Principle (cf. (49)). Therefore, if a speaker uses a discursive reflexive, the hearer immediately infers that some atypical meaning is being conveyed as an implicature (cf. (50)). This is the logophoric reading. In deriving the

---

9　Other structures that Reinhart and Reuland (1993) analyzed as instances of obligatory discursive reflexives – e.g. cases involving locative PPs, such as *Julie stepped on {*her / herself}* – turn out not to be discursive reflexives on my account. As we briefly saw in Chapter 4, there is independent motivation for assigning a covert "subject" GF to the PREDs linked to locative preposition, in addition to the GF linked to the overt P-complement (e.g. we get raising in *Julie seems in a bad mood*). In that case, the reflexive does turn out to correspond to a variable bound by a higher GF coargument (namely, the covert "subject" GF, which is controlled by *Julie*). This makes it a plain reflexive that instantiates ERC in a straightforward way.

logophoricity implicature, the hearer is guided by the association between the R alternatives of reflexives (e.g. pronominals) and their respective degrees of accessibility (cf. Figure 5).

By proposing a pragmatic derivation of logophoric readings, I do not mean to take the position that these readings are not conventionalized in any way. In fact, I think a plausible position is that they *are* conventionalized in a way external to the grammar proper: i.e. as a CONVENTION OF USAGE, rather than as CONVENTION OF MEANING, in the sense of Searle (1975). This would mean that the association between reflexives and logophoric readings, though not part of the grammar, may still be part of knowledge of language in a broader sense.

A typical example of a convention of usage which is not a convention of meaning is the practice of greeting someone by inquiring about the person's well-being (e.g. *How are you?*). As Morgan (1978, pg. 269) argues, these kinds of conventions are only incorporated into the grammar as conventions of meaning – i.e. as constructions in the SiSx sense – when the relation between what is said and what is conveyed becomes entirely arbitrary.

This is not yet the case with logophoric interpretations for discursive reflexives in English. Though these readings are plausibly conventionalized, there is still, as I have argued in this chapter, a pragmatic rationale that explains *why* they emerge in the contexts they do: in Gricean terms, they are CALCULABLE, though not necessarily calculated in real time (MORGAN, 1978, pg. 263). A convention of usage liberates speakers from *having* to perform on-line inferences to derive logophoricity every time they hear discursive reflexives.

Logophoric readings, are, therefore, examples of SHORT-CIRCUITED IMPLICATURES, as defined by Morgan (1978, pg. 274). This grants them a status akin to indirect speech acts (e.g. the use of questions to perform requests), while preserving a unified analysis, where ERC is the sole grammatical constraint governing the interpretation of reflexives in English.[10] In the next section, I briefly go over examples of other languages where the behavior of reflexives and logophoric anaphors differ from what we see in English. This variability is a motivation for the constructional approach adopted in this Thesis.

6.4 ANAPHORS IN OTHER LANGUAGES

The English Reflexive Construction (ERC) says that English reflexives must consist of a pronominal (NP[PRON]) and a *self* predicate (N[SELF]), with the variable that corresponds to

---

[10] I also do not wish to rule ou the view that logophoric readings are grammaticalized for some speakers. The discussion in Zribi-Hertz (1989), Baker (1995) and Golde (1999) suggests that this may be the case in varieties of British English. For such varieties, a unified analysis along the lines sketched here would not be pertinent.

NP[PRON] bound by a higher GF coargument of NP[REFL] as a whole. These properties are by no means universal to all anaphors (or even all reflexive anaphors). There are languages, such as Brazilian Portuguese (BP) where the primary grammaticalized 3rd person reflexive takes the form of a proclitic *se*, which is morphosyntactically simple and is necessarily bound by the subject of a verbal projection (the first GF of a PRED linked to a verb), regardless of whether the clitic itself is a direct object (cf. (56b)) or a secondary object (cf. (56c)).

(56)   a.   *O Roberto se* odeia.
             the Roberto SE hates
             '*Roberto* hates *himself*.'

       b.   *O Roberto se* apresentou para a   Marta.
             the Roberto SE presented   to   the Marta
             '*Roberto* presented *himself* to Marta.'

       c.   *O Roberto se* deu um presente.
             the Roberto SE gave a   present
             '*Roberto* gave a present to *himself*.'

Unlike what we see in English, the core grammaticalized reflexive in BP cannot be bound by a non-subject (even if the non-subject precedes it in linear order) nor within a non-verbal predicate (e.g. a PRED linked to a PP or an AP):[11]

---

[11]   For binding by non-subjects and within non-verbal predicates, BP resorts either to a secondary reflexive form – an oblique tonic pronoun *si* – or to the plain pronominals *ele/ela*, both of which are usually accompanied by the emphatic modifier *mesmo* ('same'), whose properties I examined in Chapter 3.

(i)   a.   Foi *o Roberto* que a   Joana protegeu {de *si mesmo* / d*ele   mesmo*}.
             was the Roberto that the Joana protected {of SE same   / of-PRON same}
             'It was *Roberto* that Joana protected from *himself*.'
       b.   O Roberto rolou *o   tapete* sobre {?*si mesmo* / *ele   mesmo*}
             the Roberto rolled the carpet over {SE same   / PRON same}
             '*Roberto* rolled the carpet over *itself*.'
       c.   *O Roberto* parece orgulhoso de {*si (mesmo)* / *ele   (mesmo)*}.
             the Roberto seems proud   of {SE (same)   / PRON (same)}
             '*Roberto* seems proud of *himself*.'

These non-clitic reflexive forms are obligatory in the examples above, but they can also optionally replace the clitic *se* in structures involving coargument binding by a subject:

(ii)   a.   *O Roberto* odeia {a *si (mesmo)* / *ele   (mesmo)*} .
             the Roberto hates {to SE (same)   / PRON (same)}
             '*Roberto* hates *himself*.'
       b.   *O Roberto* apresentou {? a *si (mesmo)* / *ele   (mesmo)*} para a   Marta.
             the Roberto presented   {   to SE (same)   / PRON (same)} to   the Marta
             '*Roberto* presented *himself* to Marta.'
       c.   *O Roberto* deu um presente para {*si (mesmo)* / *ele   (mesmo)*}.
             the Roberto gave a   present to   {SE same   / PRON same}
             '*Roberto* gave a present to *himself*.'

I will not explore the behavior of the reflexive *si*; my focus in what follows will be the clitic *se*, which belongs to the widely discussed class of Romance SE clitics (BAAUW; DELFITTO, 2005; LABELLE, 2008).

(57)   a.   *Foi *o Roberto* que a   Joana *se* protegeu.
            was the Roberto that the Joana SE protected
            'It was *Roberto* that Joana protected from *himself*.'

       b.   *O   Roberto rolou *o   tapete se* sobre
            the Roberto rolled the carpet SE over
            'Roberto rolled *the carpet* over *itself*.'

       c.   *O   *Roberto* parece *se* orgulhoso.
            the Roberto seems   SE proud
            '*Roberto* seems proud of *himself*.'

The constraints that account for the properties of BP reflexives will, therefore, look quite different from those that are responsible for licensing reflexives in English. In addition to superficial phonological discrepancies (e.g. the difference between *himself* and *se*), there are also morphological, syntactic and semantic differences between the reflexive forms in each language: the subject orientation, the fact that *se* is a clitic, etc.

The lexical entry in (58) represents the fact that the reflexive *se* clitic in BP (like all pronouns) is an NP that corresponds to a variable. Specifically, it corresponds to a variable of a peculiar type: a GAP VARIABLE. This is the same type of variable that is associated with gaps in unbounded dependency constructions (cf. Chapter 4).

(58)   REFLEXIVE CLITIC LEXEME

$$\begin{bmatrix} \text{PHON} & se_1 \\ \text{SYN} & NP_1 \\ \text{SEM} & x_1^{gap} \end{bmatrix}$$

In contrast to ordinary variables, gap variables only receive a semantic value if they are logically bound by some operator. This explains why BP reflexive clitics (as opposed to other kinds of pronouns, including, as we will see, English reflexives) cannot occur free in *any context*. The following example, for instance, only receives a sloppy reading, which is an indication that the variable that corresponds to *se* is necessarily bound:

(59)   *O Roberto se* odeia e   *o Lauro* também.
        the Roberto SE hates and the Lauro too
        '*Roberto* hates *himself* and so does Lauro.'   (sloppy reading only)

The following construction captures the core properties associated with the binding of the variable that corresponds to BP reflexive clitics:

(60) BRAZILIAN PORTUGUESE REFLEXIVE CLITIC CONSTRUCTION

$$\begin{bmatrix} \text{PRON} & se_2 \prec \varphi_3 \\ \text{SYN} & [_{\text{VP}} \text{ NP}_2 \ V_3] \\ \text{GF} & \text{PRED}_3{:}\langle GF_1(, GF_4)\rangle \oplus \langle GF_2{:}[3rd]\rangle \oplus (\langle \dots \rangle) \\ \text{SEM} & \sigma_1(\lambda x[\sigma_3(\sigma_1/x \dots \sigma_2/x)]) \end{bmatrix}$$

The construction above (which entails the more general construction for BP clitics) states that reflexive *se* must correspond to the second or third GF (the direct or oblique object) of a PRED. The SEM of this PRED, must, in turn, be such that the first GF within it (the subject) binds the SEM that independently corresponds to the clitic. Abstracting away from the morphophonogical details of each construction, (60) represents a specific sub-case of the structures licensed by ERC. However, (60) is more specific than ERC, because it requires the binder of the reflexive to be a subject (i.e. the SEM term that corresponds to GF_1).

The assumption that reflexive clitics are pronouns (i.e. NPs that correspond to variables) is not uncontroversial. According to Reinhart and Reuland (1991) and Baauw and Delfitto (2005), Romance *se* clitics are not NPs at all, but mere markers of a lexical bundling operation, which has the effect of reducing the arity of the predicate to which the clitic attaches to. If this is correct, the predicates in (56) would actually be intransitive.

However, even though this analysis might work for other Romance languages, I reject it for BP for the same reasons as Labelle (2008) and Sportiche (2020) reject it for French. One of these reasons has to do with the fact that *se* clitics in French and in BP, like the reflexive in English, allow PROXY READINGS – i.e. they can refer to entities that are representationally related to the denotations of their antecedents (JACKENDOFF, 1992; LIDZ, 2001; SAFIR, 2004; VARASCHIN, 2020). (61) can mean that Ringo saw his statue on the wax museum:

(61) *O Ringo se* viu no   museu de cera.
      the Ringo SE saw in-the museum of wax
      '*Ringo* saw *himself* in the wax museum.'   (himself = Ringo or statue)

Such readings are unavailable for predicates prefixed with *self-* in English (*self-hate*) and for the reflexive *si* clitic in Italian (GIORGI, 2007; BRUNETTO; ROEPER, 2018):

(62) *Ringo si* ammira nel   museo   delle cere.
      Ringo SE admires in-the museum of   wax
      '*Ringo* admires *himself* in the wax museum.   (himself = Ringo, *statue)

Lidz (2001, pg. 128) reports that proxy readings are also unavailable for the Dutch anaphor *zich* and for the verbal reflexive-marking strategy in Kannada, which is realized by means of the non-reflexive anaphor *tannu* along with the verbal reflexive morpheme *-kond*:

(63)   *Ringo* scheert *zich*.
        Ringo shaves SE
        "*Ringo* shaves *himself*."   (himself = Ringo, *statue)

(64)   *Hari tann*-annu nood-i-kond-a.
        Hari SE-ACC   see-PP-REFL-PST-3SG.MASC
        '*Hari* saw *himself*.'   (himself = Hari, *statue)

The reason why proxy readings are unavailable in (62)-(64) is arguably due to the fact that Italian *si*, Dutch *zich* and Kannada *tannu* are not NPs that correspond to independent variables in SEM. In other words, these items do not have a counterpart to the lexeme in (58). This makes these anaphors impossible targets for the proxy construction, whose SEM component is a function over *e*-typed variables (rather than λ-predicates). A simplified statement of this Proxy Construction, based on Varaschin (2020), is given below – where $\mathcal{R}$ is a meta-variable for a representational relation like **statue-of′**, **picture-of′**, etc. and ^ is Montague's (1974) cap operator, which signals that *x* is interpreted as an individual concept):[12]

(65)   PROXY CONSTRUCTION

$$\begin{bmatrix} \text{PHON} & \varphi_1 \\ \text{SYN} & NP_1 \\ \text{SEM} & \lambda x[\iota y[\mathcal{R}(y, {}^\wedge x)]](\sigma_1) \end{bmatrix}$$

The following is an example of an English construct licensed by the Proxy Construction:

---

[12]   Even in English, controlled subjects and null objects do not allow proxy readings (LIDZ; IDSARDI, 1998, pg. 122). This is expected if we take controlled subjects and null objects to be directly licensed as bound variables, rather than as full NPs that could be targeted by the Proxy Construction (see also Hornstein (2001, pg. 164)):

(i)   Ringo and I were in the wax museum. When I accidentally bumped against the statues of The Beatles ...

    a.   *Ringo* expected *himself* to fall.   (himself = Ringo or statue)
    b.   Ringo expected to fall.   ($\neq$ *Ringo* expected his statue to fall)

(ii)   Ringo and I were in the wax museum. All of a sudden...

    a.   *Ringo* began to undress *himself*   (himself = Ringo or statue)
    b.   Ringo began to undress.   ($\neq$ *Ringo* began to undress his statue)

   The fact that reflexive clitics in BP accept proxy readings is, therefore, *prima facie* evidence the view that they are genuine NPs that correspond to autonomous variables in SEM.

(66)
$$
\left[
\begin{array}{ll}
\text{PHON} & \text{Ringo}_4 \text{ saw}_3 \text{ him}_1\text{-self}_2 \\
\text{SYN} & [_S \text{ NP}_4 \; [_{VP} \text{ V}_3 \; [_{NP[REFL]} \text{ NP}[\text{PRON}]_1 \text{ N}[\text{SELF}]_2]]] \\
\text{GF} & \text{PRED}_3{:}\langle \text{GF}_4, \text{GF}_2 \rangle \; \text{PRED}_2{:}\langle \text{GF}_1 \rangle \\
\text{SEM} & \lambda P[P(\text{ringo})]_4(\lambda x[\textbf{see}'_3(\lambda P[P(\text{ringo})]_4/x, \iota y[\textbf{statue-of}'(y, \hat{}z_1)]/x)]) \\
& = \lambda P[P(\text{ringo})](\lambda x[\textbf{see}'(x, \iota y[\textbf{statue-of}'(y, \hat{}x)])])
\end{array}
\right]
$$

Lidz (2001) calls anaphors that impose strict semantic identity between coarguments PURE REFLEXIVES. A schematic outline of the construction that licenses pure reflexives (abstracting from the phonology of individual languages) is given in (67):

(67) PURE REFLEXIVE CONSTRUCTION (ITALIAN, DUTCH, KANNADA)
$$
\left[
\begin{array}{ll}
\text{SYN} & [_{VP} \text{ NP}[\text{PURE-REFL}]_2 \; V_3] \\
\text{GF} & \text{PRED}_3{:}\langle GF_1(, GF_4) \rangle \oplus \langle GF_2 \rangle \oplus (\langle \dots \rangle) \\
\text{SEM} & \sigma_1(\lambda x[\sigma_3(\sigma_1/x \dots x_2)])
\end{array}
\right]
$$

Note that (67) does not stipulate an order in PHON between the pure reflexive anaphor and the verb, since this has to be specified on a language by language basis (e.g. *si* precedes the verb in Italian, but *zich* generally follows the verb in Dutch). The reason why pure reflexives do not receive proxy readings becomes clear: they are directly licensed as bound variables in SEM, rather than being replaced by a bound variable like English and BP reflexives are. To put matters differently, pure reflexives are licensed as part of the construction (67), rather than being independently licensed by a lexical entry and plugged into a larger binding construction.

All such anaphors also share with BP *se* the property of SUBJECT ORIENTATION (i.e. they must be bound by the subject GF). This is a common trait of reflexives in various languages. Another example of this is the Norwegian reflexive *seg selv* (DALRYMPLE, 1993, pg. 93):

(68) a. *Jon* fortalte meg om *seg selv.*
          John talked about REFL
          '*Ola* talked about *himself*.'

      b. *Vi fortalte *Ola* om *seg selv.*
          we told   Jon about REFL
          'We told *Ola* about *himself*.'

Interestingly, Norwegian has an alternative reflexive form *ham selv* that shows an *opposite* behavior – i.e. it can be locally bound, but not by a subject (DALRYMPLE, 1993, pg. 138):

(69) a. **Ola* snakket om   *ham selv*.
          Ola talked   about REFL
          '*Ola* talked about *himself*.'

      b. Vi fortalte *Ola* om   *ham selv*.
          we told   Ola about REFL
          'We told *Ola* about *himself*.'

The contrast between *seg selv* and *hem selv* reveals that subject orientation cannot be fixed, once and for all, as a parametric choice for a language as a whole. What we need are lexeme-specific binding requirements, as in the simplified constructions below:[13]

(70)   NORWEGIAN REFLEXIVE CONSTRUCTIONS

a.
$$\begin{bmatrix} \text{PRON} & \text{seg} \prec \text{selv}_1 \\ \text{SYN} & \text{NP[REFL]}_1 \\ \text{GF} & \text{PRED}_3{:}\langle GF_2, (GF_4,)\, \text{GF}_1 \rangle \\ \text{SEM} & \sigma_2(\lambda x[\sigma_3(\sigma_2/x \ldots \sigma_1/x)]) \end{bmatrix}$$

b.
$$\begin{bmatrix} \text{PHON} & \text{ham} \prec \text{selv}_1 \\ \text{SYN} & \text{NP[REFL]}_1 \\ \text{GF} & PRED_3{:}\langle GF_2, GF_4, \text{GF}_1 \rangle \\ \text{SEM} & \sigma_4(\lambda x[\sigma_3(\ldots \sigma_4/x, \sigma_1/x)]) \end{bmatrix}$$

The subject orientation of *seg selv* also demonstrates another important point: namely, that the binding properties of particular anaphoric forms cannot all be deduced from their morphosyntactic composition, as more ambitious statements of the reflexivity-and-chains approach seem to imply (REINHART; REULAND, 1993). The Dutch reflexive anaphor *zichzelf*, though morphosyntactically quite similar to *seg selv*, does not display subject orientation:

(71) a. *Max* sprak met Lucie over *zichzelf*.
          Max spoke with Lucie about REFL
          '*Max* spoke with Lucie about *himself*.'   (REULAND, 2011, pg. 114)

      b. Ik vroeg *Piet* over *zichzelf*.
          I asked Peter about REFL
          'I asked *Peter* about *himself*.'   (HELLAN, 1991, pg. 41)

What we need, to use the the jargon of the Classical Binding Theory, is a separate binding condition for each pronoun in each language. Different binding conditions (i.e. constructions, in the present framework) share a basic vocabulary, consisting of notions such as PREDs, rankings of GFs within a PRED, and A-binding. Within this vocabulary, however, binding constraints can vary with a certain degree of freedom (MANZINI; WEXLER, 1987; DALRYMPLE, 1993).

As a further example of cross-linguistic variation, consider the case of *zibun* in Japanese.

---

[13] Even within BP a subject oriented reflexive clitic *se* coexists with a non-subject oriented tonic form *si*.

Like BP *se*, *seg selv* and pure reflexives, *zibun* is a subject oriented oriented reflexive:

(72) a. *John*-wa Bill-ni *zibun*-ni-tuite hanasi-ta.
          John-TOP Bill-DAT REFL-about   tell-PAST
          '*John* told Bill about *himself*.'   (KISHIDA, 2011, pg. 11)

      b. *John-wa *Bill*-ni *zibun*-ni-tuite hanasi-ta.
          John-TOP Bill-DAT REFL-about   tell-PAST
          'John told *Bill* about *himself*.   (KISHIDA, 2011, pg. 11)

Like the BP reflexive *se*, *zibun* also allows proxy readings (cf. (73)) (NISHIGAUCHI; KISHIDA, 2008, pg. 71). Unlike the BP reflexive, though, *zibun* also exhibits an animacy restriction: i.e. it cannot refer to an inanimate NP, as (74) shows:

(73) *John*-ga *zibun*-o   home-ta.
      John-NOM REFL-ACC praise-PAST
      '*John* praised *himself*.'   (himself = Ringo or statue)

(74) *Sono sinbun-ga*   kaze-ni *zibun*-o   hiroge-ta.
      that newspaper-NOM wind-in REFL-ACC unfold-PAST
      '*The newspaper* unfolded *itself* in the wind.'   (KISHIDA, 2011, pg. 11)

Since *zibun* does tolerate proxy readings, it is not a pure reflexive, in the sense of Lidz (2001). In practical terms, this means that it does correspond to an autonomous variable in SEM. Unlike BP *se*, however, *zibun* encodes, as a presupposition, the requirement that the variable it corresponds to be assigned an animate entity as a value.

(75)   ZIBUN LEXEME

$$\begin{bmatrix} \text{PHON} & \text{zibun}_1 \\ \text{SYN} & \text{NP}_1 \\ \text{SEM} & x^{gap}\{+\text{ANIMATE}(x)\}_1 \end{bmatrix}$$

Ignoring the PHON-SYN correspondence, the construction specifying the reflexive binding possibilities of *zibun* is identical to the BP reflexive clitic construction in (60):

(76)   JAPANESE REFLEXIVE CONSTRUCTION

$$\begin{bmatrix} \text{PHON} & \text{zibun}_1 \\ \text{SYN} & \text{NP}_1 \\ \text{GF} & PRED_3{:}\langle GF_2 \ldots \ GF_1 \ldots \ \rangle \\ \text{SEM} & \sigma_2(\lambda x[\sigma_3(\sigma_2/x \ldots \sigma_1/x)]) \end{bmatrix}$$

However, in addition to binding by a local subject, Japanese *zibun* also famously allows long-distance binging by a logophoric antecedent (KURODA, 1973; KUNO, 1987; SELLS, 1987; OSHIMA, 2004; KISHIDA, 2011):

(77) a. *Takasi*-wa Taroo-ni [Yosiko-ga *zibun*-o nikundeiru koto]-o
   Takasi-TOP Taroo-DAT [Yosiko-NOM REFL-ACC be-hating COMP]-ACC
   hanasita.
   told
   '*Takasi* told Taroo that Yosiko hated *him*.' (SELLS, 1987, pg. 452)

   b. *John*-wa [*zibun*-ga Bill-o tasuke-ta] to omot-teiru.
   John-TOP [REFL-NOM Bill-ACC help-PAST] COMP believe-ASP.PRES
   '*John* believes that *he* helped Bill.' (KISHIDA, 2011, pg. 25)

Unlike reflexive binding, logophoric binding of *zibun* is not subject-oriented (cf. (78)) and can even operate across separate clauses (cf. (79)). All that seems to be required is that the antecedent of *zibun* counts as a suitable perspective bearer in the context. The examples below come from Sells (1987, pg. 453-455):

(78) a. [Yosiko-ga *zibun*-o nikundeiru koto]-ga *Mitiko*-o zetuboo e
   [Yosiko-NOM REFL-ACC be-hating COMP]-gaNOM Mitiko-ACC desperation to
   oiyatta.
   drove
   'That Yosiko hated *her* drove *Mitiko* to desperation.'

   b. Taroo-wa *Takasi* kara [Yosiko-ga *zibun*-o nikundeiru to] kiita.
   Taroo-TOP Takasi from [Yosiko-NOM REFL-ACC be-hating COMP] heard
   'Taroo heard from *Takasi* that Yosiko hated *him*.'

(79) *Taroo*-wa kanasigat-tei-ta. Yosiko-ga Takasi-ga *zibun*-o hihansita noni
   Taroo-TOP sad-PROG-PAST Yosiko-NOM Takasi-NOM REFL-ACC criticized though
   bengosi-nakat-ta kara da.
   defend-NEG-PAST because COPULA
   '*Taroo* was sad. It is because Yosiko did not defend *him* though Takasi criticized *him*.'

When the referent of *zibun* does not count as a perspective bearer in the context, binding is impossible, as Sells's (1987, pg. 464-465) contrast in (80) illustrates. The problem with (80b) is that *it* ('go') indicates that the sentence is reporting the perspective of Yosiko:

(80) a. *Takasi*-wa [Yosiko-ga *zibun*-o tazunete-ki-ta node] uresigatta.
   Takasi-TOP [Yosiko-NOM REFL-ACC visit-come-PAST because] happy
   '*Takasi* was happy because Yosiko came to visit *him*.'

   b. *Takasi-wa [Yosiko-ga *zibun*-o tazunete-it-ta node] uresigatta.
   Takasi-TOP [Yosiko-NOM REFL-ACC visit-go-PAST because] happy
   '*Takasi* was happy because Yosiko went to visit *him*.'

These properties justify a separate construction for logophoric binding in Japanese. I assume something along the lines of (81), where the material in curly brackets represents a presupposition that the variable stands for a perspective bearer in the discourse.[14]

(81) JAPANESE LOGOPHORIC CONSTRUCTION

$$\begin{bmatrix} \text{PHON} & \text{zibun}_1 \\ \text{SYN} & NP_1 \\ \text{SEM} & \sigma_2\{+\text{PERSP}(\sigma_2)\}(\lambda x[\sigma_3(\ldots \sigma_2/x \ldots \sigma_1/x)]) \end{bmatrix}$$

In a case where *zibun* is bound extra-sententially (cf. (79)), σ_2 need not be present in sentence, but can be provided by the discourse context.

This discussion of Japanese serves to illustrate that it is not only the case that binding conditions need to be relativized to individual items (as noted above), but also that a single lexical item may participate in distinct binding constructions. Whenever this occurs, a unified analysis like the one I sketched for English is not desirable.

Returning to English, an important part of my proposal in Chapter 5 was the claim that reflexives contain pronominals as part of their internal structure. This yields the following prediction: if a particular reflexive does not comply with ERC, the pronominal within it should, like pronominals in general, be inherently ambiguous. It can either wind up bound in virtue of the general English Bound Anaphora Construction (cf. (82)) or remain free and be assigned an arbitrary referent, which the hearer will infer to be an empathy locus, by means of the neo-Gricean principles proposed in Section 6.3. (The construction licensing the free occurrence of the pronominal *him* is given in (83).)

(82) ENGLISH BOUND ANAPHORA CONSTRUCTION

$$\begin{bmatrix} \text{SYN} & [Y_3 \ldots [X_2 \ldots NP[\text{PRON}]_1]] \\ \text{GF} & PRED_3{:}\langle \ldots GF_4 \ldots \rangle \; PRED_2{:}\langle \ldots GF_1 \ldots \rangle \\ \text{SEM} & \sigma_4'(\lambda x[\sigma_3'(\ldots \sigma_4'/x \ldots \sigma_1'/x \ldots)]) \end{bmatrix}$$

---

[14] I do not attempt to formalize the particular concept of perspecive which is relevant for the interpretive constraint that holds of logophoric reflexives in Japanese. Oshima (2004) proposes that *zibun* is subject to two distinct kinds of point-of-view binding: logophoric binding and empathic binding. He argues that these two modes of binding are syntactically and semantically distinct. If his analysis is correct, we would need three, rather than simply two distinct binding constructions associated with *zibun*.

(83)   FREE PRONOUN CONSTRUCTION

$$\begin{bmatrix} \text{PHON} & him_1 \\ \text{SYN} & \text{NP[PRON]}_1 \\ \text{GF} & \textit{PRED}: \langle \textit{GF}, \dots \text{ GF}_1: \textit{[3st, sing, masc]} \rangle \\ \text{SEM} & x\{+\text{MASC}(x) \wedge -\text{SPEAKER}(x) \wedge -\text{ADDRESSEE}(x)\}_1 \end{bmatrix}$$

The prediction that discursive reflexives are ambiguous is in fact borne out, as the example in (84) makes clear (REINHART; REULAND, 1993, pg. 674):

(84)   Only *Bobby* enjoys looking at pictures of *himself*.

The sentence in (84) has two readings: one where Bobby is the only person who enjoys looking at his own pictures (the bound reading) and another where Bobby is the only person who enjoys looking at pictures of Bobby (the free reading). Only the former would be true of a situation where other people enjoy looking at pictures of Bobby; only the latter would be true of a situation where some other individual likes looking at their own picture.

I belabor this point in order to illustrate the fact that structures containing discursive reflexives, despite violating ERC, still instantiate the constructions associated with English pronominals. In other words, the pronominal part of the reflexive retains its semantic contribution as a free or bound variable even in when ERC fails to apply. Logophoric readings piggyback, as it were, on the pronominals contained within the reflexives that evoke them. This suggests the following hypothesis about the cross-linguistic manifestation of logophoric reflexives: a logophoricity implicature (as opposed to a genuinely grammaticalized logophoric reading like we see in Japanese) will only be possible for reflexive forms which include, within their structure, a pronominal that can support the the logophoric interpretation.

This is explains why we don't find logophoric interpretations for the reflexive clitic in languages like BP – even in contexts where anti-reflexivity is at stake:

(85)   *A *Maria* ama o marido dela   porque [ele *se* parece].
        the Maria loves the husband of-PRON because he   SE resembles.
        'Maria loves her husband because he resembles herself.'

Despite containing an ostensive violation of the BP Reflexive Clitic Construction (60), (85) does not yield a logophoricity implicature. The reason is arguably that the logophoric reading has nothing to latch onto: since *se* is not a pronominal, it does not have a long-distance

binding interpretation independently specified by the grammar.

This rationale also explains why, among Dutch anaphors, only *hemzelf* allows logophoric readings. The data below comes from Rooryck and Vanden Wyngaerd (2011, pg. 147-148):

(86) a. *Max* pochte dat de koningin Marie en {*hem-zelf*   / **zich-zelf* / **zich*} had
          Max boasted that the queen   Marie and {PRON-SELF / SE-SELF / SE}   had
          uitgenodigd voor een glas.
          invited   for a   glass
          '*Max* boasted that the Queen invited Mary and *himself* for a drink.'

      b. Het stoorde   hem dat ze een man zoals {hem-zelf   / *zich-zelf / *zich}
          it disturbed him that she a   man like {PRON-SELF / SE-SELF / SE}
          probeerde te versieren.
          tried   to attract
          'It angered *him* that she tried to attract a man like *himself*.'

      c. *Clara* had voldoende tijd om   te zien dat er,   behalve   {*haar-zelf*   /
          Clara had sufficient time COMP to see that there, apart.from {PRON-SELF /
          *zich-zelf / **zich*}, ook een man van de BBC aanwezig was.
          SE-SELF / SE}   also a   man from the BBC present   was
          '*Clara* found time to check that apart from *herself* there was a man from the

          BBC.'

Another important observation is that pragmatically derived logophoric readings, as we see in English, are only predicted to exist in languages where a reflexive binding construction is, in fact, embodied by a particular lexical item. In languages lacking a lexical form that grammatically encodes a reflexivizing function, there is nothing to be ostensibly violated in contexts which give rise to discursive occurrences. This seems to be the case in Old English (KÖNIG; SIEMUND, 2000a) and Traditional Jambi Malay (COLE; HERMON; YANTI, 2015). Therefore, my prediction for these kinds of languages is that there should be no forms that are consistently associated with logophoric interpretations in specific contexts (e.g. anti-reflexive PREDs), like we see in English.

Consider Traditional Jambi Malay, as discussed by Cole, Hermon and Yanti (2015). Though the pronominal *dio* can be used to signal local binding in (87), local binding is more strongly enforced if the intensifier element *dewe*ʔ follows the pronoun, as in (88):

(87) *Yanti* neŋoʔ *dio*ʔ di tipi.
      Yanti saw   PRON on TV
      '*Yanti* saw *herself* on TV.'

(88)   *Arna* cinto dɲan *dio*   deweʔ.
        Arna love with PRON INT
        '*Arna* loves *herself*.'

This could be taken to suggest that *deweʔ*, like the *-self* morpheme in English, is the phonological manifestation of a reflexive construction. Furthermore, also like English (and unlike BP), this construction would be one that contains a pronominal element (*dio*). On the basis of such facts, we could expect logophoric readings to emerge in the specific circumstances where they do in English. But this is not what we observe. In the following example, *dio* has a non-c-commanding antecedent, despite the fact that the predicate *cinto* is not anti-reflexive:

(89)   Mamaʔ *Nik* cinto *dio*   deweʔ.
        mother Nick love PRON INT
        '*Nick*'s mother loves only *him*.'

The reason why we don't see the same logophoric pattern in English and Traditional Jambi Malay is because, as suggested above, (*dio*) *deweʔ* is not, in fact, part of a grammaticalized reflexive construction. So nothing is in fact violated in an example like (89).

Likewise, if a language has grammaticalized logophoric interpretations for specific lexical item – as seems to be the case with the Japanese reflexive *zibun* – there should be no requirement that logophoric readings emerge only in the narrow circumstances where they emerge in English. This is also correct, as the data in (90) shows. Note that *zibun* is bound across an intervening subject despite the fact that the embedded predicate is not anti-reflexive.

(90)   Taroo-wa *Takasi*-kara [Yosiko-ga   *zibun*-o nikundeiru to]   kiita.
        Taro-TOP Takasi-OBL Yoshiko-NOM self-ACC hate   COMP heard
        'Taro heard from *Takasi* that Yoshiko hated *him*.'   (YASHIMA, 2015, pg. 136)

An important feature of my account is that there is no such thing as an EXEMPT ANAPHOR. Both in languages like Japanese, where logophoric readings are grammaticalized, as well as in English, where they emerge pragmatically, anaphors are always subject to some grammatical constraint or other. This is a corollary of my definition of well-formedness in Chapter 4: pieces of structure always need to instantiate some constraint in the grammar in order to be well-formed. Logophoric reflexives in English are instances of ACCEPTABLE UNGRAMMATICALITY, in the sense of Chapter 4, which means they are still subject to the ERC; that is, ERC is the only construction that would be relevant to license them, but they are not licensed by ERC.

What is somewhat paradoxical is that the failure to be licensed is a crucial part of what makes the structures containing discursive reflexives acceptable to most speakers.

6.5 CONCLUSION

The present chapter presented a proposal to derive the correlation between English discursive reflexives and logophoric readings. For this purpose, I employed the constructional approach developed in the preceding chapters and an independently motivated neo-Gricean theory of pragmatics, where the notions of informativeness and complexity play a central role (HORN, 1984; RETT, 2020). The outcome of this exercise can summarized in the following theses – many of which merit further typological and experimental investigation:

(i) Reflexives in SiSx are constructionally licensed, which means that there is fine-grained variation cross-linguistically, and this variation is not problematic.

