REVISITING VARIATION BETWEEN NULL SUBJECT LANGUAGES: THE VIEW FROM OVERT SUBJECT PRONOUNS

By

Daniel Charles Greeson

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Linguistics – Master of Arts

2021

ABSTRACT

REVISITING VARIATION BETWEEN NULL SUBJECT LANGUAGES: THE VIEW FROM OVERT SUBJECT PRONOUNS

By

Daniel Charles Greeson

Taking a micro-comparative approach within Romance null subject languages, I argue that cross-linguistic variation in the distributions of null subjects can be attributed to differences in the featural composition of overt subject pronouns. I use data from Brazilian Portuguese and Spanish to show that more semantically and syntactically complex overt pronouns are restricted in their distributions when compared to less complex overt pronouns; as a result, *pro*-drop languages with highly restricted overt subject pronouns will omit these forms in favor of *pro* more often. I will explain a number of interpretative and frequency-based contrasts associated with pronouns in both languages in terms of Schlenker (2005)'s pragmatic principle *Minimize Restrictors!*. I will show that this approach has an empirical advantage over the standard analyses of variation between null subject languages, which posit parameters in the inflectional domain that over-predict categorical syntactic differences in the availability of *pro* in different languages.

Copyright by
DANIEL CHARLES GREESON
2021

ACKNOWLEDGMENTS

To the extent that this thesis is at all interesting, thought-provoking, or readable, I owe it to a very large number of people who have gotten me here.

To Cristina, thank you for drawing me into MSU Linguistics in your Intro to Syntax class, and for welcoming me into the lab. You have always believed in me more than I felt made sense and for that I'm extremely grateful. You have invested a truly astronominal amount of time and enthusiasm into my education. You are also the funniest/wittiest person on the entire campus, so thanks for making me laugh. You have managed to be the most involved mentor I know of while also supporting my total intellectual independence. Somewhere in an alternate universe I'm an angry actuary because I didn't have a Cristina in my life. Alan, thank you for treating me like a grown-up syntactician even when I didn't (don't) know what I was doing. Thanks for bringing the sharpest questions around so that I never feel unprepared in the least. Thank you for all the intellectual generosity and the conversations spanning far more than linguistics, and for appreciating and generating good puns. To both of you, thanks for being by my side through some of the strangest moments of my life and for helping me to ambulate through them.

Adolfo, thank you for being the expert linguist I could go to in the Spanish department and the Spanish expert I could go to in the Linguistics program, and for helping me to keep one foot in both myself. I remember the time you very generously offered to be my reading partner, and in my discussions with you the idea for this thesis was born. Thank you for keeping me around even as I must have broken some record for number of courses taken with the same professor, and thanks for being the fun, creative, and knowledgeable instructor that I kept coming back to. I owe a lot of this work and my development as a linguist to you.

Brian, thank you for your attentive feedback as an instructor, semantics group leader, and committee member. Thank you for carefully pointing out details that no one else notices and for helping me to start to regain some confidence in semantics/pragmatics after too much time away. I'm sad that our time at MSU didn't overlap more.

Marcin, thank you for introducing me to formal semantics in your fun and unique way and for having me at Awkward Time even when I was an awkward undergrad who didn't know anything. Thanks for making me feel like semantics was something I could do.

Karthik, thank you for making data analysis fun and for the advice about life beyond linguistics. Suzanne, thank you for the help and support as DGS and for teaching me to care about the sociolinguistic aspects of theoretical problems in linguistics.

Hannah, thank you for mentorship, advice (about writing, life, academia, and everything else). I enjoyed being your travel buddy and I'm incredibly grateful that you trusted me to code, write, and research with you. I developed a lot as a researcher thanks to your willingness to work with me and help me along the way.

I also couldn't have gotten through all the late night writing sessions, term papers, and the rest of my time at MSU without the other Linguistics students, and I certainly wouldn't have had as much fun either.

Rachel, thanks for always being a Zoom call away and for being my first friend in the Linguistics department. I cherish our friendship deeply and hope we stay friends for a long time to come!

Natasha, thank you for being my syntax buddy, summer school travel companion, and my fellow mathematician-turned-linguist (although you were always more of a mathematician than I was). I'm glad we're now on the same coast!

Lucia, thanks for the laughs we shared watching SpongeBob on the gigantic screen in Wells Hall, for the genius wordplay that I still use to remember Binding Theory, the crosswords, and for many other wonderful memories we shared in math and linguistics. You are extremely missed.

Several linguists outside of MSU have also been part of my development so far. I'm grateful to Ivona Kučerová for her extremely generous and constructive feedback after CamCoS 9 and for having me virtually at her Syntax Lab. I learned a lot and am extremely grateful for your generosity and the opportunity to exchange ideas with someone whose work has had a large impact on my own. Naomi Shin, thank you for your extremely kind and thoughtful response to my email several years ago asking about your null subjects coding scheme. You could have simply forwarded me to

the source you used, but you were kind enough to invite Hannah and I to meet with you at HLS and the conversations we had were informative and enjoyable. Heather Taylor, although we also overlapped at MSU for a bit, your linguistics course at UM Flint was the first time I ever learned what linguistics was, and I have been hooked ever since. Thank you for that. To Ivo, Zeineb, Ian, Angelica, Jonathan, Bryn, Hannah, Jeanette, Giovanni, and other summer school friends, thanks for the shared mischief, linguistics, and laughs. I hope we see each other again soon.

Finally, there are many, many non-linguists who I couldn't be here without either.

Thank you to my family for material, emotional, and mental support you've provided for 23 years. Thank you Mom for always listening to me explain what I'm writing about and asking great questions, even as I drone on. Thank you Megan and Emily for being born before me so that I'm lucky enough to absorb some of your experience and wisdom.

Thanks to my non-linguistics friends Allison, Anna, Jon, Kate, Laura, Megan (also a linguistics friend!), Sarah, and Wendy for keeping me sane and for making sure I still have hobbies outside of school.

Thank you Gustavo for always being there for me and for being my biggest supporter even as I launch a battery of grammaticality judgment requests at you in your free time.

Thank you also to Dr. Testa and many nurses for getting me to this point largely in one piece.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Originally laid out in Perlmutter (1971), the so-called *pro*-drop parameter was initially formulated to reflect the fact that some languages, like English (1) or French, require subjects of sentences to be expressed overtly, while others, like Spanish (2) and Italian (the *pro*-drop, or null subject languages (NSLs)), allow these subjects to be covert (glossed henceforth as *pro*).

(1) *Am happy.

(2) *pro* estoy feliz. (Spanish)
pro be-1S-PRES happy
"I am happy."

The syntactic variation captured by the null subject parameter has since established itself as one of the key empirical domains in parametric syntactic theory. In the Principles & Parameters era (Chomsky 1981), the *pro*-drop parameter was further formalized to be related to properties of "rich" inflection in languages like Spanish & Italian that enabled Infl to formally identify the empty category in subject position (Jaeggli 1986, Rizzi 1986). Under contemporary parametric analyses, the role of inflection in licensing null subjects persists in the form of a D(efiniteness)-feature on T. This D-feature is hypothesized to license definite, referential null subjects in the canonical or 'consistent' NSLs, like Italian, Spanish, Greek, Turkish, & many others (Holmberg 2005).

Holmberg's (2005) system for classifying null subject languages also crucially extends beyond an explanation for languages like Italian & Spanish (consistent NSLs): it also accounts for the distributions of subject pronouns in so-called *partial* NSLs (e.g. Finnish, Hebrew, & Brazilian Portuguese) which allow for null subjects, but only in a significantly reduced set of contexts relative to the consistent NSLs. Holmberg attributes this reduced distribution of null subjects in partial NSLs to a lack of the D-feature on T that is present and licenses null subjects in consistent NSLs. Thus, while languages like Spanish & Italian license definite null subjects in all tensed clauses via this D-feature, partial NSLs may only license definite null subjects when some other mechanism is

available, such as a control relation between an embedded null subject and a c-commanding matrix antecedent.

Holmberg's distinction between consistent & partial NSLs has a distinct empirical advantage over previous iterations of the null subject parameter in accounting for more than one type of distribution of null subjects. Nonetheless, it too runs into a significant empirical issue: it overestimates the magnitude of syntactic difference between null subject languages of different types. Recent research into Romance NSLs of both the partial and consistent varieties (see Fernandes et al (2018) for Portuguese, and Frascarelli & Jiménez (2019) for Spanish) has painted a much more probabilistic picture of how null subjects are distributed, with partial NSLs on one end of a gradient spectrum and consistent NSLs on another. In this thesis I will account for this probabilistic distribution of null subjects by suggesting a different type of syntactic variation than is traditionally assumed to exist between NSLs of different types. More specifically, I will show that that what differs between languages like Brazilian Portuguese and Spanish is not how/whether null subjects are licensed, but rather what the structural and semantic properties of *overt* subject pronouns are, & how they compare to the alternative null variant.

This thesis is organized as follows: chapter 1 gives an overview of standard accounts of the difference between partial and consistent null subject languages and highlights some of the empirical problems of this analysis before proposing an overt pronoun-centered alternative. Chapter 2 establishes the empirical basis of this thesis by demonstrating that overt pronouns in Brazilian Portuguese and Spanish differ from each other in their featural compositions; a syntactic analysis for how these differences arise is also provided. Chapter 3 accounts for differences with respect to Montalbetti (1984)'s Overt Pronoun Constraint by making use of the different properties of overt pronouns in Spanish and Brazilian Portuguese as discussed in chapter 2. Chapter 4 accounts for the different frequencies and interpretations associated with *pro* and overt pronouns in Spanish and Brazilian Portuguese in terms of Schlenker (2005)'s *Minimize Restrictors!*; in this chapter I also revisit my semantic analysis of Spanish overt pronouns and adopt Esipova (2019)'s account of biosocial gender. Chapter 5 extends the empirical scope to several additional Romance languages

that pattern with Brazilian Portuguese and Spanish in ways the strengthen the core arguments of the thesis. Chapter 6 considers the implications of my analysis for how the respective distributions of null and overt subjects are acquired in different types of null subject languages. Finally, chapter 7 addresses remaining issues and concludes.

1.1 Standard accounts of partiality

An emerging view in the null subjects literature is that partial NSLs like Brazilian Portuguese differ from consistent NSLs like Spanish in terms of a D(efiniteness)-feature on T (Holmberg 2005, Holmberg, Nayudu & Sheehan 2009, Barbosa 2019, and many others), as mentioned above. In consistent NSLs such as Spanish, Italian, and Greek there is hypothesized to be a D-feature on T; this D-feature licenses definite null subjects in essentially all contexts. In Brazilian Portuguese, a partial NSL, there is, by hypothesis, no D-feature on T; definite null subjects are extremely limited in their distribution and must be licensed by some other mechanism. From now on, I will refer to this view as the *D-feature hypothesis*. An important prediction made under the D-feature hypothesis is that definite third-person[1] null subjects in Brazilian Portuguese are only possible in a very limited set of contexts, namely when the null subject is (i) in an embedded clause & (ii) has a c-commanding controller in the next highest clause (Rodrigues 2002, 2004, Sheehan 2006, Holmberg, Nayudu & Sheehan 2009, Barbosa 2019, Nunes 2008, 2019, Roberts 2019). This c-command requirement explains the im/possible coindexations in (3) & (4) (judgments from Rodrigues 2004):

(3) O João$_i$ disse que *pro*$_{i/*j}$ embarcou no trem. (Brazilian Portuguese (BrPt))
The John said that *pro* embarked on-the train
"John$_i$ said that he$_i$ boarded the train."

(4) [O pai do João$_i$]$_j$ disse que *pro*$_{*i/j}$ é inteligente. (BrPt)
the dad of-the John said that *pro* is intelligent
"[John$_i$s dad]$_j$ said that he$_j$s intelligent."

[1]The acceptability of null first- and second-person pronouns in Brazilian Portuguese is the subject of considerable debate and appears to be subject to a great deal of interspeaker variation. As such, I will be setting aside the issue of second and first & second person and focusing exclusively on third-person pronouns for this thesis, which are empirically more than sufficient to support the general premise.

This contrast seems to support the D-feature hypothesis (i.e. that the hypothesized lack of a D-feature on T in Brazilian Portuguese means definite null subjects can only be licensed by a c-commanding matrix controller). However, the c-command requirement seems to be too strong. A search for similar examples on social media (slightly modified for anonymity) reveals a number of examples like the following:

(5) [O pai d-ela$_i$]$_j$ disse que *pro* só vai namorar depois dos trinta. (BrPt)
the father of-her says that *pro* only goes-to date-INF after of-the thirty
"[Her$_i$ father]$_j$ says that she$_i$ can only date after she's thirty."

Here, a non-c-commanding antecedent is fine; the intended coindexation is clear from context.

Additional evidence against the c-command requirement comes from experimental results from Fernandes et al. (2018). Their team presented Brazilian Portuguese-speaking participants with sentences such as the following:

(6) O atleta consultou o médico depois que *pro* regressou da viagem à Itália.
the athlete consulted the doctor after that *pro* returned from-the trip to-the Italy
(Brazilian Portuguese)

"The athlete consulted the doctor after he returned from the trip to Italy."

When asked to disambiguate the reference of the embedded null subject, Brazilian Portuguese speakers chose the non-c-commanding antecedent, *o médico* ('the doctor'), approximately 26% of the time.

Finally, one other piece of evidence against a strict c-command requirement comes from the following split-antecedent example from Holmberg, Nayudu & Sheehan (2009):

(7) O Zé$_i$ convenceu os meninos$_j$ que *pro*$_{i+j}$ tinham que ir embora. (BrPt)
the Zé convinced the kids that *pro* had that go-INF away
'Zé$_i$ convinced the kids$_j$ that they$_{i+j}$ had to go away".

In this example, a perfectly accessible interpretation of this string is one in which the embedded null subject has a split antecedent in the form of *os meninos* & *o Zé*, even though *os meninos* does not c-command the embedded subject.

1.2 Overt pronouns as the source of variation

Based on the evidence against the c-command requirement, I want to explore the idea that null subjects in Brazilian Portuguese *are* in fact licensed in the same syntactic contexts as in Spanish. What differs may be merely the strength of the bias towards a subject antecedent reading in sentences like (6), which is relatively stronger in Brazilian Portuguese than in Spanish. This leaves open the possibility that there is no *syntactic* difference in the licensing conditions for null subjects in Brazilian Portuguese & Spanish. The varying strength of the subject antecedent preference will be further explored in chapter 4. For now I will merely note that the difference between these languages does not amount to a syntactic requirement for a c-commanding antecedent, and I will proceed to explore alternative dimensions according to which Brazilian Portuguese and Spanish may differ in their properties as null subject languages.

Even if the syntactic licensing of null subjects does not differ in categorical terms from Brazilian Portuguese to Spanish, there are clearly differences between the distributions of null and overt subjects in each language. These include sensitivity to the Overt Pronoun Constraint (section 3), in addition to relative rates of subject omission and the anaphoric biases of null and overt subjects (section 4). The task now is to account for these differences without positing a categorical syntactic difference in how null subjects are licensed by inflection in each language, given the undesirable empirical consequences discussed earlier in this chapter. One analytical option is to appeal to a neo-Gricean account of the distributions of different nominal arguments, such as *Minimize Restrictors!* (Schlenker 2005). Nonetheless, even under such an account it's not immediately clear where the variation between interpretations of subjects in partial and consistent NSLs could arise *outside* of the syntax, assuming speakers of both language types obey general Gricean pragmatic constraints of the same type. A potential solution to this issue is that there is, in fact, a parametric syntactic difference between consistent and partial null-subject languages, but in a domain other than verbal inflection. In this thesis I will suggest a version of this. Namely, I will argue that the syntactic variation between Spanish and Brazilian Portuguese lies in the pronominal domain, specifically within each language's set of overt pronouns; these differences in turn lead to differences in the

relative distributions of null and overt subjects.