(ii) There is only one type of grammatically licensed reflexive in English, which correlates semantic reflexivity and morphosyntactic reflexivity.

(iii) Complex reflexives like those in English are subject to licensing constraints on pronominals. Simple reflexives (e.g. BP SE clitics, Dutch *zich*) are not.

(iv) For this reason, it is possible to use reflexives in English in ways that do not meet the licensing conditions for reflexives but are nevertheless acceptable – these uses are marked.

(v) These reflexives depart from canonical interpretations in various ways. What they share is the markedness property, which is correlated with atypical interpretations.

(vi) In the case of discursive reflexives in English, markedness is interpreted as signaling a high degree of accessibility, which entails perspective (i.e. logophoricity).

(vii) Languages that do not have a reflexive-marking construction of the kind found in English cannot trigger logophoricity implicatures for reflexives like English does either.

(viii) The fact that logophoric interpretations are pragmatically motivated in English does not preclude them from being conventionalized in some way. However, the convention in question is a convention of usage rather than a convention of meaning established within the grammar (SEARLE, 1975; MORGAN, 1978; BACH; HARNISH, 1979).

(ix) When logophoric interpretations are genuinely grammaticalized by a particular item (e.g. as with Japanese *zibun* (OSHIMA, 2004)), they are not restricted to specific pragmatic contexts, as they are in languages where logophoricity emerges as an implicature.

There are certainly many details to be fleshed out in this overall picture. I have only scratched the surface of the typological diversity of reflexive forms. If I were to investigate this diversity with the same detail I dedicated to English, I could have arguably made an even stronger case for the constructional approach, where languages are allowed to differ extensively in how they grammatically encode of semantic notions like reflexivity and A-binding.

In this Thesis, I have looked mainly to languages where reflexivity is marked by particular forms used as arguments of predicates (i.e. languages that have reflexive pronouns). I have not looked languages where where reflexivity is marked on predicates themselves: e.g. by means of infixes, different roots or argument structure alternations (FALTZ, 1985; VOLKOVA; REULAND, 2014). All of these reflexive-marking strategies could presumably be characterized in constructional terms: i.e. as correspondences between the form of linguistic objects (PHON, SYN and GF), which varies widely, and their meaning (SEM), which is more constant.

Note that neither the ERC nor any other of the contructions mentioned throughout this section directly represent the properties the properties of referential (in)dependence and re-flexivizing function, which form the basis for the typology of pronouns I have been assuming throughout this Thesis. The reason for this is that these properties are not intrinsic properties of lexical items, as Reinhart and Reuland (1993) proposed, but, rather, semantic or functional in nature. This is precisely what gives them a universal basis. Referential dependency is a consequence of the processing factors that enter into the association between different types of NPs and degrees of accessibility, in the sense of Ariel (1990); reflexivizing function, in turn, as we will see in the next chapter, is a consequence of (effort-based) complexity – i.e. the same factor that plays a role in the neo-Gricean pragmatic approach explored in this chapter.

7 A SIMPLER SYNTAX OF PRONOMINALS: A CLOSER LOOK AT BRAZILIAN PORTUGUESE

Chapter 5 begun with an attempt to incorporate the basic insights of Reinhart's (1983) theory of bound anaphora into the constraint-based constructional theory of Simpler Syntax (SiSx). A crucial part of this exercise was the idea that bound anaphora is licensed by means of a construction associated with pronominals, rather than pronouns in general:

(1)   ENGLISH BOUND ANAPHORA CONSTRUCTION

$$\begin{bmatrix} \text{SYN} & [Y_3 \ldots [X_2 \ldots NP[\text{PRON}]_1]] \\ \text{GF} & PRED_3{:}\langle\ldots GF_4\ldots\rangle \; PRED_2{:}\langle\ldots GF_1\ldots\rangle \\ \text{SEM} & \sigma'_4(\lambda x[\sigma'_3(\ldots \sigma'_4/x \ldots \sigma'_1/x \ldots)]) \end{bmatrix}$$

The English Bound Anaphora Construction entails that pronominals can only be bound by GF antecedents outside of their local PRED. This approximates the basic idea underlying Condition B (CB) of the Classical Binding Theory. In this chapter, I discuss the extent to which something like CB is part of my theory of anaphora, attempting to motivate the claim that the syntactic anti-locality part of (1) is, not, in fact, universal, but, rather, a property of languages like English. I focus specifically on data concerning pronominals in Brazilian Portuguese (BP).

Recall that CB of the Classical Binding Theory is a syntactic constraint that says pronominals cannot be bound by their antecedents in a local domain (LEES; KLIMA, 1963; LANGACKER, 1969; CHOMSKY, 1981; CHOMSKY, 1986b; REINHART, 1983; POLLARD; SAG, 1994). The following are typical illustrations of phenomena that motivate it:

(2)   a. *Paul saw him in the mirror.

        b. *Joanne forgot to include her in the guest list.

        c. *Bobby thinks of him first, then others.

Over the years, there have been several attempts to derive CB effects from more principled assumptions (BURZIO, 1989, 1991; SAFIR, 2004; KAYNE, 2005; HICKS, 2009, REULAND, 2011); including, as we saw in Chapter 2, Reinhart and Reuland's (1993) proposal to reduce the syntactic residues of CB effects from constraints on movement (i.e. their Chain Condition). All of these approaches, however, assume that CB as it was originally stated embodies a deep truth about the syntax of anaphoric dependencies across the world's languages, reflecting an innate property of the human faculty of language.

As we saw in Chapter 3, this view is challenged by data from Brazilian Portuguese (BP). Counterparts of (2) containing 3rd person singular pronominals *ele/ela* (henceforth BP-PRONOUNS) are fully acceptable in certain dialects of BP, spoken mostly in the state of Minas Gerais, but also found in São Paulo (cf. Grolla (2011) and Grolla and Bertolino (2011), i.a.):

(3)   a. *O Paulo* viu *ele* no   espelho.
            the Paulo saw him in-the mirror

      b. *A Joana* esqueceu de incluir *ela* na   lista de convidados.
            the Joana forgot   to include her in-the list of guests

      c. *O Roberto* pensa primeiro n*ele*,   depois nos   outros.
            the Roberto thinks first   on-him, then   on-the others

Ever since Moreira da Silva (1983) reported these facts, three basic strategies to salvage the classical Binding Theory have emerged in the literature. The first of these, which I call the NOT-A-PRONOMINAL STRATEGY attempts to reconcile the data in (3) with CB by claiming that BP-PRONOUNS are not pronominals in Chomsky's (1981) sense, but essentially anaphors – i.e. defective forms governed by Condition A (LEMLE, 1985). We have seen an instance of this approach applied to Jambi and Frisian phenomena in Chapter 3 (REULAND, 2011, 2016). A second approach, which I call the NOT-BOUND STRATEGY, denies that (3) contains instances bound anaphora, suggesting, alternatively, that what we see there is mere accidental coreference (GALVES, 1986). The third strategy, which I call the NOT-LOCAL STRATEGY argues that, due independent properties of BP, the relationship between BP-PRONOUNS and their antecedents in (3) is not local, contrary to appearances (MENUZZI, 1999).

These accounts all purport to offer general explanations for why BP-PRONOUNS can take local antecedents on the basis of data like (3). However, few of them note that slight modifications of these examples make the local disjointness effects associated with CB reappear:

(4)   a. **O Paulo* viu *ele*.
            the Paulo saw him
            '*Paulo* saw *him(self)*.'

      b. **A Joana* esqueceu de elogiar *ela* na   festa.
            the Joana forgot   to praise her in-the party
            '*Joana* forgot to praise *her(self)* in the party.'

      c. **O Roberto* bateu primeiro n*ele*,   depois nos   outros.
            the Roberto hit   first   on-him, then   on-the others
            '*Roberto* hit *him(self)* first, then other people.'

This DUALITY OF BP-PRONOUNS presents a major puzzle for all existing approaches:

theorists can neither affirm that BP-PRONOUNS are run-of-the mill CB-abiding forms like their English counterparts, nor that these items are completely exempt from CB effects.

In order to address this issue, one must first confront why BP-PRONOUNS are capable of being locally bound *at all* and then, on the basis of a general answer to that, figure out what filters out cases like (4). None of the efforts mentioned above adequately meet the first these desiderata. The second is is rarely even addressed – the main exception to this trend being Menuzzi (1999, chap. 3). In this paper, I propose a fourth strategy (already suggested in Chapter 3) where the work traditionally done by CB (in various languages) is distributed into two separate factors – neither of which is a *syntactic* universal:

(i) Language-specific constructions that license bound-variable interpretations for pronominals only for antecedents that occupy different PREDs (e.g. (1)).

(ii) A pragmatic principle that associates simple unmarked forms with PROTOTYPICAL IN-TERPRETATIONS. Since pronominals are unmarked in contrast to anaphors, they cannot express reflexivity in predicates for which non-reflexive readings are prototypical (cf. (4)) (FALTZ, 1985; LEVINSON, 1991; COMRIE, 1999; MATTAUSCH, 2004).

It is the universal presence of the second factor and the occasional absence of the first that explains the peculiar binding patterns we observe across different languages. In particular, the dual behavior of pronominals in the Minas Gerais/São Paulo variety of BP (sometimes appearing to obey CB, sometimes not) follows from the fact that these expressions are only subject to the second factor. This means that, when there is no PROTOTYPICAL non-reflexivity involved, BP-PRONOUNS are free to be bound wherever they please. Similar patterns have been shown to hold in languages like Middle English (KÖNIG; SIEMUND, 2000a) and French (ZRIBI-HERTZ, 1995). English pronominals, on the other hand, are much more constrained, since, in addition to (ii), they must also obey (i).

In Section 7.1, I summarize the relevant BP data. Section 7.2 argues against previous efforts to explain why BP-PRONOUNS are capable of being locally bound: the NOT-A-PRONOMINAL STRATEGY, the NOT-BOUND STRATEGY and the NOT-LOCAL STRATEGY. What all of these have in common is the impetus to preserve the core of CB as a property of Universal Grammar (UG). I show that this is incompatible with the full range of data. Section 7.3 proposes that CB effects in languages like English are reducible to a set of lexical properties. Section 7.4 offers a pragmatic account for the residue of CB effects we see in BP. I conclude by exploring some of the broader consequences of the picture sketched in this chapter.

7.1 A SUMMARY OF THE DATA

The purpose of this section is to briefly recollect some basic concepts that will be useful throughout this discussion and to situate the data concerning locally bound BP-PRONOUNS within the broader system of anaphora in Brazilian Portuguese (BP). The BP judgments come from the 40-year literature on BP-PRONOUNS, as well as from formal experiments conducted by Grolla (2011), Grolla and Bertolino (2011) and Lacerda, Oliveira and Leitão (2014).

7.1.1 Basic concepts

Throughout this thesis, I have been using the term ANAPHORA as a general label for the phenomenon of coconstrual between linguistic expressions – where COCONSTRUAL refers to semantic dependencies formed by means of intended coreference and binding relations alike.

The logical notion of binding holds between variables and certain operators (e.g. $\forall, \exists, \lambda$). However, the linguistically relevant concept – which I called A-BINDING – concerns a relation between a variable and its antecedent. Reinhart (2006, pg. 171) defines the latter in terms of the logical notion as follows:

(5) α (A-)BINDS β iff α is the sister of a λ-predicate whose operator logically binds β.

This definition does not pertain to narrow syntax (SYN), but to the level of semantic structure (SEM) – i.e. a representation which supports inference and is subject to a model-theoretic interpretation (REULAND, 2011, pg. 34). As we saw in Chapter 2, binding in SEM is typically licensed by syntactic configurations where an antecedent is more syntactically prominent than a pronoun, where prominence is defined by c-command, o-command or some similar notion (REINHART, 1983; POLLARD; SAG, 1994; BÜRING, 2005). I have been assuming that only binding relations (in the sense of (5)) are regulated by grammatical constraints, and that coreference is essentially a pragmatic phenomenon (REINHART, 1983) .

In order to be bound, NPs must, first and foremost, correspond to variables in SEM. From this it follows without stipulation that r-expressions are never bound.[1] The only kinds of expressions that can be bound are PRONOUNS. Following Chomsky (1981), Pollard and Sag (1994) and many others, pronouns can be sub-classified into ANAPHORS and PRONOMINALS. I

---

[1]   By taking the notion of binding to be a semantic one, I am rejecting the syntactic definition of binding that is central to the Classical Binding Theory (CBT) (CHOMSKY, 1981), consistent with the argument made in Varaschin, Culicover and Winkler (in press) that there is no need for Condition C *per se* as a syntactic constraint.

view Conditions A and B of the CBT as well as the Chain Condition of Reinhart and Reuland (1993) as empirical hypotheses about the syntactic environments under which anaphors and pronominals can be bound. As Burzio (1991) notes, this presupposes that each of those pronoun types can be distinguished on the basis of features that are independent of those which are prescribed by the binding conditions themselves.

The distinction between anaphors and pronominals can be established on the basis of the semantic property of REFERENTIAL INDEPENDENCE: anaphors are referentially dependent while pronominals are, in principle *in*dependent (CHOMSKY, 1986b; ZRIBI-HERTZ, 1995; SAFIR, 2004; REULAND, 2011; SPORTICHE, 2013). As we saw in Chapter 1, the inherent dependency of anaphors is reflected in the fact that they cannot be used as deictics (SAFIR, 2003). Since pronominals are not referentially dependent in the same way, they can be used deictically to refer to entities that are not previously given in the discourse. The behavior of particular pronouns under deixis can, therefore, serve as an independent test to detect their membership in the class of anaphors or pronominals.

In addition to the distinction between anaphors and pronominals, I assume a further cross-classification of anaphors into REFLEXIVES and NON-REFLEXIVES, as per Reinhart and Reuland (1993). I interpret this distinction as being of a different nature than the one between anaphors and pronominals. The difference lies in the fact that the notions of reflexive and non-reflexive are essentially interdependent: an anaphor is defined as reflexive only in virtue of standing in a *marked* opposition with another pronoun, which thereby acquires a non-reflexive status. In other words, a reflexive anaphor is any referentially dependent form that counts as more *marked* or *complex* than any alternative pronoun that could carry the same meaning.

This can be understood in terms of neo-Gricean approach outlined in Chapter 6 (HORN, 1984; RETT, 2020). The relevant concepts are repeated below:

(6)   For any constructs μ and τ, $\mu <_c \tau$ (i.e. μ is LESS COMPLEX than τ) *iff* using (e.g. processing, articulating or understanding) μ consistently involves less effort than using τ. If neither $\mu <_c \tau$ nor $\tau <_c \mu$, then $\mu =_c \tau$ (i.e. μ and τ are equally complex)

(7)   For any constructs μ and τ with corresponding SEM terms μ' and τ', $\mu =_i \tau$ (i.e. μ is INFORMATIONALLY EQUIVALENT to τ) *iff* every SEM term M of type t containing μ' entails a semantic term T of type t containing τ' and vice versa, where M is identical to T except for the presence of μ' and τ'.

(8)   R ALTERNATIVES

For any construct μ, the set of R alternatives for μ is $A_R(\mu)$

$$A_R(\mu) := \{\tau : \tau =_i \mu\}$$

On the basis of these concepts, we can define a general concept of reflexive as follows (ZRIBI-HERTZ, 1995, pg. 359):

(9)   A anaphor μ counts as a REFLEXIVE in a language $\mathscr{L}$, if there is another pronoun τ in $\mathscr{L}$ such that $\tau \in A_R(\mu)$ and $\tau <_c \mu$.[2]

Basically, a reflexive is an anaphor that shares a basic semantic contribution with a pronominal, but which is, in some way (which might vary from language to language) more complex than a pronominal. This higher complexity can be phonological, morphosyntactic, or even be due to extra-grammatical factors that contribute to effort of use. Regardless of its causes, we will see in section 7.4 that complexity is what allows reflexives to signal coconstrual between coarguments of a typical transitive predicate. This definition is meant to derive the Reflexive Marking Universal 1 of Chapter 3. Table 4 displays some examples:

(10)   REFLEXIVE MARKING UNIVERSAL 1

Forms that signal reflexivity are at least as complex than forms that do not.

	Reflexive Forms	Non-reflexive Forms
Dutch	zichzelf	zich / hem
English	herself	her
Greek	ton eaftó tu	ton
Hebrew	(et) ʔacmo	oto
Turkish	kendini	onu
Oriya	nijaku	taaku
Lezgian	wič	am
Japanese	zibun	∅
Chinese	ziji	tā
German	sich	ihn
French	se	le

Table 6 – Reflexive Forms *vs.* Non-reflexive Forms (HASPELMATH, 2008, pg. 48)

---

[2] Recall from Chapter 3 that the only place where we actually need to appeal specifically to the notion of a reflexive was in the statement of Reflexivity Condition B, which the present chapter reinterprets as a pragmatic principle. It is no wonder, thus, that the notion of a reflexive is defined here in essentially pragmatic terms. Reflexivity Condition A was merged together with the Chain Visibility part of the Chain Condition and assumed to apply solely to the class of *grammaticalized* reflexives.

Though the two notions coincide in English, non-reflexives are not always pronominals: Dutch *zich* is referentially dependent (an anaphor) but, since it is no more complex than the pronominal *hem* (and also less complex than *zichzelf*), it does not function as a reflexive. In some languages, the higher complexity of reflexives comes to be systematically associated with an interpretive effect that enforces local binding between coarguments of a predicate. Such languages develop a reflexive construction, in the sense of Chapter 5. The forms associated with reflexive constructions are what I called GRAMMATICALIZED REFLEXIVES in Chapter 3.

The distinction between anaphors and pronominals is not pragmatic in the same way as the one between reflexives and non-reflexives. However, there have also been, as we have seen in Chapter 2, several attempts to construe the semantic difference between anaphors and pronominals as more than just a primitive semantic property of lexical items. Following Bouchard (1984), Reuland and Reinhart (1995) hypothesize that referential (in)dependence is determined by the inherent ϕ-feature specification of NPs – i.e. a morphosyntactic property:

(11)   An NP is referentially independent *iff* it is fully specified for ϕ-features.

Another line of reasoning says that the ϕ-deficiency of anaphors derives, not from their inherent features, but from the way they are syntactically composed. Anagnostopoulou and Everaert (1999) and Menuzzi (1999), following Helke (1979), argue that English reflexive forms are compounds which are headed by (and, thus, inherit their ϕ-deficiency from) the *-self* morpheme, which is only specified for number (see also Chomsky (1981, pg. 102)).

Zribi-Hertz (1995), in turn, proposes that at least some instances of referential defectiveness should be explained on prosodic grounds. She claims that the fact that the French intensive pronoun *lui-même* and the English complex possessive *his own* cannot be used as a deictics is due to the fact that adjunction of an emphatic modifier causes deaccenting on the pronoun, thereby signaling that its referent is GIVEN in the discourse context (ROCHEMONT, 2016):

(12)   Donnez un livre à lui, lui et {lui / *lui-même}.   [pointing gestures]
          give   a   book to him, him and {him / him-same}
(13)   Take her coat, not {his / *his own}!   [pointing gesture]

Regardless of how one conceptualizes referential (in)dependence in general (as lexically, morphosyntactically, or prosodically determined), the point remains that it serves as a semantic property which is independent of any specific theory regulating the conditions under which NPs can be bound. This allows us to say that the property of being an anaphor, for instance, carries no

implication about whether anaphoric NPs must have antecedents that obey particular structural requirements (e.g. locality). In this framework, the correlation between anaphors/pronominals and structural requirements is an empirical matter, rather than a stipulation. This is, in fact, the way the Binding Theory has traditionally been interpreted (CHOMSKY, 1981, 1986b).

In this light, we can take Condition B as making the empirical claim that referentially independent pronouns – i.e. forms such as *him*, but not *himself* – cannot be bound by other NPs in a local domain (roughly, the domain of the closest subject). This is also the claim made by the Chain Economy part of Reinhart and Reuland's (1993) Chain Condition. We will see below that data from BP appear to contradict this observation.

7.1.2 Pronouns in Brazilian Portuguese

Let us now turn to how different sorts of pronouns in BP relate to the taxonomy discussed above, focusing on the divide between anaphors and pronominals. There is considerable diachronic and synchronic variation in the BP system of anaphora (see Moura Neves (2008) for a useful survey). The following discussion deals with the standard dialect spoken by educated speakers in major urban centers in the southeast of Brazil, encompassing, specifically, the states of São Paulo and Minas Gerais. Despite not being a speaker of this dialect, I also share all of the judgments reported here. However, I leave open the extent to which the phenomena I discuss exist in other varieties spoken across the country.

My choice rests on the assumption that the facts concerning the repertoire of pronouns and their behavior with respect to anaphora is sufficiently coherent within this São Paulo/Minas Gerais variety of BP. This is not uncontroversial. Many of the phenomena reported here – in particular, locally bound BP-PRONOUNS – have often been described as specific properties of the dialect spoken in Minas Gerais (the *mineiro* dialect), which are claimed to set it apart from *all* other varieties of BP (LEMLE, 1985; MENUZZI; LOBO 2016).

Other studies, however, suggest that these alleged idiosyncrasies of the *mineiro* dialect are also found (perhaps to a slightly lesser extent) in the São Paulo dialect (GALVES, 1986; PEREIRA, 2007; GROLLA; BERTOLINO, 2011; GROLLA, 2011). Forms that have been claimed to be present solely in the non-*mineiro* varieties (e.g. SE clitics) have also been shown to not be entirely absent from the *mineiro* dialect either (D'ALBUQUERQUE, 1984; ROCHA, 1999; MELO, 2005; LACERDA; OLIVEIRA; LEITÃO, 2014). With the proviso that we are dealing with a dynamic system in flux and that some of the tendencies I discuss might be more

entrenched in Minas Gerais than in São Paulo, the idealization of a hybrid São Paulo/Minas Gerais system of anaphora seems reasonably justified.

Table 7 provides an approximate sketch of the pronoun forms used in this particular variety of BP, which I will henceforth refer to simply as BP. The person and number features reflect the semantics of the pronouns, rather than their morphosyntactic properties. Forms that stand in free variation within the same semantically defined paradigm are listed vertically, while different gendered realizations of a single slot are separated by a forward slash, as in MASC/FEM.

		Singular			Plural		
		1st	2nd	3rd	1st	2nd	3rd
NON-REFLEXIVE	SUBJ	eu	você	**ele/ela**	a gente nós	vocês	**eles/elas**
	OBJ	me-	te- você	**ele/ela**	a gente nos-	vocês	**eles/elas**
	PREPOSITION GOVERNED	mim	você	**ele/ela**	a gente	vocês	**eles/elas**
REFLEXIVE	OBJ	me-	se-	se-	se-	se-	se-
	PREPOSITION GOVERNED	mim	si	si	si	si	si

Table 7 – The BP system of anaphora (adapted from Menuzzi and Lobo (2016, pg. 341))

What I have been calling BP-PRONOUNS are simply the 3rd person non-reflexive forms that appear in boldface in Table 7. In order to unravel the nature of these expressions, it will be useful to compare them with the other 3rd person forms capable of expressing local binding, in particular, the reflexive SE forms (the clitic *se* and its preposition-governed counterpart *si*, which is generally followed by the emphatic modifier *mesmo*).

The reason for focusing on 3rd person instead of 1st and 2nd person forms is twofold. First, the status of 1st and 2nd person pronouns as anaphors is harder to pin down due to the ubiquitous presence of a speaker and a hearer in any discourse context (ROSS, 1970; STIRLING, 1993; SCHLENKER, 2005a; KISS, 2009; VARASCHIN, 2020). Second, locally bound 3rd person pronominals are much more of a novelty in the literature than 1st or 2nd person ones, which are attested in all Germanic languages except English (FALTZ, 1985; HASPELMATH, 2008; REULAND, 2011). Consider the following examples from German and Dutch (ROORYCK; VANDEN WYNGAERD, 2011, pg. 18, 31):[3]

---

[3] Some accounts of binding phenomena draw significant conclusions about the nature of CB from the observation that locally bound pronominals are only (if ever) acceptable in the 1st and 2nd person (REULAND, 2011; CULICOVER, 2021). If BP-PRONOUNS are indeed pronominals, accounts which restrict CB exemption to 1st and 2nd person forms can't be the whole story.

(14)   a.   *Ich* liebe *mich*.   (German)
            I   love me

       b.   *Ik* heb *me* gewassen.   (Dutch)
            I   have me washed

I will not attempt to address here why local disjointness is often suspended for 1st and 2nd person pronominals. However, I believe a plausible answer is suggested by Culicover (2021): unless there is a specific grammatical construction enforcing anti-locality, local binding for 1st and 2nd person is allowed because these forms are not ambiguous in the same way 3rd person forms are (see also Dowty (1980)).

In addition to SE forms, BP also expresses local 3rd person binding with the complex emphatics (e.g. *ele mesmo*), whose syntactic and semantic properties are similar to those of the French pronoun *lui-même* (ZRIBI-HERTZ, 1995). As we saw in Chapter 3, these forms present their own set of problems for the Classical Binding Theory: e.g. they are referentially deficient (cf. (15)), but, unlike anaphors, they can appear as subjects and need not have local c-commanding antecedents, as shown in (16) (GROLLA, 2011, pg. 78-80):

(15)   a.   *Dê um livro pra ele, pra ele e   pra ele mesmo.   [pointing gestures]
            give a   book to him to him and to him same

(16)   a.   *O Roberto* disse pra   Joana que *ele mesmo* comeu o   bolo.
            the Roberto said to-the Joana that him same   ate   the cake
            '*Roberto* said to Joana that *he himself* ate the cake.'

       b.   A atitude da   *Amy* prejudicou *ela mesma*.
            the attitude of-the Amy damaged   her same
            '*Amy*'s attitude damaged *her(self)*'

Since *ele* and *ele mesmo* have the same informative content (*ele =$_i$ ele mesmo*), but *ele mesmo* is more complex than *ele* (*ele <$_c$ ele mesmo*), *ele mesmo* counts as a reflexive, according to (9). However, unlike *himself* in English, *ele mesmo* does not *enforce* reflexivity: i.e. its semantic contribution as a reflexivizer of predicates is not specified by the grammar. What this means, in terms of SiSx, is that there is no construction that obliges a reflexive interpretation for the predicate where *ele mesmo* appears as an argument like the English Reflexive Construction does for *himself*.[4] In light of this somewhat dubious behavior, I refrain from discussing emphatic forms and focus on contrasting BP-PRONOUNS with SE pronouns, whose status with respect to the Classical Binding Theory can be established less controversially.

---

[4]   Vieira (2015) argues at length for the same point. He provides evidence from acquisition that suggests *mesmo* is a scalar focus particle in BP, rather than the morphological manifestation of a reflexive construction.

Among the two kinds of pronouns considered here, only BP-PRONOUNS pass the deictic test for referential independence:

(17) a. Olha pra ele!   [pointing gesture]
          look at   him

      b. *Se olha!   [pointing gesture]
          SE look

      c. *Olha pra si mesmo!   [pointing gesture]
          look at   SE same

This is the behavior expected under Reinhart and Reuland's (1993) reduction of referential (in)dependence to ϕ-feature specification, since, as Table 7 shows, SE forms are unspecified for person, number and gender, while BP-PRONOUNS are fully specified (i.e. *ele* uniquely stands for 3rd person masculine singular).

The data in (17) also makes sense from the point of view of Zribi Hertz's (1995) prosodic account of referential dependency because SE forms are either clitics (cf. (17b)) or preposition-governed (cf. (17c)), in which case they are generally accompanied by the emphatic modifier *mesmo*. In both of these cases, SE forms lack nuclear stress and are, thus, arguably too prosodically weak to support independent reference. (See Ariel (1990, 2001) for more on the correlation between prosodic strength and referential autonomy.)

This strongly suggest that, by independent criteria, SE forms count as ANAPHORS and BP-PRONOUNS as PRONOMINALS. If the Binding Theory provides an accurate picture of the interpretation of such items, BP-PRONOUNS should, thus, differ from SE forms with respect to the domain in which they can be bound.

Condition A predicts that SE forms should be unacceptable in non-local binding configurations. There should be no such restriction for BP-PRONOUNS. This is in fact what we observe:

(18) a. *A Amy disse que o   Paulo *se* ama.
          the Amy said that the Paulo SE loves
          '*Amy* said that Paulo loves *her(self)*.'

      b. A Amy disse que o   Paulo ama *ela*.
          the Amy said that the Paulo loves *her*
          '*Amy* said that Paulo loves *her*.'

By the same token, Condition B (CB) predicts BP-PRONOUNS to be deviant in local binding environments. This seems to be correct if we look at (19) (GALVES, 1986; MENUZZI, 1999; GROLLA; BERTOLINO, 2011; MENUZZI; LOBO, 2016; CARVALHO, 2019):

(19)   a.   *O *Paulo* viu *ele*.
                the Paulo saw him
                '*Paulo* saw *him(self)*.'

        b.   *A *Joana* esqueceu de elogiar *ela* na   festa.
                the Joana forgot   to praise her in-the party
                '*Joana* forgot to praise *her(self)* at the party.'

        c.   *A *Amy* bateu primeiro *nela*, depois nos   outros.
                the Amy hit first   on-her, then   on-the others
                '*Amy* hit *her(self)* first, then other people.'

        d.   *A *Amy* ficou com ódio d*ela* depois de ter   ido   mal na   prova.
                the Amy was with hate of-her after   of having gone badly in-the exam
                '*Amy* got angry at *her(self)* after having done badly in the the exam.'

        e.   *O *Pedro* não defendeu *ele* na   festa.
                the Pedro not defended him in-the party
                '*Pedro* didn't defend *him(self)* at the party.'

        f.   *A *Susana* ficou discordando d*ela* a   noite toda.
                the Susana stayed disagreeing of-her the night all
                '*Susana* kept disagreeing with *her(self)* all night.'

        g.   *O *Roberto* deu um carro pra *ele*.
                the Roberto gave a   car to him
                '*Roberto* gave a car to *him(self)*.'

The problem, however, as we saw above, is that slight modifications of sentences like those in (19) make binding of a BP-PRONOUN by a local coargument fully acceptable. The examples below are all adapted from Moreira da Silva (1983), Lemle (1985), Galves (1986), Menuzzi (1999), Grolla (2011), Grolla and Bertolino (2011) and Carvalho (2019):

(20)   a.   *O Paulo* viu *ele* no   espelho.
                the Paulo saw him in-the mirror
                '*Paulo* saw *him(self)* in the mirror.'

        b.   *A Joana* esqueceu de incluir *ela* na   lista de convidados.
                the Joana forgot   to include her in-the list of guests
                '*Joana* forgot to include *her(self)* in the guest list.'

        c.   *A Amy* pensa primeiro n*ela*, depois nos   outros.
                the Amy thinks first   on-her, then   on-the others
                '*Amy* thinks of *her(self)* first, then of others.'

        d.   *A Amy* ficou com vergonha d*ela* depois de ter   ido   mal na   prova.
                the Amy was with shame   of-her after   of having gone badly in-the exam
                '*Amy* was ashamed of *her(self)* after having done badly the exam.'

        e.   *O Pedro* não reconheceu *ele* na   foto.
                the Pedro not recognized him in-the photo
                '*Pedro* didn't recognize *him(self)* in the photo'.

    f.   *A Susana* ficou falando d*ela* a   noite toda.
         the Susana stayed talking of-her the night all
         '*Susana* kept talking about *her(self)* all night.'

    g.   *O Roberto* comprou um carro pra *ele*.
         the Roberto bought   a   car   for him
         '*Roberto* bought a car for *him(self)*.'

This DUALITY OF BP-PRONOUNS is difficult to characterize in syntactic terms. There appears to be no syntactic generalization that distinguishes (19) from (20) in a general way (see Section 7.2.3 for more discussion). Rather, the difference here seems to be related to semantic/pragmatic aspects of the contexts in which BP-PRONOUNS appear (MENUZZI, 1999). So, for instance, whatever makes the BP-PRONOUN acceptable in (20c) in contrast to (19c) has to hinge on a semantic distinction between the predicates *pensar* ('think') and *bater* ('hit'). As I discuss in Section 7.4, the crucial issue is the PROTOTYPICALITY of a (non-)reflexive interpretation. The more prototypical a non-reflexive interpretation is for a predicate, the less acceptable are the locally bound BP-PRONOUNS contained within it.

Experiments by Grolla (2011) and Grolla and Bertolino (2011) confirm that speakers accept local binding of BP-PRONOUNS in some cases and reject it in others. In a study using Chien and Wexler's (1990) "Simon says" methodology, Grolla (2011) found that adults accept 40% of locally bound occurrences of BP-PRONOUNS with the predicates *coçar* ('scratch'), *abanar* ('fan'), *beliscar* ('pinch') and *cheirar* ('smell'). In a subsequent truth-value-judgment task experiment with 40 adult speakers from São Paulo, Grolla and Bertolino (2011) found a 60% acceptance rate for locally bound BP-PRONOUNS across a wider range of predicates.[5]

---

5   Bertolino and Grolla (2012) and Bertolino (2013) found lower acceptance rates for locally bound BP-PRONOUNS in the object position of transitive predicates. However, these subsequent experiments adopted a slightly different methodology, where the test sentences were immediately preceded by lead-ins containing another possible referent for the BP-PRONOUN, other than the local coargument subject. Alternative referents were also salient for the test sentences in Grolla and Bertolino (2011), but these were not evoked by immediately preceding lead-ins. The purpose of this change in experimental design was to make a locally free reading for BP-PRONOUNS more contextually plausible to participants. It is not clear, however, whether this is as relevant as the authors take it to be. Even in the absence of salient sentence external referents, adult English speakers overwhelmingly judge sentences corresponding those in (20) as unacceptable (NICOL; SWINNEY, 1989; CHIEN; WEXLER, 1990; ASUDEH; KELLER, 2001; KELLER; ASUDEH, 2001):

(i)   a.   *Paul saw *him* in the mirror.
      b.   *Joanne forgot to include *her* in the guest list.
      c.   *Amy felt ashamed of *her* after having done badly in the exam.

In other words, the disjointness effects enforced by CB in languages like English are nearly impossible to override, even if the interpretation that violates CB is the most salient one available in the context. Furthermore, it is possible that, by adding a lead-in sentence explicitly invoking sentence-external discourse referent, Bertolino and Grolla (2012) and Bertolino (2013) made the sentence-external discourse referent *too* salient to participants, who simply chose the more plausible among two grammatically licensed interpretations.