CHAPTER 2

ESTABLISHING DIFFERENT FEATURAL COMPOSITIONS

2.1 Overview of Chapter 2

In this chapter I will show that the overt subject pronouns in Brazilian Portuguese and Spanish have different featural compositions. More specifically, I will show that Spanish overt subject pronouns are more semantically restricted than their Brazilian Portuguese counterparts.

2.2 Core data

First, Spanish overt subject pronouns (8) are categorically restricted to animate referents only (Enríquez 1984), while those in Brazilian Portuguese (9) may refer to either inanimates or animates:

(8) Ella es bonita. (Spanish)
3S-FEM is beautiful
"She is beautiful". [Cannot mean "It is beautiful"].

(9) Ela é bonita. (Brazilian Portuguese)
3S-FEM is beautiful
"She/it is beautiful".

The second way in which Spanish overt subject pronouns carry more semantic restrictions relative to Brazilian Portuguese concerns gender. When there is a mismatch between the biosocial gender of the real-world individual being discussed (e.g. a man named John) and the arbitrary syntactic gender of a noun phrase antecedent (e.g. *persona*/*pessoa* 'person', invariably syntactically feminine), Spanish overt subject pronouns must realize the gender that corresponds to the biosocial gender of the real-world individual (12). On the other hand, Brazilian Portuguese pronouns may agree in gender with the noun *pessoa*, in this case licensing a feminine pronoun even when talking about a male referent (13).

(10) Una persona que realmente cree que {$^{??}$ella misma$_i$ / ^{OK}él mismo} va
a.FEM person.FEM that really thinks that 3S-FEM same / 3S-MASC same goes
a marcar una diferencia en la sociedad, es Juan$_i$. (Spanish)
to mark-INF a difference in the society, is John.
Intended: "A person x who really thinks that x is going to make a difference in society, it's John."

(11) Uma pessoa que realmente acha que {^{OK}ela mesma$_i$ / $^{?}$*ele mesmo}
a.FEM person.FEM that really thinks that 3S-FEM same / 3S-MASC same
vai fazer a diferença na sociedade, é o João$_i$. (BrPt)
goes make-INF the difference in-the society, is the John.
"A person x really thinks that he x is going to make a difference in society, it's John."

The same pattern holds when we consider an invariably syntactically masculine antecedent such as *indivíduo* ('individual') while the person being discussed is a female individual who, in general, uses feminine pronouns. Brazilian Portuguese allows for a mismatch between real-life biosocial gender and pronominal gender while Spanish does not[1]:

(12) Un individuo que realmente cree que {$^{??}$él mismo$_i$ / ^{OK}ella
a.MASC individual.MASC that really thinks that 3S-MASC same / 3S-FEM
misma} va a marcar una diferencia en la sociedad, es María$_i$. (Spanish)
same goes to mark-INF a difference in the society, is Maria.
Intended: "A person x really thinks that x is going to make a difference in society, it's María."

(13) Um indivíduo que realmente acha que {^{OK}ele mesmo$_i$ / $^{?}$*ela
a.MASC individual.MASC that really thinks that 3S-MASC same / 3S-FEM
mesma} vai fazer a diferença na sociedade, é a Maria$_i$. (BrPt)
same goes make-INF the difference in-the society, is the Maria.
"A person x who really thinks that x is going to make a difference in society, it's Maria."

Similarly, while pronouns in Brazilian Portuguese can always reflect the arbitrary syntactic gender of their linguistic antecedent (14a), Spanish pronouns may actually take a conflicting gender

[1]Note: I report here the judgments that were attested by three of the four Spanish speakers I consulted. However, Adolfo Ausín reports that it may be possible for him to obtain a felicitous interpretation of (12) with a feminine pronoun. It is highly possible that there is significant inter-speaker variation with examples like these. In any case though, the important thing is that there is a contrast with Spanish speakers (even if some accept the feminine pronoun in (12) and Brazilian Portuguese speakers, who uniformly reported to me a preference for the feminine pronoun and a dispreference for the masculine pronoun in equivalent sentences.

value that does not match the syntactic gender of their antecedent (15a). In the Spanish sentence in (15a), the plural pronoun *ellos* (which is both the masculine and neuter plural pronoun in Spanish) is used over a feminine pronoun, reflecting the fact that the feminine DP *las personas* ('the people') refers to people of all genders and not just women.

(14) *Brazilian Portuguese* (Davies & Ferreira 2016):

a. [As pessoas]$_i$ não percebem que [elas mesmas]$_i$ acabam
the-PL person.FEM-PL NEG perceive that 3PL-FEM same-FEM-PL end-up
"contaminando" o ambiente com seu mau humor. (Brazilian Portuguese)
contaminating the environment with their bad mood
"People don't perceive that they themselves end up "contaminating" the atmosphere with their bad mood."

(15) *Spanish* (Davies 2016):

a. [Cuántas personas]$_i$ piensan que [ellos mismos]$_i$ no son
howmany-PL person.FEM-PL think that 3PL-MASC same-MASC-PL NEG are
responsables! (Spanish)
responsible-PL
'How many people think that they themselves are not responsible!'

Although the corpus example above used a quantificational antecedent, the same facts hold when we use a definite determiner like in the Portuguese example:

(16) [Las personas]$_i$ piensan que [ellos mismos]$_i$ no son
the-PL person.FEM-PL think that 3PL-MASC same-MASC-PL NEG are
responsables! (Spanish)
responsible-PL
'The people think that they themselves are not responsible!'

These data suggest that gender is obligatorily semantically interpreted and must correspond to a real-world correlate for Spanish overt subject pronouns, while gender on BP pronouns need not be. Keeping these contrasts in mind, we can consider different proposals for the sets of semantic restrictions applying to Brazilian Portuguese and Spanish overt pronouns. When considering gender, I will follow Ahn (2019)'s assumption that only semantically interpreted -features, and not

narrowly syntactic grammatical features, count as semantic restrictions on pronouns. This assumption is justified by the fact that while the biosocial gender of a referent is relevant for determining whether the predicate associated to a particular pronoun is true or not, purely grammatical gender on pronouns merely arises when the arbitrary, lexically-specified gender of a nominal is copied to the pronoun; it has no truth-conditional content. I also adopt her denotation of *pro*, which is similarly licensed in Spanish and Brazilian Portuguese with referents of any animacy or gender:

(17) ella/él (Spanish):

x: entity(x) & animate(x) & (fe)male(x)

(18) ela/ele (BP), *pro* (BP and Spanish):

x: entity(x)

2.3 Syntactic analysis

On the basis of several minimal pairs, we have motivated different denotations for overt subject pronouns in BP and Spanish. The next step is to determine how this semantic difference relates to the syntactic makeup of these forms. I propose that the animacy and biosocial gender restrictions on Spanish overt subject pronouns comes from the presence of PERSON, while Brazilian Portuguese overt pronouns lack PERSON, or at least do not obligatorily realize it; this mirrors Rooryck (2000)'s PERSON-less approach to analyzing the French pronoun *il* which may refer to inanimates. An argument for a connection between PERSON and biosocially interpreted gender comes from Kučerová (2018). She follows Longobardi (2008), Landau (2010), and Sudo (2012) in positing a connection between PERSON and a semantic referential index. Kučerová then proposes that the difference between contextually-determined 'biosocial' gender and arbitrary grammatical gender is that the former arises due to the association of grammatical gender with a particular referent via PERSON. Although Kučerová does not argue for an additional connection between PERSON and animacy, I believe that positing this would be advantageous for a number of reasons.

First, extending the connection between PERSON and biosocial gender to a connection between PERSON and animacy captures the fact that only animate individuals may be associated with a

biosocial gender. Secondly, first- and second-person forms are already inherently animate-denoting, so positing that only animate 'third-person' referents truly host PERSON unites the three persons as a natural class. Third, in language families where PERSON plays an extensive role for a number of morphosyntactic phenomena, such as Algonquian and Dene, there is overt evidence that inanimate third-person entities pattern against animate ones (Lochbihler, Oxford and Welch 2015). For example, Algonquian transitive verbs realize the morpheme *-*ekw* whenever the subject outranks the object according to the following hierarchy: 1/2 person > 3rd-person animate proximate > 3rd-person animate obviative > 3rd-person inanimate. Lochbihler, Oxford and Welch point out that the inclusion of animacy in the person hierarchy is unexpected if animacy is treated like gender or another unrelated phi-feature, but expected if animacy is inherent to PERSON. Under this view, "3rd-person" in animates rank lowest as they fail to realize the feature PERSON at all. The authors also highlight similar patterns from Algonquian object agreement and Dene agreement in which PERSON agreement targets 1/2 person and 3rd-person animates but fails to target inanimates entirely. Following their analysis, I will assume going forward that PERSON not only gives rise to a biosocial gender restriction on pronouns, but also to an animacy restriction on pronouns.

2.4 Summary of Chapter 2

In this chapter I argued that Brazilian Portuguese overt subject pronouns are specified for less semantically interpreted features than their Spanish counterparts. I then proposed that this semantic difference stems from the fact that Spanish overt subject pronouns host PERSON while Brazilian Portuguese overt subject pronouns do not. I paid special attention to the ties between the feature PERSON and the gender and animacy properties of pronouns.

With these connections to gender and animacy in mind, we are in a position to examine some of the effects the absence or presence of PERSON has on the grammatical systems of different null subject languages. In the next chapter we will explore the differing behavior of Brazilian Portuguese and Spanish with respect to the Overt Pronoun Constraint, and I will show how BP's failure to obey this constraint follows from a lack of PERSON on its overt subject pronouns.

CHAPTER 3

THE OVERT PRONOUN CONSTRAINT

3.1 Overview of Chapter 3

A good test case for re-locating NSL variation to the domain of overt pronouns is the difference between Brazilian Portuguese, a partial NSL, and Spanish, a consistent NSL, with respect to Montalbetti's (1984) Overt Pronoun Constraint (19). This constraint is typically taken to be a direct result of the licensing of null subjects, although the constraint itself is about the interpretative possibilities of overt pronouns:

(19) Overt pronouns cannot occur under the scope of a quantifier iff the alternation overt/empty obtains.

I will show that the Overt Pronoun Constraint (OPC) is *not* strictly a constraint that governs the alternation between null & overt pro-forms as is standardly assumed. Rather, the OPC is a side effect of structural & semantic differences between pronouns of all types within & across languages.

To argue this, I will (i) demonstrate that there are OPC effects even in a language that does not allow null pronouns (English), and (ii) show how the different OPC facts that have been described for two different null subject languages, Brazilian Portuguese & Spanish, are best explained chiefly through properties of *overt* pronouns in each language.

3.2 The OPC in Brazilian Portuguese & Spanish

Data such as (20a) shows that Spanish is sensitive to the OPC, as an overt third-person singular feminine pronoun is unavailable with the bound interpretation (despite its antecedent being feminine and singular). Compare this to the grammatical sentence in (20b) with a referential antecedent binding an overt pronoun.

(20) *Spanish*

a. Toda persona$_i$ cree que *ella$_i$ / *pro$_i$* es inteligente. (Spanish)
every person.FEM thinks that 3S-FEM / *pro* is intelligent
Intended: "For every person x, x thinks that x is intelligent."

b. María$_i$ cree que ella$_i$ / *pro$_i$* es inteligente. (Spanish)
María thinks that 3S-FEM / *pro* is intelligent
"María thinks that she's intelligent"

Sensitivity to the OPC has also been reported in a number of other unrelated languages that allow null subjects, including Japanese (Okuma 2015a) & Turkish (Çnar and Çakr 2019). Due to the OPC's application in a wide variety of unrelated null subject languages, it is standardly assumed to be a universal of languages that robustly allow null subjects. A consequence of this view is that if a language doesn't obey the OPC, it does not truly license null subjects. For this reason, when it was discovered that Brazilian Portuguese deviates from the OPC (see (21b), adapted from Negrão 1999), this was taken as empirical support for the view that Brazilian Portuguese is not a consistent null subject language (Ferreira 2000, 2004, Roberts 2019):

(21) *Brazilian Portuguese*

a. Toda pessoa$_i$ acha que ela$_i$ / *pro$_i$* é inteligente. (Brazilian Portuguese)
Every person.FEM thinks that 3S-FEM / *pro* is intelligent
"For every person x, x thinks that x is intelligent."

b. Maria$_i$ acha que ela$_i$ / *pro$_i$* é inteligente. (Brazilian Portuguese)
Maria thinks that 3S-FEM / *pro* is intelligent
"Maria thinks that she is intelligent."

Recall that contemporary accounts for null subjects in Brazilian Portuguese (e.g. Nunes 2004, Rodrigues 2002, 2004, Sheehan 2006, Holmberg, Nayudu and Sheehan 2009, Barbosa 2019, Nunes 2008, 2019, Roberts 2019) posit that third-person definite null subjects in Brazilian Portuguese are generally illicit. Under these accounts, the only contexts in which third-person definite null subjects do superficially appear licit is in the subject position of an embedded clause, where the antecedent is a c-commanding controller in the next highest clause. Then, what appears to be a null subject pronoun here is actually a trace from control-as-movement. True alternation between null & overt pronominal forms is a precondition for the OPC, hence its failure to apply in Brazilian Portuguese.

So far, then, these analyses of null pronouns in Brazilian Portuguese are consistent with the idea that OPC effects are tied to the licensing of null pronouns.

Nonetheless, if we view the OPC as related to the licensing of null subjects, we miss out on the broader cross-linguistic picture. If the OPC effects observed for null subject languages are described in more general terms, it may be possible to uniformly account for a version of OPC in both NSLs and non-NSLs without any reference to null pronouns. I will show there are indeed OPC effects in English, a non-*pro*-drop language, suggesting the need for a more general approach.

3.3 Bound-variable singular *they* in English

The preference for English singular *they* with a quantified DP antecedent like *every person* is directly parallel to OPC effects in Spanish, despite the fact that English is not an NSL[1]:

(22) Every person$_i$ thinks that {she$_{*i/j}$'s / they$_{i/j}$'re} the most intelligent one.

English speakers' intuition is that *she* is unacceptable because it must refer to a specific woman (that is, it's specified for 'biosocial' gender, to use Ackerman (2019)'s term), while *every person* is gender-neutral (not specified for biosocial gender) & refers to no specific individual whatsoever. Aligning with previous syntactic accounts that place biosocial gender high on the nominal spine (e.g. Steriopolo & Wiltschko 2010, Matushansky 2013, Pesetsky 2013, Landau 2016, Kučerová 2018, Sigurðsson 2019), a number of authors have proposed that the 'gender-neutrality' of singular *they* relates to properties of D. For example, Bjorkman (2017) posits that biosocial gender is either located at D or PhiP and that bound-variable singular *they* realizes a smaller structure (e.g. a D- or Phi-less structure) than referential (i.e. non-bound variable) instances of gendered pronouns such as *he* and *she*. Conrod (2017) takes a slightly different approach, but similarly bases their analysis around biosocial gender on D, assuming that bound-variable singular *they* fails to undergo to N-to-D movement. Konnelly & Cowper (2020), on the other hand, do not assume that singular *they* lacks D, but instead posit that the copying of gender features from *n* to D is optional.

[1] I am not the first to make the observation that the OPC appears operative even in non-NSLs; Cardinaletti & Starke note similar examples in French and St-Galler German.