Neither of these experiments had the type of predicate as an independent variable. This gives the appearance of chance-level performance.[6] The crucial observation, however, is that adult English speakers don't even *appear* to perform at chance-level in these settings: they reject locally bound pronominals regardless of predicate-type, as predicted by CB (NICOL; SWINNEY, 1989; CHIEN; WEXLER, 1990; ASUDEH; KELLER, 2001).

Another relevant fact is that SE forms are acceptable (with no detectable meaning difference) in all of the local contexts in (19)-(20) – including, crucially, those where BP-PRONOUNS can also be locally bound, as (21) illustrates:

(21)   a. *O Paulo se viu no espelho.*
          the Paulo SE saw in-the mirror

        b. *A Joana esqueceu de se incluir na lista de convidados.*
          the Joana forgot to SE include in-the list of guests

        c. *A Amy pensa primeiro em si mesma,* depois nos outros.
          the Amy thinks first on SE same, then on-the others

        d. *A Sara ficou com vergonha de si mesma* depois da prova.
          the Sara was with shame of SE same after the exam

        e. *O Pedro não se reconheceu na foto.*
          the Pedro not SE recognized in-the photo

        f. *A Susana ficou falando de si mesma a noite toda.*
          the Susana stayed talking of SE same the night all

        g. *O Roberto comprou um carro pra si mesmo.*
          the Roberto bought a car for SE same

This data is particularly important because it undermines competition-based accounts of CB (LEVINSON, 1987; MENUZZI, 1999; HORNSTEIN, 2001; SAFIR, 2004; MENUZZI; LOBO, 2016). These kinds of approaches are more flexible than the classical Binding Theory because they do not posit CB as a syntactic primitive. According to them, locally bound pronominals are not excluded in principle – they are allowed in cases where anaphors are not available as alternative ways to express coconstrual. This can happen either when the language

---

6   In another experiment, which involved 51 children (ages between 4 and 6), Grolla and Bertolino (2011) controlled for predicate-type and did find a statistically relevant effect: locally bound BP-PRONOUNS were accepted 72% of the time for the prototypically reflexive predicate *coçar* ('scratch') and at lower rates for prototypically non-reflexive predicates like *bater* ('hit') (33%) and *cheirar* ('smell') (28%). However, this effect was not replicated in Lacerda, Oliveira and Leitão's (2014) off-line experiment conducted with a group of 25 adult students from Minas Gerais. These subjects accepted locally bound BP-PRONOUNS 62% of the time with prototypically non-reflexive predicates and 54,2% of the time for neutral or prototypically reflexive predicates. This discrepancy shows that more studies are needed to experimentally confirm the relevance of predicate-type, which is attested in the theoretical literature (MENUZZI, 1999) and in corpus-based studies (PEREIRA, 2007).

as a whole lacks a dedicated reflexive anaphor or when the anaphors of the language cannot be used in a particular construction for independent reasons.

However, neither of these conditions are met in BP. What we see in (20)/(21) is a context where *both* pronominals and anaphors are equally acceptable for the same speakers. This is directly at odds with the idea that BP-PRONOUNS are mere elsewhere forms that apply whenever anaphors are not allowed by the grammar.

In their competition-based account, Menuzzi and Lobo (2016) state that locally bound BP-PRONOUNS in transitive structures (i.e. in (20a), (20b) and (20e), but not in (20c) and (20d) and (20f)) are only used in dialects where accusative SE clitics are no longer available. Certain dialects of BP do indeed seem to be in the midst of losing their reflexive SE forms. It is also true, as I argue in section 7.3, that this tendency is correlated with the acceptability of locally bound BP-PRONOUNS. However, Menuzzi and Lobo's (2016) statement is simply too strong because there is, in reality, no variety of BP where reflexive SE is entirely absent.

Pereira's (2007) extensive corpus study failed to observe a categorical omission of the object SE clitic in urban speakers from São Paulo (see also Moura Neves (2008)). In transitive reflexive clauses, speakers from São Paulo retain locally bound SE 75% of the time. Though other studies report lower rates of reflexive clitic use in a corpora from São Paulo (around 50% in Nunes (1995)) and Minas Gerais (51% in Melo (2005)), these figures also do not warrant the conclusion that the SE reflexive is completely absent from these varieties of BP (see also Rocha (1999)), contrary to what Menuzzi and Lobo (2016, pg. 342) claim.[7]

The upshot is that since BP-PRONOUNS are, to all appearances, pronominals, the kind of local binding we see in (20) directly contradicts the predictions of CB – regardless of whether CB is conceived as a primitive or not. The co-existence of locally bound pronominals and of grammaticalized reflexives (the SE forms) is also particularly problematic for competition-based approaches. Similar situations have also been reported in languages like French (ZRIBI-HERTZ, 1995) and Middle English (VAN GELDEREN, 2000). In the next section, I discuss previous attempts to deal with this problem within the framework of the Classical Binding Theory and Reinhart and Reuland's (1993) Chain Theory. I will argue that these attempts do not succeed in explaining the possibility of local binding for BP-PRONOUNS.

---

[7]   In fact, comprehension experiments conducted by Lacerda, Oliveira and Leitão (2014) show that speakers from Minas Gerais accept locally bound SE reflexives between 78% and 93% of the time (depending on the type of predicate within which the reflexive appears). This is precisely the group of speakers who are widely recognized as accepting locally bound BP-PRONOUNS in transitive clauses (e.g. (20a), (20b) and (20e)) according to experiments and informal surveys (LEMLE, 1985; GALVES, 1986; EVERETT, 2000; GROLLA; BERTOLINO, 2011; LACERDA; OLIVEIRA; LEITÃO, 2014; MENUZZI; LOBO, 2016; CARVALHO, 2019).

7.2 PREVIOUS ACCOUNTS OF LOCALLY BOUND BP-PRONOUNS

The previous section left us with two main facts in want of explanation: (a) locally bound BP-PRONOUNS are sometimes acceptable; (b) locally bound BP-PRONOUNS are sometimes *not* acceptable. In this section, I discuss three previous efforts to tackle the first of these facts, which is the one that more directly threatens the empirical validity of CB's claim to universality. None of the theories reviewed in this section address why certain contexts disallow the local binding of BP-PRONOUNS. A thorough discussion of these cases is postponed until Section 7.4.

7.2.1 The Not-a-Pronominal Strategy

The first binding-theoretic account of locally bound BP-PRONOUNS was the NOT-A-PRONOMINAL STRATEGY advocated by Lemle (1985). As the name implies, this is an attempt to reconcile the BP data with CB by claiming that BP-PRONOUNS are not really pronominals, but anaphors – i.e. defective forms governed by Condition A.

A version of this idea was proposed by Ronat (1982) as an explanation for the local binding of the pronoun *lui* in French, illustrated with his examples below:

(22) a. *Victor* parle souvent de *lui*.
          Victor talks often   about him
          '*Victor* talks about *him(self)*.'

      b. *Victor* a honte de *lui*.
          Victor is shame of him
          '*Victor* is ashamed of *him(self)*.'

The NOT-A-PRONOMINAL STRATEGY also became popular as an explanation for the local binding of pronouns in Old and Middle English (VAN GELDEREN, 2000). Middle English in particular is similar to BP in that apparent pronominals (e.g. *hym*) alternate with reflexives (e.g. the newly created form *hymself*) in local binding contexts, contradicting competition-based theories of anaphora (FALTZ, 1985; PEITSARA, 1997):

(23) a. *He* cladde *hym* as a pore laborer.

          '*He* dressed *him(self)* as a poor laborer.'   (Chaucer, *CT*, 551)

      b. And softe unto *hymself* *he* seyde.

          'And softly, *he* said to *himself*.'   (Chaucer, *CT*, 1773)

As with the examples involving BP-PRONOUNS, *hym* is less prone to appear locally bound

in predicates like (23b), which are stereotypicall non-reflexive, than in predicates like (23a), which are more likely to be reflexive (FALTZ, 1985; KÖNIG; SIEMUND, 2000a).

More recently, as we saw in Chapter 2, Reuland (2016) has argued for a version of the NOT-A-PRONOMINAL STRATEGY as a solution to the puzzle posed by locally bound pronouns in Traditional Jambi Malay. Example (24) is from Cole, Hermon and Yanti (2015, pg. 148):

(24)   *Yanti* neŋoʔ *dio* di tipi.
       Yanti saw   her on TV

What is attractive about this approach is that it predicts the odd behavior of BP-PRONOUNS and similar forms in other languages while leaving the Classical Binding Theory essentially intact. If the forms in question are not pronominals, but anaphors, their capacity to be locally bound ceases to be a mystery.

Leaving aside the fact that the NOT-A-PRONOMINAL STRATEGY does not explain cases where local binding is *not* acceptable (cf. (19)), the most obvious objection against applying this proposal to BP is the observation that BP-PRONOUNS blatantly violate Condition A of the Binding Theory in every imaginable way (GALVES, 1986). As discussed in Section 7.1, BP-PRONOUNS can be used as deictics (cf. (25a)), as well with non-local (cf. (25b)) and non-c-commanding antecedents (cf. (25c)):

(25)   a.   Dê um livro pra ela e   pra ele.   [pointing gestures]
            give a book to her and to   him
            'Give a book to her and to him.'

       b.   *O Roberto* disse que a   Joana quer   conversar com *ele*.
            the Roberto said   that the Joana wants talk   with him
            '*Roberto* said that Joana wants to talk to *him*.'

       c.   Os amigos do   *Roberto* adoram o   estilo d*ele*.
            the friends of-the Roberto adore   the style of-him
            '*Roberto*'s friends adore *his* style.'

The data above make it clear that BP-PRONOUNS have almost none of the properties associated with anaphors and basically all of the properties associated with pronominals – including, crucially, referential independence.

It is, of course, possible to claim that BP-PRONOUNS have two homophonous entries, one of which is an anaphor that can be locally bound in certain predicates and the other of which is a pronominal that figures in contexts like (25). But this only underscores the stipulative nature of the NOT-A-PRONOMINAL STRATEGY: the main justification for the claim that BP-PRONOUNS

are not pronominals when they are locally bound is the fact that they fail to comply with CB in precisely these contexts – a reasoning that presupposes (rather than establishes) the truth of CB. The homophony allegation effectively makes CB unfalsifiable in BP.

Lemle (1985) attempts to make the NOT-A-PRONOMINAL STRATEGY look a somewhat more principled by hypothesizing that the recent impoverishment of verbal inflection in BP (particularly, the loss of PERSON distinctions) caused a corresponding impoverishment in the ϕ-feature specification of BP-PRONOUNS, effectively turning them into anaphors.

Note that this story tacitly replaces the definition of anaphors as referentially deficient forms by a (weaker) purely morphosyntactic definition in terms of ϕ-feature specification.[8] However, Table 7 shows that there is no positive evidence even for this weaker claim: BP-PRONOUNS are, in fact, fully specified for ϕ-features (including PERSON).

A version of this argument is proposed more explicitly in Carvalho (2019). It seems to rest on the following two premises: (i) the ϕ-specification of inflection (and possibly other agreement markers) reflects the ϕ-specification of nominative forms (NUNES, 2020); (ii) the ϕ-specification of nominative forms determines the ϕ-specification of object forms (i.e. the forms that appear in local binding configurations like (20)). Table 9 illustrates two stages of the inflectional paradigm for the regular first conjugation the verb *amar* ('love').

	Early 20th Century BP		Contemporary BP	
	Nominative Pronouns	Present Tense	Nominative Pronouns	Present Tense
1SG	eu	amo	eu	amo
2SG	tu	amas	você	ama
3SG	ele/ela	ama	ele/ela	
1PL	nós	amamos	a gente	
2PL	vocês	amam	vocês	amam
3PL	eles/elas		eles/elas	

Table 8 – The BP inflectional system (adapted from Duarte (1993, pg. 109))

I see no independent motivation to accept the premises that license the inferential leap from the impoverishment of PERSON features on inflection to impoverishment of PERSON features on object pronouns.[9] However, if we accept it on purely theory-internal grounds, we

---

[8] In Chapter 3, we saw a similar move in Reuland and Reinhart's (1995) attempt to account for the local binding of *him* in Frisian and in Reuland's (2016) analysis of Traditional Jambi Malay *dio* (cf. (24)).

[9] In fact, from a functional perspective, quite the opposite conclusion seems to be warranted: given that PERSON has to marked for interpretive purposes, the lack of PERSON specification on inflection *requires* the presence of PERSON specification elsewhere in the structure – i.e. on nominative forms. This is a popular explanation for why BP is losing its pro-drop status and resorting to overt subject pronouns more often (DUARTE, 1993).

can infer that the 2nd and 3rd singular and 1st plural forms are unspecified for PERSON (and arguably NUMBER too) in Contemporary BP.

The problem is that it is not clear why a similar rationale would not support classifying personal pronouns in English as anaphors too, given that English is even poorer than Contemporary BP in its inflectional paradigm. For instance, given that there is no specific verbal inflection in English for 3rd person plural, we would expect the pronoun *them* to be an anaphor.

	English	
	Nominative Pronouns	Present Tense
1SG	I	
2SG	you	
1PL	we	love
2PL	you	
3PL	they	
3SG	he/she	loves

Table 9 – The English inflectional system

But *them* clearly cannot be locally bound in contexts like (26), which parallel those in (20):

(26) a. **Paul and Amy* saw *them* in the mirror.

   b. **Larry and Joanne* forgot to include *them* in the guest list.

   c. **Harry and Sarah* were ashamed of *them* after the exam.

The flaw in the argument is even more serious. Lemle (1985) and Carvalho (2019) are apparently assuming that an NP is never specified for a particular ϕ-feature unless this ϕ-feature enters into overt agreement relations with a head. Another consequence of this is that every N in English would be unspecified for gender, since English has no gender agreement. If lack of specification in one feature suffices to count a form as an anaphor (as Lemle and Carvalho presuppose in connection with PERSON in BP), all NPs in English would count as anaphors – obviously a wrong conclusion.

I conclude, therefore, that the NOT-A-PRONOMINAL STRATEGY does not work as explanation for why BP-PRONOUNS tolerate local binding.

7.2.2 The Not-Bound Strategy

The second approach to BP-PRONOUNS that emerged was the NOT-BOUND STRATEGY, first proposed by Bouchard (1984) as an analysis of French and later adapted for BP by Galves

(1986). This approach denies that locally anteceded BP-PRONOUNS are interpreted as bound variables. The suggestion is that what we see in such cases is mere ACCIDENTAL COREFERENCE, which, following Reinhart (1983), is taken to not be governed by the Binding Theory.

This is an attempt to reduce new problematic data (apparent local binding of pronominals) to a well-understood phenomenon which is also attested in English (LASNIK, 1976; REINHART, 1983; GRODZINSKY; REINHART, 1993). Heim (1998, pg. 213) gives the example in (27), where *her* also appears to be locally bound:

(27)   A: Is this speaker *Zelda*?

        B: How can you doubt it? *She* praises *her* to the sky.   (*she=her=Zelda*)

B's reply in (27) is not ruled out by CB because it involves an interpretation where *she* and *her* independently pick out the same entity in the context, rather than one where *her* acts as a variable bound by *she*. The assumption is that these two NPs bear distinct indices, which just happen to be anchored to the same real-world individual in the particular context of (27). This accidental coreference reading is licensed on purely pragmatic grounds.

If accidental coreference is a pragmatic phenomenon, we should not expect to see any major differences between English and BP. A puzzle for the NOT-BOUND STRATEGY is, therefore, to determine why the alleged local coreference of pronominals in BP is not restricted to peculiar contexts like the identity debate scenario in (27) (ZRIBI-HERTZ, 1995). There must be some property of BP which allows BP-PRONOUNS to accidentally corefer with a local antecedent in more neutral contexts like (28), where accidental coreference is ruled out in English:

(28)   a.   *O Paulo* viu *ele* no   espelho.
               the Paulo saw him in-the mirror

      b.   **Paul* saw *him* in the mirror.

Galves's (1986) hypothesis is that this cross-linguistic difference is a consequence of the topic-prominent character of BP (LI; THOMPSON, 1976; PONTES, 1987; GALVES, 1998). Galves assumes that anaphoric reference to topics is always acceptable in virtue of their high degree of discourse accessibility. If main clause subjects in BP are topics by default, using a BP-PRONOUN to refer to a topic would give rise to an illusion of local binding by the subject.

However, this hypothesis runs into difficulties. It wrongly predicts that BP-PRONOUNS can only have subjects as local antecedents. This is falsified by (29):

(29) a. A Joana tentou falar   com *o   Roberto* sobre *ele*.
          the Joana tried   to.talk with the Roberto about him
          'Joana tried to talk to *Roberto* about *him(self)*.'

      b. A Joana pegou o   álbum de fotografias e   mostrou pr*o   Roberto ele*
          the Joana took   the album of photos   and showed to.the Roberto him
          quando *ele* era criança.
          when   he was child
          'Joana got the photo album and showed to *Roberto him(self)* when *he* was a child.'

Moreover, the assumption that coreference with topics suffices to exempt pronominal forms from CB effects (as Galves (1986) seems to be implying) does not explain why accidental coreference is ruled out in English topic constructions like (30):

(30) a. *As for *Bobby*, *he* saw *him* in the mirror.

      b. *As for *Joanne*, *she* forgot to include *her* on the guest list.

      c. *As for *Amy*, *she* was ashamed of *her* after the exam.

If the pronouns in (30) are free to pick out topics as their antecedents, the examples in (30) should, in principle be just as good as the corresponding BP cases in (20). They are not.

The most basic problem for the NOT-BOUND STRATEGY, however, is the supposition that the apparent violations of CB in BP are restricted to coreference readings. This is simply not the case, as the acceptability of non-referential antecedents in (31) and the presence of sloppy readings in (32) illustrate:

(31) a. *Todo ator* pensa primeiro n*ele*,   depois nos   outros.
          every actor thinks first   on-him, then   on-the others
          '*Every actor* thinks of *him(self)* first, and then others.'

      b. *Nenhuma garota* esqueceu de incluir *ela* na   lista de convidados.
          no   girl   forgot   to include her on-the list of guests
          '*No girl* forgot to include *her(self)* in the guest list.'

      c. *Todo político*   quer ficar falando d*ele*   em vez do   País.
          every politician wants stay talking of-him in turn of-the country
          '*Every politician* wants to talk about *him(self)* rather the country.'

(32) a. *O Paulo* viu *ele* no   espelho e   o Roberto também.
          the Paulo saw him in-the mirror   and the Roberto also
          '*Paulo* saw *him(self)* in the mirror, and so did Roberto.'   (sloppy reading)

      b. *O Pedro* reconheceu *ele* na   foto e   o David também.
          the Pedro recognized him in-the photo and the David also
          '*Pedro* recognized *him(self)* in the photo and David did too'.   (sloppy reading)

    c.   *A Amy* comprou um carro pra *ela* e   a Susana também.
          the Amy bought a   car for her and the Susana also
          '*Amy* bought a car for *her(self)* and so did Susana.'   (sloppy reading)

The data above are confirmed by Grolla and Bertolino's (2011) experiments, which found no statistically relevant difference between the acceptability of referential and quantified antecedents for BP-PRONOUNS with local antecedents. Grolla and Bertolino (2011) also attested that speakers accept sloppy readings of BP-PRONOUNS under VP ellipsis (a standard diagnostic for bound variable readings) just as much as they accept local binding with referential antecedents in non-elliptical contexts.

These facts suggest that the NOT-BOUND STRATEGY does not succeed in making locally bound BP-PRONOUNS compatible with CB of the Binding Theory.[10]

7.2.3 The Not-Local Strategy

The final putative explanation for the acceptability of locally bound BP-PRONOUNS to be considered here is the NOT-LOCAL STRATEGY suggested in Menuzzi (1999). This proposal argues that, due to independent properties of BP, locally bound BP-PRONOUNS are not, in fact, in a *local* relationship to their antecedents.

In order to make this work theoretically, Menuzzi (1999) appeals to Reinhart and Reuland's (1993) version of the Binding Theory we discussed in Chapter 2, which reduces the syntactic residue of CB to a well-formedness condition on chains, stated as (33):

(33)   CHAIN CONDITION: A maximal A(rgument)-chain $(\alpha_1, ..., \alpha_n)$ contains exactly one link – the head link α_1 – which is referentially independent.

The CHAIN CONDITION expresses the relatedness between A-movement and A-binding by resorting to a generalized definition of CHAIN, which encompasses not only dependencies formed by movement, but also purely representational dependencies formed by way of coconstrual. A version of Reinhart and Reuland's (1993) definition is given in (34):

---

[10] Galves (1986) considers in passing a another explanation for why BP-PRONOUNS can have local antecedents: the hypothesis that BP-PRONOUNS, unlike English pronominals, are 'specialized' for referential readings (MONTALBETTI, 1984; NEGRÃO; MULLER, 1996). One way to implement this is by having BP-PRONOUNS not correspond to variables in SEM. If this hypothesis is true, BP-PRONOUNS would be exempt from CB (which only governs bound variable readings). However, this hypothesis is plainly *not* true, as (31)(32) indicate (see also Ferreira (2000, pg. 23)). Furthermore, not being interpretable as variable does not spare an expression from being subject to non-coreference, as the existence of Condition C effects shows (e.g *She$_i$ saw Amy$_i$'s dad*). I assume coreference is governed by Rule I from Grodzinsky and Reinhart (1993).

(34)   C is a CHAIN *iff* C is the maximal sequence of links $(\alpha_1, ..., \alpha_n)$ such that, for all j, α_j antecedent governs α_{j+1}, where δ ANTECEDENT GOVERNS γ *iff*:

   a. δ and γ are coindexed;

   b. δ c-commands γ, and

   c.   there is no barrier between δ and γ.

One of the practical effects of (33)/(34) is a prohibition against the occurrence of referentially independent NPs as non-head links of chains. As we saw in Chapter 3, this entails CB, since being a non-head link implies being c-commanded by an antecedent in a local domain (i.e. the domain of the closest barrier). However, unlike standard statements of CB, an account based on (33)/(34) lends itself more readily to a kind of parametrization of locality, insofar as what counts as a barrier for chain formation may vary from language to language.

Menuzzi (1999) exploits this feature of the CHAIN CONDITION to explain some of the facts about the local binding of pronominals across different languages – in particular, the contrast between the BP-PRONOUNS in (35) and their English counterparts in (36):

(35)   a.   **Amy* thinks of *her* first, then of others.

   b.   **Sara* was ashamed of *her* after the exam.

   c.   **Susan* kept talking about *her* all night.

   d.   **Bobby* bought a car for *him*.

(36)   a.   *A Amy* pensa primeiro n*ela*,   depois nos   outros.
          the Amy thinks first   on-her, then   on-the others

   b.   *A Sara* ficou com vergonha d*ela*   depois da prova.
          the Sara was   with shame   of-her after   the exam

   c.   *A Susana* ficou falando d*ela*   a   noite toda.
          the Susana stayed talking of-her the night all

   d.   *O Roberto* comprou um carro pra *ele*.
          the Roberto bought   a   car   for him

Menuzzi (1999) interprets the absence of P-stranding in a language as evidence for the barrierhood of PPs. Since there is P-stranding in English (cf. (37)), pronominals and their antecedents in (35) are not separated by a barrier and constitute an ill-formed A-chain. In BP, on the other hand, there is no P-stranding (cf. (38)), so PPs do count as barriers.

(37)   a.   Who$_1$ did Mary think of t_1?

        b.   Who$_1$ was Sara ashamed of t_1?

        c.   Who$_1$ did Susan keep talking about t_1 all night?

        d.   Who$_1$ did Bobby buy a car for t_1?

(38)   a.   *Quem$_1$ a   Amy pensa primeiro em t_1?
           who   the Amy thinks first   on t_1

        b.   *Quem$_1$ a   Sara ficou com vergonha de t_1?
           who   the Sara was with shame   of t_1

        c.   *Quem$_1$ a   Susana ficou falando de t_1 a noite toda?
           who   the Susana stayed talking of t_1 the   night all

        d.   *Quem$_1$ o   Roberto comprou um carro para t_1?
           who   the Robert bought a   car   for t_1

This means that binding a BP-PRONOUN across a PP boundary, as in (36), does not form an A-chain and, consequently, does not violate anything. The relationship between BP-PRONOUNS within PPs and their antecedents is never truly a local one.[11]

Carvalho (2019) extends the CHAIN CONDITION approach to BP-PRONOUNS in transitive structures. In cases like (39), she argues that BP-PRONOUNS form a Small Clause (SC) with the PPs that follow them and that these SCs count as barriers for the purposes of A-chain formation:

(39)   a.   *O Paulo* viu [$_{SC}$ *ele* no   espelho].
           the Paulo saw   him in-the mirror

        b.   *A Amy* esqueceu de incluir [$_{SC}$ *ela* na   lista].
           the Amy forgot   to include   her in-the list

        c.   *O Pedro* não reconheceu [$_{SC}$ *ele* na   foto].
           the Pedro not recognized   him in-the photo

But there are serious flaws in this proposal. The first is that, since SE forms would be fully acceptable in (39) (cf. (21)), such structures now become a threat to Condition A of the Classical Binding Theory: they imply that anaphors in BP can be non-locally bound, contrary to fact (cf. ***A Amy* disse que o Paulo *se ama*** in (18a)).

Furthermore, unlike the PPs in (36) the structures in (39) are ones from which A-chain

---

[11] Reuland (2011) proposes a similar account for French (cf. (22)) which could also apply to BP. Unlike Menuzzi (1999), Reuland claims that PPs *always* count as local domains for the purposes of pronominal binding (HUANG, 1983; HESTVIK, 1991). What happens in English is that prepositions undergo obligatory reanalysis with their verbs in LF (HORNSTEIN; WEINBERG, 1981). As a consequence, pronominals and their antecedents end up being coarguments in (35), violating CB. In French and BP, on the other hand, there is no reanalysis and, therefore, no CB violation for pronominals inside PPs. The problem with Reuland's account is that, unlike Menuzzi's, it cannot be extended to locally bound BP-PRONOUNS in transitive structures like (39).

formation is possible, as the passive example in (40) illustrates. These facts mean that either the CHAIN CONDITION is false or SCs are not barriers after all:

(40) a.   O Paulo$_1$ foi visto [$_{SC}$ t_1 no   espelho].
          the Paulo   was seen   in-the mirror

      b.   A Amy$_1$ foi incluída [$_{SC}$ t_1 na   lista].
          the Paulo was included   in-the list

      c.   O Pedro$_1$ foi reconhecido [$_{SC}$ t_1 na   foto].
          the Paulo   was recognized   in-the photo

Another problem for this non-locality proposal is that it requires *ad hoc* stipulations to explain why English counterparts of (39) are bad. We'd have to say that the PPs in (41) don't form SCs or that SCs are not barriers in all languages. Both of these alternatives seem arbitrary.

(41) a.   *Paul* saw [$_{SC}$ *him* in the mirror].

      b.   *Amy* forgot to include [$_{SC}$ *her* on the list].

      c.   *Peter* didn't recognize [$_{SC}$ *him* on the photo].

Lastly, the assumption that the presence of a PP is essential to the acceptability of BP-PRONOUNS in transitive structures like (39) is simply wrong. Even though it is in principle possible to stipulate that there is an invisible prepositional phrase forming a SC in cases like like (42)-(44) below, there is no independent motivation for this other than the desire to make the data compatible with the CHAIN CONDITION.

(42)   A: Quem o Paulo viu no espelho? ('Who did Paulo see in the mirror?')

       B: *O Paulo* viu *ele*. ('*Paulo* saw *him*.')

(43)   *A Amy* fez   a lista de convidados, mas esqueceu de incluir *ela*.
       the Amy made the list of guests,   but forgot   to include her
       '*Amy* made a guest list, but forgot to include *her(self)*'.

(44)   Quando eu mostrei a foto para os meus amigos, *o Pedro* não reconheceu *ele*.
       when   I showed the photo to the my   friends, the Pedro not recognized him
       'When I showed my friends the photo, *Pedro* didn't recognize *him(self)*.'

All in all, even though the NOT-LOCAL STRATEGY bears an interesting deductive structure, it ultimately does not work – especially for cases where BP-PRONOUNS are bound in the object position of transitive clauses. In the next section, I propose my own solution for why BP-PRONOUNS are acceptable in local binding configurations.

7.3 THE NATURE OF CONDITION B

The previous section showed that there is no plausible way to avoid the conclusion that BP-PRONOUNS are pronominals which, on occasion, can be locally bound by their antecedents. I want to argue that the simplest account of this phenomenon involves giving up on the idea that BP abides by a syntactic Condition B (CB) altogether. This implies rejecting the view that CB is a syntactic universal, which, in turn, raises the question of how the CB effects that do exist in languages like English come about.

What I propose, in a nutshell, is that CB of the Classical Binding Theory boils down to language-specific binding constructions that only license bound-variable interpretations for pronominals if their binders occupy different PREDs, thereby blocking local binding. Since BP lacks a construction of this sort, there is no *grammatical* constraint against the local binding of BP-PRONOUNS. This constraint only emerges in languages in which alternative locally bound forms become frequent to the point of statistically preempting locally bound pronominals.

There are various ways to implement this, but the suggestion here is stated within the Simpler Syntax (SiSx) framework I have been developing in previous chapters (CULICOVER; JACKENDOFF, 2005; JACKENDOFF; AUDRING, 2020; CULICOVER, 2021). As a model-theoretic framework, SiSx makes a fundamental distinction between the objects that are licensed by a grammar (i.e. its models in the logician's sense) and the set of constraints that license these objects (PULLUM; SCHOLZ, 2001; PULLUM, 2013).

The models employed within SiSx are the relational structures defined in Chapter 4, which consist in tuples of parallel phonological (PHON), syntactic (SYN) and semantic (SEM) representations. SiSx also includes an independent tier for Grammatical Functions (GFs), which represents the syntactic argument structure of predicates. The basic units of the GF tier are PREDs (short for syntactic predicates), which are composed of a sequence of ranked positions. These positions are not explicitly labeled as SUBJECT or OBJECT; rather, such notions can be relationally defined as first GF of PRED, second GF of PRED, etc. Only arguments governed by a head are represented as GFs of the PRED that corresponds to that head and its projections. The main function of the GF tier is to aid in the mapping from SEM structures to SYN structures. The ranking of GFs follows a hierarchy like the one in Keenan and Comrie (1977).

An SiSx grammar is a set of constraints (i.e. partial descriptions of structures on PHON, SYN, GF and SEM and of the correspondences between each level) which expressions have to satisfy in order to be well-formed (i.e. grammatical). Among these constraints, I include a

set of BINDING CONDITIONS which specify well-formed correspondences between particular configurations of SYN and GF tier structures and bound-variable representations in SEM. The English Reflexive Construction of Chapter 5 and CB are both constraints of this sort.

The account in this section will inevitably be too permissive, insofar as it does not purport to explain the cases where locally bound BP-PRONOUNS are bad (cf. (19)). These examples will be dealt with in Section 7.4 by means of a purely pragmatic mechanism. This section is solely concerned with sketching what a syntactic CB in SiSx looks like when it is not interpreted as a principle of Universal Grammar.

7.3.1   A constructional account of Condition B

Instead of stating CB as a universal modular constraint, as the Classical Binding Theory typically does, I will attempt to push CB into a SiSx-style expanded lexicon, treating it as a construction. Since the lexicon contains a large bulk of what is learnable and variable in language, it will come as no surprise that a language like BP might come to lack CB.

As we have seen in Chapter 4, SiSx includes a richly structured lexicon that includes both traditional lexical items as well as more regular and productive patterns, often thought of as the only genuine "rules of grammar". There is a *continuum* from stereotypical words, which specify fully linked phonology, syntax, and semantics (e.g. (45)), through idioms with a few variable slots in them (e.g. (46)), constructional idioms with *nothing but* variables (e.g. (47), which licenses structures like *kissed Will unconscious*), to fully general rules (e.g. (48), which licenses 3rd person singular agreement in English). All of these are stated in the same format: as declarative schemas.

(45)
$$\begin{bmatrix} \text{PHON} & \text{cow}_1 \\ \text{SYN} & \text{N}_1 \\ \text{SEM} & \lambda x[\mathbf{cow}'_1(x)] \end{bmatrix}$$

(46)
$$\begin{bmatrix} \text{PHON} & [\text{catch}_1\ \varphi_2\text{'s}_3\ \text{eye}_4]_5 \\ \text{SYN} & [_{\text{VP}}\ \text{V}_1\ [_{\text{NP}}\ NP_2\text{-GENITIVE}_3\ \text{N}_4]]_5 \\ \text{SEM} & \lambda x[\mathbf{notice}'(\text{EXPERIENCER}{:}\sigma_2, \text{THEME}{:}x)]_5 \end{bmatrix}$$

(47)
$$\begin{bmatrix} \text{PHON} & [\varphi_1\ \varphi_2\ \varphi_3]_4 \\ \text{SYN} & [_{VP}\ V_1\ NP_2\ AP_3]_4 \\ \text{SEM} & \lambda x[\textbf{cause}'(\text{AGENT}{:}x, \text{THEME}{:}\sigma_3(\sigma_2), \text{MEANS}{:}\sigma_1)]_4 \end{bmatrix}$$

(48)
$$\begin{bmatrix} \text{PHON} & \varphi_1 \prec s_3 \\ \text{SYN} & [_V\ V_1\ \text{Affix}_3] \\ \text{GF} & PRED_1{:}\langle GF_2{:}[\textit{3rd, sing}], \dots \rangle \end{bmatrix}$$

In a similar spirit, I proposed in Chapter 5 that bound variable readings in English are licensed by means of a construction associated with pronominals, rather than pronouns in general. Since reflexives (e.g. **her***self*, **them***selves*) include pronominals as part of their structure, they are also partly licensed by this construction. Due to its generality, I called this the English Bound Anaphora Construction. The formulation previously given in (1) is repeated below:

(49)   ENGLISH BOUND ANAPHORA CONSTRUCTION
$$\begin{bmatrix} \text{SYN} & [Y_3 \dots [X_2 \dots NP[\text{PRON}]_1]] \\ \text{GF} & PRED_3{:}\langle\dots GF_4\dots\rangle\ PRED_2{:}\langle\dots GF_1\dots\rangle \\ \text{SEM} & \sigma_4'(\lambda x[\sigma_3'(\dots \sigma_4'/x\dots \sigma_1'/x\dots)]) \end{bmatrix}$$

The constraint in (49) says that a pronominal may be interpreted as a variable bound by an antecedent σ_4' only if σ_4' corresponds to a GF that occupies a PRED whose corresponding SYN node c-commands the SYN node of PRED containing the pronominal.[12] This merges the dominance-sensitivity of Reinhart's bound anaphora condition with a predicate-based notion of anti-locality. The latter approximates the basic idea behind Condition B (CB).