The analyses reviewed above suggest the possibility of analyzing singular *they* in a similar manner to Brazilian Portuguese overt pronouns, namely by positing that both types of overt pronouns fail to obey the OPC due to a lack of PERSON. Since at least Ritter (1995), the feature PERSON has been associated with the D head, so a D-less bound-variable singular *they* also lacks PERSON. Like Brazilian Portuguese overt pronouns, English bound-variable *they* fails to encode biosocial gender. One notable difference, however, is that BP pronouns are compatible with inanimate referents, which I associated with a lack of a PERSON feature in the previous chapter, while singular *they* is restricted to animate referents. Despite this difference, I believe a PERSON-less analysis of singular *they* is still entirely tenable. The animacy restriction of singular *they* likely follows from the presence of a specialized inanimate singular pronoun *it* in the English pronominal paradigm which blocks *they* from referring to inanimate singular referents. Indeed, when there is no specialized inanimate form, as is the case with *they* used to refer to plural referents, reference to inanimate pluralities is perfectly licit.

Turning to English gendered pronouns *he* and *she*, which are restricted to biosocially gendered animate referents just like Spanish overt subject pronouns, I suggest analyzing these as pro-DP's that realize a PERSON feature, leading to the animacy and gender restrictions they impose.

Now the task that remains is to establish exactly how a lack of PERSON leads to the acceptability of a given pronoun in a bound-variable context (i.e. disobedience of the OPC) and likewise how the presence of PERSON leads to the unacceptability of pronouns in the same contexts. For antecedents such as *every person*, my account provides an obvious explanation for the acceptability of *they* and Brazilian Portuguese *ela*: since PERSON is the feature responsible for creating a biosocial gender interpretation, these forms' lack of this feature allow them to be licit with antecedents that lack biosocial gender, such as *every person*. On the other hand, Spanish *ella* and English *she* and *he* host a PERSON feature, which (following Kučerová 2018) associates the feature GENDER feature with a referential index at the CI-interface, leading to a biosocial gender interpretation which is incompatible with an antecedent like *every person*, which does not specify any biosocial gender.

In our discussion so far we've looked at what happens when a gendered pronoun has a gender-

neutral quantified antecedent (e.g. *every person*); this is also the type of case most thoroughly discussed in the literature about the Overt Pronoun Constraint. However, we must still address the question of quantified DPs that *are* specified for biosocial gender. My Spanish-speaking informants' intuitions regarding *toda persona* is that it makes a very poor antecedent for *ella* (23); interestingly, though, they report that *toda mujer* ('the woman') is considerably less degraded as an antecedent:

(23) Toda persona$_i$ cree que ella$_{*i/j}$ es inteligente. (Spanish)
Every-FEM person.FEM thinks that 3S-FEM is intelligent
"Every person thinks that she is intelligent."

(24) Toda mujer$_i$ cree que ella$_{?i/j}$ es inteligente. (Spanish)
Every-FEM woman.FEM thinks that 3S-FEM is intelligent
"Every woman thinks that she's intelligent".

To the extent that (24) is still degraded with the bound reading, we can attribute to the fact that the semantic contribution of PERSON is association with a referential index (Longobardi 2008, Landau 2010, Sudo 2012, Kučerová 2018)[2] Then, although *toda mujer* does not conflict with the biosocial gender of *ella*, the subject pronoun *ella* is inherently referential and must pick out a specific individual in the world, making it a poor bound variable even with a biosocially gendered antecedent. Thus, even though a quantified DP like *toda mujer* is biosocially/semantically gendered just like *María*, the latter is a better antecedent for *ella* due to its referential nature.[3]

(25) María$_i$ cree que ella$_{i/j}$ / *pro*$_{i/j}$ es inteligente. (Spanish)
María thinks that 3S-FEM / *pro* is intelligent
"María said that she is intelligent".

This contrast aligns nearly exactly with the fact that many English speakers prefer singular *they* even with quantified gender-specific antecedents like *every woman* (Bjorkman 2017, Conrod 2019,

[2]If the Overt Pronoun Constraint is a real effect that has to do with quantificational antecedents, then there should be some sort of contrast between the acceptability of an overt pronoun with gendered quantified DP antecedents like *toda mujer* ('every woman') and referential DPs like proper names. I follow Montalbetti (1984) and the subtle intuitions of my consultants in assuming that there is a slight contrast here. However, as far as I know, the vast majority of the cases discussed in the literature do not carefully distinguish between biosocially gendered DPs like 'every woman' and neutral ones like 'every person' or 'every student', so there is a major gap in empirical coverage.

[3]Since I have not drawn a hard line between Spanish pronouns like *ella* and R-expressions in this thesis, it may be necessary to say that coindexiation between R-expressions like *María* and overt subjects like *ella* arises via accidental coreference rather than binding. In any case, correference of any type is facilitated by the fact that the featural compositions of *María* and *ella* do not clash.

Konnelly & Cowper 2020), although the gendered alternative *she* is less degraded than with *every person* (this indeed corresponds to my judgment):

(26) Every woman$_i$ thinks that {she$_{?i/j}$'s / they$_{i/j}$'re} the most intelligent one. *(Some speakers)*

This suggests that gendered pronouns in English closely parallel their counterparts in Spanish with respect to the OPC, despite the fact that English does not permit null subjects (or null pronouns of any sort): the most strongly disfavored coindexation occurs when a biosocial gender mismatch occurs between an overt gendered pronoun and a gender-neutral quantified antecedent, followed by a weaker dispreference for a gendered overt pronoun even with a gendered quantified antecedent.

In Brazilian Portuguese, unlike in Spanish, we see no strong contrast between these three sentence types:

(27) Toda pessoa$_i$ acha que ela$_i$ / *pro*$_i$ é inteligente. (Brazilian Portuguese)
Every-FEM person.FEM thinks that 3S-FEM / *pro* is intelligent
"Every person$_i$ thinks that she$_i$ is intelligent.'

(28) Toda mulher$_i$ acha que ela$_i$ / *pro*$_i$ é inteligente. (Brazilian Portuguese)
Every-FEM woman.FEM thinks that 3S-FEM / *pro* is intelligent
"Every woman$_i$ thinks that she$_i$ is intelligent."

(29) Maria$_i$ acha que ela$_i$ / *pro*$_i$ é inteligente. (Brazilian Portuguese)
Maria thinks that 3S-FEM / *pro* is intelligent
"Maria$_i$ thinks that she is intelligent."

This is what we expect if Brazilian Portuguese pronouns lack PERSON and are neither restricted to a biosocial gender interpretation nor inherently referential.

3.4 Summary of Chapter 3

To recap so far: if we assume the OPC follows from the presence/absence of PERSON on pronouns rather than from the possibility of null subjects, we can explain a number of otherwise surprising facts. These include (i) exact parallels between OPC effects in English and Spanish, and (ii) an asymmetry between co-reference of an overt pronoun with biosocially gendered quantified DPs like

toda mujer vs. 'gender-neutral' ones like *toda persona*, which is not predicted by Montalbetti's formulation.

CHAPTER 4

ACCOUNTING FOR DIFFERENT DISTRIBUTIONS OF NULL SUBJECTS

4.1 Overview of Chapter 4

In this chapter I will argue that differences between BP and Spanish in rates of subject omission as well as interpretative differences between overt and null subjects in both languages can be attributed to *Minimize Restrictors!* (Schlenker 2005). This constraint is given fully below:

(30) A definite description *the A B* [where the order of *A* vs. *B* is irrelevant] is deviant if *A* is redundant, i.e. if:
(i) *the B* is grammatical and has the same denotation as *the A B* (=Referential Irrelevance), and
(ii) A does not serve another purpose (=Pragmatic Irrelevance).

For example, Schlenker notes that, if John has exactly one father, then uttering *John's father* to refer to him is perfectly felicitous. However, *John's blond father* would be infelicitous, as this expression uses an additional restrictor *blond* which is not required to establish reference to the intended referent or to achieve any special pragmatic purpose.

I propose that *Minimize Restrictors!* also determines the competition between *pro* and overt pronouns in null subject languages, and will proceed to explain differences between the distributions of null subjects in Brazilian Portuguese and Spanish accordingly. Before doing so, I will defend the viability of applying *Minimize Restrictors!* to gendered pronouns.