However, as opposed to other predicate-based versions of CB, like the one in Pollard and Sag (1994), (49) is not a negative constraint; that is, rather than saying that pronominals *cannot* be bound in the domain of their syntactic predicates, (49) makes a positive statement about when pronominals *can* be bound. Since this positive statement requires binding to take place from a separate PRED, local disjointness is implied. What gives (49) its universal negative force is the absence of other constructions licensing local binding of pronominals in the grammar of English. Since anti-locality is hardwired into (49), binding between the boldfaced items in (50)

---

[12]   Some fine-grained details of the constraint are obscured somewhat by the convention established in Chapter 5 of only annotating correspondences to GF on terminal nodes in SYN. Since Y_3 corresponds to $PRED_3$, all of the higher projections of Y_3 are also linked to $PRED_3$, and likewise for the projections of X_2. So there isn't only *one* SYN node linked to each PRED. In order for (49) to be satisfied, it suffices that one of the SYN nodes linked to $PRED_3$ c-commands at least one of the SYN nodes linked to $PRED_2$.

is correctly ruled out:

(50) a.
$$\begin{bmatrix} \text{PHON} & \textbf{Every} \prec \textbf{actress}_1 \text{ admires}_2 \textbf{ her}_3 \\ \text{SYN} & [_S \textbf{NP}_1 [_{VP} V_2 \textbf{NP[PRON]}_3]] \\ \text{GF} & \text{PRED}_2: \langle \textbf{GF}_1, \textbf{GF}_3 \rangle \end{bmatrix}$$

b.
$$\begin{bmatrix} \text{PHON} & \textbf{She}_1 \text{ admires}_2 \textbf{ every} \prec \textbf{actress}_3 \\ \text{SYN} & [_S \textbf{NP[PRON]}_1 [_{VP} V_2 \textbf{NP}_3]] \\ \text{GF} & \text{PRED}_2: \langle \textbf{GF}_1, \textbf{GF}_3 \rangle \end{bmatrix}$$

The construction in (49) only licenses bound variable readings if pronominals and their antecedents occupy different PREDs. The problem with (50) is that the GFs that correspond the pronominal and its prospective quantified antecedent occupy the same local PRED.

Binding in (51) would also violate (49), but for a different reason: the problem is that the GF of the quantifier binder is a member of a PRED whose corresponding SYN nodes (V, VP and the embedded S) do not c-command a SYN node of the PRED containing the pronominal.

(51)
$$\begin{bmatrix} \text{PHON} & \textbf{She}_1 \text{ said}_2 \text{ Franklin}_3 \text{ loves}_4 \textbf{ every} \prec \textbf{actress}_5 \\ \text{SYN} & [_S \textbf{NP[PRON]}_1 [_{VP} V_2 [_S \text{NP}_3 [_{VP} V_4 \textbf{NP}_5]]]] \\ \text{GF} & \text{PRED}_2: \langle \textbf{GF}_1 \rangle \text{ PRED}_4: \langle \textbf{GF}_3, \textbf{GF}_5 \rangle \end{bmatrix}$$

The structure in (52) does not have these problems. Therefore, binding is licensed:

(52)
$$\begin{bmatrix} \text{PHON} & \textbf{Every} \prec \textbf{actress}_1 \text{ said}_2 \text{ Franklin}_3 \text{ loves}_4 \textbf{ her}_5 \\ \text{SYN} & [_S \textbf{NP}_1 [_{VP} V_2 [_S \text{NP}_3 [_{VP} V_4 \textbf{NP[PRON]}_5]]]] \\ \text{GF} & \text{PRED}_2: \langle \textbf{GF}_1 \rangle \text{ PRED}_4: \langle \textbf{GF}_3, \textbf{GF}_5 \rangle \\ \text{SEM} & \forall(\textbf{actress}')_1(\lambda z[\textbf{say}'_2(\forall(\textbf{actress}')_1/z, \textbf{loves}'_4(\text{franklin}_3, y_5/z))]) \\ & = \forall(\textbf{actress}')(\lambda z[\textbf{say}'(z, \textbf{love}'(\text{franklin}, z))]) \end{bmatrix}$$

The variable associated with *her* (y_5) is replaced by a variable which is bound by the antecedent *every actress*. This structure is a model of (49) because the PRED that includes GF_1 (PRED_2) corresponds to a node that c-commands the PRED that includes GF_5 (PRED_4).
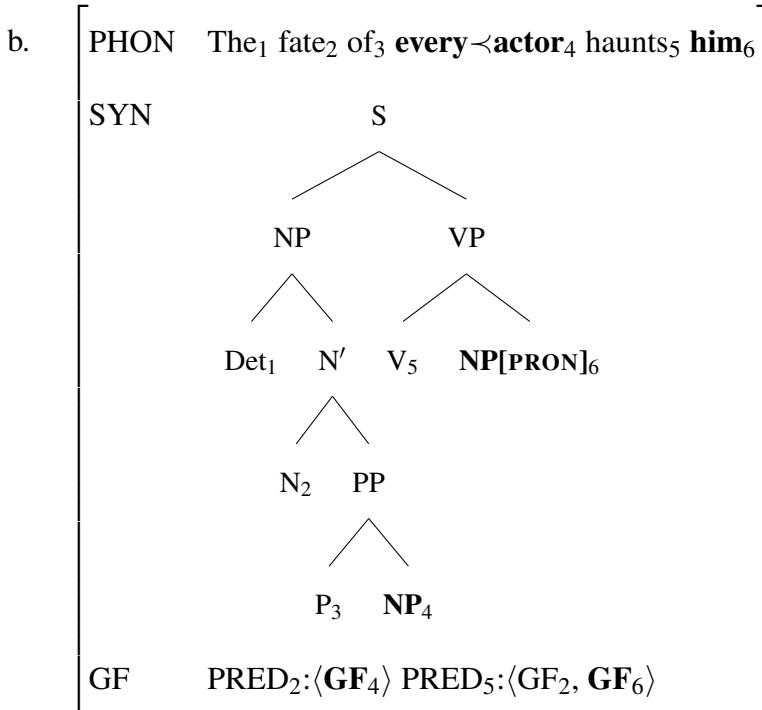
Binding is not licensed in structures like (53), where the putative antecedent of the pronominal is a subconstituent of an embedded relative clause. As in (51), the crucial factor blocking binding is the fact that the antecedent is a member of a PRED whose SYN nodes do not c-

command any of the SYN nodes of the PRED containing the pronominal.

(53)

$$
\begin{bmatrix}
\text{PHON} & \text{The}_1 \text{ director}_2 \text{ who}_3 \text{ knows}_4 \textbf{ every} \prec \textbf{actress}_5 \text{ admires}_6 \textbf{ her}_7 \\
\text{SYN} & \\
\text{GF} & \text{PRED}_4:\langle \text{GF}_3, \textbf{GF}_5 \rangle \text{ PRED}_6:\langle \text{GF}_2, \textbf{GF}_7 \rangle
\end{bmatrix}
$$

SYN tree:
- S
  - NP
    - Det$_1$
    - N$'$
      - N$_2$
      - S
        - NP$_3$
        - VP
          - V$_4$
          - **NP$_5$**
  - VP
    - V$_6$
    - **NP[PRON]$_7$**

All of our predictions so far coincide with those of Reinhart's (1983) Bound Anaphora Condition coupled with CB. Interestingly, however, (49) does license pronominal binding in (54), previously discussed as counterexamples to c-command in Chapter 2:

(54) a.

$$
\begin{bmatrix}
\text{PHON} & \textbf{Every} \prec \textbf{boy's}_1 \text{ mother}_2 \text{ loves}_3 \textbf{ him}_4 \\
\text{SYN} & \\
\text{GF} & \text{PRED}_2:\langle \textbf{GF}_4 \rangle \text{ PRED}_5:\langle \text{GF}_2, \textbf{GF}_6 \rangle
\end{bmatrix}
$$

SYN tree:
- S
  - NP
    - **NP[GEN]$_1$**
    - N$_2$
  - VP
    - V$_3$
    - **NP[PRON]$_4$**

b.

$$\begin{bmatrix} \text{PHON} & \text{The}_1 \text{ fate}_2 \text{ of}_3 \textbf{ every} \prec \textbf{actor}_4 \text{ haunts}_5 \textbf{ him}_6 \\[2em] \text{SYN} & \\[8em] \text{GF} & \text{PRED}_2 \text{:} \langle \textbf{GF}_4 \rangle \text{ PRED}_5 \text{:} \langle \text{GF}_2, \textbf{GF}_6 \rangle \end{bmatrix}$$

SYN tree:
- S
  - NP
    - Det₁
    - N′
      - N₂
      - PP
        - P₃
        - **NP₄**
  - VP
    - V₅
    - **NP[PRON]₆**

What distinguishes (53) from (54) is the fact that, in the latter, the antecedent is a syntactic argument of the NP. This entails that the quantifier GF belongs to a PRED which corresponds to at least one SYN node (i.e. the one realized by the subject NP) that does c-command the SYN node of the PRED containing the pronominal. The effect of (49) is, therefore closer to Langacker's (1969) command condition on binding (BRUENING, 2014) or Kayne's (1994) proposal, rather than to Reinhart's (1983) c-command condition.

We saw in Chapter 2 that there is some evidence that Reinhart's (1983) stricter c-command condition is in fact correct and that coconstrual in cases like (54) is not semantically equivalent to ordinary variable binding. There is an easy way to incorporate this into our present account by simply revising the English Bound Anaphora Construction as follows:

(55) ENGLISH BOUND ANAPHORA CONSTRUCTION (REVISED)

$$\begin{bmatrix} \text{SYN} & [_{Y(P)} \; Y \ldots [_{X(P)} \; X \ldots NP[\text{PRON}]_1]_2]_3 \\[0.5em] \text{GF} & PRED_3 \text{:} \langle \ldots GF_4 \ldots \rangle \; PRED_2 \text{:} \langle \ldots GF_1 \ldots \rangle \\[0.5em] \text{SEM} & \sigma'_4(\lambda x[\sigma'_3(\ldots \sigma'_4/x \ldots \sigma'_1/x \ldots)]) \end{bmatrix}$$

The constraint in (55) now says that a pronominal may be interpreted as a variable bound by an antecedent σ'_4 only if σ'_4 corresponds to a GF that occupies a PRED whose corresponding maximal SYN node dominates the SYN node of PRED containing the pronominal. This condition is not met in the cases lacking c-command in (54): the maximal projections linked to the

PREDs containing the binders are the subject NPs, which do not dominate any node linked to the PREDs containing the pronominals.[13]

I will not attempt to decide between the two formulations of the English Bound Anaphora Construction given above. Regardless of whether one sticks to (49) or (55) – and there are empirical facts pointing in both directions (BÜRING, 2004; BRUENING, 2014; MOULTON; HAN, 2018) – the point remains that we can use the SiSx formalism to state a binding construction that guarantees a non-local syntactic relationship between English pronominals and their binders. Both of the constructions proposed above do this rather straightforwardly. SiSx has no trouble in capturing anti-locality in constructions within its richly structured lexicon.[14]

An advantage of this approach, which is shared with other predicate-based versions of the Binding Theory (REINHART; REULAND, 1993; POLLARD; SAG, 1994), is that it provides a simple explanation for why local binding is acceptable when pronominals appear within predicative PPs without requiring any special phrase-structure representation for adjuncts:

(56)   a.
$$\begin{bmatrix} \text{PHON} & \textbf{Curly}_1 \text{ put}_2 \text{ the} \prec \text{book}_3 \text{ near}_4 \textbf{ him}_5 \\ \text{SYN} & [_S \textbf{ NP}_1 [_{VP} \text{ V}_2 \text{ NP}_3 [_{PP} \text{ P}_4 \textbf{ NP[PRON]}_5]]] \\ \text{GF} & \text{PRED}_2 : \langle \textbf{GF}_1, \text{GF}_2 \rangle \text{ PRED}_4 : \langle \textbf{GF}_5 \rangle \end{bmatrix}$$

   b.
$$\begin{bmatrix} \text{PHON} & \textbf{Curly}_1 \text{ saw}_2 \text{ Laurie}_3 \text{ near}_4 \textbf{ him}_5 \\ \text{SYN} & [_S \textbf{ NP}_1 [_{VP} \text{ V}_2 \text{ NP}_3 [_{PP} \text{ P}_4 \textbf{ NP[PRON]}_5]]] \\ \text{GF} & \text{PRED}_2 : \langle \textbf{GF}_1, \text{GF}_2 \rangle \text{ PRED}_4 : \langle \textbf{GF}_5 \rangle \end{bmatrix}$$

The idea is that, since such PPs project their own PREDs, their arguments are not part of the PRED list of the verb, which ends up not including the GF of the pronominal. Since the GFs of the binder and bindee are not members of the same PRED and the PRED corresponding to the former is hierarchically superior than the one that corresponds to the latter, the English Bound Anaphora Construction (either (49) or (55)) licenses binding in these cases.

---

[13] This constraint reverses the convention of only annotating correspondences to GF on terminal nodes in SYN. However, since the constituent *Y(P)*₃ corresponds to *PRED₃*, all of its lower projections (including *Y*) will also be linked to *PRED₃*. This statement has the same effect as the previous one, except for cases where the antecedent of the pronominal is embedded within an NP argument (cf. (54)). For instance, binding is still impossible in (51) because GF of the putative binder (*every actress*) is a member of a PRED whose corresponding maximal SYN node (i.e. the subject NP) does not dominate the SYN node of the PRED containing the pronominal. Binding is still possible in (52) because the PRED that includes the antecedent GF_1 (PRED_2) corresponds to at least one node (e.g. the matrix S) that dominates the node of PRED_4, which GF_5.

[14] Neither (49) nor (55) say nothing about cases where a pronominal merely corefers with a local coargument. As mentioned above, I assume that coreference (i.e. the assignment of an arbitrary individual to a free variable) is governed by extra-grammatical constraints, such as Rule I from Grodzinsky and Reinhart (1993).

Constructions (49)/(55) also predict anti-locality for PREDs linked to adjectives and nouns:

(57) a. $\left[\begin{array}{ll} \text{PHON} & \textbf{Fosca}_1 \text{ seems}_2 \text{ ashamed}_3 \text{ of}_4 \textbf{ her}_5 \\ \text{SYN} & [_S \textbf{NP}_1 [_{VP} V_2 [_{AP} A_2 [_{PP} P_4 \textbf{NP[PRON]}_5]]]] \\ \text{GF} & \text{PRED}_2: \langle \textbf{GF}_1 \rangle \text{ PRED}_3: \langle \textbf{GF}_1, \textbf{GF}_5 \rangle \end{array}\right]$

    b. $\left[\begin{array}{ll} \text{PHON} & \textbf{Fosca's}_1 \text{ letter}_2 \text{ to}_3 \textbf{ her}_4 \\ \text{SYN} & [_{NP} \textbf{NP[GEN]}_1 [_{N'} N_2 [_{PP} P_3 \textbf{NP[PRON]}_4]]] \\ \text{GF} & \text{PRED}_2: \langle \textbf{GF}_1, \textbf{GF}_4 \rangle \end{array}\right]$

In both structures above, binding between *Fosca* and *her* is impossible because the NPs correspond to GF arguments of the same PRED. The general consequence of the PRED-based constructional approach (regardless of whether it commits specifically to (49) or (55)) is that any kind of binding between syntactic coarguments is ruled out (RUNNER; KAISER, 2005).

Unlike the CB approach (or its Chain Condition variant in Reinhart and Reuland (1993)), however, this story makes it easier to account for languages like BP, which don't exhibit the effects of CB in the same way English. What we have to say, given the ample attestation of locally bound BP-PRONOUNS, is that the general constraint associated with pronominal binding in BP is more permissive than the one operative in English in that it does not require anti-locality. This constraint and the definition of GF-COMMAND it invokes (which is based on Pollard and Sag's (1994, pg. 253) concept of O-COMMAND) follow below:

(58) BP BOUND PRONOMINAL CONSTRUCTION

$\left[\begin{array}{ll} \text{SYN} & [\dots NP[\text{PRON}]_2 \dots]_3 \\ \text{GF} & GF_1 >_{gf} GF_2 \\ \text{SEM} & \sigma_1'(\lambda x[\sigma_3'(\dots \sigma_1'/x \dots \sigma_2'/x \dots)]) \end{array}\right]$

(59) For any two GFs, GF_1 and GF_2, GF_1 GF-COMMANDS GF_2 (i.e. $GF_1 >_{gf} GF_2$) *iff*

    (i) GF_1 precedes GF_2 within a PRED or;

    (ii) a node that corresponds to the PRED that includes GF_1 c-commands a node that corresponds to the PRED that includes GF_2.

The constraint in (58) does not say that a bound pronominal and its potential binder need to inhabit separate PREDs. All that is required for bound variable readings to be licensed for BP-PRONOUNS is that the latter should be GF-COMMANDED by their binders. The concept of

GF-COMMAND includes non-local binding as a possibility, but it also subsumes local binding. This allows (58) to license structures like (60) (corresponding to *Every actress said that Frank loves her*) as well as those in (61) (corresponding to *Every actor thinks of him(self)*):

(60)
$$
\begin{array}{ll}
\text{PHON} & \textbf{Toda} \prec \textbf{atriz}_1 \text{ disse}_2 \text{ que}_3 \text{ o} \prec \text{Frank}_4 \text{ ama}_5 \ \textbf{ela}_6 \\
\text{SYN} & [_S \ \textbf{NP}_1 \ [_{VP} \ V_2 \ [_S \ \text{Comp}_3 \ \text{NP}_4 \ [_{VP} \ V_5 \ \textbf{NP[PRON]}_6 \ ] \ ]] \\
\text{GF} & \text{PRED}_2{:} \ \langle \textbf{GF}_1 \rangle \ \text{PRED}_4{:} \ \langle \text{GF}_3, \textbf{GF}_5 \rangle \\
\text{SEM} & \forall(\textbf{actress}')_1(\lambda z[\textbf{say}'_2(\forall(\textbf{actress}')_1/z, \textbf{loves}'_4(\text{frank}_3, y_5/z))]) \\
& = \forall(\textbf{actress}')(\lambda z[\textbf{say}'(z, \textbf{love}'(\text{frank}, z))])
\end{array}
$$

(61)
$$
\begin{array}{ll}
\text{PHON} & \textbf{Todo} \prec \textbf{ator}_1 \text{ pensa}_2 \text{ n}_3 \textbf{ele}_4 \\
\text{SYN} & [_S \ \textbf{NP}_1 \ [_{VP} \ V_2 \ [_{PP} \ P_3 \ \textbf{NP[PRON]}_4]]] \\
\text{GF} & \text{PRED}_2{:} \ \langle \textbf{GF}_1, \textbf{GF}_4 \rangle \\
\text{SEM} & \forall(\textbf{actor}')_1(\lambda z[\textbf{thinks}'_2(\text{EXPERIENCER}{:} \ \forall(\textbf{actor}')_1/z, \text{THEME}{:} \ y_3/z)]) \\
& = \forall(\textbf{actor}')(\lambda z[\textbf{thinks}'(\text{EXPERIENCER}{:} \ z, \text{THEME}{:} \ z)])
\end{array}
$$

The structure in (60) is licensed in the same way as a parallel structure in English (cf. (52)) would be licensed: the GF of the antecedent is a member of a PRED whose corresponding SYN node c-commands the PRED that includes the GF of the pronominal. The local structure in (61), however, is licensed simply because the GF of the antecedent precedes the GF of the pronominal within their PRED list. This illustrates how, in BP, there is no principle to block local binding between a pronominal and a higher coargument of its syntactic predicate.

Regardless of details in implementation, this view preserves a characterization of pronoun types which is independent of the Binding Theory. BP-PRONOUNS and their English counterparts can all be classified as PRONOMINALS, given that they correspond to variables that may function both as deictics and as bound arguments. The only difference is that BP lacks the anti-locality lexical stipulation encapsulated by (49)/(55). This allows BP-PRONOUNS to be locally bound. The same presumably applies to other languages where syntactic CB effects are absent, e.g. French, Old English, Middle English and possibly even Traditional Jambi Malay.

7.3.2 Learning the Condition B pattern

We saw that a simple construction establishing a pairing between structures in SYN, GF and SEM (i.e. the English Bound Anaphora Construction in (49)/(55)) suffices to bring about

a prohibition against the local binding of pronominals in English. Let's call the prohibition of local binding in general the CB PATTERN. By reducing the CB PATTERN to a constructional property that can vary from language to language, depending on particularities of the lexicon, my account predicts that it has to be learned in languages in which it exists.

This indeed seems to be the case, as evidenced by the well-attested fact that English-speaking children do not display robust adult-level knowledge of CB until around the age of seven (CHIEN; WEXLER, 1990; UTAKIS, 1995; ELBOURNE, 2005; HAMANN, 2011; MAT-TAUSCH; GÜZLOW, 2011; ROORYCK; VANDEN WYNGAERD, 2015; BAAUW, 2018). If the CB PATTERN we see in English has to be learned, it is not surprising that children might not know it at some point.[15]

A long-standing challenge for this kind of approach has always been to determine how a negative constraint ruling out local pronoun binding could come to be learned in the absence of overt negative evidence. That is, given that caregivers do not overtly correct children when they produce sentences with locally bound pronominals (or otherwise provide direct evidence that such structures are ungrammatical), how can children learn that local binding of pronominals is *not allowed*? Any particular well-formed sentence the English learner encounters is, in principle, compatible with a grammar that does not encode the CB PATTERN (i.e. a grammar that allows local binding of pronominals). Instances of locally free pronominals, however numerous, do not contradict a rule system which allows pronominals to be locally bound. These poverty-of-stimulus issues were an important driving force behind the postulation of CB as an innate bias in acquisition (ELBOURNE, 2005; HAMANN, 2011).

As we have seen, the SiSx approach does not encode the negative force of CB directly in terms of the English Bound Anaphora Construction. The negative force of the CB PATTERN follows, rather, from the fact that the English Bound Anaphora Construction is the only constraint licensing binding for pronominals in English and it only licenses binding in non-local environments. The poverty-of-stimulus dilemma characteristic of negative constraints can still be raised: given the unavailability direct negative evidence, how do English learners retreat from their incorrect initial generalization that pronominal binding is allowed in local environments? Or, to put it differently, how do English learners replace the general GF-command condition

---

[15] Some experiments report that this delay of CB effects only occurs for pronominals with referential antecedents (GRODZINSKY; REINHART, 1993; THORNTON; WEXLER, 1999). Elbourne (2005) provides an extensive methodological critique of these studies and suggests a reinterpretation of their data which is compatible with the hypothesis explored here. He also points to other work which failed to replicate an asymmetry between types of antecedents in the acquisition of CB effects.

(e.g. the condition that governs BP-PRONOUNS) by a more restrictive *non-local* GF-command condition that holds of pronominals in the variant of English spoken by adults?[16]

Most of this discussion on the acquisition of CB appears to overemphasize the absence of *direct* negative evidence, inferring from that the need for an innate bias, plausibly subject to some sort of maturation in order to account for the absence of CB effects in children. However, a simpler assumption is that CB is not innate and that the acquisition of the CB PATTERN is an example of learning via STATISTICAL PREEMPTION, which is a type of *indirect* negative evidence (BOYD; GOLDBERG, 2011; GOLDBERG, 2011, 2016; AMBRIDGE et. al., 2013).

In statistical preemption, persistently hearing a particular formulation B in a context where an alternative formulation A is expected may count as indirect evidence to the effect that A is not appropriate in that context. To be more specific, we say that a formulation B statistically preempts A *iff* the conditional probability of B given a context where A is expected approximates 1 (GOLDBERG, 2011, pg. 135):

(62)   B STATISTICALLY PREEMPTS A when P(B|context where A is expected)≈1

When (62) holds, it is plausible to assume that learners have enough evidence posit a negative constraint against using A in the contexts where B appears – or, equivalently, a positive constraint that only licenses A in the complement of such contexts.

What likely occurs for CB in English is that children first hypothesize a general rule that allows pronominals to be bound everywhere (i.e. something like the BP constraint in (58)). They then retreat from this generalization and posit a constraint that restricts pronominal binding to non-local contexts (i.e. (49)/(55)) if they are consistently faced with positive evidence for other forms (e.g. reflexives) that occur in local binding contexts, where pronominals are expected. In the case of English, the prohibition against local binding entailed by the English Bound Anaphora Construction is induced when P(reflexives|local binding contexts)≈1.

In more concrete terms, the fact that the child does not hear local binding of pronominals where it is expected given a fully general constraint like (58) (e.g. ***Bobby*** *bought a car for* ***him***, ***Bobby*** *saw* ***him*** *in the mirror*) counts as indirect negative evidence against the local binding of pronominals. The idea, then, is that children first formulate an overly general rule that allows pronominals to be bound under GF-COMMAND. Later, this looser GF-COMMAND requirement

---

[16] Elbourne (2005) raises a similar question with respect to his theory, where CB is stated as a parameter whose value is initially set as *off*. The question there is to determine what kind of evidence do learners have for switching the value of this parameter to *on* in languages like English.

is replaced by a stricter NON-LOCAL GF-command requirement (clause (ii) in (63)) in virtue of the fact that reflexives start to take hold of local GF-command contexts (clause (i) in (63)).

(63) For any two GFs, GF_1 and GF_2, GF_1 GF-COMMANDS GF_2 (i.e. $GF_1 >_{gf} GF_2$) *iff*

   (i) GF_1 precedes GF_2 within a PRED or;

   (ii) a node that corresponds to the PRED that includes GF_1 c-commands a node that corresponds to the PRED that includes GF_2.

The historical development of the basic CB PATTERN from Middle English to Modern English can shed light on this overall process (PEITSARA, 1997; GOLDE, 1999; KÖNIG; SIEMUND, 2000a; VAN GELDEREN, 2000, MATTAUSCH, 2004). Old English did not have a dedicated reflexive anaphor (KEENAN, 2002). Once reflexive anaphors were established in the early Middle English period, they began to spread and pronominals gradually started to disappear in local binding environments:

	1150-1250	1250-1350	1350-1420	1420-1500	1500-1570	1570-1640	1640-1710
pronominal %	50.3	88.9	76.2	72.0	33.3	12.0	1.9
anaphor %	49.7	11.1	23.8	28.0	66.7	88.0	98.1

Table 10 – Forms used for local binding English (PEITSARA, 1997, pg. 288)

The main force driving the development and expansion of reflexive anaphors is addressed in Section 7.4. Notice, however, that the appearance of reflexives did not automatically expel locally bound pronominals from the language, as competition-based variants of the Binding Theory imply (MENUZZI, 1999; SAFIR, 2004). As Levinson (1991, pg. 139) put it, the emergence of reflexives "long preceded the acquisition of [CB] patterns outlawing the reflexive use of ordinary pronouns, which survived well into Shakespeare".

This gradual emergence of the CB PATTERN falls out neatly from the statistical preemption scenario. Learners only posit a negative constraint against binding pronominals in a local domain (or, equivalently, a positive constraint restricting pronominal binding to non-local domains) when the evidence for an alternative way to express binding in local contexts becomes strong and consistent enough – i.e. when reflexives statistically preempt pronominals in local binding contexts. In the case of English, this only happened sometime in the 17th century.

An opposite trajectory has occurred in the history of BP. Diachronic studies report a growing tendency to suppress anaphoric uses of SE forms in favor of locally bound BP-PRONOUNS and null objects (MOREIRA DA SILVA, 1983, NUNES, 1995):
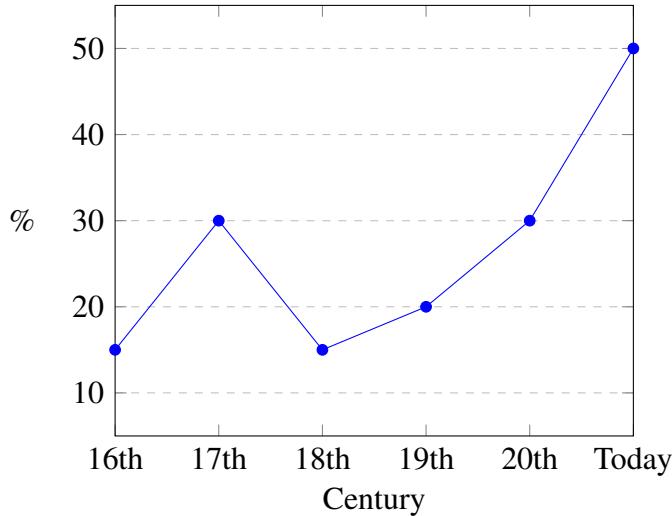
Figure 6 – Supression of locally bound SE anaphors in BP (NUNES, 1995, pg. 211)

The tendency to drop SE forms is plausibly related to the generalized loss of anaphoric clitics in BP (CYRINO, 2003; CARVALHO; CALINDRO, 2018). Regardless of its causes, this has the effect of eliminating a strong preemptive alternative to locally bound BP-PRONOUNS. As a consequence, BP learners come to lack sufficient indirect evidence to posit the anti-locality constraint associated with the CB PATTERN.

In other words, SE forms reached a point where they are not frequent enough in local binding contexts to *drive out* BP-PRONOUNS via statistical preemption. Assuming Nunes' (1995) data, we get P(SE|local binding contexts)≈0.5, which does not suffice to yield the CB PATTERN. As a result, BP-PRONOUNS retain a general possibility of being bound anywhere – despite the fact that a more specific form dedicated for the expression local binding is still available.[17]

We can now understand what goes amiss in competition-based theories (MENUZZI, 1999; SAFIR, 2004). These accounts say that pronominals can only be bound in environments where anaphors are absent. This correctly predicts that, if reflexive anaphors are altogether absent from a language, pronominals can be bound everywhere (see Levinson (1991)). The problem, however, is that the notion of 'absence' implied by these theories is too absolute. What we see in BP (and in Middle English) is that it is not necessary for anaphors to be *completely* absent for locally bound pronominals to be acceptable. It suffices that anaphors not be frequent to the extent required for the CB PATTERN to be posited through statistical preemption.

---

[17] This may not be the case in all dialects of BP. The comprehension study conducted by Lacerda, Oliveira and Leitão (2014) shows that speakers of the *paraibano* dialect of BP, spoken in the northeastern part of Brazil, are more conservative and less prone to accept locally bound pronominals than speakers of the mineiro dialect of BP, which is a part of the southeastern variety examined in this chapter. If the statistical preemption story is correct, this should be related to the stronger prevalence of SE forms in the *paraibano* dialect.

In sum, we find that making the CB PATTERN come out as a construction not only accounts for the typological differences between BP and English, but also carries some degree of plausibility from the perspective of language acquisition.

7.4 A PRAGMATIC ACCOUNT OF THE DUALITY OF BP-PRONOUNS

So far I have argued that the grammar of BP does not include any constraint enforcing non-local binding for pronominals. This explains why locally bound BP-PRONOUNS are acceptable in some contexts. However, this story still falls short of an explanation for the cases where local binding of BP-PRONOUNS is not acceptable – i.e. we cannot yet account for the DUALITY OF BP-PRONOUNS, illustrated below:

(64) a. *O Paulo* viu *ele* no   espelho.
          the Paulo saw him in-the mirror
          '*Paulo* saw *him(self)* in the mirror.'

      b. *O Paulo viu *ele*.
          the Paulo saw him
          '*Paulo* saw *him(self)*.'

(65) a. *A Susana* ficou   falando d*ela*   a   noite toda.
          the Susana stayed talking of-her the night all
          '*Susana* kept talking about *her(self)* all night.'

      b. *A Susana ficou   falando com *ela* a   noite toda.
          the Susana stayed talking to   her the night all
          '*Susana* kept talking to *her(self)* all night.'

This section proposes a strategy for addressing this last piece of the puzzle based on work by Levinson (1991, 2000) and Mattausch (2004, 2007). This will essentially be a pragmatic reinterpretation of Reinhart and Reuland's (1993) Reflexivity Condition B, couched in terms of the neo-Gricean approach laid out in the previous chapter (HORN, 1984; RETT, 2020).

What I want to suggest, in a nutshell, is that BP-PRONOUNS are, in fact, not subject to any kind of syntactic anti-locality effect analogous to CB. The cases where locally bound BP-PRONOUNS are not acceptable are ruled out by an entirely different principle: a universal constraint on the morphosyntactic encoding of semantically reflexive predicates.

(66) CONSTRAINT ON REFLEXIVE PREDICATES (CRP)
      If a semantic predicate *P* is reflexive and *P* is prototypically non-reflexive, then *P* must be reflexive-marked.

(67)   a.   the SEMANTIC PREDICATE formed of *P* is *P* and all arguments that realize a semantic role associated with *P* (i.e. the SEMANTIC ARGUMENTS of *P*);

        b.   *P* is NON-REFLEXIVE *iff* none of its arguments are bound by the same λ-operator;

        c.   *P* is REFLEXIVE-MARKED iff one of *P*'s arguments is a reflexive.

The notion of prototypical (non-)reflexivity is familiar in the functionalist literature on anaphora (HAIMAN, 1983; FALTZ, 1985; LEVINSON, 1991; ZRIBI-HERTZ, 1995; COMRIE 1999; KÖNIG; SIEMUND, 2000a, 2000b; ARIEL, 2008; HASPELMATH, 2008). It is based on the intuition that reflexive interpretations are less expected, natural or likely for some predicates (e.g. *kiss*, *be angry*) than for others (e.g. *shave*, *be ashamed*). A typical statement of the distinction is found in König and Siemund (2000b, pg. 61):

> The semantic property that plays a fundamental role in the selection of a reflexivizing strategy concerns the question whether the situation denoted by the verb or adjective is typically or conventionally directed at others or not. All processes of grooming (washing, dressing, shaving, etc.) are typically performed by a person on himself or herself, with the exception of very young or very old people. All violent actions are typically directed against others. In the domain of attitudes and emotions 'being proud of' and 'being ashamed of' are typical examples of attitudes relating only to a person's own sphere: we can be proud of our achievements, those of our family, perhaps also those of our countrymen, but we are hardly ever proud of the achievements ascribed to the leader of a party in a foreign country. Love, hate and jealousy, by contrast, are standardly directed towards others.