The first issue is that gender on pronouns is standardly understood to be presuppositional (e.g. Heim & Kratzer 1998), in which case pronouns should be subject to *Maximize Presupposition!* (Heim 1991) rather than *Minimize Restrictors!*. If this were the case, we would expect the competition between *pro* and an overt subject pronoun like Spanish *ella* to favor *ella*, as the former only presupposes for some individual *x* that *x* is an entity, while, under the standard view, the latter

presupposes that x is an entity, x is animate, and x is female. This appears to give us the exact opposite of what is borne out by the data, as it is well-established within the *pro*-drop literature that null subjects are generally favored over overt subjects absent a special pragmatic motivation for an overt form (see e.g. Cardinaletti and Starke 1999). In what follows, I will review and defend Esipova (2019)'s account of pronominal gender and argue for applying to Spanish pronouns, and *against* applying it to Brazilian Portuguese pronouns.

4.2 Spanish pronominal gender as a form indexical

Esipova (2019) argues that gender on human-referring pronouns is a 'form indexical' whose not-at-issueness stems from its status as a non-restrictive pronoun-internal modifier, not from its status as a presupposition. For Esipova, 'projection' of gender is better understood as "the process of selecting a context-appropriate form, affected by a range of utterance-internal and -external factors, including social cost of using the wrong form". Noting that even binary-gendered pronouns do not map one-to-one to the biosocial categories 'male' and 'female,' Esipova gives the following denotation for a gendered pronoun like *she*:

(31) x: form(*she*, x, c)

This denotation can be taken to mean that the speaker c_s believes *she* to be an appropriate way to refer to x in c. Let us then revise our denotations from (17) in Chapter 2 as follows, replacing *male/female*, which do not map perfectly onto pronominal forms, with *form(ella/él)*[1]:

(32) ella/él (Spanish):

x: entity(x) & animate(x) & form(*ella/él*, x, c)

Since *pro* does not impose any restrictions related to animacy or biosocial gender, we may continue with the denotation in (40), repeated from above. Since Brazilian Portuguese pronouns

[1]Note that adopting Esipova's analysis of form indexicals instead of the standard mapping of gender features to the categories *male/female* should not affect the syntactic analysis given in chapter 2.1: just as PERSON has been previously tied to a biosocial gender interpretation corresponding to either a male or female individual (Kučerová 2018) I assume PERSON is responsible for establishing a biosocial gender interpretation corresponding to an individual that uses a given set of pronouns.

do not obligatorily encode any information about animacy or biosocial gender[2], we also maintain the denotation from above in (18), repeated below:

(33) *pro* (BP & Spanish), BP ela/ele (BP):

x: entity(x)

Adopting this denotation for Brazilian Portuguese pronouns under Esipova's account raises some important questions that need to be addressed. First, recall that the motivation for positing this denotation for Brazilian Portuguese pronouns in chapter 2 came from the following data:

(34) Uma pessoa que realmente acha que ela mesma$_i$ pode fazer a
a person.FEM that really believes that 3S-FEM same can make-INF a
diferença na sociedade, é o João$_i$. (Brazilian Portuguese)
difference in-the society, is the João
"A person x who really thinks that x can make a difference in society is John."

In this example, despite the fact that the speaker is discussing a human being of known male gender, John, who generally uses masculine pronouns (e.g. *ele*), the feminine pronoun *ela* is licensed by the syntactically feminine (but biosocially gender neutral) DP *a pessoa*, ('the person'). In the equivalent Spanish example, (12), we saw that the masculine pronoun must be used despite the presence of the feminine DP *la persona*, as the individual, John, being discussed in the context was male and generally used masculine pronouns. It was suggested in chapter 2 that this reflects a difference between Spanish and Brazilian Portuguese, which is reflected in our denotations in this chapter: gender on Spanish subject pronouns is always contextually determined, while gender on Brazilian Portuguese pronouns is narrowly syntactic. However, one could contend that perhaps gender on Brazilian Portuguese pronouns is only supplied syntactically in cases with a clear antecedent as in (34), while it is contextually determined in other cases, such as when the gender of the pronoun *does* correspond to the real-life biosocial gender of the individual being discussed by the speaker, as in (35):

[2]As discussed in chapter 2, I continue to follow Ahn in assuming that purely grammatical gender does not factor into competition between pro-forms.

(35) O João, ele é muito bonito. (Brazilian Portuguese)
The.MASC João, 3-MASC is very beautiful-MASC.
"John, he's very beautiful."

In the example above, the masculine gender on the pronoun *ele* corresponds to real-world biosocial gender, assuming that 'João' (a masculine given name equivalent to English 'John') refers to a male individual who generally uses masculine pronouns. As such, this instance of *ele* is compatible with a contextually determined form-indexical approach whereby the speaker determines from context that 'João' is appropriately referred to with masculine pronouns. However, this requires the stipulation that there are two distinct *ele* forms: one form which carries a contextually determined form indexical and one which does not, and is only valued for gender syntactically. I argue that this stipulation is unnecessary, as we can also account for cases where gender on BP pronouns corresponds to contextually-determined information about the referent in narrowly syntactic terms. Since the DP *o João* is syntactically masculine, as evidenced by the masculine article *o*, I assume that *ele* in (35) realizes masculine gender simply because it agrees in gender with the antecedent DP *o João*. Further support for this approach comes from cases where an individual who ordinarily uses a particular set of pronouns, e.g. masculine pronouns, is referred to using a pronoun of a different gender, e.g. a feminine pronoun, as is often done by LGBTQ Brazilian Portuguese-speaking individuals to playfully refer to a friend. For example, the man referred to as *o João* in the previous example may be referred to with feminine pronouns in a playful social context. The following example illustrates this, and is adapted from Twitter:

(36) A João, ela é muito afetada. (Brazilian Portuguese)
the.FEM João 3S-FEM is very affected-FEM
"João, she's really affected (artificial)".

In this tweet, the gender on the pronoun used to refer to João is feminine (likely for mocking/humorous effect), but so is the name *João*, as evidenced by its occurrence with the feminine article *a*. Although the gender on the pronoun in this case aligns with the speaker's assessment of what would be a contextually appropriate form to refer to João, I argue that this is only an indirect result; the locus of contextually determined gender is the name *a João*, and the pronoun

ela merely agrees in gender with this DP. If gender were entirely contextually determined by an Esipova-style form indexical on the pronoun, we might predict the possibility mismatch between a grammatically masculine name and a feminine pronoun whose form was deemed by the speaker to be appropriate for the given context. However, my BP-speaking consultants find this degraded, suggesting a tight relationship between the gender of the antecedent nominal and the gender of the pronoun in Brazilian Portuguese:

(37) **?**O João, ela é muito afetada. (Brazilian Portuguese)
the.MASC João 3-FEM is very affected-FEM
"João, she's really affected (artificial)".

I have defended the position that even when gender on Brazilian Portuguese pronouns matches real-world information about a referent's gender, this gender is still obtained in the narrow syntax, not from context. The fact that this gender may correspond to contextually available information about the referent is obtained indirectly from the antecedent nominal. In the case of antecedents such as proper names, e.g. *o/a João*, I leave open the question of whether there are two different roots that can be selected for based on context, or one root whose gender is determined from context at the interface (see Kučerová 2018 for an implementation of the latter approach).

The remaining case to be accounted for is that in which there is no overt antecedent, and an out-of-the-blue pronoun is used to refer to someone. Consider the following example, uttered by a speaker who observes someone they assume to be a male individual:

(38) Ele parece triste. (Brazilian Portuguese)
3-MASC seems sad.
"He seems sad".

Despite the lack of an overt antecedent for *ele* in this example, I argue that we can still conceive of pronominal gender on this instance of *ele* as arising through purely syntactic means. Following Sigurðsson (2019), who treats gender as an 'edge-linker' i.e. a feature that links extended verbal projection with the left periphery (e.g. Sigurðsson 2019), I propose that this instance of *ele* is a pronoun with no lexical content that copies the gender features of an antecedent syntactically present in the C domain.