The idea that prototypical reflexivity is associated with particular lexical items rather than to utterances as a whole, is not quite accurate for reasons mentioned in Chapter 3. However, it serves as a rough approximation of the facts. Table 11 summarizes characteristic examples of prototypically non-reflexive and non-prototypically non-reflexive predicates:

PROTOTYPICALLY REFLEXIVE OR NEUTRAL PREDICATES	PROTOTYPICALLY NON-REFLEXIVE PREDICATE
grooming predicates (*shave*, *wash*, *dress*)	violent action predicates (*kill*, *destroy*, *hit*)
prepare, *protect*	emotion predicates (*love*, *hate*)
think of, *talk about*	communication predicates (*talk to*, *speak to*, *argue with*)
proud of, *ashamed of*	*jealous of*, *angry with*, *pleased with*

Table 11 – Examples adapted from König and Siemund (2000b, pg. 61)

There is debate about what causes this distinction, but it seems likely that the class of prototypically non-reflexive predicates indicates inductive regularities in speakers' experience of the world (LEVINSON, 2000). In most Western cultures, people experience less self-directed instances of actions like *speaking to* than of actions like *shaving*.

This is obviously a gradient phenomenon. The distinction between predicate types is also plausibly reflected in frequency of reflexive use: given a large corpus of utterances, prototypically non-reflexive predicates like *jealous of* and *hit* will occur less often with reflexive pronouns (signaling reflexive readings) than more neutral or prototypically reflexive predicates like *proud of* or *dress*, which will tend to occur more often with non-reflexive pronouns (signaling non-reflexive readings). This is confirmed by the following data collected from the British National Corpus (BNC) and the Longman Spoken American Corpus (LSAC):[18]

	Non-reflexive Pronoun	Reflexive Pronoun
proud of	212 (84%)	39 (16%)
jealous of	41 (100%)	0 (0%)

Table 12 – Reflexive *vs.* non-reflexive readings in the BNC (HASPELMATH, 2008, pg. 47)

	Non-reflexive Pronoun	Reflexive Pronoun
hit	109 (99.1%)	1 (0.09%)
dress	4 (6.2%)	60 (93.7%)

Table 13 – Reflexive *vs.* non-reflexive readings in the LSAC (ARIEL, 2008, pg. 231-232)

The CRP is similar to the Reflexivity Condition B of Reinhart and Reuland's (1993) Reflexivity Theory. Unlike Reinhart & Reuland's principle, however, (66) should not be seen a primitive, but as a consequence of a universal pragmatic constraint that associates unmarked forms with prototypical interpretations. This constraint is a byproduct of the neo-Gricean framework I explored in the previous chapter. Some relevant definitions are repeated below:

(68)   For any constructs μ and τ, $\mu <_c \tau$ (i.e. μ is LESS COMPLEX than τ) *iff* using (e.g. processing, articulating or understanding) μ consistently involves less effort than using τ. If neither $\mu <_c \tau$ nor $\tau <_c \mu$, then $\mu =_c \tau$ (i.e. μ and τ are equally complex).

(69)   R ALTERNATIVES

    For any construct μ, the set of R alternatives for μ is $A_R(\mu)$

    $A_R(\mu) := \{\tau : \tau =_i \mu\}$

---

[18]   See Ariel (2008) and Haspelmath (2008) for some further frequency counts that confirm this prediction.

(70)   R PRINCIPLE (REVISED)

Do not use μ if there is another expression $\tau \in A_R(\mu)$ such that both (i) $\tau <_c \mu$, and (ii) τ is assertable (i.e. the speaker believes τ is applicable and relevant in the context).

The basic idea is that, whenever a speaker employs an unmarked form τ, they rely on the hearer's capacity to informationally enrich the truth-conditional content conventionally associated with τ in accordance to salient prototypes about the kinds of situations τ describes (ATLAS; LEVINSON, 1981; LEVINSON, 2000). Given that a structure that explicitly encodes of the prototype would typically be more complex than one which relies on the hearer's inferential abilities, this inference to the prototype is an instance of the the R Principle in (70). Alternatively, whenever a speaker employs a marked form μ when a truth-conditionally equivalent unmarked form τ is available, the hearer will use the R Principle to infer that whatever makes μ marked with respect to τ was, in fact, needed in order to convey their intended message: i.e. that the prototypical interpretation usually inferred from τ does not hold.

This association between marked forms and atypical meanings was summarized, in the previous chapter, in the form of Rett's (2020) Marked Meaning Principle. What interests me for the purpose of deriving the CRP is the dual of this, which I will call, accordingly, the Unmarked Meaning Principle. Both of these principles are given below:

(71)   MARKED MEANING PRINCIPLE

For any constructs μ and τ such that $\tau \in A_R(\mu)$ and $\tau <_c \mu$, using μ carries the implicature **atypical$'$(μ')**.

(72)   UNMARKED MEANING PRINCIPLE

For any constructs μ and τ such that $\mu \in A_R(\tau)$ and $\tau <_c \mu$, using τ carries the implicature **prototypical$'$(μ')**.

In order to derive the CRP (repeated as (66)) from (72), I will assume that pronominals (PRON) and reflexives (REFL) share roughly the same truth-conditional contribution (i.e. REFL $\in A_R$(PRON)) insofar that both can function as variables in SEM.

(73)   CONSTRAINT ON REFLEXIVE PREDICATES (CRP)

If a semantic predicate P is reflexive and P is prototypically non-reflexive, then P must be reflexive-marked.

Recall from section 7.1 that reflexives are *defined* as forms that stand in a markedness contrast to pronominals. For the case of BP, the clitic SE and pronominal forms combined with the emphatic marker *mesmo* count as reflexives in this sense. Both are (for different reasons) more marked/complex than standard argument pronominals: SE is more marked because it is a clitic (and clitics are inherently marked in BP); emphatic forms are more marked because they are morphosyntactically more complex than bare pronominals (*ele* vs. *ele mesmo*).

Since pronominals always unmarked in contrast to reflexives (i.e. PRON$<_c$REFL), the Unmarked Meaning Principle predicts that they will convey a prototypical interpretation for each predicate within which they occur. This means that if a non-reflexive interpretation is prototypical for a predicate *P*, pronominals arguments of *P*, *qua* unmarked forms, will trigger an R-based inference to a non-reflexive interpretation for *P*. The only way to signal that *P* is reflexive in such cases is resorting to specialized reflexive-marking.[19]

The reason is this. Since reflexive marking involves using a form more complex than a pronominal (i.e. a reflexive), the Marked Meaning Principle predicts that, for any predicate *P*, reflexive marking *P* carries the negation of the prototypical reading associated with *P* as an implicature (i.e. an *atypical* reading of *P*). If the prototypical reading associated with *P* is a non-reflexive one, a reflexive-marked variant *P* will carry a reflexive reading as an implicature.[20] The CRP follows, therefore, from an interaction between (71) and (72).

Logically, the CRP is a conditional statement of the form *If A and B, then C*, where C is the reflexive marking requirement. If either one of the conjuncts of the antecedent (A or B) are false, reflexive marking is not necessary. This gives us basically two logically possible scenarios where a locally bound non-reflexive pronoun ψ may avoid violating the CRP:

(74)   a.   When the predicate that includes ψ is not prototypically non-reflexive.

        b.   When the predicate that includes ψ is not reflexive.

---

[19] The Unmarked Meaning Principle thus creates a functional pressure for languages to develop strategies for signaling reflexivity in order to satisfy the CRP. Since non-reflexivity is arguably prototypical for *most* linguistic predicates (see Levinson (1991, pg. 127)), the tendency is for reflexive marking strategies to become fully grammaticalized over time, starting with prototypically non-reflexive predicates (e.g. *kill*), reaching a point where they spread even to cases where reflexivity *is* prototypical (e.g. *shave*). This explains why English developed reflexive anaphors in the first place (FALTZ, 1985; KEENAN, 2002; MATTAUSCH, 2007; ARIEL, 2008). BP, however, seems to be an exception, since it appears to be losing its dedicated reflexive SE form. What this shows is that the pressure to mark reflexivity can be overridden by other propensities in the development of languages. In the case of BP, this is plausibly the tendency to fully expurgate clitics from the grammar. However, it seems plausible that the fact that the CRP holds is part of the reason why SE clitics seem to be dying off more slowly than other object clitics in BP.

[20] Notice that reflexive marking does not have to be effected by a *grammaticalized* reflexive (like *se* in BP or *himself* in English): any form that is semantically equivalent to a pronominal but more complex suffices.

Consider how this accounts the DUALITY OF BP-PRONOUNS mentioned in Section 7.1, starting with contrasts like (75):

(75)   a.   *A   Amy* pensa primeiro n*ela*,   depois nos   outros.
                the Amy thinks first   on-her, then   on-the others
                '*Amy* thinks of *her(self)* first, then of others.'

        b.   *A   Amy* bateu primeiro n*ela*,   depois nos   outros.
                the Amy hit   first   on-her, then   on-the others
                '*Amy* hit *her(self)* first, then other people.'

The CRP does not require reflexive-marking in (75a) because the predicate is one where a non-reflexive interpretation is not prototypical – in people's daily life, thinking about oneself is not a particularly remarkable act. This is an instance of (74a). Since there is no particular prototype associated with the predicate *pensar em* ('think of'), the Unmarked Meaning Principle does not derive non-reflexivity as an implicature. Therefore, the speaker must rely on independent features of the context to determine the interpretation of the pronoun. In (75b), on the other hand, reflexivity is highly atypical, so using a pronoun triggers an R-based inference of disjointness. Therefore, reflexive marking is necessary to signal reflexivity.

The same explanation extends to (76)-(77), as Zribi-Hertz (1995) suggests for similar examples in French. The states of affairs in (76b)- (77b) are prototypically non-reflexive in a way that than those in (76a)-(77a) are not:

(76)   a.   *A   Sara* ficou com vergonha d*ela* depois da prova.
                the Sara was with shame   of-her after   the exam
                '*Sara* was ashamed of *her(self)* after the exam.'

        b.   *A   Sara* ficou com ódio d*ela*   depois da prova.
                the Sara was with hate of-her after   the exam
                '*Sara* got angry at *her(self)* after the exam.'

(77)   a.   *A   Susana* ficou falando d*ela*   a   noite toda.
                the Susana stayed talking of-her the night all
                '*Susana* kept talking about *her(self)* all night.'

        b.   *A   Susana* ficou discordando d*ela*   a   noite toda.
                the Susana stayed disagreeing of-her the night all
                '*Susana* kept disagreeing with *her(self)* all night.'

Though the CRP also contributes to the unacceptability of English counterparts of (75b)/ (76b)/(77b), translations of all of the sentences above are categorically ruled out in English by English Bound Anaphora Construction (cf. (49)/(55)), which is a purely grammatical constraint, not sensitive to the particular prototypes associated with the semantics of the predicate.

The contrast in (78) is also ascribable to the CRP, but for a slightly different reason. Cases like (78a) arguably involve PROXY READINGS, in the sense of section 6.4 (CARVALHO, 2019). These occur when the meaning of an NP is shifted to denote a proxy or representation of its usual denotation: e.g. when one uses a name like *Ringo* to refer to a statue of Ringo.

(78) a. *O Paulo* viu *ele* no espelho.
    the Paulo saw him on-the mirror
    '*Paulo* saw *him(self)* in the mirror.'

   b. **O Paulo* viu *ele*.
    the Paulo saw him
    '*Paulo* saw *him(self)*.'

The BP-PRONOUN *ele* receives a proxy reading in (78a) because it does not denote Paulo, but a visual representation of Paulo. This means that reflexive-marking by SE is not required, because the predicate *ver* ('see') is simply not reflexive in that example: one of the predicate's semantic roles is attributed to Paulo, and the other is attributed to a proxy of Paulo (Paulo's image in the mirror). In (78b), on the other hand, there is no proxy reading to spare the predicate from the obligation to be reflexive-marked.

The Proxy Construction which licenses proxy readings is repeated in (79) (where $\mathscr{R}$ is a meta-variable over representational relations like **image-of′**, **statue-of′**, **story-about′**):

(79) PROXY CONSTRUCTION

$$\begin{bmatrix} \text{PHON} & \varphi_1 \\ \text{SYN} & NP_1 \\ \text{SEM} & \lambda x[\iota y[\mathscr{R}(y, {}^\wedge x)]](\sigma_1) \end{bmatrix}$$

We can detect proxy readings by exploiting a property of representational relations I mentioned in Chapter 5: the fact that they introduce a kind of referential opacity, which potentially gives rise to non-veridical readings. A sentence like *Paul saw a man with three heads in the mirror* (and its counterpart in BP) does not entail that a three-headed man exists in the actual world. What blocks existential generalization, in this case, is precisely the representational relation **image-of′**, which is licensed by (79). See Varaschin (2020) for more details.

The structures underlying each example in (78) are summarized below, ignoring irrelevant details like the semantics of the PP in (78a):

(80) a.

$$
\begin{array}{ll}
\text{PHON} & \text{O} \prec \text{Paulo}_1 \ \text{viu}_2 \ \text{ele}_3 \ \text{no} \prec \text{espelho}_4 \\[4pt]
\text{SYN} & [_S \ NP_1 \ [_{VP} \ V_2 \ NP[\text{PRON}]_3 \ PP_4]] \\[4pt]
\text{GF} & \text{PRED}_2 : \langle GF_1, GF_3 \rangle \\[4pt]
\text{SEM} & \lambda P[P(\text{paulo})]_1 (\lambda z[\mathbf{saw}'_2 (\lambda P[P(\text{paulo})]_1/z, \\[4pt]
& \quad \lambda x[\iota y[\mathbf{image\text{-}of}'(y, \hat{}x)]](u_3)/z)]) \\[4pt]
& = \lambda P[P(\text{paulo})](\lambda z[\mathbf{saw}'(\lambda P[P(\text{paulo})]/z, \iota y[\mathbf{image\text{-}of}'(y, \hat{}u)]/z)]) \\[4pt]
& = \lambda P[P(\text{paulo})](\lambda z[\mathbf{saw}'(z, \iota y[\mathbf{image\text{-}of}'(y, \hat{}z)])])
\end{array}
$$

b.

$$
\begin{array}{ll}
\text{PHON} & \text{O} \prec \text{Paulo}_1 \ \text{viu}_2 \ \text{ele}_3 \\[4pt]
\text{SYN} & [_S \ NP_1 \ [_{VP} \ V_2 \ NP[\text{PRON}]_3]] \\[4pt]
\text{GF} & \text{PRED}_2 : \langle GF_1, GF_3 \rangle \\[4pt]
\text{SEM} & \lambda P[P(\text{paulo})]_1 (\lambda z[\mathbf{saw}'_2 (\lambda P[P(\text{paulo})]_1/z, u_3/z)]) \\[4pt]
& = \lambda P[P(\text{paulo})](\lambda z[\mathbf{saw}'(z, z)])
\end{array}
$$

The example in (78a)/(80a) is exempt from the CRP not because the predicate is not prototypically non-reflexive, but because it is not reflexive at all – i.e. it is an instance of the scenario in (74b). The structure in (78b)/(80b), on the other hand, is ruled out by the CRB because it conveys a reflexive semantic predicate (i.e. $\lambda z[\mathbf{saw}'(z, z)]$)) which is prototypically non-reflexive: most acts of seeing involve a relation between two distinct entities.

Example (81a) avoids the need for reflexive marking for a similar reason: the predicate *incluir* ('include') expresses a relation between Joana and a proxy of Joana (namely, *her name*). The semantic predicate linked to *elogiar* ('praise') in (81b), on the other hand, is genuinely reflexive (and also prototypically non-reflexive), so marking by a reflexive argument (either the SE reflexive clitic or the emphatic *ela mesma*) is needed.

(81) a. *A Joana* esqueceu de incluir *ela* na   lista de convidados.
    the Joana forgot   to include her in-the list of guests
    '*Joana* forgot to include *her(self)* in the guest list.'

b. **A Joana* esqueceu de elogiar *ela* na   festa.
    the Joana forgot   to praise her in-the party
    '*Joana* forgot to praise *her(self)* at the party.'

Exactly the same thing happens in (82), with the difference that the proxy in (82a) is an image in a photograph:

(82) a. *O Pedro* não reconheceu *ele* na foto.
          the Pedro not recognized him in-the photo
          '*Pedro* didn't recognize *him(self)* in the photo'.

      b. **O Pedro* não defendeu *ele* na festa.
          the Pedro not defended him in-the party
          '*Pedro* didn't defend *him(self)* at the party.'

Similar contrasts can be multiplied for other sorts of proxy readings. The contrast in (83) is due to the fact that the pronoun in (83a) is interpreted as an auditory proxy of Amy (e.g. a recording) and not as Amy herself.

(83) a. *A Amy* ouviu *ela* no rádio.
          the Amy heard her in-the radio
          '*Amy* heard *her(self)* on the radio'.

      b. **A Amy* ouviu *ela*.
          the Amy heard her
          '*Amy* heard *her(self)*.'

It is important to note that proxy readings to not alter the syntactic properties of the NPs to which they apply (JACKENDOFF, 1992). Crucially, they do not introduce a covert PRED which takes the GF linked to the proxy NP as an argument.

This can be seen in cases of subject-verb agreement. In a scenario involving a single statue of four oxen, the agreement pattern reveals the NUMBER feature of the GF linked to *the four oxen* has *plural* as its value, despite the fact that the statue is semantically singular.

(84) The ⟨*statue-of*⟩ four oxen {are / * is} cute

This suggests that the SEM predicate introduced by (79) (i.e. $\lambda x[\iota y[\mathscr{R}(y, \hat{} x)]]$) does not correspond to a separate PRED on the GF tier. In other words, proxy readings leave the relationship between the NPs to which they apply and their PREDs on the GF tier unchanged.

This last observation is important because it explains why proxy readings in structures analogous to (78a)/(81a)/(83a) would not be sufficient to exempt English pronominals from disjointness effects, as (85) exemplifies.

(85) a. **Paul* saw *him* in the mirror.

      b. **Joanne* forgot to include *her* on the guest list.

      c. **Amy* heard *her* on the radio.

Even though the pronominals in these examples denote proxies and are, therefore, not coextensional with their antecedents, they are still bound by their GF antecedents in a local domain (i.e. the domain of their local PRED), as the representation for (85c) in (87) makes clear. This is forbidden by the English Bound Anaphora Construction, repeated below.

(86)   ENGLISH BOUND ANAPHORA CONSTRUCTION

$$
\begin{bmatrix}
\text{SYN} & [Y_3 \ldots [X_2 \ldots NP[\text{PRON}]_1]] \\
\text{GF} & PRED_3{:}\langle \ldots GF_4 \ldots \rangle \; PRED_2{:}\langle \ldots GF_1 \ldots \rangle \\
\text{SEM} & \sigma'_4(\lambda x[\sigma'_3(\ldots \sigma'_4/x \ldots \sigma'_1/x \ldots)])
\end{bmatrix}
$$

(87)

$$
\begin{bmatrix}
\text{PHON} & \text{Amy}_1 \; \text{heard}_2 \; \text{her}_3 \; \text{on} \prec \text{the} \prec \text{radio}_4 \\
\text{SYN} & [_S \, NP_1 \, [_{VP} \, V_2 \, NP[\text{PRON}]_3 \, PP_4]] \\
\text{GF} & PRED_2{:} \; \langle GF_1, GF_3 \rangle \\
\text{SEM} & \lambda P[P(\text{amy})]_1(\lambda z[\textbf{hear}'_2(\lambda P[P(\text{amy})]_1/z, \\
& \quad \lambda x[\iota y[\textbf{recording-of}'(y, \hat{\;}x)]](u_3)/z)]) \\
& = \lambda P[P(\text{amy})](\lambda z[\textbf{hear}'(\lambda P[P(\text{amy})]/z, \iota y[\textbf{recording-of}'(y, \hat{\;}u)]/z)]) \\
& = \lambda P[P(\text{amy})](\lambda z[\textbf{hear}'(z, \iota y[\textbf{recording-of}'(y, \hat{\;}z)])])
\end{bmatrix}
$$

The structure (87) is not licensed by the grammar of English because (86) only licenses the replacement of a free pronominal variable by a bound variable in cases where the GF of the antecedent occupies a separate PRED. This is not the case in (87), regardless of the semantic effects of the Proxy Construction.

Another interesting test case for the validity of the claim that BP is subject to the CRP and not to the kind of syntactic anti-locality enforced by (86) are raising to object structures. Since the CRP is stated in terms of SEMANTIC PREDICATES, local binding of "raised" pronominal objects should not be excluded in BP. This is in fact what we observe (CARVALHO, 2019):

(88)   a.   *O Roberto* imaginou *ele* casado.
             the Roberto imagined him married
             '*Roberto* imagined *him(self)* married.'

        b.   *A Joana* estranhou   *ela* cansada.
             the Joana found.strange her tired
             '*Joana* found it strange that *she* was tired.'

Since the BP-PRONOUNS in (88) are merely *syntactic* arguments of the matrix verbs, no

reflexive semantic predicate is formed: it is not the case that two semantic roles of the verbs are assigned to arguments bound by the same λ-operator, as the structure for (88a) below shows.

$$(89) \quad \begin{bmatrix} \text{PHON} & \text{O} \prec \text{Roberto}_1 \text{ imaginou}_2 \text{ ele}_3 \text{ casado}_4 \\ \text{SYN} & [_S \text{ NP}_1 \ [_{VP} \text{ V}_2 \text{ NP}[\text{PRON}]_3 \text{ AP}_4]] \\ \text{GF} & \text{PRED}_2: \langle \text{GF}_1, \text{GF}_3 \rangle \ \text{PRED}_4: \langle \text{GF}_3 \rangle \\ \text{SEM} & \lambda P[P(\text{roberto})]_1(\lambda z[\textbf{imagined}'_2(\lambda P[P(\text{roberto})]_1/z, \textbf{married}'_3(x_3/z))]) \\ & = \lambda P[P(\text{roberto})](\lambda z[\textbf{imagined}'(z, \textbf{married}'(z))]) \end{bmatrix}$$

These raising to object cases, like the earlier proxy examples, are further instances of exemption due to (74b): i.e. they are cases where the CRP does not apply because no reflexive semantic predicate is formed. As a consequence, these structures could only be ruled out by a purely syntactic constraint on locally bound pronominals (a constructional analogue of CB such as (86)). But a constraint of this sort simply does not exist in BP.

7.4.1 Further illustrations of the Constraint on Reflexive Predicates

The CRP is compatible all of the relevant facts so far. In particular, it covers all of the cases where local binding of BP-PRONOUNS is not allowed in BP without invoking a grammatical constraint like CB (or any construction that entails CB). Since the CRP grounded in the pragmatics, it should be universal. So we expect to see some of its effects in other languages where CB absent, as well as in English constructions that are exempt from CB effects.

The Frisian pronoun *him* has been widely claimed to not be subject to CB effects on the basis of examples like (90a) (HOEKSTRA, 1994; REULAND; REINHART, 1995). However, it seems to be subject to the CRP, as (90b) illustrates:[21]

---

[21] Interestingly, my judgment is that local binding of BP-PRONOUNS in grooming predicates like (90) is less than fully acceptable (but see Grolla and Bertolino (2011) for different judgments):

(i)   a.   ?*O Roberto* banhou *ele*.
           the Roberto bathed him
     b.   ?*O Roberto* barbeou *ele*.
           the Roberto shaved him

   A possible explanation for this is that BP tends to use periphrastic light verb constructions to express these meanings: reflexive variants of *banhar* ('bathe') and *barbear* ('shave') are more often expressed as *tomar um banho* ('have a bath') and *fazer a barba* ('do the beard'). If frequency of reflexive use influences the prototypically of reflexive readings (as Haspelmath (2008) suggests), then it is possible that the cases in (i) are deviant because they are too infrequently used to express reflexivity, given larger frequency of a dedicated inherently reflexive periphrastic form.

(90) a. *Max* wasket *him*.
          Max washes him

      b. **Max* hatet *him*.
          Max hates him

Since *hatet* ('hates') corresponds to a semantic predicate for which non-reflexivity is prototypical, special reflexive-marking is necessary. This is not the case for *wasket* ('washes').

Zribi-Hertz (1995) argues that French personal pronoun *lui* is not subject to the CB PATTERN when it is within a PP (RONAT, 1982; PICA, 1984):

(91) a. *Pierre* parle souvent de   *lui*.
          Pierre often talks   about him
          '*Pierre* often talks about *him(self)*.'

      b. *Pierre* est fier   de *lui*.
          Pierre is   proud of him
          '*Pierre* is proud of *him(self)*.'

      c. *Pierre* a honte   de *lui*.
          Pierre is ashamed of him
          '*Pierre* is ashamed of *him(self)*.'

      d. *Pierre* pense souvent à *lui*.
          Pierre thinks often   of him
          '*Pierre* often thinks of *him(self)*.'

However, when a PP-internal *lui* is an argument of a semantic predicate for which a non-reflexive interpretation is prototypical (cf. (92)), the CRP makes reflexive marking by the emphatic *même* ('same') necessary (e.g. **Pierre** *bavarde avec* **lui-même**) (ZRIBI-HERTZ, 1995).

(92) a. **Pierre* bavarde   avec *lui*.
          Pierre is.chatting with him
          '*Pierre* is talking to *him(self)*.'

      b. **Pierre* est jaloux de *lui*.
          Pierre is jealous of him
          '*Pierre* is jealous of *him*.'

      c. **Pierre* a   besoin de *lui*.
          Pierre has need   of him
          '*Pierre* needs *him(self)*.'

      d. **Pierre* se confie   à *lui*.
          Pierre CL confides to him
          '*Pierre* confides in *him(self)*.'

A similar reasoning explains the dubious behavior of what I called NON-REFLEXIVE ANAPHORS in Chapters 1-3. The following data illustrate the case of the Dutch anaphor *zich*:

(93)   a.   *Max* wast   {*zichzelf* / *zich* / **hem*}.
               Max washes {REFL   / SE   / PRON}

        b.   *Henk* hoorde {*zichzelf* / *zich* / **hem*} zingen.
             Henk heard   {REFL   / SE   / PRON} sing.

(94)   a.   *Max* haat {*zichzelf* / **zich* / **hem*}.
               Max hates {REFL   / SE   / PRON}

        b.   *Henk* overreedde {*zichzelf* / **zich* / **hem*} te zingen.
             Hank persuaded {REFL   / SE   / PRON} to sing

The paradox consists in the fact that *zich* behaves like a pronominal in contexts like (93) but like a reflexive in contexts like (94). We can now speculate a deeper reason for why this is the case. As we saw in previous chapters, *zich* is a genuine anaphor (it is a referentially deficient form). It also not subject to a syntactic anti-locality constraint that prohibits it from being locally bound, as the examples in (93) show (REINHART; REULAND, 1993; REULAND, 2011).

However, since *zich* is no more complex than the pronominal *hem* (while at the same time being informationally equivalent to it), *zich* does not count as a reflexive according to the functional definition provided in section 7.1 and repeated below:

(95)   A anaphor μ counts as a REFLEXIVE in a language $\mathscr{L}$, if there is another pronoun τ in $\mathscr{L}$ such that $\tau \in A_R(\mu)$ and $\tau <_c \mu$

The absence of a (semantically equivalent) form which is less complex than *zich*, is what makes *zich* a NON-REFLEXIVE anaphor in Dutch. Given that *zich* is not a reflexive, the CRP entails that it can only be locally bound within semantic predicates that are either not prototypically non-reflexive or not reflexive at all, as per (74).[22]

This is precisely what happens in (93): (93a) is a case where the predicate is prototypically reflexive and (93b) instantiates a raising to object construction where no reflexive semantic predicate is formed in the first place. In (94), on the other hand, *zich* violates the CRP because it does not provide appropriate marking to drive the hearers away from inferring the prototypical non-reflexive interpretation associated with its predicate. This analysis, unlike Reinhart and Reuland's (1993) dual-entry account, is compatible with the results obtained in studies of

---

[22] As we saw on the previous chapter, *zich* cannot be subject to a proxy reading because it does not correspond to an autonomous variable in SEM. So the only way it can be locally bound in a predicate which is not reflexive is if there is some kind of mismatch between semantic predicates and PREDs, which is precisely what happens in raising to object structures like (93b). As we saw in Chapter 3, *zich* can also be used with predicates like *schaamt* ('shame') as in *Max schaamt zich* ('Max is ashamed'). In this case, since the verb is semantically monadic (albeit arguably syntactically transitive), reflexive marking is also not necessary.

pronoun choice in Dutch using large-scale robust empirical data, collected in corpus-based and experimental studies (SMITS; HENDRIKS; SPENADER, 2007; HENDRIKS; SPENADER; SMITS, 2008; BOUMA; SPENADER, 2008; HENDRIKS; HOEKS; SPENADER, 2014).[23]

The effects of the CRP can also be perceived in English constructions where, for one reason or another, pronominals are exempt from the syntactic anti-locality effects enforced by the English Bound Anaphora Construction. An interesting example of this, as Reinhart and Reuland (1993) note, are conjoined NPs.

Pronominals within conjoined NPs are always exempt from CB effects because they are only members of the PRED list of the conjunction itself. As long as they are not bound to their fellow conjunct NP within the PRED linked to the conjunction (and they can't be because the PRED linked to the conjunction is anti-reflexive), conjoined pronominals cannot violate the English Bound Anaphora Construction (cf. (86)). Consider the sample structure for a conjunction like *Larry and her* below:

(96)
$$
\begin{bmatrix}
\text{PHON} & \text{Larry}_1 \text{ and}_2 \text{ her}_3 \\
\text{SYN} & [_{NP} \text{ NP}_1 \text{ Conj}_2 \text{ NP[PRON]}_3] \\
\text{GF} & \text{GF}_3\text{:}[plur] \text{ PRED}_2\text{:}\langle \text{GF}_1, \text{GF}_3 \rangle \\
\text{SEM} & \lambda z[\lambda Q[Q(z,\text{larry})]]_1 \sqcap_2 \lambda y[\lambda S[S(y,x)]]_3 \\
& = \lambda w[\lambda z[\lambda Q[Q(z,\text{susan})]](w) \sqcap \lambda y[\lambda S[S(y,x)]](w)] \\
& = \lambda w[\lambda Q[Q(w,\text{susan})] \sqcap \lambda S[S(w,x)]] \\
& = \lambda R[\lambda w[\lambda Q[Q(w,\text{susan})](R) \sqcap \lambda S[S(w,x)](R)]] \\
& = \lambda R[\lambda w[R(w,\text{larry}) \wedge R(w,x)]]
\end{bmatrix}
$$

In spite of never being GF-coarguments of their antecedents, pronominals within conjoined NPs do seem to be sensitive to the CRP:

(97)   a.   *Joanne* forgot to include Larry and *her* in the guest list.

        b. ??*Joanne* forgot to praise Larry and *her* at the party.

---

[23] Bouma and Spenader (2008), for instance, found a positive correlation between the overall rate of reflexive usage for a verb (which reflects the prototypicality of a reflexive interpretation) and the use of the weak anaphor *zich* in a large-scale parsed Dutch corpus. Verbs with a strong preference for *zichzelf* over *zich* were used non-reflexively 97% of the time. This suggests prototypical non-reflexivity, which requires a marked form to signal the atypical reflexive meaning, as per the CRP. In contrast, verbs where *zich* is preferred to *zichzelf* were used non-reflexively only 72% of the time. This suggests neutrality *vis-à-vis* reflexivity; unmarked *zich* can, therefore, be used reflexively in such cases due to the absence of a non-reflexive prototype. These results were confirmed and refined in a subsequent experimental study conducted by Hendriks, Hoeks and Spenader (2014).

(98) a. *Harry* saw Sarah and *him* on TV.

    b. ??*Harry* saw Sarah and *him*.

(99) a. *Paul* thinks of Amy and *him* first, then others.

    b. ??*Paul* praised Amy and *him* first, then others.

In (97a)/(98a), the pronominals refer to proxies of their antecedents. As a consequence, no reflexive semantic predicate is formed. In (99a), we do find a reflexive predicate (the AGENT and the THEME of *think* are bound by the same λ-operator), but this is a predicate for which a non-reflexive interpretation is not prototypical, so no special marking is required and the hearer relies on the context to determine the interpretation of the pronominal. The reflexive predicates in (97b)/(99b)/(98b), on the other hand, are all prototypically non-reflexive. Since these instances of prototypical non-reflexivity are not accompanied by reflexive marking, they are excluded by the CRP.[24]

Like BP-PRONOUNS, conjoined English pronominals in the object position of raising to object predicates are also never excluded by the CRP. This contrasts with conjoined pronominals in object control configurations, as Reuland's (2017, pg. 22) examples in (100) illustrate:

(100) a. In the end, *John* expected Mary and *him* to leave the country.

    b. *In the end, *John* convinced Mary and *him* to leave the country.

None of the sentences in (100) are excluded by the English Bound Anaphora Construction in (86) for the reasons mentioned above (*him* is not bound within the PRED linked to the conjunction). The only difference between them is that, in (100b), a prototypically non-reflexive predicate is formed. Since this reflexive predicate is not properly marked, a disjointness inference based on the Unmarked Meaning Principle is triggered.

What we find, in summary, is that the CRP explains not only the ambivalent behavior of

---

[24] To return to a point made in the previous chapter, it is interesting to note that reflexive marking by *himself* in (97b)/(99b)/(98b) would occur in violation of the core English Reflexive Construction (ERC), since it implies the use of a reflexive NP which is not bound by a higher coargument of the same PRED:

(i) a. *Joanne* forgot to praise Larry and *herself* at the party.
    b. *Harry* saw Sarah and *himself*.
    c. *Paul* praised Amy and *himself* first, then others.

However, as we saw, the violation of ERC in such cases is a crucial part of what makes the structures in (i) marked and, therefore, prone to generating an inference which denies the non-reflexive prototype associated with the predicates in these examples. In other words, in contexts like (i) *self*-forms function as reflexives not because they are grammatically prescribed to do so (they are not), but because they contribute to complexity, which yields an implicature suggesting that the situation described by (i) does not conform to what is generally the case (the prototype), as the Marked Meaning Principle (cf. (71)) predicts.

pronominals in BP and other languages which lack lacks CB effects altogether, but also facts about English, which does have an analogue to CB (the English Bound Anaphora Construction).

7.5 FINAL REMARKS

The facts surveyed throughout this chapter motivated a division of the work traditionally done by CB into two different principles:

(i) Language-specific constructions that condition bound-variable interpretation of pronominals to non-local contexts (e.g. the English Bound Anaphora Construction)

(ii) A pragmatically-motivated constraint on the morphosyntactic encoding of semantically reflexive predicates (i.e. the CRP).

I have argued that it is the occasional absence of (i) and the universal presence of (ii) that explains the patterns we observe across different languages. The puzzling DUALITY OF BP-PRONOUNS follows from the fact that pronominals in BP are only subject to (ii).

As we saw in Chapter 2, this division of labor is also a feature of the theory championed by Reinhart and Reuland (1993), who distinguish between a condition on syntactic dependencies formed by pronouns and their antecedents (their CHAIN CONDITION) and an interface condition on reflexive predicates (their REFLEXIVITY CONDITION B). Despite the conceptual problems noted in Chapter 3, the latter provides a useful way of thinking about what I called the DUALITY OF BP-PRONOUNS – it was, in fact, the main inspiration for the statement of the CRP. The former, however, is ultimately incompatible with locally bound BP-PRONOUNS for the same reason as the traditional CB.

Another difference between the proposal I arrived at in this chapter and Reinhart and Reuland's theory surveyed in the first part of this Thesis is the the fact that I am assuming a pragmatic definition of the notion of reflexive, rather than the purely morphosyntactic definition suggested in Reinhart and Reuland (1993, pg. 658). On this approach, there is no essential difference between reflexives and other kinds of pronouns: reflexives are simply anaphors that stand in some sort of markedness contrast with other pronoun forms in a language. It remains to be seen whether this can be extended to languages that mark reflexivity on verbs, rather than on arguments (FALTZ, 1985).

This approach suggests a story about how anaphoric systems develop in languages which is similar to the one offered by Levinson (1991, 2000) and Mattausch (2004). In a first stage,

speakers come up with marked forms to express semantic reflexivity in order to avoid the R-based inference of disjointness between coarguments, which occurs for most predicates (given that most predicates are arguably prototypically non-reflexive). At this stage, these marked forms are still somewhat somewhat *ad hoc* and are usually co-opted from the class of emphatic or intensifier markers in the language (e.g. Old English *self*, BP *mesmo*, French *même*).

Some languages reach a second stage where these reflexive marking strategies become grammaticalized, e.g., in the form of dedicated reflexive anaphors. This is when Condition A patterns start to emerge: i.e. we start to see forms that are grammatically restricted to signaling binding in local contexts (KEENAN, 2002). After this happens, another subset of languages move to a third stage, where the reflexive marking strategies become so frequently used in local binding contexts that they statistically preempt locally bound pronominals, giving rise to a CB PATTERN. As we saw in Section 7.3, BP possibly switched from the third to the second stage, while English reached the third stage sometime in the 17th century.

These three stages can be summarized as follows:

STAGE 1	CONDITION A AND CONDITION B PATTERNS ABSENT: Languages where grammaticalized reflexives are absent but the CRP requires *ad hoc* marked forms (e.g. emphatics) to expresses reflexivity for most predicates
STAGE 2	CONDITION A PATTERNS PRESENT CONDITION B PATTERNS ABSENT: Languages where reflexives are grammaticalized but still not frequent to the point of statistically preempting pronominals in local binding contexts
STAGE 3	CONDITION A AND CONDITION B PATTERNS PRESENT: Languages where reflexives are grammaticalized and widespread, statistically preempting pronominals in local binding contexts

Table 14 – Stages in the development of anaphoric systems

The overall picture we end up proposes to eliminate the CB as a syntactic anti-locality principle from Universal Grammar. Though my general proposal was stated within the framework of Simpler Syntax, this conclusion is a welcome one within the Minimalist Program as well, since the latter seeks to minimize the role of domain-specific innate principles in the explanation of linguistic phenomena, with the purpose of seeking a smoother integration with

plausible scenarios concerning language evolution (CHOMSKY, 2005; HORNSTEIN, 2018). My approach reduces the forces driving local disjointness effects to learned pairings of form and meaning (i.e. constructions acquired on the basis of statistical preemption) and to the CRP The genuine source of invariance behind local disjoints effects across languages are the pragmatic principles behind the CRP and the statistical learning mechanisms behind preemption – both of which arguably stem from domain-general cognitive principles.

CONCLUSION

Throughout this Thesis, I have argued that interpretive constraints on individual pronoun forms are either language-specific constraints establishing well-formed correspondences between SYN/GF and variable binding in SEM or epiphenomena of extra-grammatical factors. According to this conjecture, the main possible sources of anaphora universals are:

 (i) the vocabulary of syntactic dominance and grammatical functions (notions like PRED and GF-command) in terms of which the language-particular anaphora constructions are stated;

 (ii) the SEM concepts and relations that are typically encoded by language particular anaphora constructions (basically, A-binding, reflexive predicates, logophoricity);

(iii) extra-grammatical principles related to discourse, pragmatics and processing.

The first of these is explored in Dalrymple's (1993) LFG account as well as as well as Manzini and Wexler's (1987) parameterized Binding Theory. The second was the conceptual core of Reinhart's (1983, 2006) theory of anaphora over the years (SELLS, 1987; CHIERCHIA, 1989; REULAND, 2011; KOORNNEEF; REULAND, 2016; KRIFKA, 2018). The second is often discussed in functionally-oriented approaches (FALTZ, 1985; COMRIE, 1999; LEVINSON, 2000; MATTAUSCH, 2004). To my knowledge, however, these three factors have never been brought together in a systematic way within a formal theory of grammar.

The constructional approach, insofar as it reduces general principles of grammars to lexical facts, turns out to be the right one in allowing for the full-range of cross-linguistic diversity found in anaphoric systems. It naturally accounts for the fact that different languages may express the universal semantic category of reflexivity or A-binding by radically different means – e.g. by dependent-marking (as in English) or by head-marking (as in Kinyarwanda) (FALTZ, 1985; DALRYMPLE, 1993; VOLKOVA; REULAND, 2014).

The idea that language-specific constructions are violable (with violation possibly giving rise to marked meanings as implicatures) is a useful addition to this constructional stance, since it helps to account for the diverse behavior of specific pronoun forms within an individual language as well, as we saw in connection to English reflexives in Chapters 5-6. My analysis of English discursive reflexives in terms of acceptable ungrammaticality draws inspiration from suggestions first made in Menuzzi (1999, 2004). However, the theoretical framework I assume

is both more formally oriented and also more constructional than Menuzzi's peculiar blend of late Government and Binding theory with Optimality intuitions.

By reducing the syntactic anti-locality effect for pronominals typically handled by Condition B to a constructional property, my theory predicts that syntactic disjointness must be learned. This provides a novel way to interpret the well-attested fact that children do not display robust adult-level knowledge of Principle B until the age of seven (ELBOURNE, 2005; HAMANN, 2011; BAAUW, 2018). It also comes as no surprise that many languages simply lack a syntactic disjointness principle for pronominals of the kind seen in English. This hypothesis was instrumental in providing an explanation for the peculiar patterns of disjointness we find for pronominals in the São Paulo/Minas Gerais variety of Brazilian Portuguese.

In spite of its fairly new underpinnings, my proposal also turns have very traditional flavor. In this respect, it follows Pollard and Sag (1992) and Reinhart and Reuland (1993), who have similarly espoused going "back to basics" in their theoretical characterization of anaphoric forms. Much like Jespersen (1933), both of these accounts claim that the relevant domain in which grammar regulates the interpretation of pronouns is the SYNTACTIC PREDICATE – i.e. the domain over which valence and grammatical functions like Subject and Object can be defined. This reformulation accounts for the locality of reflexivization and the anti-locality of pronominalization (in languages where the latter exists) in a more natural and straightforward way than approaches based primarily on configurational licensing conditions (CHOMSKY, 1981; HORNSTEIN, 2001; CHARNAVEL, 2019). The challenge of making this compatible with non-local discursive reflexives and with the cross-linguistically ubiquitous urge to use special reflexive forms to signal local binding is turned over to pragmatics.

The common thread that runs through most of these arguments is the goal of making the account of pronouns (in English and other languages) as simple as possible. Since many theoretical devices and abstract phrase-structure configurations were postulated on the basis of facts about anaphora, a simpler theory of anaphora can contribute to a simpler overall theory of grammar – a welcome goal to many diverse contemporary grammatical theories (DOWTY, 1996; CULICOVER; JACKENDOFF, 2005; HORNSTEIN, 2009; MIHALIČEK; POLLARD, 2012; KUBOTA; LEVINE, 2020).

BIBLIOGRAPHY

ABBOTT, B. Support for individual concepts. *Linguistic and Philosophical Investigations*, n. 10, p. 23–44, 2011.

ACUÑA-FARIÑA, J. C. Opportunistic processing of language. *Language Sciences*, v. 57, p. 34–48, 2016.

ALMOR, A. Noun-phrase anaphora and focus: the informational load hypothesis. *Psychological Review*, v. 106, n. 4, p. 748–765, 1999.

ALMOR, A. Constraints and mechanisms in theories of anaphor processing. In: CROCKER, M. W.; PICKERING, M. J.; CLIFTON, C. (Ed.). *Architectures and mechanisms for language processing*. Cambridge, MA: MIT Press, 2000. p. 341–354.

ALSINA, A. *The Role of Argument Structure in Grammar*. Stanford, CA: CSLI Lecture Notes Series, 1996.

AMBRIDGE, B. et al. The retreat from overgeneralization in child language acquisition: Word learning, morphology, and verb argument structure. *Wiley Interdisciplinary Reviews: Cognitive Science*, Wiley Online Library, v. 4, n. 1, p. 47–62, 2013.

ANAGNOSTOPOULOU, E.; EVERAERT, M. Toward a more complete typology of anaphoric expressions. *Linguistic Inquiry*, v. 30, n. 1, p. 97–119, 1999.

ANDERSON, S. R. The typology of anaphoric dependencies: Icelandic (and other) reflexives. In: HELLAN, L.; CHRISTENSEN, K. K. (Ed.). *Topics in Scandinavian Syntax*. Dordrecht: Kluwer Academic Publishers, 1986. p. 65–88.

AOUN, J. *Generalized Binding: The Syntax and Logical Form of Wh-interrogatives*. Berlin: De Gruyter, 1986.

ARIEL, M. *Accessing Noun-Phrase Antecedents*. London: Routledge, 1990.

ARIEL, M. Accessibility theory: an overview. In: SANDERS, T.; SCHILPEROORD, J.; SPOOREN, W. (Ed.). *Text Representation: Linguistic and Psycholinguistic Aspects*. Amsterdam: John Benjamins, 2001. p. 29–87.

ARIEL, M. *Pragmatics and Grammar*. Cambridge: Cambridge Univeristy Press, 2008.

ARNOLD, J.; GRIFFIN, Z. The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, v. 56, n. 4, p. 521–536, 2007.

ARNOLD, J. E. Reference production: Production-internal and addressee-oriented processes. *Language and cognitive processes*, Taylor & Francis, v. 23, n. 4, p. 495–527, 2008.

ARNOLD, J. E. How speakers refer: The role of accessibility. *Language and Linguistics Compass*, v. 4, n. 4, p. 187–203, 2010.

ARNOLD, J. E.; LAO, S.-Y. C. Effects of psychological attention on pronoun comprehension. *Language, cognition and neuroscience*, v. 30, n. 7, p. 832–852, 2015.

ASUDEH, A. Direct compositionality and the architecture of LFG. In: BUTT, M.; DALRYMPLE, M.; KING, T. H. (Ed.). *Intelligent Linguistic Architectures: Variations on Themes by Ronald M. Kaplan*. Stanford: CSLI Publications, 2006.

ASUDEH, A.; DALRYMPLE, M.; TOIVONEN, I. Constructions with lexical integrity. *Journal of Language Modelling*, v. 1, n. 1, p. 1–54, 2013.

ASUDEH, A.; GIORGOLO, G.; TOIVONEN, I. Meaning and valency. In: CSLI PUBLICATIONS. *Proceedings of the LFG14 Conference*. [S.l.], 2014. p. 68–88.

ASUDEH, A.; KELLER, F. Experimental evidence for a predication-based binding theory. *Chicago Linguistic Society*, v. 37, p. 1–14, 2001.

ATLAS, J. D.; LEVINSON, S. C. It-clefts, informativeness and logical form. In: COLE, P. (Ed.). *Radical pragmatics*. San Diego, CA: Academic Press, 1981. p. 1–62.

BAAUW, S. The interpretation of pronouns across Spanish-Speaking Populations. *Languages*, Multidisciplinary Digital Publishing Institute, v. 3, n. 2, p. 11, 2018.

BAAUW, S.; DELFITTO, D. New views on reflexivity: Delay effects in Romance. *Probus*, v. 17, n. 2, p. 145–184, 2005.

BACH, E. Problominalization. *Linguistic Inquiry*, v. 1, n. 1, p. 121–123, 1970.

BACH, E. Control in Montague Grammar. *Linguistic Inquiry*, v. 10, n. 4, p. 515–531, 1979.

BACH, E.; PARTEE, B. H. Anaphora and semantic structure. In: KREIMAN, J.; OJEDA, A. E. (Ed.). *Papers from the Parasesion on Pronouns and Anaphora*. Chicago Illinois: University of Chicago, 1980. p. 1–28.

BACH, K.; HARNISH, R. *Linguistic communication and speech acts*. Cambridge, MA: MIT Press, 1979.

BAILYN, J. F. A (purely) derivational approach to Russian scrambling. In: *Formal approaches to Slavic linguistics*. [S.l.: s.n.], 2003. v. 11, p. 41–62.

BAKER, C. L. Contrast, discourse prominence, and intensification, with special reference to locally free reflexives in British English. *Language*, v. 71, n. 1, p. 63–101, 1995.

BAKER, M. Thematic roles and syntactic structure. In: HAEGEMAN, L. (Ed.). *Elements of Grammar*. Dordrecht: Kluwer Academic Publishers, 1997. p. 73–137.

BAKER, M. C. The macroparameter in a microparametric world. In: BIBERAUER, T. (Ed.). *The Limits of Syntactic Variation*. Philadelphia: John Benjamins, 2008.

BALTIN, M.; POSTAL, P. M. More on reanalysis hypothesis. *Linguistic Inquiry*, v. 27, n. 1, p. 127–145, 1996.

BARKER, C. Quantificational binding does not require c-command. *Linguistic inquiry*, MIT Press, v. 43, n. 4, p. 614–633, 2012.

BARSS, A. *Chains and Anaphoric Dependence: On Reconstruction and its Implications*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, 1986.

BARSS, A.; LASNIK, H. A note on anaphora and double objects. *Linguistic Inquiry*, v. 17, n. 2, p. 347–354, 1986.

BEAVERS, J. On affectedness. *Natural Language & Linguistic Theory*, Springer, v. 29, n. 2, p. 335–370, 2011.

BELLETTI, A.; RIZZI, L. Psych-verbs and θ-theory. *Natural Language and Linguistic Theory*, v. 6, n. 3, p. 291–352, 1988.

BENEDICTO, E. Latin long-distance anaphora. In: KOSTER, J.; REULAND, E. (Ed.). *Long-Distance Anaphora*. Cambridge: Cambridge Univeristy Press, 1991. p. 171–184.

BERNSTEIN, L. *The Unanswered Question: six talks at Harvard*. Cambride, MA: Harvard University Press, 1976.

BERTOLINO, K. G. *Restrições sobre a interpretação da proforma ele com antecedente local do Português Brasileiro: Um estudo experimental*. Phd Thesis (PhD Thesis) — Universidade de Sao Paulo, Sao Paulo, Brazil, 2013.

BERTOLINO, K. G.; GROLLA, E. O pronome" ele" está sujeito ao Princípio B? uma discussão sobre resultados experimentais. *Revista Linguistica*, v. 8, n. 2, 2012.

BERWICK, R. C.; CHOMSKY, N. *Why only us: Language and evolution*. Cambride, MA: MIT press, 2016.

BLOOMFIELD, L. *Language*. New York and London: Henry Holt, 1933.

BOECKX, C. *Bare Syntax*. New York, NY: Oxford University Press, 2008.

BOECKX, C. Not only us. *Inference: International Review of Science*, v. 3, n. 1, 2017.

BOECKX, C.; HORNSTEIN, N.; NUNES, J. Overt copies in reflexive and control structures: a movement analysis. *University of Maryland working papers in linguistics*, v. 15, n. 1, p. 1–46, 2007.

BOER, B. de et al. Evolutionary dynamics do not motivate a single-mutant theory of human language. *Scientific reports*, Nature Publishing Group, v. 10, n. 1, p. 1–9, 2020.

BOLINGER, D. Transitivity and spatiality: The passive of prepositional verbs. In: MAKKAI, A.; MAKKAI, V. B.; HEILMANN, L. (Ed.). *Linguistics at the Crossroads*. Lake Bluff, IL: Jupiter Press, 1977. p. 57–78.

BOSCH, A. R. Syllable-internal structure. In: OOSTENDORP, M. et al. (Ed.). *The Blackwell companion to phonology*. Malden, MA: Blackwell, 2011. p. 1–18.

BOUCHARD, D. *On the content of Empty Categories*. Dordecht: Foris, 1984.

BOUMA, G.; SPENADER, J. The distribution of weak and strong object reflexives in Dutch. *LOT Occasional Series*, v. 12, p. 103–114, 2008.

BOYD, J. K.; GOLDBERG, A. E. Learning what not to say: The role of statistical preemption and categorization in a-adjective production. *Language*, JSTOR, p. 55–83, 2011.

BRANCO, A.; MARRAFA, P. Long-distance reflexives and the binding square of opposition. In: WEBELHUTH, G.; KOENIG, J.-P.; KATHOL, A. (Ed.). *Lexical and Constructional Aspects of Linguistics Explanation*. Stanford, CA: CSLI Publications, 1999. p. 163–177.

BRENNAN, S. E. Centering attention in discourse. *Language and Cognitive processes*, Taylor & Francis, v. 10, n. 2, p. 137–167, 1995.

BRESNAN, J. Control and complementation. *Linguistic Inquiry*, v. 13, n. 3, p. 343–434, 1982.

BRESNAN, J. The passive in lexical theory. In: BRESNAN, J. (Ed.). *The Mental Representation of Grammatical Relations*. Cambridge: M.I.T. Press, 1982. p. 3–86.

BRESNAN, J. Linear order, syntactic rank, and empty categories: On weak crossover. In: DALRYMPLE, M.; KAPLAN RONALD M, J. T. M. I.; ZAENEN, A. (Ed.). *Formal Issues in Lexical-Functional Grammar*. [S.l.]: Center for the Study of Language (CSLI), 1995.

BRESNAN, J. *Lexical-Functional Syntax*. Oxford: Blackwell, 2001.

BRESNAN, J.; KANERVA, J. M. Locative inversion in Chichewa: a case study of factorization in grammar. *Linguistic Inquiry*, v. 20, n. 1, p. 1–50, 1989.

BRESNAN, J.; KAPLAN, R. Lexical-Functional Grammar: A Formal System for Grammatical Representation. In: BRESNAN, J. (Ed.). *The Mental Representation of Grammatical Relations*. Cambridge, MA: The MIT Press, 1982. p. 173–281.

BRESNAN, J.; KAPLAN, R. M. Introduction: Grammars as mental representations of language. In: *The Mental Representation of Grammatical Relations,*. Cambridge, MA: MIT Press, 1982. p. xvii–lii.

BRESNAN, J.; MCHOMBO, S. A. The lexical integrity principle: Evidence from Bantu. *Natural Language and Linguistic Theory*, v. 13, n. 2, p. 181–254, 1995.

BRITO, D. B. S. d.; SEDRINS, A. P. Sistemas de concordância e de correferência no paradigma pronominal de primeira pessoa do plural no português brasileiro. *Atas do V SIMELP*, p. 13–28, 2017.

BROWN, P.; LEVINSON, S. C. *Politeness: some universals in language usage*. Cambridge: Cambridge University Press, 1987.

BRUENING, B. Precede-and-command revisited. *Language*, Linguistic Society of America, v. 90, n. 2, p. 342–388, 2014.

BRUENING, B. The lexicalist hypothesis: Both wrong and superfluous. *Language*, Linguistic Society of America, v. 94, n. 1, p. 1–42, 2018.

BRUNETTO, V.; ROEPER, T. The acquisition path of near-reflexivity. In: SYRETT, K.; ARUNACHALAM, S. (Ed.). *Semantics in Language Acquisition*. Philadelphia, PA: John Benjamins Publishing Company, 2018. v. 24, p. 351–377.

BÜRING, D. Crossover situations. *Natural Language Semantics*, v. 12, n. 1, p. 23–62, 2004.

BÜRING, D. *Binding Theory*. Cambridge: Cambridge University Press, 2005.

BURKHARDT, P. Logophors: looking outside of syntax. In: *Proceedings of console IX*. [S.l.: s.n.], 2002. p. 14–27.

BURZIO, L. On the non-existence of disjoint reference principles. *Rivista di Grammatica Generativa, vol. 14 (1989), p. 3-27*, Unipress, 1989.

BURZIO, L. The morphological basis of anaphora. *Journal of Linguistics*, JSTOR, v. 27, n. 1, p. 81–105, 1991.

BURZIO, L. Anaphora and soft constraints. In: BARBOSA, P. et al. (Ed.). *Is the Best Good Enough?* Cambridge, MA: MIT Press, 1998. p. 93–114.

BYBEE, J.; DAHL, Ö. The creation of tense and aspect systems in the languages of the world. *Studies in Language*, v. 1, n. 13, p. 51–103, 1989.

BYBEE, J.; PERKINS, R.; PAGLIUCA, W. *The Evolution of Grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press, 1994.

CANTRALL, W. R. *On the nature of the reflexive in English*. Phd Thesis (PhD Thesis) — University of Illinois, Urbana-Champaign, Illinois, 1969.

CANTRALL, W. R. *Viewpoint, Reflexives, and the Nature of Noun Phrases*. The Hague: Mouton, 1974.

CARSTON, R. Enrichment and loosening: Complementary processes in deriving the proposition expressed? *Linguistische Berichte*, v. 1, n. 8, p. 103–127, 1997.

CARVALHO, J. Is 'ele' an anahor in Brazilian Portuguese? a study of small clause subjects. Handout. 2019.

CARVALHO, J.; CALINDRO, A. R. A unified account for the loss of third person clitics in Brazilian Portuguese. In: CARVALHO, D.; BRITO, D. (Ed.). *Pronomes: Morfossintaxe e semântica*. Salvador: EdUFBA, 2018. p. 91–110.

CASASANTO, L. S.; SAG, I. A. The advantage of the ungrammatical. In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. [S.l.: s.n.], 2008. v. 30, n. 30.

CHAMBERS, C. G.; SMYTH, R. Structural parallelism and discourse coherence: a test of Centering Theory. *Journal of Memory and Language*, v. 39, n. 4, p. 593–608, 1998.

CHARNAVEL, I. *Locality and Logophoricity: A Theory of Exempt Anaphora*. New York, NY: Oxford University Press, 2019.

CHARNAVEL, I.; SPORTICHE, D. Anaphor binding: What french inanimate anaphors show. *Linguistic Inquiry*, v. 47, n. 1, p. 35–87, 2016.

CHARNAVEL, I. C.; ZLOGAR, C. D. English reflexive logophors. In: *Proceedings of the 51st annual meeting of the Chicago Linguistic Society*. Chicago, IL: [s.n.], 2015.

CHAVES, R. P. An expectation-based account of subject islands and parasitism. *Journal of Linguistics*, Cambridge Univeristy Press, v. 49, n. 2, p. 285–327, 2013.

CHAVES, R. P.; DERY, J. E. Which subject islands will the acceptability of improve with repeated exposure? In: SANTANA-LABARGE, R. E. (Ed.). *Proceedings of the Thirty-First Meeting of the West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project, 2014. p. 96–106.

CHAVES, R. P.; DERY, J. E. Frequency effects in subject islands. *Journal of linguistics*, Cambridge University Press, v. 55, n. 3, p. 475–521, 2019.

CHIEN, Y.-C.; WEXLER, K. Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition*, v. 1, n. 3, p. 225–295, 1990.

CHIERCHIA, G. Anaphora and attitudes de se. In: BARTSCH, R.; BENTHEM, J. F. A. K. van; BOAS, P. van E. (Ed.). *Semantics and Contextual Expression*. Dordecht: Foris Publications, 1989.

CHIERCHIA, G. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In: BELLETTI, A. (Ed.). *Structures and Beyond: The Cartography of Syntactic Structures*. Oxford: Oxford University Press, 2004. v. 3, p. 39–103.

CHIERCHIA, G. *Logic in grammar: polarity, free choice and intervention*. Oxford: Oxford University Press, 2013.

CHIOU, M. *NP-Anaphora in Modern Greek: A Partial Neo-Gricean Pragmatic Approach*. Newcastle: Cambridge Scholars Publishing, 2010.

CHOMSKY, N. *Morphophonemics of Modern Hebrew*. Master's Thesis (Master's Thesis) — University of Pennsylvania, 1951.

CHOMSKY, N. Logical structure of linguistic theory. Mimeo. 1955.

CHOMSKY, N. *Syntactic structures*. The Hague: Mouton, 1957.

CHOMSKY, N. On certain formal properties of grammars. *Information and control*, Elsevier, v. 2, n. 2, p. 137–167, 1959.

CHOMSKY, N. Some methodological remarks on generative grammar. *Word*, Taylor & Francis, v. 17, n. 2, p. 219–239, 1961.

CHOMSKY, N. *Aspects of the theory of syntax*. Cambridge: MIT Press, 1965. 251 p.

CHOMSKY, N. Remarks on nominalization. In: JACOBS, J.; ROSENBAUM, P. (Ed.). *Readings in English Transformational Grammar*. Waltham, Massachusetts: Ginn, 1970. p. 184–221.

CHOMSKY, N. Deep structure, surface structure, and semantic interpretation. In: STEINBERG, D.; JACOBOVITS, L. (Ed.). *Semantics in generative grammar*. Cambridge: Cambridge Univeristy Press, 1971. p. 183–216.

CHOMSKY, N. Conditions on transformations. In: ANDERSON, S.; KIPARSKY, P. (Ed.). *A Festschrift for Morris Halle*. New York, NY: Holt, Reinhart & Winston, 1973. p. 232–286.

CHOMSKY, N. *Reflections on language*. New York, NY: Pantheon Books, 1975.

CHOMSKY, N. On *wh*-movement. In: CULICOVER, P. W.; WASOW, T.; AKMAJIAN, A. (Ed.). *Formal syntax*. New York, NY: Academic Press, 1977. p. 71–132.

CHOMSKY, N. On binding. *Linguistic Inquiry*, v. 11, n. 1, p. 1–46, 1980.

CHOMSKY, N. *Lectures on government and binding*. Dordecht: Foris, 1981.

CHOMSKY, N. *Some concepts and consequences of the theory of government and binding.* Cambridge, MA: MIT Press, 1982. 110 p.

CHOMSKY, N. *Barriers.* Cambridge, MA: MIT Press, 1986.

CHOMSKY, N. *Knowledge of language.* New York, NY: Praeger Publishers, 1986.

CHOMSKY, N. Some notes on economy of derivation and representation. In: FREIDIN, R. (Ed.). *Principles and Parameters in Comparative Grammar.* Cambridge, MA: MIT Press, 1991. p. 417–454.

CHOMSKY, N. A minimalist program for linguistic theory. In: HALE, K.; KEYSER, J. (Ed.). *The view from Building 20.* Cambridge, MA: MIT Press, 1993. p. 1–52.

CHOMSKY, N. *The minimalist program.* Cambridge, MA: The MIT Press, 1995. 420 p.

CHOMSKY, N. Derivation by phase. In: KENSTOWICZ, M. (Ed.). *Ken Hale: A Life in Linguistics.* Cambridge, MA: MIT Press, 2001. p. 1–52.

CHOMSKY, N. Three factors in language design. *Linguistic inquiry*, MIT press, v. 36, n. 1, p. 1–22, 2005.

CHOMSKY, N. Approaching UG from below. In: SAUERLAND, U.; GÄRTNER, H.-M. (Ed.). *Interfaces + Recursion = Language?* 1–30: De Gruyter, 2007.

CHOMSKY, N. On phases. In: FREIDIN, R.; OTERO, C. P.; ZUBIZARRETA, M. L. (Ed.). *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud.* Cambridge, MA: MIT Press, 2008. p. 133–66.

CHOMSKY, N. Problems of projection. *Lingua*, Elsevier, v. 130, p. 33–49, 2013.

CHOMSKY, N.; GALLEGO, Á. J.; OTT, D. Generative grammar and the faculty of language: Insights, questions, and challenges. *Catalan Journal of Linguistics*, p. 229–261, 2019.

CHOMSKY, N.; HALLE, M. *The Sound Pattern of English.* New York, NY: Haper & Row, Publishers, 1968. 470 p.

CHOMSKY, N.; LASNIK, H. Principles and parameter theory. In: STECHOW, A. von; STERNEFELD, W.; VENNEMANN, T. (Ed.). *Syntax: an International Handbook of Contemporary Research.* Berlin: Walter de Gruyter, 1993. p. 506–569.

CHRISTENSEN, K. R. The dead ends of language: The (mis) interpretation of a grammatical illusion. In: VIKNER, S.; JORGENSEN, H.; GELDEREN, E. van (Ed.). *Let us have articles betwixt us: Papers in Historical and Comparative Linguistics in Honour of Johanna L. Wood.* Aarhus: Aarhus University, 2016. p. 129–159.

CINQUE, G. *Adverbs and functional heads: a cross-linguistic perspective.* New York, NY: Oxford University Press, 1999.

CLARK, A.; LAPPIN, S. *Linguistic Nativism and the poverty of the stimulus.* Oxford: Blackwell, 2011.

CLEMENTS, G. N. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages*, v. 10, n. 2, p. 141–177, 1975.

COLE, P.; HERMON, G.; YANTI. Grammar of binding in the languages of the world: Innate or learned? *Cognition*, v. 141, p. 138–160, 2015.

COLE, P.; SUNG, L.-M. Head movement and long-distance reflexives. *Linguistic Inquiry*, v. 25, n. 4, p. 335–406, 1994.

COLLINS, C.; STABLER, E. A formalization of minimalist syntax. *Syntax*, Wiley Online Library, v. 19, n. 1, p. 43–78, 2016.

COMRIE, B. Reference-tracking: description and explanation. *STUF-Language Typology and Universals*, De Gruyter, v. 52, n. 3-4, p. 335–346, 1999.

CONROD, K. Pronouns and gender in language. In: HALL, K.; BARRETT, R. (Ed.). *The Oxford Handbook of Language and Sexuality*. Oxford: Oxford University Press, 2018.

COOPER, R. The interpretation of pronouns. In: HENY, F.; SCHNELLE, H. (Ed.). *Syntax and Semantics 10: Selections from the Third Groningen Round Table*. New York: Academic Press, 1979. p. 61–92.

COOPER, R. *Quantification and Syntactic Theory*. Dordecht: Reidel, 1983.

COPESTAKE, A.; BRISCOE, T. Lexical operations in a unification-based framework. In: SPRINGER. *Workshop of SIGLEX (Special Interest Group within ACL on the Lexicon)*. [S.l.], 1991. p. 101–119.

CORBALLIS, M. C. The evolution of language. In: CALL, J. et al. (Ed.). *APA handbook of comparative psychology: Basic concepts, methods, neural substrate, and behavior*. Washington, DC: American Psychological Association, 2017. p. 273–297.

CULCOVER, P. W. The Minimalist impulse. In: CULICOVER, P. W.; MCNALLY, L. (Ed.). *The Limits of Syntax*. San Diego: Academic Press, 1998, (Syntax and Semantics, v. 29).

CULICOVER, P. W. A constraint on coreferentiality. *Foundations of Language*, v. 14, n. 1, p. 109–118, 1976.

CULICOVER, P. W. *Principles and Parameters*. Oxford: Oxford University Press, 1997.

CULICOVER, P. W. *Syntactic nuts: hard cases, syntactic theory, and language acquisition*. Oxford: Oxford University Press, 1999.

CULICOVER, P. W. *Natural language syntax*. New York, NY: Oxford University Press, 2009.

CULICOVER, P. W. English (zero-)relatives and the competence-performance distinction. *International Review of Pragmatics*, v. 5, p. 253–270, 2013.

CULICOVER, P. W. *Grammar and Complexity: Language at the Intersection of Competence and Performance*. Oxford: Oxford University Press, 2013.

CULICOVER, P. W. Simpler Syntax and explanation. In: Müller, S. (Ed.). *The 20th International Conference on Head-Driven Phrase Structure Grammar*. Stanford: CSLI, 2013. p. 263–283.

CULICOVER, P. W. Constructions, complexity and word order variation. In: NEWMEYER, F. J.; PRESTON, L. B. (Ed.). *Measuring Linguistic Complexity*. Oxford: Oxford University Press, 2014. p. 148–178.

CULICOVER, P. W. Simpler Syntax and the mind. In: TOIVONEN, I.; CSURI, P.; ZEE, E. V. D. (Ed.). *Structures in the mind: Essays on Language, Music, and Cognition in Honor of Ray Jackendoff*. Cambride, MA: MIT Press, 2015. p. 3–20.

CULICOVER, P. W. *Language Change, Variation, and Universals: A constructional approach*. Oxford: Oxford University Press, 2021.

CULICOVER, P. W.; JACKENDOFF, R. Something else for the binding theory. *Linguistic Inquiry*, v. 26, n. 2, p. 249–276, 1995.

CULICOVER, P. W.; JACKENDOFF, R. *Simpler syntax*. New York, NY: Oxford University Press, 2005.

CULICOVER, P. W.; JACKENDOFF, R.; AUDRING, J. Multiword constructions in the grammar. *Topics in cognitive science*, v. 9, n. 3, p. 552–568, 2017.

CULICOVER, P. W.; WINKLER, S. *Subject islands aren't islands, and parasitic gaps aren't parasitic or, The Case of the Uninvited Guest*. 2018. Unpublished manuscript. The Ohio State University and University of Tübingen.

CUNNINGS, I.; PATTERSON, C.; FELSER, C. Structural constraints on pronoun binding and coreference: evidence from eye movements during reading. *Frontiers in psychology*, v. 6, p. 840, 2015.

CURRY, H. B. Some logical aspects of grammatical structure. In: JACOBSON, R. (Ed.). *Structure of Language and its Mathematical Aspects: Proceedings of the Twelfth Symposium in Applied Mathematics*. [S.l.]: American Mathematical Society, 1963. p. 56–68.