To recap this chapter, I have argued that gender on overt pronouns in Spanish is a contextually-determined form indexical, following Esipova (2019)'s analysis of gendered pronouns in English. For Brazilian Portuguese, I reiterated and defended the position that gender is *not* contextually determined on Brazilian Portuguese pronouns and is instead always obtained through narrowly syntactic means. This leaves us with the following denotations for pro-forms in BP and Spanish, listed side-by-side here:

(39) ella/él (Spanish):

x: entity(x) & animate(x) & form(*ella/él*, x, c)

(40) *pro* (BP & Spanish), BP ela/ele (BP):

x: entity(x)

4.3 Returning to *Minimize Restrictors!*

Having now argued that Esipova (2019)'s *non*-presuppositional form indexicals are the source of semantically interpreted gender on Spanish pronouns, this opens up the possibility that *Minimize Restrictors!*, rather than *Maximize Presupposition!*, is the relevant pragmatic constraint governing competition between pro-forms. With this in mind, we can explain why rates of *pro*-drop in Spanish, approximately 80% (Enríquez 1984) are higher than in Brazilian Portuguese, where the rate of *pro*-drop has been reported at approximately 26% (Duarte 1993, 2000). Since *Minimize Restrictors!* favors expressions that impose fewer semantic restrictions whenever possible, Spanish speakers should favor *pro* over gendered pronouns across the board, as the latters' denotations involve an animacy restriction and a gendered form-indexical restriction not imposed by *pro*. This correctly predicts high rates of *pro*-drop in Spanish.

For Brazilian Portuguese, neither *pro* nor gendered forms *ele* and *ela* impose more semantic restrictions than the other, so *Minimize Restrictors!* does not dictate that one form should be favored over the other. However, this fact should not be taken to suggest that the rate of *pro*-drop is predicted to be exactly 50%. There are a number of other factors that may contribute to a

significantly lower rate of *pro*-drop in favor of more overt pronouns, such as priming. Mayol (2011) in particular discusses a phenomenon by which one instance of an overt subject pronoun primes further instances of overt subject pronouns, which in turn may prime other instances, leading to 'snowballing' of overt pronoun usage with time. In any case, *Minimize Restrictors!* correctly predicts significantly lower rates of *pro*-drop in Brazilian Portuguese than in Spanish.

We can additionally account for interpretative differences between overt pronouns in Spanish and Brazilian Portuguese using *Minimize Restrictors!*, more specifically the 'Pragmatic Irelevance' portion of the constraint, which states that including a restrictor *A* in *the A B* is illicit if *A* does not serve another purpose. For Spanish, where *Minimize Restrictors!* actively governs the choice between *pro* and overt subject pronouns, this means that an overt pronoun is generally illicit *unless there is a particular pragmatic purpose* for using one. As a result, we predict that the use of an overt pronoun in Spanish is associated with some sort of special pragmatic intent. This is indeed what we find: when there is pragmatic reason to draw attention to a particular referent, such as when making continued reference to the same entity, *pro* is the favored option and an overt subject is infelicitous. By contrast, when there is a special pragmatic reason to mark the subject, such as when the referent is different than the preceding topic, or when the subject is contrastive, an overt pronoun is felicitous (see de la Fuente 2015 and references therein). To illustrate, consider the following example:

(41) Juan$_i$ se-fue porque {él$_{\#i/j}$ / *pro*$_{i/j}$} tenía hambre. (Spanish)
Juan left because he / *pro* had hunger.
"Juan left because he was hungry."

In this example, which can be paraphrased as "Juan left because he was hungry", if the subject of the embedded clause is an overt pronoun, it can only be felicitously used to signal a change in reference to some referent other than *Juan*, the previous subject and topic. *Pro*, on the other hand, can be used to refer either to *Juan* or to a different referent.

For Brazilian Portuguese, the prediction made by *Minimize Restrictors!* is that there should be little to no pragmatic requirement for an overt subject pronoun to be licit, as the overt pronoun does

not carry more restrictors than *pro* in this language. This prediction is borne out: unlike in Spanish, overt subject pronouns in Brazilian Portuguese can be felicitously used to refer to salient discourse referents, including recently mentioned referents and current topics. Consider the BP equivalent to the Spanish sentence given above:

(42) O João$_i$ foi-embora porque {ele$_{i/j}$ / *pro*$_{i/\#j}$} tava com fome. (BrPt)
The.MASC João left because he / *pro* was with hunger.
"João left because he was hungry."

In this example, an overt subject in the embedded clause is perfectly acceptable when used to refer either to a new referent or to the previous subject and topic. On the other hand, also unlike in Spanish, *pro* generally sounds unnatural relative to an overt pronoun when used to refer to a new referent (e.g. someone other than João). As discussed in chapter 1, this is not a categorical contrast, and embedded *pro* can refer to a non-matrix subject antecedent under the right pragmatic conditions. While some scholars have made the empirically problematic choice to describe this contrast as the result of a syntactic requirement that *pro* have a c-commanding antecedent, I will explain this contrast in pragmatic terms without reference to c-command in the following section.

4.4 C-command revisited

In this chapter, I will examine Brazilian Portuguese sentences such as (47), repeated from (6) in chapter 1 (adapted from Fernandes et al. 2018). It has previously been reported that only a c-commanding antecedent is possible for *pro* in Brazilian Portuguese, meaning that only a subject antecedent *o atleta* should be possible for the embedded null subject, and not the object antecedent *o médico* (or any other non-subject antecedent).

(43) O atleta consultou o médico depois que *pro* regressou da viagem à Itália.
the athlete consulted the doctor after that *pro* returned from-the trip to-the Italy
(Brazilian Portuguese)

"The athlete consulted the doctor after he returned from the trip to Italy."

As discussed previously, a strict c-command requirement makes the wrong prediction, as Fernandes et al. (2018) found for sentences like (47) that Brazilian Portuguese may indeed associate

the embedded null subject with an object antecedent, even if they choose a subject antecedent a higher percentage of the time. In what follows I will argue that what has been previously described as a syntactic c-commanding antecedent requirement for *pro* in Brazilian Portuguese is actually a strong pragmatic inference that *pro* in (47) refers to the previous subject. I will argue that the strength of this inference relative to Spanish, where *pro* can relatively easily refer to a non-subject antecedent, is due to differences in the two languages' overt pronouns and not to differences in how *pro* is licensed in both languages. To do this, I will make use of a simplified model of pronoun resolution that suffices to illustrate my point.

As mentioned previously in this chapter, a well-established generalization about null subject languages is that *pro* is more likely interpreted as referring to an entity that is 'salient' or 'prominent' in the discourse, while the overt pronoun is associated with less prominent/salient entities. A subcase of this is that subject *pro* is more likely to be associated with a subject antecedent relative to an overt pronoun, which is more likely to be associated with an object antecedent (Carminati 2002, Alonso-Ovalle et al. 2002, Filiaci et al. 2013, de la Fuente 2015, and many others).[3]. Building on this, let us assume a simple model of pronoun resolution in which potential antecedents can be ordered on a scale of how prominent or salient they are in the discourse, from 0 to 3 (0 being the least prominent, e.g. completely a completely new referent, and 3 being the most prominent, e.g. a unique contextually salient referent). First, I will assign a fixed range of antecedent salience to overt subject pronouns in Brazilian Portuguese and Spanish. Then, I will assume that *pro* in both languages is in principle semantically compatible with referents of any salience, but that the resolution of *pro* is enriched by way of an inferential reasoning process. By doing so, we can see how *pro* ends up being more restricted in Brazilian Portuguese than in Spanish[4].

[3]This relates to prominence as subjects in null subject languages tend to be topics and therefore make prominent antecedents.