CYRINO, S. M. L. Para a história do português brasileiro: a presença do objeto nulo e a ausência dos clíticos. *Letras de Hoje*, v. 38, n. 1, 2003.

D'ALBUQUERQUE, A. da C. R. C. A perda dos clíticos num dialeto mineiro. *Tempo Brasileiro*, n. 78/79, 1984.

DALRYMPLE, M. *The syntax of anaphoric binding*. Stanford, CA: CSLI, 1993.

DALRYMPLE, M. *Lexical Functional Grammar*. San Diego, CA: Academic Press, 2001.

DALRYMPLE, M.; KAPLAN, R.; KING, T. H. Linguistic generalizations over descriptions. In: BUTT, M.; KING, T. H. (Ed.). *Proceedings of the LFG04 Conference*. Stanford, CA: [s.n.], 2004. p. 199–208.

DAVIS, A.; KOENIG, J.-P.; WECHSLER, S. Argument structure and linking. In: MÜLLER, S. et al. (Ed.). *Head-Driven Phrase Structure Grammar: The handbook*. Berlin: Language Science Press, 2021.

DAVISON, A. Peculiar passives. *Language*, v. 56, n. 1, p. 42–66, 1980.

DÉCHAINE, R.-M.; WILTSCHKO, M. A formal typology of reflexives. *Studia Linguistica*, v. 71, n. 1-2, p. 60–106, 2017.

DENNETT, D. C. *From bacteria to Bach and back: The evolution of minds*. New York, NY: WW Norton & Company, 2017.

DOUGHERTY, R. C. An interpretive theory of pronominal reference. *Foundations of Language*, p. 488–519, 1969.

DOWTY, D. Comments on the paper by Bach and Partee. In: KREIMAN, J.; OTEDA, A. (Ed.). *Papers from the Parasession on Pronouns and Anaphora*. [S.l.: s.n.], 1980.

DOWTY, D. R. On recent analyses of the semantics of control. *Linguistics and Philosophy*, v. 8, n. 3, p. 291–331, 1985.

DOWTY, D. R. Toward a minimalist theory of syntactic structure. In: BUNT, H.; HORCK, A. van (Ed.). *Discontinuous Constituency*. Berlin: Mouton De Gruyter, 1996. v. 6, p. 11–62.

DUARTE, M. E. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do brasil. In: ROBERTS, I.; KATO, M. A. (Ed.). *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993. p. 107–128.

DUBINSKY, S.; HAMILTON, R. Epithets as antilogophoric pronouns. *Linguistic Inquiry*, v. 29, n. 4, p. 685–693, 1998.

EDMONDSON, J. A.; PLANK, F. Great expectations: An intensive self analysis. *Linguistics and Philosophy*, Springer, v. 2, n. 3, p. 373–413, 1978.

ELBOURNE, P. On the acquisition of Principle B. *Linguistic Inquiry*, v. 36, n. 3, p. 333–365, 2005.

EMONDS, J. E. *Root and structure-preserving transformations.* Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, 1970.

ERTESCHIK-SHIR, N. *On the Nature of Island Constraints*. Phd Thesis (PhD Thesis) — MIT, Cambridge, MA, 1973.

ERTESCHIK-SHIR, N. Bridge phenomena. In: EVERAERT, M.; RIEMSDIJK, H. van (Ed.). *Blackwell Companion to Syntax*. [S.l.]: Wiley Online Library, 2006. p. 284–294.

EVANS, G. Pronouns. *Linguistic Inquiry*, v. 11, n. 2, p. 337–362, 1980.

EVERAERT, M. Contextual determination of the anaphor/pronominal distinction. In: KOSTER, J.; REULAND, E. (Ed.). *Long-Distance Anaphora*. Cambridge: Cambridge University Press, 1991. p. 77–118.

EVERETT, D. Why there are no clitics: On the storage, insertion and form of the ϕ-features. In: COOPMANS, P.; EVEREART, M.; GRIMSHAW, J. (Ed.). *Lexical Specification and Insertion*. Philadelphia, PA: John Benjamins, 2000. p. 91–114.

FALK, Y. *Lexical-Functional Grammar: An Introduction to Parallel Constraint-Based Syntax*. Stanford: CSLI Publications, 2001.

FALTZ, L. M. *Reflexivization: A study in universal syntax*. New York, NY: Garland, 1985.

FANSELOW, G. On the sentential nature of prenominal adjectives in German. *Folia Linguistica*, v. 20, n. 3-4, p. 341–380, 1986.

FARMER, A. K.; HARNISH, R. M. Communicative reference with pronouns. In: VERSCHUEREN, J.; BERTUCCELLI-PAPI, M. (Ed.). *The pragmatic perspective*. Amsterdam: John Benjamins, 1987. p. 547–565.

FERREIRA, F.; PATSON, N. D. The 'good enough'approach to language comprehension. *Language and Linguistics Compass*, Wiley Online Library, v. 1, n. 1-2, p. 71–83, 2007.

FERREIRA, M. B. *Argumentos Nulos em Portugues Brasileiro*. Phd Thesis (PhD Thesis) — Universidade Estadual de Campinas, Campinas, 2000.

FIENGO, R.; MAY, R. *Indices and identity*. Cambridge, MA: MIT Press, 1994.

FINDLAY, J. Mapping theory without argument structure. *Journal of Language Modelling*, v. 4, n. 2, p. 293–338, 2017.

FITCH, W. T. Empirical approaches to the study of language evolution. *Psychonomic bulletin & review*, Springer, v. 24, n. 1, p. 3–33, 2017.

FLACH, S. Constructionalization and the Sorites Paradox. In: SOMMERER, L.; SMIRNOVA, E. (Ed.). *Nodes and Networks in Diachronic Construction Grammar*. Philadelphia, PA: John Benjamins, 2020. v. 27.

FODOR, J. A. *The modularity of mind : an essay on faculty psychology*. Cambridge, MA: MIT Press, 1983. 145 p.

FODOR, J. A.; BEVER, T. G.; GARRETT, M. F. *The Psychology of Language*. New York, NY: McGraw Hill, 1974.

FOX, D. Chain and binding: a modification of Reinhart and Reuland's 'Reflexivity'. Unpublished Manuscript. Cambridge, MA. 1993.

FOX, D. Locality in variable binding. In: BARBOSA, P. et al. (Ed.). *Is the Best Good Enough?* Cambridge, MA: MIT Press, 1998. p. 129–156.

FRAZIER, L. Syntactic complexity. In: DOWTY, D.; KARTTUNEN, L.; ZWICKY, A. M. (Ed.). *Natural language parsing: Psychological, computational, and theoretical perspectives*. Cambridge: Cambridge Univeristy Press, 1985. p. 129–189.

FROUD, K. Prepositions and the lexical/functional divide: Aphasic evidence. *Lingua*, Citeseer, v. 111, n. 1, p. 1–28, 2001.

GAJEWSKI, J. L-analycity in natural language. *Unpublished manuscript: MIT*, 2002.

GALVES, C. A interpretação "reflexiva" do pronome no português do Brasil. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada*, v. 2, n. 2, 1986.

GALVES, C. Tópicos, sujeitos, pronomes e concordância no português brasileiro. *Cadernos de estudos lingüísticos*, v. 34, 1998.

GELDEREN, E. van. *A history of English reflexive pronouns: Person, self, and interpretability*. Amsterdam: John Benjamins, 2000.

GEURTS, B. Weak and strong reflexives in Dutch. In: SCHLENKER, P.; KEENAN, E. (Ed.). *Proceedings of the ESSLLI workshop on semantic approaches to binding theory*. Nancy: Université Henri Poincaré, 2004.

GIBSON, E. The dependency locality theory: a distance-based theory of linguistic complexity. In: MARANTZ, A.; MIYASHITA, Y.; O'NEIL, W. (Ed.). *Image, language, brain*. Cambridge, MA: MIT Press, 2000. p. 95–126.

GIBSON, E.; THOMAS, J. Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes*, v. 14, n. 3, p. 225–248, 1999.

GIORGI, A. On the nature of long-distance anaphors. *Linguistic Inquiry*, v. 38, n. 2, p. 321–342, 2007.

GIVÓN, T. *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins Publishing Company, 1983.

GIVÓN, T. Markedness in grammar: Distributional, communicative and cognitive correlates of syntactic structure. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, John Benjamins, v. 15, n. 2, p. 335–370, 1991.

GOLDBERG, A. E. *Constructions: a construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press, 1995.

GOLDBERG, A. E. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press, 2006.

GOLDBERG, A. E. Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics*, De Gruyter Mouton, v. 22, n. 1, p. 131–153, 2011.

GOLDBERG, A. E. Constructionist approaches. In: HOFFMANN, T.; TROUSDALE, G. (Ed.). *Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press, 2013. p. 15–31.

GOLDBERG, A. E. Partial productivity of linguistic constructions: Dynamic categorization and statistical preemption. *Language and cognition*, Cambridge University Press, v. 8, n. 3, p. 369–390, 2016.

GOLDBERG, A. E.; JACKENDOFF, R. The English resultative as a family of constructions. *Language*, v. 80, n. 3, p. 532–568, 2004.

GOLDE, K. E. *Binding theory and beyond: An investigation into the English Pronominal System*. Phd Thesis (PhD Thesis) — The Ohio State University, Columbus, OH, 1999.

GOLDSMITH, J. A. *Autosegmental phonology*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, Cambride, MA, 1976.

GORDON, P. C.; HENDRICK, R. The representation and processing of coreference in discourse. *Cognitive Science*, Elsevier, v. 22, n. 4, p. 389–424, 1998.

GRICE, P. Logic and conversation. In: *Studies in the Way of Words*. Cambridge, MA: Harvard University Press, 1989. p. 22–40.

GRINDER, J.; POSTAL, P. M. Missing antecedents. *Linguistic Inquiry*, v. 2, n. 3, p. 269–312, 1971.

GRODZINSKY, Y.; REINHART, T. The innateness of binding and coreference. *Linguistic Inquiry*, v. 24, n. 1, p. 69–102, 1993.

GROLLA, E. The acquisition of contrastive and non-contrastive anaphoric forms in Brazilian Portuguese. In: *Selected Proceedings of the 4th Conference on Generative Approaches to Language Acquisition North America*. [S.l.: s.n.], 2011. p. 78–89.

GROLLA, E.; BERTOLINO, K. G. A proforma 'ele' com antecedente local em portugues brasileiro adulto e infantil. In: *Estudos da Linguagem: casamento entre temas e perspectivas*. Joao Pessoa: Editora Universitária da UFPB, 2011.

GUNDEL, J.; HEDBERG, N.; ZACHARSKI, R. Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, v. 69, n. 2, p. 274–307, 1993.

HAGÈGE, C. Les pronoms logophoriques. *Bulletin de la Société de Linguistique*, v. 69, n. 1, p. 287–310, 1974.

HAIDER, H.; OLSEN, S.; VIKNER, S. Introduction. In: HAIDER, H.; OLSEN, S.; VIKNER, S. (Ed.). *Studies in comparative Germanic syntax*. Dordecht: Kluwer Academic Publishers, 1995. p. 1–45.

HAIMAN, J. Iconic and economic motivation. *Language*, v. 59, n. 4, p. 781–819, 1983.

HALE, K.; KEYSER, S. J. On argument structure and the lexical expression of syntactic relations. In: HALE, K.; KEYSER, S. J. (Ed.). *The view from building 20*. Cambridge, MA: MIT Press, 1993. p. 53–110.

HALLE, M.; MARANTZ, A. Some key features of distributed morphology. *MIT working papers in linguistics*, v. 21, n. 275, p. 88, 1994.

HAMANN, C. Binding and coreference: Views from child language. In: VILLIERS, J. de; ROEPER, T. (Ed.). *Handbook of generative approaches to language acquisition*. Dordrecht: Springer, 2011. p. 247–290.

HANKAMER, J. Morphological parsing and the lexicon. In: MARSLEN-WILSON, W. (Ed.). *Lexical Representation and Process*. Cambridge: MIT Press, 1989.

HARRIS, Z. Co-occurrence and transformation in linguistic structure. *Language*, v. 33, n. 3, p. 283–340, 1957.

HASPELMATH, M. Against markedness (and what to replace it with). *Journal of Linguistics*, p. 25–70, 2006.

HASPELMATH, M. A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery*, Dartmouth College, v. 6, n. 1, p. 40–63, 2008.

HAWKINS, J. A. *Efficiency and complexity in grammars*. Oxford: Oxford University Press, 2004.

HEIM, I. Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. In: PERCUS, O.; SAUERLAND, U. (Ed.). *MIT Working Papers in Linguistics The Interpretive Tract*. Cambridge, MA: MIT, Department of Linguistics, 1998. v. 25, p. 205–246.

HEIM, I.; KRATZER, A. *Semantics in generative grammar*. Malden, MA: Blackwell, 1998.

HEINAT, F. *Probes, Pronouns and Binding in the Minimalist Program*. Phd Thesis (PhD Thesis) — Lund University, Lund, 2006.

HELKE, M. *The Grammar of English Reflexives*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, Cambride, MA, 1970.

HELKE, M. *The Grammar of English Reflexives*. New York, NY: Garland, 1979.

HELLAN, L. Containment and connectedness anaphors. In: KOSTER, J.; REULAND, E. (Ed.). *Long-Distance Anaphora*. Cambridge: Cambridge Univeristy Press, 1991.

HENDRIKS, P.; HOEKS, J. C.; SPENADER, J. Reflexive choice in dutch and german. *The Journal of Comparative Germanic Linguistics*, Springer, v. 17, n. 3, p. 229–252, 2014.

HENDRIKS, P.; SPENADER, J.; SMITS, E.-J. Frequency-based constraints on reflexive forms in Dutch. In: *Proceedings of the 5th International Workshop on Constraints and Language Processing*. [S.l.: s.n.], 2008. p. 33–47.

HESTVIK, A. Subjectless binding domains. *Natural Language & Linguistic Theory*, Springer, v. 9, n. 3, p. 455–496, 1991.

HESTVIK, A. Reflexives and ellipsis. *Natural Language Semantics*, v. 3, n. 2, p. 211–237, 1995.

HICKS, G. *The Derivation of Anaphoric Relations*. Philadelphia, PA: John Benjamins, 2009.

HIGGINBOTHAM, J. Pronouns and bound variables. *Linguistic Inquiry*, v. 11, p. 679–708, 1980.

HIGGINBOTHAM, J. Logical form, binding, and nominals. *Linguistic Inquiry*, v. 14, n. 3, p. 395–420, 1983.

HIGGINBOTHAM, J. On semantics. *Linguistic inquiry*, v. 16, n. 4, p. 547–593, 1985.

HINTERWIMMER, S. A unified account of the properties of German demonstrative pronouns. In: GROSZ, P.; PATEL-GROSZ, P.; YANOVICH, I. (Ed.). *he proceedings of the workshop on pronominal semantics at NELS 40*. Amherst, MA: GLSA Publications, 2015.

HINTERWIMMER, S. Prominent protagonists. *Journal of Pragmatics*, v. 154, p. 79–91, 2019.

HINTERWIMMER, S.; BROCHER, A. An experimental investigation of the binding options of demonstrative pronouns in German. *Glossa: a journal of general linguistics*, Ubiquity Press, v. 3, n. 1, 2018.

HOEKSTRA, J. Pronouns and case: On the distribution of frisian 'harren' and 'se' "them". *Leuvense Bijdragen*, v. 83, p. 47–65, 1994.

HOFMEISTER, P.; CASASANTO, L. S.; SAG, I. A. Islands in the grammar? Standards of evidence. In: SPROUSE, J.; HORNSTEIN, N. (Ed.). *Experimental syntax and the islands debate*. Cambridge: Cambridge University Press, 2013. p. 42–63.

HOFMEISTER, P. et al. Locality and accessibility in *wh*-questions. In: FEATHERSTON, S.; STERNEFELD, W. (Ed.). *Roots: Linguistics in search of its evidential base*. Berlin: Mouton de Gruyter, 2007. p. 185–206.

HOFMEISTER, P.; SAG, I. Cognitive constraints and island effects. *Language*, v. 86, p. 366–415, 2010.

HORN, L. Current issues in neo-Gricean pragmatics. *Intercultural Pragmatics*, Walter de Gruyter, v. 2, n. 2, 2005.

HORN, L. R. *On the Semantic Properties of Logical Operators in English*. Phd Thesis (PhD Thesis) — University of California at Los Angeles, Los Angeles, CA, 1972.

HORN, L. R. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicatures. In: SCHIFFRIN, D. (Ed.). *Meaning, Form and Use in Context*. Washington, DC: Georgetown University Press, 1984. p. 11–42.

HORNSTEIN, N. *Move! A minimalist theory of construal*. Malden, MA: Blackwell, 2001.

HORNSTEIN, N. *A theory of syntax: Minimal operations and universal grammar*. Cambridge: Cambridge University Press, 2009.

HORNSTEIN, N. The minimalist program after 25 years. *Annual Review of Linguistics*, Annual Reviews, v. 4, 2018.

HORNSTEIN, N.; WEINBERG, A. Case theory and preposition stranding. *Linguistic Inquiry*, v. 12, n. 1, p. 55–91, 1981.

HUANG, C.-T. J. A note on the binding theory. *Linguistic Inquiry*, v. 14, n. 3, p. 554–561, 1983.

HUANG, Y. *Anaphora: A cross-linguistic approach*. New York, NY: Oxford University Press, 2000.

HUDDLESTON, R.; PULLUM, G. K. *The Cambridge Grammar of the English language*. Cambridge: Cambridge University Press, 2002.

HUKARI, T. E.; LEVINE, R. D. On SLASH termination and valence-based binding theory. Paper presented at the LSA Meeting, January 1995. 1995.

HYAMS, N.; SIGURJÓNSDÓTTIR, S. The development of "long-distance anaphora": A cross-linguistic comparison with special reference to Icelandic. *Language Acquisition*, v. 1, n. 1, p. 57–94, 1990.

JACKENDOFF, R. Quantifiers in English. *Foundations of Language*, v. 4, p. 422–442, 1968.

JACKENDOFF, R. *Some rules of semantic interpretation for English*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, Cambridge, MA, June 1969.

JACKENDOFF, R. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press, 1972.

JACKENDOFF, R. Morphological and semantic regularities in the lexicon. *Language*, v. 51, p. 639—671, 1975.

JACKENDOFF, R. *X' Syntax*. Cambridge, MA: MIT Press, 1977.

JACKENDOFF, R. On beyond zebra: The relation of linguistic and visual information. *Cognition*, v. 26, n. 2, p. 89–114, 1987.

JACKENDOFF, R. On Larson's treatment of the double object construction. *Linguistic Inquiry*, v. 21, n. 3, p. 427–455, 1990.

JACKENDOFF, R. *Semantic Structures*. Cambridge, MA: MIT Press, 1990.

JACKENDOFF, R. Mme. Tussaud meets the Binding Theory. *Natural Language and Linguistic Theory*, v. 10, n. 1, p. 1–32, 1992.

JACKENDOFF, R. *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press, 1997.

JACKENDOFF, R. Possible stages in the evolution of the language capacity. *Trends in cognitive sciences*, Elsevier, v. 3, n. 7, p. 272–279, 1999.

JACKENDOFF, R. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press, 2002.

JACKENDOFF, R. *Language, consciousness and culture*. Cambridge: The MIT Press, 2007.

JACKENDOFF, R. A parallel architecture perspective on language processing. *Brain research*, v. 1146, p. 2–22, 2007.

JACKENDOFF, R. Construction after construction and its theoretical challenges. *Language*, v. 84, n. 1, p. 8–28, 2008.

JACKENDOFF, R. Your theory of language evolution depends on your theory of language. In: LARSON, R. K.; DÉPREZ, V.; YAMAKIDO, H. (Ed.). *The evolution of human language: Biolinguistic perspectives*. Cambridge: Cambridge Univeristy Press, 2010. p. 63–72.

JACKENDOFF, R. Alternative minimalist visions of language. In: BORSLEY, R. D.; BÖRJARS, K. (Ed.). *Nontransformational Syntax*. Oxford: Wiley-Blackwell, 2011. p. 268–296.

JACKENDOFF, R. What is the human language faculty?: Two views. *Language*, Linguistic Society of America, v. 87, n. 3, p. 586–624, 2011.

JACKENDOFF, R.; AUDRING, J. *The Texture of the Lexicon: Relational Morphology and the Parallel Architecture*. Oxford: Oxford University Press, 2020.

JACKENDOFF, R.; CULICOVER, P. A reconsideration of dative movements. *Foundations of Language*, v. 7, n. 3, p. 397–412, 1972.

JACKENDOFF, R.; PINKER, S. The nature of the language faculty and its implications for evolution of language (reply to fitch, hauser, and chomsky). *Cognition*, v. 97, n. 2, p. 211–225, 2005.

JACOBSON, P. Binding connectivity in copular sentences. In: HARVEY, M.; SANTELMANN, L. (Ed.). *SALT IV*. Ithaca, NY: Cornell University, 1994. v. 4, n. 1, p. 161–178.

JACOBSON, P. Paychecks, stress, and variable free semantics. In: MATTHEWS, T.; JACKSON, B. (Ed.). *Proceedings of SALT X*. Cornell University, Ithaca, NY: CLC Publications, 2000. p. 65–82.

JESPERSEN, O. *Essentials of English Grammar*. London: Edward Arnold, 1933.

KALLESTINOVA, E. D. *Aspects of word order in Russian*. 165 p. Phd Thesis (PhD Thesis) — University of Iowa, 2007.

KAPLAN, R. M. Three seductions of computational psycholinguistics. In: WHITELOCK, P. (Ed.). *Linguistic theory and computer applications*. London: Academic Press, 1987. p. 149–181.

KAPLAN, R. M. The formal architecture of lexical-functional grammar. In: *Formal Issues in Lexical-Functional Grammar*. Stanford: CSLI Publications, 1995.

KAPLAN, R. M.; ZAENEN, A. Long-distance dependencies, constituent structure, and functional uncertainty. In: DALRYMPLE, M. et al. (Ed.). *Formal Issues in Lexical-Functional Grammar*. [S.l.]: CSLI Lecture Notes Series, 1995. p. 137–165.

KATZIR, R. Structurally-defined alternatives. *Linguistics and Philosophy*, Springer, v. 30, n. 6, p. 669–690, 2007.

KAYNE, R. Pronouns and their antecedents. In: KAYNE, R. (Ed.). *Movement and Silence*. Oxford: Oxford University Press, 2005. p. 105–135.

KAYNE, R. S. *French Syntax: the Transformational Cycle*. Cambridge, MA: MIT Press, 1975.

KAYNE, R. S. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press, 1994.

KEENAN, E. Names, quantifiers, and the sloppy identity problem. *Papers in Linguistics*, v. 4, p. 211–232, 1971.

KEENAN, E. L. Complex anaphors and bind alpha. In: *Papers from the 24th Anual Regional Meeting of the Chicago Linguistic Society*. Chicago, IL: [s.n.], 1988. p. 216–232.

KEENAN, E. L. Explaining the creation of reflexive pronouns in English. In: TRAUGOTT, E.; KORTMANN, B. (Ed.). *Studies in the History of the English Language*. Berlin, New York: Mouton De Gruyter, 2002. p. 325–354.

KEENAN, E. L.; COMRIE, B. Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, v. 8, n. 1, p. 63–99, 1977.

KELLER, F.; ASUDEH, A. Constraints on linguistic coreference: Structural vs. pragmatic factors. In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. [S.l.: s.n.], 2001. v. 23, n. 23.

KIM, J.-B. Three types of english pseudo-passives: A lexicalist perspective. *Korean Journal of Linguistics*, v. 34, n. 1, p. 1–24, 2009.

KISHIDA, M. *Reflexives in Japanaese*. Phd Thesis (PhD Thesis) — University of Maryland, College Park, MD, 2011.

KISS, T. Reflexivity and dependency. In: ALEXIADOU, A.; KISS, T.; MÜLLER, G. (Ed.). *Local Modelling of Non-Local Dependencies in Syntax*. Berlin: De Gruyter, 2009. p. 155–185.

KLEIN, E.; SAG, I. A. Type-driven translation. *Linguistics and Philosophy*, v. 8, n. 2, p. 163–201, 1985.

KLIMA, E. S. Negation in English. In: FODOR, J. A.; KATZ, J. J. (Ed.). *The Structure of Language: Readings in the Philosophy of Language*. Englewood Cliffs, NJ: Prentice Hall, 1964. p. 246–323.

KLUENDER, R. *Cognitive Constraints on Variables in Syntax*. Phd Thesis (PhD Thesis) — University of California, San Diego, La Jolla, CA, 1991.

KLUENDER, R. Deriving islands constraints from principles of predication. In: GOODLUCK, H.; ROCHEMONT, M. (Ed.). *Island constraints: Theory, acquisition and processing*. Dordrecht: Kluwer, 1992. p. 223–258.

KLUENDER, R. Are subject islands subject to a processing account. In: CASCADILLA PRESS SOMERVILLE, MA. *Proceedings of WCCFL*. [S.l.], 2004. v. 23, p. 475–499.

KLUENDER, R.; KUTAS, M. Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, v. 5, n. 2, p. 196–214, 1993.

KÖNIG, E.; GAST, V. Reflexive pronouns and other uses of self-forms in english. *Zeitschrift für Anglistik und Amerikanistik*, v. 50, n. 3, p. 1–14, 2002.

KÖNIG, E.; SIEMUND, P. The development of complex reflexives and intensifiers in english. *Diachronica*, John Benjamins, v. 17, n. 1, p. 39–84, 2000.

KÖNIG, E.; SIEMUND, P. Intensifiers and reflexives: A typological perspective. In: FRAJZYNGIER, Z.; CURL, T. (Ed.). *Reflexives: Forms and Functions*. Amsterdam: Benjamins, 2000. p. 41–74.

KOOPMAN, H.; SPORTICHE, D. Variables and the bijection principle. *The Linguistic Review*, v. 2, p. 139–160, 1983.

KOOPMAN, H.; SPORTICHE, D. The position of subjects. *Lingua*, v. 85, p. 211–258, 1991.

KOORNNEEF, A.; REULAND, E. On the shallow processing (dis) advantage: Grammar and economy. *Frontiers in Psychology*, v. 7, p. 82, 2016.

KORNAI, A.; PULLUM, G. K. The X-bar theory of phrase structure. *Language*, v. 66, n. 1, p. 24–50, 1990.

KOSTER, J. *Locality Principles in Syntax*. Dordrecht, The Netherlands: Foris Publications, 1978.

KOSTER, J. On binding and control. *Linguistic Inquiry*, v. 15, n. 3, p. 417–459, 1984.

KOSTER, J. *Domains and Dynasties*. Dordrecht: Foris Publications, 1987.

KOSTER, J.; REULAND, E. *Long-distance anaphora*. Cambridge: Cambridge University Press, 1991.

KRATZER, A. Severing the external argument from its verb. In: ROORYCK, J.; ZARING, L. (Ed.). *Phrase Structure and the Lexicon*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1996. p. 109–137.

KRATZER, A. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry*, v. 40, n. 2, p. 187–237, 2009.

KRIFKA, M. A direct compositionality approach to Condition C effects under reconstruction and their exceptions. In: KRIFKA, M.; SCHENNER, M. (Ed.). *Reconstruction Effects in Relative Clauses*. Berlin: De Gruyter, 2018.

KROCH, A. Amount quantification, referentiality, and long *wh*-movement. In: DIMITRIADIS, A. et al. (Ed.). *Penn Working Papers in Linguistics, Vol. 5.2*. Philadelphia: University of Pennsylvania, 1998.

KUBOTA, Y.; LEVINE, R. D. *Type-Logical Syntax*. Cambride, MA: MIT Press, 2020.

KUNO, S. *Functional syntax: Anaphora, discourse and empathy*. Chicago, IL: University of Chicago Press, 1987.

KUNO, S.; KABURAKI, E. Empathy and syntax. *Linguistic Inquiry*, v. 8, n. 4, p. 627–672, 1977.

KURODA, S.-Y. Where epistemology, grammar and style meet: A case study. In: ANDERSON, S. R.; KIPARSKY, P. (Ed.). *A Festschrift for Morris Halle*. [S.l.]: Holt, Rinehart and Wilson, 1973. p. 377–391.

KUSH, D.; LIDZ, J.; PHILLIPS, C. Relation-sensitive retrieval: evidence from bound variable pronouns. *Journal of memory and language*, v. 82, p. 18–40, 2015.

LABELLE, M. The french reflexive and reciprocal se. *Natural Language & Linguistic Theory*, Springer, v. 26, n. 4, p. 833, 2008.

LACERDA, M. C. M.; OLIVEIRA, R. C. de; LEITÃO, M. M. O processamento da anáfora "se" em português brasileiro e a influência da variação dialetal. *Fórum Linguístico*, v. 11, n. 3, p. 243–260, 2014.

LAKOFF, G. Pronouns and reference. Unpublished manuscript. 1968.

LAKOFF, R. Politeness, pragmatics and performatives. In: *Proceedings of the Texas Conference on Performatives, Presuppositions, and Implicatures*. Austin, Texas: [s.n.], 1977. p. 79–106.

LANDAU, B.; JACKENDOFF, R. "What" and "where" in spatial language and spatial cognition. *Behavioral and brain sciences*, v. 16, n. 2, p. 217–238, 1993.

LANGACKER, R. W. On pronominalization and the chain of command. In: REIBEL, D. A.; SCHANE, S. A. (Ed.). *Modern Studies in English*. Englewood Cliffs, NJ: Prentice Hall, 1969. p. 160–186.

LANGENDOEN, T. D.; POSTAL, P. M. *The vastness of natural languages*. Oxford: Blackwell, 1984.

LARSON, R. On the double object construction. *Linguistic Inquiry*, v. 19, n. 3, p. 335–392, 1988.

LARSON, R. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry*, v. 21, n. 4, p. 589–632, 1990.

LASCARIDES, A. et al. Order independent and persistent typed default unification. *Linguistics and Philosophy*, v. 19, n. 1, p. 1–89, 1996.

LASNIK, H. Remarks on coreference. *Linguistic Analysis*, v. 2, p. 1–22, 1976.

LASNIK, H. On two recent treatments of disjoint reference. *Journal of Linguistic Research*, v. 1, n. 4, p. 48–58, 1981.

LASNIK, H. On the necessity of binding conditions. In: *Essays on anaphora*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1989. p. 149–167.

LASNIK, H.; SAITO, M. On the subject of infinitives. In: DOBRIN, L. M.; NICHOLS, L.; RODRIGUEZ, R. M. (Ed.). *Chicago Linguistics Society 27*. Chicago University: [s.n.], 1991. v. 27, p. 324–343.

LEBEAUX, D. A distributional difference between reciprocals and reflexives. *Linguistic Inquiry*, v. 14, n. 4, p. 723–730, 1983.

LEBEAUX, D. Locality and anaphoric binding. *The Linguistic Review*, v. 4, n. 4, p. 343–363, 1985.

LEES, R. B.; KLIMA, E. S. Rules for English pronominalization. *Language*, JSTOR, v. 39, n. 1, p. 17–28, 1963.

LEIVADA, E.; WESTERGAARD, M. Acceptable ungrammatical sentences, unacceptable grammatical sentences, and the role of the cognitive parser. *Frontiers in Psychology*, v. 11, p. 364, 2020.

LEMLE, M. Pronomes, anáforas, zero: observações sobre uma mudança lingüística. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada*, v. 1, n. 1/2, 1985.

LEVIN, B.; HOVAV, M. R. *Unaccusativity*. Cambridge, MA: MIT Press, 1995. 336 p.

LEVINE, R. D. The ass camouflage construction: Masks as parasitic heads. *Language*, v. 86, n. 2, p. 265–301, 2010.

LEVINSON, S. *Pragmatics*. Cambridge: Cambridge University Press, 1983.

LEVINSON, S. C. Pragmatics and the grammar of anaphora: A partial pragmatic reduction of binding and control phenomena. *Journal of Linguistics*, v. 23, n. 2, p. 379–434, 1987.

LEVINSON, S. C. Pragmatic reduction of the binding conditions revisited. *Journal of Linguistics*, v. 27, n. 1, p. 107–161, 1991.

LEVINSON, S. C. *Presumptive meanings*. Cambridge, MA: MIT Press, 2000.

LI, C.; THOMPSON, S. Subject and topic: A new typology of language. In: LI, C. (Ed.). *Subject and Topic*. New York, NY: Academic Press, 1976.

LIDZ, J. Condition R. *Linguistic Inquiry*, v. 32, n. 1, p. 123–140, 2001.

LIDZ, J.; IDSARDI, W. J. Chains and phono-logical form. *University of Pennsylvania Working Papers in Linguistics*, ERIC, v. 5, n. 1, p. 109–125, 1998.

MAJOR, T.; OZKAN, S. Anaphora in two Turkic languages: Condition A is not enough. *IULC Working Papers*, v. 18, n. 2, 2018.

MALING, J. Non-clause-bounded reflexives in Modern Icelandic. *Linguistics and Philosophy*, Springer, v. 7, n. 3, p. 211–241, 1984.

MANNING, C.; SAG, I. Dissociations between Argument Structure and Grammatical Relations. In: WEBELHUTH, G.; KOENIG, J.-P.; KATHOL, A. (Ed.). *Lexical and Constructional Aspects of Linguistic Explanation*. Stanford: CSLI, 1999. p. 63–78.

MANNING, C. D.; SAG, I. A. Argument structure, valence, and binding. *Nordic Journal of Linguistics*, v. 21, p. 107–144, 1998.

MANZINI, M. R. *Locality*. Cambridge, MA: MIT Press, 1992. 171 p.

MANZINI, M. R.; WEXLER, K. Parameters, binding theory, and learnability. *Linguistic Inquiry*, v. 18, n. 3, p. 413–444, 1987.

MARANTZ, A. *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press, 1984.

MARANTZ, A. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics*, v. 4, n. 2, p. 201–225, 1997.

MARR, D. *Vision*. San Francisco: W.H. Freeman and Company, 1982.

MATSUMOTO, Y. The conversational condition on Horn scales. *Linguistics and philosophy*, Springer, v. 18, n. 1, p. 21–60, 1995.

MATTAUSCH, J. Optimality theoretic pragmatics and binding phenomena. In: BLUTNER, R.; ZEEVAT, H. (Ed.). *Optimality theory and pragmatics*. New York, NY: Palgrave Macmillan, 2004. p. 63–90.

MATTAUSCH, J. Optimality, bidirectionality & the evolution of binding phenomena. *Research on language and computation*, Springer, v. 5, n. 1, p. 103–131, 2007.

MATTAUSCH, J.; GÜLZOW, I. A note on acquisition in frequency-based accounts of binding phenomena. In: GÜLZOW, I.; GAGARINA, N. (Ed.). *Frequency Effects in Language Acquisition*. Berlin: De Gruyter Mouton, 2011. p. 331–358.

MAY, R. *The Grammar of Quantification*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, Cambride, MA, 1977.

MAY, R. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press, 1985.

MCCAWLEY, J. D. Concerning the base component of a transformational grammar. *Foundations of Language*, v. 4, n. 3, p. 243–269, 1968.

MCCAWLEY, J. D. Conversational implicature and the lexicon. In: COLE, P. (Ed.). *Pragmatics*. San Diego, CA: Academic Press, 1978, (Syntax and Semantics). p. 245–259.

MCCAWLEY, J. D. *Everything that Linguists have Always Wanted to Know about Logic*...*but were ashamed to ask*. Chicago, IL: The University of Chicago Press, 1981.

MCKILLEN, A. *On the interpretation of reflexive pronouns*. Phd Thesis (PhD Thesis) — McGill University, Montreal, 2016.

MELO, N. da S. S. *O clítico "se" com valor reflexo ou recíproco: uma abordagem sociolinguística*. Master's Thesis (Master's Thesis) — Universidade Federal de Uberlândia, Uberlândia, 2005.

MENUZZI, S. *Binding theory and pronominal anaphora in Brazilian Portuguese*. Utrecht: Holland Academic Graphics, 1999.

MENUZZI, S. Sobre as opções anafóricas para antecedentes genéricos e para variáveis ligadas: comentários a ana müller. *Letras de Hoje*, v. 38, n. 1, 2003.

MENUZZI, S. Non-conflicting violations of grammatical constraints?: Logophoric reflexives, peculiar passives, and Gricean implicatures. *Ilha do Desterro*, v. 1, n. 47, p. 129–168, 2004.

MENUZZI, S.; LOBO, M. Binding and pronominal forms in Portuguese. In: WETZELS, W. L.; MENUZZI, S.; COSTA, J. (Ed.). *The Handbook of Portuguese Linguistics*. Malden, MA: Wiley-Blackwell, 2016. p. 338–355.

MIHALIČEK, V.; POLLARD, C. Distinguishing phenogrammar from tectogrammar simplifies the analysis of interrogatives. In: GROOTE, P. D.; NEDERHOF, M.-J. (Ed.). *Formal Grammar*. Berlin / Heidelberg, 2012. p. 130–145.

MILIORINI, R. O que cortar com a navalha: economia na Gramática Gerativa e na Sintaxe mais Simples. *Revista do GELNE*, v. 20, n. 1, p. 57–73, 2018.

MILLER, G.; CHOMSKY, N. Finitary models of language users. In: LUCE, R. D.; BUSH, R. R.; GALANTER, E. (Ed.). *Handbook of mathematical psychology*. New York, NY: Wiley, 1963. v. 2, p. 419–492.

MILLER, P. Usage preferences: The case of the English verbal anaphor do so. In: MÜLLER, S. (Ed.). *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI, 2013. p. 121–139.

MINKOFF, S. A. *How some so-called "Thematic Roles" that selact animate arguments are generated and how these roles inform control*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, Cambride, MA, 1994.

MINKOFF, S. A. Consciousness, backward coreference, and logophoricity. *Linguistic Inquiry*, MIT Press, v. 35, n. 3, p. 485–494, 2004.

MIYAGAWA, S. et al. The integration hypothesis of human language evolution and the nature of contemporary languages. *Frontiers in psychology*, Frontiers, v. 5, p. 564, 2014.

MONTAGUE, R. *Formal Philosophy*. New Haven, CT: Yale University Press, 1974.

MONTALBETTI, M. *After Binding*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, Cambridge, MA, 1984.

MORGAN, J. L. Two types of convention in indirect speech acts. In: COLE, P. (Ed.). *Pragmatics*. New York, NY: Academic Press, 1978, (Syntax and Semantics, v. 9).

MORRILL, G. *Categorial Grammar: Logical Syntax, Semantics, and Processin*. New York, NY: Oxford University Press, 2011.

MOULTON, K.; HAN, C.-H. C-command vs. scope: An experimental assessment of bound-variable pronouns. *Language*, v. 94, n. 1, p. 191–219, 2018.

MÜLLER, A. Anaphora and genericity in Brazilian Portuguese. *Letras de Hoje*, v. 38, n. 1, 2003.

MÜLLER, S. *Grammatical theory: From transformational grammar to constraint-based approaches*. Berlin: Language Science Press, 2020.

MUSKENS, R. Language, lambdas, and logic. In: KRUIJFF, G.-J.; OEHRLE, R. (Ed.). *Resource-Sensitivity, Binding and Anaphora*. Dordrecht: Kluwer Academic Publishers, 2003. p. 23–54.

NAPOLI, D. J. *Predication theory: a case study in indexing theory*. Cambridge: Cambridge University Press, 1989.

NAPPA, R.; ARNOLD, J. E. The road to understanding is paved with the speaker's intentions: Cues to the speaker's attention and intentions affect pronoun comprehension. *Cognitive Psychology*, v. 70, p. 58–81, 2014.

NEGRÃO, E. V.; MULLER, A. L. As mudanças na sistema pronominal do portugues brasileiro: substituição ou especialização de formas? *DELTA*, v. 12, n. 1, 1996.

NESPOR, M.; VOGEL, I. *Prosodic phonology*. Dordrecht: Foris, 1986. xiv, 327 p.

NEVES, M. H. d. M. Os pronomes. In: ILARI, R.; NEVES, M. H. d. M. (Ed.). *Gramática do Português Culto Falado no Brasil*. Campinas: Editora da Unicamp, 2008. v. 3, p. 507–622.

NEWMEYER, F. J. Nonsyntactic explanations of island constraints. *Annual Review of Linguistics*, v. 2, p. 187–210, 2016.

NICOL, J.; SWINNEY, D. The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, v. 18, n. 1, p. 5–19, 1989.

NISHIGAUCHI, T.; KISHIDA, M. Reflexive binding and attitudes de se. *Theoretical and applied linguistics at Kobe Shoin*, n. 11, p. 67–89, 2008.

NORDLINGER, R. *Constructive case: evidence from Australian languages*. Stanford, CA: CSLI, 1998.

NUNES, J. Ainda o famigerado SE. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada*, v. 11, n. 2, 1995.

NUNES, J. Especificação morfológica de pronomes nominativos, concordância verbal e sujeitos nulos em português brasileiro. *Fórum Linguístico*, v. 17, n. Esp., p. 4658–4672, 2020.

OEHRLE, R. T. Term-labeled categorial type systems. *Linguistics and Philosophy*, v. 17, n. 6, p. 633–678, 1994.

OKADA, S. Reflexive pronouns with split antecedents. *Journal of Pragmatics*, v. 30, n. 1, p. 59–84, 1998.

OSHIMA, D. On empathic and logophoric binding. *Research on Language and Computation*, v. 5, n. 1, p. 19–35, 2007.

OSHIMA, D. Y. Zibun revisited: empathy, logophoricity, and binding. *Proceedings of the 20th NWLC*, v. 23, p. 175–190, 2004.

OTERO, C. Acceptable ungrammatical sentences in Spanish. *Linguistic Inquiry*, v. 3, n. 2, p. 233–242, 1972.

PARSONS, T. *Events in the Semantics of English: A Study in Subatomic Semantics*. Cambridge, MA: MIT Press, 1990.

PARTEE, B. H. *Subject and Object in Modern English*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, Cambride, MA, 1965.

PARTEE, B. H. Noun phrase interpretation and type-shifting principles. In: GROENENDIJK, J.; JONGH, D. de; STOKHOF, M. (Ed.). *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. Dordecht: Foris, 1986. p. 115–143.

PARTEE, B. H. The Garden of Eden period for Deep Structure and semantics. In: GALLEGO, Á.; OTT, D. (Ed.). *50 Years Later: Reflections on Chomsky's Aspects*. Cambride, MA: MIT Press, 2015. p. 187–198.

PARTEE, B. H.; BORSCHEV, V. Integrating lexical and formal sematics: Genitives, relational nouns, and type-shifting. In: *Proceedings of the Second Tbilisi Symposium on Language, Logic, and Computation*. [S.l.: s.n.], 1998.

PARTEE, B. H.; BORSCHEV, V. Genitives, relational nouns, and argument-modifier ambiguity. In: LANG, E.; MAIENBORN, C.; FABRICIUS-HANSEN, C. (Ed.). *Modifying Adjuncts*. [S.l.]: Mouton de Gruyter, 2003. p. 67–112.

PARTEE, B. H.; MEULEN, A. G. B. t.; WALL, R. E. *Mathematical methods in linguistics*. Dordrecht ; Boston: Kluwer Academic, 1990. xx, 663 p.

PARTEE, B. H.; ROOTH, M. Generalized conjunction and type ambiguity. In: BÄUERLE, R.; SCHWARZE, C.; STECHOW, A. von (Ed.). *Meaning, Use and Interpretation of Language*. [S.l.]: de Gruyter, 1983. p. 362–383.

PATEJUK, A.; PRZEPIÓRKOWSKI, A. Reducing grammatical functions in LFG. In: *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*. Stanford, CA: CSLI, 2016. p. 541–559.

PATEL-GROSZ, P. The Principle A problem. *Journal of South Asian Linguistics*, v. 6, 2013.

PATEL-GROSZ, P.; GROSZ, P. On the typology of donkeys: Two types of anaphora resolution. In: *Proceedings of Sinn und Bedeutung*. [S.l.: s.n.], 2010. v. 14, p. 339–355.

PEITSARA, K. The development of reflexive strategies in English. In: RISSANNE, M. K. M.; HEIKKONEN, K. (Ed.). *Grammaticalization at Work: Studies of Long-term Developments in English*. Berlin / New York: Walter de Gruyter, 1997. v. 24, p. 277–370.

PEREIRA, D. C. *Variação e mudança no uso dos pronomes reflexivos no português popular da capital paulista*. Phd Thesis (PhD Thesis) — Universidade de Sao Paulo, Sao Paulo, Brazil, 2007.

PERLMUTTER, D. M.; POSTAL, P. M. Toward a universal characterization of passivization. In: *Annual Meeting of the Berkeley Linguistics Society*. Berkeley: [s.n.], 1977. v. 3, p. 394–417.

PERLMUTTER, D. M.; POSTAL, P. M. Some proposed laws of basic clause structure. In: *Studies in Relational Grammar*. Chicago, IL: The University of Chicago Press, 1983. v. 1, p. 81–128.

PESETSKY, D. Binding problems with experiencer verbs. *Linguistic Inquiry*, v. 18, n. 1, p. 126–140, 1987.

PESETSKY, D. *Zero syntax: Experiencers and Cascades*. Cambridge, MA: MIT Press, 1995.

PESETSKY, D.; TORREGO, E. The syntax of valuation and the interpretability of features. In: KARIMI, S.; SAMIIAN, V.; WILKINS, W. K. (Ed.). *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*. Amsterdam: John Benjamins Publishing Company, 2007. p. 262–294.

PETERS, S. P.; RITCHIE, R. W. On the generative power of transformational grammars. *Information Sciences*, v. 6, p. 49–83, 1973.

PHILLIPS, C.; WAGERS, M. W.; LAU, E. F. Grammatical illusions and selective fallibility in real-time language comprehension. In: RUNNER, J. (Ed.). *Syntax and Semantics: Experiments at the Interfaces*. [S.l.]: Brill, 2011. v. 37, p. 147–180.

PICA, P. Liage et contiguité. In: *Recherches sur l'anaphore*. [S.l.]: Paris IV, 1984.

PICA, P. On the nature of the reflexivization cycle. In: *Proceedings of the North East Linguistic Society*. [S.l.: s.n.], 1987. v. 17, n. 2, p. 485–500.

PINKER, S.; BLOOM, P. Natural language and natural selection. *Behavioral and brain sciences*, Cambridge University Press, v. 13, n. 4, p. 707–727, 1990.

POLLARD, C. The nature of constraint-based grammar. In: *PACLIC conference*. [S.l.: s.n.], 1996. v. 98.

POLLARD, C. Lectures on the foundations of HPSG. Unpublished Manuscript. 1997.

POLLARD, C. Remarks on binding theory. In: Müller, S. (Ed.). *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar*. Stanford, CA: [s.n.], 2005. v. 12, p. 561–577.

POLLARD, C.; SAG, I. A. *Information-based syntax and semantics: volume 1*. Chicago, IL: Chicago University Press, 1987.

POLLARD, C.; SAG, I. A. Anaphors in English and the scope of binding theory. *Linguistic Inquiry*, v. 23, n. 2, p. 261–304, 1992.

POLLARD, C.; SAG, I. A. *Head-Driven Phrase Structure Grammar*. Chicago, IL: University of Chicago Press and CSLI Publications, 1994.

POLLARD, C.; XUE, P. Chinese reflexive *ziji*: Syntactic reflexives vs. nonsyntactic reflexives. *Journal of East Asian Linguistics*, v. 7, n. 4, p. 287–318, 1998.

PONTES, E. *O tópico no português do Brasil*. Campinas: Pontes, 1987.

POST, E. L. Formal reductions of the general combinatorial decision problem. *American Journal of Mathematics*, v. 65, n. 2, p. 197–215, 1943.

POSTAL, P. *On raising*. Cambridge, MA: MIT Press, 1974.

POSTAL, P.; PULLUM, G. Expletive noun phrases in subcategorized positions. *Linguistic Inquiry*, v. 19, p. 635–679, 1988.

POSTAL, P. M. A note on 'understood transitively'. *International Journal of American Linguistics*, v. 32, n. 1, p. 90–93, 1966.

POSTAL, P. M. On so-called 'pronouns' in English. In: *Monograph Series on Languages and Linguistics: 17th Annual Round Table*. Washington, D. C.: Georgetown University Press, 1966. v. 1, n. 19, p. 177–206.

POSTAL, P. M. *Crossover phenomena*. Yorktown Heights, NY: IBM Corporation, 1968.

POSTAL, P. M. On Coreferential Complement Subject Deletion. *Linguistic Inquiry*, v. 1, n. 4, p. 439–500, 1970.

POSTAL, P. M. *Cross-Over Phenomena*. New York, NY: Holt, Rinehart and Winston, 1971.

POSTAL, P. M. *Skeptical Linguistic Essays*. New York, NY: Oxford University Press, 2004.

POSTAL, P. M. Remarks on English long-distance anaphora. *Style*, v. 40, n. 1-2, p. 7–18, 2006.

PREMINGER, O. The Anaphor Agreement Effect: further evidence against binding-as-agreement. Unpublished ms., University of Maryland at College Park. 2019.

PROGOVAC, L. *Evolutionary Syntax*. Oxford: Oxford University Press, 2015.

PROGOVAC, L. A gradualist scenario for language evolution: precise linguistic reconstruction of early human (and neandertal) grammars. *Frontiers in Psychology*, v. 7, p. 1714, 2016.

PROGOVAC, L. Sudden (saltationist) approaches to language evolution. In: PROGOVAC, L. (Ed.). *A Critical Introduction to Language Evolution*. Berlin: Springer, 2019. p. 15–29.

PULLUM, G. K. *Such that* clauses and the context-freeness of English. *Linguistic Inquiry*, v. 16, n. 2, p. 291–298, 1985.

PULLUM, G. K. The central question in comparative syntactic metatheory. *Mind & Language*, v. 28, n. 4, p. 492–521, 2013.

PULLUM, G. K. Theorizing about the syntax of human language: a radical alternative to generative formalisms. *Cadernos de Linguística*, v. 1, n. 1, p. 1–33, 2020.

PULLUM, G. K.; SCHOLZ, B. C. On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In: *International Conference on Logical Aspects of Computational Linguistics*. Berlin: [s.n.], 2001. p. 17–43.

PUSTEJOVSKY, J. *The generative lexicon*. Cambridge, MA: MIT Press, 1995.

PUSTEJOVSKY, J. Type theory and lexical decomposition. In: PUSTEJOVSKY, J. et al. (Ed.). *Advances in generative lexicon theory*. New York, NY: Springer, 2012. p. 9–38.

QUINE, W. V. O. Two dogmas of empiricism. *Philosophical Review*, Harper Torchbooks, v. 60, n. 1, p. 20–43, 1951.

RECANATI, F. *Literal Meaning*. New York, NY: Cambridge Univeristy Press, 2004.

RECANATI, F. *Truth-conditional pragmatics*. Oxford: Oxford University Press, 2010.

REINHART, T. *The syntactic domain of anaphora*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, Cambridge, MA, 1976.

REINHART, T. *Anaphora and Semantic Interpretation*. Chicago, IL: University of Chicago Press, 1983.

REINHART, T. *Interface Strategies*. Cambridge, MA: MIT Press, 2006.

REINHART, T.; REULAND, E. Anaphors and logophors: An argument structure perspective. In: KOSTER, J.; REULAND, E. (Ed.). *Long-Distance Anaphora*. Cambridge: Cambridge University Press, 1991. p. 283–322.

REINHART, T.; REULAND, E. Anaphoric territories. In: *Groninger Arbeiten zur germanistischen Linguistik*. Groningen: Groningen University, 1992. v. 34.

REINHART, T.; REULAND, E. Reflexivity. *Linguistic Inquiry*, v. 24, n. 4, p. 657–720, 1993.

REINHART, T.; SILONI, T. The lexicon-syntax parameter: Reflexivization and other arity operations. *Linguistic Inquiry*, v. 36, n. 3, p. 389–436, 2005.

RETT, J. *Semantics of Evaluativity*. Oxford: Oxford University Press, 2015.

RETT, J. Manner implicatures and how to spot them. *International Review of Pragmatics*, v. 12, n. 1, p. 44–79, 2020.

REULAND, E. Primitives of binding. *Linguistic Inquiry*, v. 32, n. 3, p. 439–492, 2001.

REULAND, E. *Anaphora and Language Design*. Cambridge, MA: MIT Press, 2011.

REULAND, E. Reflexives. In: THOMASON, S. (Ed.). *Oxford Handbooks online*. [S.l.]: Oxford University Press, 2014.

REULAND, E. Grammar of binding in the languages of the world: Unity versus diversity. *Cognition*, v. 168, p. 370–379, 2016.

REULAND, E. Why is reflexivity so special? understanding the world of reflexives. *Studia linguistica*, v. 71, n. 1-2, p. 12–59, 2017.

REULAND, E.; REINHART, T. Pronouns, anaphors and case. In: HAIDER, H.; OLSEN, S.; VIKNER, S. (Ed.). *Studies in Comparative Germanic Syntax*. Dordrecht: Kluwer Academic Publishers, 1995. p. 241–268.

REULAND, E.; SIGURJÓNSDÓTTIR, S. Long distance 'binding' in Icelandic: Syntax or discourse? In: PICA, P.; ROORYCK, J. (Ed.). *Atomism and Binding*. Dordrecht: Foris, 1997. p. 323–340.

REULAND, E.; WONG, S. C. H.; EVERAERT, M. How the complexity of mandarin zi-ji simplifies the grammar. *Linguistic Inquiry*, v. 51, n. 4, p. 799–814, 2020.

RICHTER, S.; CHAVES, R. P. Investigating the role of verb frequency in factive and manner-of-speaking islands. In: *Proceedings of the 42nd Annual Virtual Meeting of the Cognitive Science Society*. Toronto: [s.n.], 2020. p. 1772–1777.

RIDDLE, E.; SHEINTUCH, G. A functional analysis of pseudo-passives. *Linguistics and Philosophy*, JSTOR, p. 527–563, 1983.

RIZZI, L. On the anaphor agreement effect. *Rivista di Grammatica Generativa*, v. 2, p. 27–42, 1990.

ROCHA, A. de F. *Clíticos Reflexivos: Uma Variante Sociolingüística na Cidade de Ouro Preto*. Master's Thesis (Master's Thesis) — Universidade Federal de Minas Gerais, Belo Horizonte, 1999.

ROCHEMONT, M. Givenness. In: FÉRY, C.; ISHIHARA, S. (Ed.). *The Oxford Handbook of Information Structure*. Oxford: Oxford University Press, 2016. p. 41–62.

RONAT, M. Une solution pour un apparent contre-exemple à la théorie du liage. *Lingvisticae Investigationes*, John Benjamins, v. 6, n. 1, p. 189–196, 1982.

ROORYCK, J. Prepositions and minimalist case marking. In: THRAÍNSSON, H.; EPSTEIN, S. D.; PETER, S. (Ed.). *Studies in comparative Germanic syntax*. Dordecht: Kluwer Academic Publishers, 1996. v. 2, p. 226–256.

ROORYCK, J.; Vanden Wyngaerd, G. *Dissolving Binding Theory*. Oxford: Oxford University Press, 2011.

ROORYCK, J.; Vanden Wyngaerd, G. Morphological transparency and the Delay of Principle B Effect. *Lingua*, Elsevier, v. 155, p. 121–139, 2015.

ROSENBAUM, P. S. *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press, 1967. 128 p.

ROSS, J. R. *Constraints on variables in syntax*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, 1967.

ROSS, J. R. On the cyclic nature of English pronominalization. In: REIBEL DAVID A. SCHANE, S. A. (Ed.). *Modern Studies in English*. [S.l.]: Prentice Hall, 1969. p. 187–200.

ROSS, J. R. On declarative sentences. In: JACOBS, R. A.; ROSENBAUM, P. S. (Ed.). *Readings in English transformational grammar*. Waltham, MA: Ginn-Blaisdell, 1970. p. 222–272.

RUNNER, J. T.; KAISER, E. Binding in picture noun phrases: Implications for Binding Theory. In: Müller, S. (Ed.). *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI Publications, 2005. p. 594–613.

SADOCK, J. M. *Toward a Linguistic Theory of Speech Acts*. New York, NY: Academic Press, 1974.

SAFIR, K. Implied non-coreference and the pattern of anaphora. *Linguistics and Philosophy*, Springer, v. 15, n. 1, p. 1–52, 1992.

SAFIR, K. Semantic atoms of anaphora. *Natural Language and Linguistic Theory*, v. 14, n. 3, p. 545–589, 1996.

SAFIR, K. Symmetry and unity in the theory of anaphora. In: BENNIS, H.; PICA, P.; ROORYCK, J. (Ed.). *Atomism and Binding*. Dordecht: Foris, 1997. p. 341–379.

SAFIR, K. Anaphors, movement and coconstrual. In: DELSING, L.-O.; et al. (Ed.). *Grammatik i fokus. Grammar in focus: A festschrift for Christer Platzack*. Lund: Lund University, Department of Scandinavian Languages, 2003.

SAFIR, K. *The syntax of anaphora*. New York, NY: Oxford University Press, 2004.

SAG, I. *Deletion and logical form*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, Cambridge, 1976.

SAG, I. A. English relative clause constructions. *Journal of Linguistics*, v. 33, p. 431–484, 1997.

SAG, I. A. Sign-based construction grammar - a synopsis. In: BOAS, H. C.; SAG, I. A. (Ed.). *Sign-Based Construction Grammar*. Stanford, CA: CSLI, 2012. p. 61–197.

SAG, I. A.; HOFMEISTER, P.; SNIDER, N. Processing complexity in Subjacency violations: the complex noun phrase constraint. In: *Proceedings of the 43 Annual Meeting of the Chicago Linguistic Society*. Chicago: University of Chicago, 2007.

SAG, I. A.; WASOW, T. Performance-compatible competence grammar. In: BORSLEY, R. D.; BÖRJARS, K. (Ed.). *Non-transformational syntax: Formal and explicit models of grammar*. [S.l.]: Wiley Online Library, 2011. p. 359–377.

SAUERLAND, U. Scalar implicatures in complex sentences. *Linguistics and Philosophy*, v. 27, n. 3, p. 367–391, 2004.

SAUERLAND, U. The silent content of bound variable pronouns. In: JOHNSON, K. (Ed.). *Topics in Ellipsis*. Cambridge, United Kingdom: Cambridge University Press, 2008. p. 183–209.

SAUERLAND, U. Presuppositions and the alternative tier. In: SNIDER, T. (Ed.). *Proceedings of SALT*. Santa Cruz, CA: [s.n.], 2013. v. 23, p. 156–173.

SAXON, L. On one's own: the semantics and pragmatics of reflexives. In: GEORGOPOULOS, C.; ISHIHARA, R. (Ed.). *Interdisciplinary Approaches to Language*. Dordrecht: Kluwer, 1991. p. 501–517.

SCHLENKER, P. Clausal equations (a note on the connectivity problem). *Natural Language and Linguistic Theory*, v. 21, n. 1, p. 157–214, 2003.

SCHLENKER, P. A plea for monsters. *Linguistics and Philosophy*, Springer, v. 26, n. 1, p. 29–120, 2003.

SCHLENKER, P. Minimize restrictors! (notes on definite descriptions, condition C and epithets). In: MAIER, E.; BARY, C.; HUITINK, J. (Ed.). *Proceedings of Sinn und Bedeutung*. [S.l.: s.n.], 2005. v. 9, p. 385–416.

SCHLENKER, P. Non-redundancy: Towards a semantic reinterpretation of Binding Theory. *Natural Language Semantics*, v. 13, n. 1, p. 1–92, 2005.

SCHÜTZE, C. *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Chicago, IL: University of Chicago Press, 1996.

SEARLE, J. R. Indirect speech acts. In: COLE, P.; MORGAN, J. L. (Ed.). *Speech Acts*. [S.l.]: Academic Press, 1975, (Syntax and Semantics, v. 3). p. 59–82.

SELLS, P. Aspects of logophoricity. *Linguistic Inquiry*, v. 18, n. 3, p. 445–480, 1987.

SHIEBER, S. M. *An introduction to Unification-based approaches to Grammar*. Stanford, CA: Center for the Study of Language and Information, 1986. CLSI Lecture Notes, No. 4.

SIEMUND, P. Reflexive and intensive self-forms across varieties of english. *Zeitschrift für Anglistik und Amerikanistik*, v. 50, n. 3, p. 250–268, 2002.

SIGURJÓNSDÓTTIR, S.; HYAMS, N. Reflexivization and logophoricity: Evidence from the acquisition of icelandic. *Language Acquisition*, Taylor & Francis, v. 2, n. 4, p. 359–413, 1992.

SILVA, S. Moreira da. *Études sur la symétrie et l'asumétrie sujet/objet dans le Portuguais du Brésil*. Phd Thesis (PhD Thesis) — Paris 7, Paris, 1983.

SIMONS, M. A gricean view on intrusive implicatures. In: PETRUS, K. (Ed.). *Meaning and Analysis: new essays on Grice*. London: Palgrave Macmillan, 2010. p. 138–169.

SMIRNOVA, A.; JACKENDOFF, R. Case assignment and argument realization in nominals. *Language*, v. 93, n. 4, p. 877–911, 2017.

SMIRNOVA, E. Constructionalization and constructional change. In: BARDHDAL, J. et al. (Ed.). *Diachronic construction grammar*. Philadelphia, PA: John Benjamins, 2015. p. 81–106.

SMITS, E.-J.; HENDRIKS, P.; SPENADER, J. Using very large parsed corpora and judgment data to classify verb reflexivity. In: BRANCO, A. (Ed.). *Anaphora: analysis, algorithms and applications*. Berlin: Springer, 2007. p. 77–93.

SOBIN, N. An acceptable ungrammatical construction. In: LIMA, S. D.; CORRIGAN, R.; IVERSON, G. K. (Ed.). *The Reality of Linguistic Rules*. Philadelphia, PA: John Benjamins Publishing, 1994. v. 26.

SPATHAS, G. Focus on reflexive anaphors. In: *Semantics and Linguistic Theory*. [S.l.: s.n.], 2010. v. 20, p. 471–488.

SPEAS, P.; TENNY, C. Configurational properties of point of view roles. In: SCIULLO, A. M. di (Ed.). *Asymmetry in Grammar*. Amsterdam: John Benjamins, 2003. v. 1, p. 315–345.

SPERBER, D.; WILSON, D. Loose talk. In: JSTOR. *Proceedings of the aristotelian society*. [S.l.], 1985. v. 86, p. 153–171.

SPERBER, D.; WILSON, D. *Relevance: communication and cognition*. Cambridge: Blackwell, 1995.

SPORTICHE, D. Binding theory - structure sensitivity of referential dependencies. *Lingua*, Elsevier, v. 130, p. 187–208, 2013.

SPORTICHE, D. Constraints on reflexivization. Unpublished Manuscript. 2020. Available at: <lingbuzz/005488>.

STEEDMAN, M. Combinators and grammars. In: OEHRLE, R.; BACH, E.; WHEELER, D. (Ed.). *Categorial grammars and natural language structures*. Dordrecht: Foris, 1988.

STEEDMAN, M. *Surface structure and interpretation*. Cambridge, MA: MIT Press, 1996.

STERNEFELD, W. *Semantic vs. Syntactic Reconstruction*. Tübingen, 2000.

STIRLING, L. *Switch-Reference and Discourse Representation*. Cambridge: Cambridge University Press, 1993.

STOWELL, T. Tense and article omission in abbreviated English. Unpublished Manuscript. 1996.

SZABOLCSI, A. Bound variables in syntax (are there any?). In: GROENENDIJK, J.; VELTMAN, F.; STOKHOF, M. (Ed.). *Sixth Amsterdam Colloquium Proceedings*. Amsterdam: University of Amsterdam, 1987.

SZABOLCSI, A. Combinatory grammar and projection from the lexicon. In: SAG, I. A.; SZABOLCSI, A. (Ed.). *Lexical Matters*. Stanford, California: Center for the Study of Language and Information, 1992. p. 241–268.

TARSKI, A. The semantic conception of truth: and the foundations of semantics. *Philosophy and phenomenological research*, v. 4, n. 3, p. 341–376, 1944.

TARSKI, A. The concept of truth in formalized languages. *Logic, semantics, metamathematics*, Oxford, v. 2, n. 152-278, p. 7, 1956.

THORNTON, R.; WEXLER, K. *Principle B, VP Ellipsis, and Interpretation in Child Grammar*. Cambridge, MA: The MIT Press, 1999. 241 p.

THRAÍNSSON, H. Long-distance reflexives and the typology ofnps. In: KOSTER, J.; REULAND, E. (Ed.). *Long-Distance Anaphora*. Cambridge: Cambridge Univeristy Press, 1991. p. 49–75.

TRAUGOTT, E. C.; TROUSDALE, G. *Constructionalization and constructional change*. Oxford: Oxford University Press, 2013.

TROTZKE, A.; ZWART, J.-W. The complexity of narrow syntax: Minimalism, representational economy, and simplest merge. In: NEWMEYER, F. J.; PRESTON, L. B. (Ed.). *Measuring Grammatical Complexity*. Oxford: Oxford University Press, 2014. p. 128–147.

UTAKIS, S. *Quantification and definiteness in child grammar*. Phd Thesis (PhD Thesis) — CUNY, New York, 1995.

VALIAN, V. Syntactic subjects in the early speech of American and Italian children. *Cognition*, v. 40, n. 1-2, p. 21–81, 1991.

VARASCHIN, G. Anti-reflexivity and logophoricity: an account of unexpected reflexivization contrasts. *Glossa: a journal of general linguistics*, v. 5, n. 1, 2020.

VARASCHIN, G. LFG and Simpler Syntax. In: DALRYMPLE, M. (Ed.). *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press, to appear.

VARASCHIN, G.; CULICOVER, P. W.; WINKLER, S. In pursuit of Condition C: (non-)coreference in grammar, discourse and processing. In: KONIETZKO, A.; WINKLER, S. (Ed.). *Information Structure and Discourse in Generative Grammar: Mechanisms and Processes*. Berlin: De Gruyter, in press.

VIEIRA, R. C. Brazilian Portuguese 'mesmo' in 'ele mesmo'as a scalar focus particle: evidence from acquisitional data. *ReVEL*, v. 9, p. 139–150, 2015.

VOLKOVA, A.; REULAND, E. Reflexivity without reflexives? *The Linguistic Review*, v. 31, n. 3-4, p. 587–633, 2014.

WALKER, H. Adjuncts and the HPSG binding theory. In: Müller, S. (Ed.). *Proceedings of the 18th International Conference on HPSG*. Stanford, CA: CSLI Publications, 2011. p. 227.

WELLWOOD, A. et al. The anatomy of a comparative illusion. *Journal of Semantics*, Oxford University Press, v. 35, n. 3, p. 543–583, 2018.

WILKINS, W. Thematic structure and reflexivization. In: WILKINS, W. (Ed.). *Syntax and semantics: Thematic relations*. San Diego, CA: Academic Press, 1988. p. 191–213.

WILLIAMS, E. Discourse and logical form. *Linguistic Inquiry*, v. 8, p. 101–139, 1977.

WILLIAMS, E. PRO and subject of NP. *Natural Language and Linguistic Theory*, v. 3, n. 3, p. 297–316, 1985.

WILTSCHKO, M. On the syntax and semantics of (relative) pronouns and determiners. *The Journal of Comparative Germanic Linguistics*, v. 2, n. 2, p. 143–181, 1998.

WOOLFORD, E. More on the anaphor agreement effect. *Linguistic Inquiry*, v. 30, n. 2, p. 257–287, 1999.

WURZEL, W. U. On markedness. *Theoretical linguistics*, v. 24, n. 1, p. 53–71, 1998.

YANTI; COLE, P.; HERMON, G. The grammar of binding in the languages of the world: A response to Reuland. *Cognition*, v. 168, p. 380–384, 2017.

YASHIMA, J. *Antilogophoricity: In Conspiracy with the Binding Theory*. Phd Thesis (PhD Thesis) — University of California, Los Angeles, CA, 2015.

ZIPF, G. K. *Human behavior and the principle of least effort: An introduction to human ecology*. Cambridge: Addison Wesley, 1949.

ZRIBI-HERTZ, A. Anaphor binding and narrative point of view: English reflexive pronouns in sentence and discourse. *Language*, v. 65, n. 4, p. 695–727, 1989.

ZRIBI-HERTZ, A. Emphatic or reflexive? on the endophoric character of French lui-même and similar complex pronouns. *Journal of Linguistics*, v. 31, n. 2, p. 333–374, 1995.

ZWART, J.-W. Issues relating to a derivational theory of binding. In: EPSTEIN, S. D.; SEELY, T. D. (Ed.). *Derivation and Explanation in the Minimalist Program*. Malden, Massachusetts: Blackwell Publishing, 2002. p. 269–304.