[4]One may reasonably object to the fact that I have fixed the possible range of salience for antecedents of overt pronouns in BP and Spanish while letting the range of salience for antecedents of *pro* be determined through an implicature, when the opposite configuration is also logically possible. I argue that fixing values for overt pronouns and letting the resolution of *pro* be determined via implicature is the empirically supported option, because while both *pro* and and overt pronouns have different behaviors w.r.t. pronoun resolution when you compare BP and Spanish, only for overt pronouns do we have evidence that the denotation varies from language to language. While there is substantial evidence for overt pronouns in Spanish being semantically more restricted than those in BP, *pro* in both languages appears minimially semantically restricted; in both languages, *pro* is compatible with referents of any gender

Now, since we have established that overt pronouns in Brazilian Portuguese and Spanish vary in terms of their interpretative properties, we can ask what values for antecedent prominence overt pronouns in each language have. For Spanish, since *Minimize Restrictors!* dictates that overt pronouns may only be deployed instead of *pro* when there is a special pragmatic reason, let us assume that overt pronouns are only licensed in order to signal reference to a very non-salient referent, e.g. those referents that lie between 0 and 1 on our prominence scale.

For Brazilian Portuguese, *Minimize Restrictors!* does not dictate that overt pronouns such as *ele* and *ela* be reserved for special pragmatic purposes, as these forms are minimally semantically restricted. However, even in BP overt pronouns are still less minimal than *pro* in purely phonological terms, and thus the overt pronoun is likely still interpreted as slightly more pragmatically marked than the null form (e.g. by the Gricean Maxim of Manner, or Levinson's (1998) Maxim of Minimization). In fact, this is what Fernandes et al.'s (2018) study of pronoun resolution in finds: even in Brazilian Portuguese, overt subjects are associated more with non-subject referents than *pro*, although the contrast is considerably weaker than in consistent null subject languages. As such, let's assume that BP overt pronouns are still reserved for the lower end of the salience scale (albeit less so than in Spanish) and are permitted with antecedents whose salience ranges from 0 to 2. Finally, following Carminati (2002), Alonso-Ovalle et al. (2002) and others, I assume that subject antecedents are more salient in the discourse than object antecedents, placing subjects at 2 on the prominence scale and objects at 1.

The prominence of subject and object antecedents, along with the range of salience of antecedents that each language's overt pronouns may be associated with, are plotted in Figure 4.1.

and animacy, and is compatible with referential and bound-variable interpretations. Thus, it makes sense to assume that any interpretative difference between *pro* in BP and in Spanish stems from something other than the denotation of *pro* itself, e.g. from an implicature.

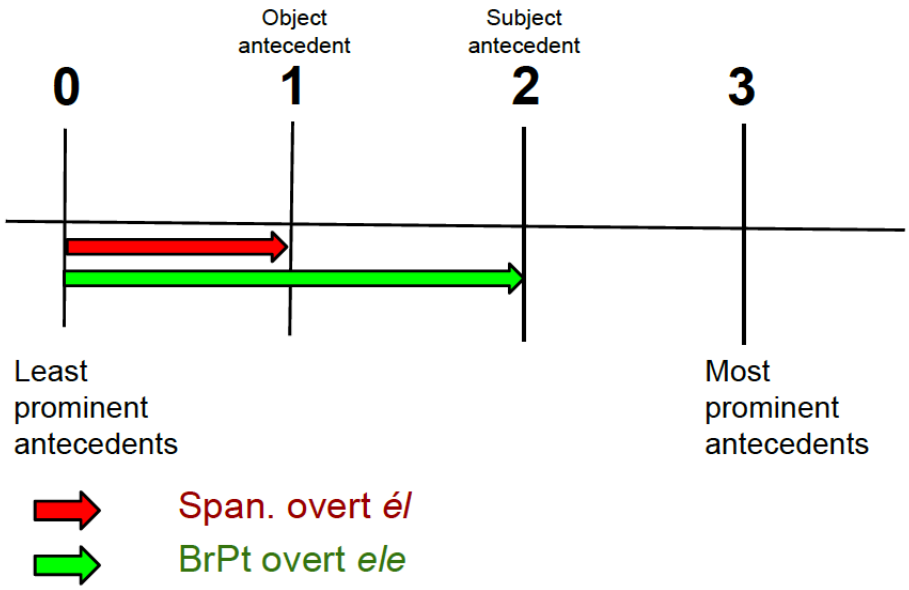

Figure 4.1: The resolution biases of overt pronouns in Spanish (red) and Brazilian Portuguese (green) plotted in terms of how prominent their antecedents may be.

This graph should be interpreted to mean that in a sentence like (44), Spanish overt *él* is natural with the object antecedent[5] *el médico* but unnatural with the subject antecedent *el atleta*:

(44) El atleta$_i$ consultó al médico$_j$ después de que él$_{\#i/j}$ regresó de su viaje a
the athlete consulted DOM-the doctor after of that he returned from his trip to
Italia. (Spanish)
Italy
"The athlete consulted the doctor after he returned from the trip to Italy."

We also see that BP overt *ele* is compatible with both a subject and an object antecedent:

(45) O atleta$_i$ consultou o médico$_j$ depois que ele$_{i/j}$ regressou da viagem à
the athlete consulted the doctor after that he returned from-the trip to-the
Itália. (Brazilian Portuguese)
Italy
"The athlete consulted the doctor after he returned from the trip to Italy."

Now I will derive the interpretative difference between *pro* in BP and Spanish. The essence of the argument is that *pro* is interpreted as most felicitous with antecedents that an overt pronoun would not have been felicitous with, i.e. in the 'blank space' of Fig. 4.1. I assume that *pro*, being

[5]Or with any other non-subject antecedent; I used an object antecedent for the sake of illustration.

semantically unrestricted in both languages, is semantically compatible with referents of any level discourse prominence, but that its resolution biases are enriched via an implicature.

I argue that the inference that takes place for Spanish is as follows: upon hearing an instance of *pro*, the Spanish-speaking hearer reasons that the speaker did not use an overt pronoun, but would have done so if it were pragmatically called for. Since an overt pronoun is the pragmatically appropriate form for non-prominent antecedents ranging from 0 to 1 on the prominence scale, the intended antecedent of *pro* must be something of prominence 1 or higher. This marginally includes an object antecedent, and clearly allows for a subject antecedent for *pro*[6]:

(46) El atleta$_i$ consultó al médico$_j$ después de que *pro*$_{i/j}$ regresó de su viaje a Italia. (Spanish)
the athlete consulted DOM-the doctor after of that *pro* returned from his trip to Italy
"The athlete consulted the doctor after he returned from the trip to Italy."

Meanwhile, in Brazilian Portuguese, the inference process is as follows: upon hearing *pro*, a Brazilian Portuguese speaker reasons that if an overt pronoun were pragmatically called for, it would have been used, so it must not have been called for. The range of possible antecedents for an overt pronoun in BP goes from 0 to 2 and completely includes object antecedents. Thus, all else equal, the BP-speaking parser concludes that *pro* is *not* intended to have an object antecedent, as this is firmly included in the range of antecedents that an overt subject pronoun would've been used for. Meanwhile, a subject antecedent is an edge case that can be associated with either null or overt form. As a result, the expected (and attested) BP judgments are as follows:

(47) O atleta$_i$ consultou o médico$_j$ depois que *pro*$_{i/\#j}$ regressou da viagem à Itália. (Brazilian Portuguese)
the athlete consulted the doctor after that *pro* returned from-the trip to-the Italy
"The athlete consulted the doctor after he returned from the trip to Italy."

[6]Note that both an object antecedent and a subject antecedent are acceptable for *pro* in Spanish, but when all else is equal, Spanish speakers' default interpretation is to associate *pro* with a subject antecedent. This is reflected graphically by the fact that a subject antecedent is clearly outside of the bounds of the possible antecedent of an overt pronoun, and thus clearly a candidate antecedent for *pro*, while an object antecedent is at the edge and compatible with either an overt or a null form. This is still quite a simplified picture; for further discussion see de la Fuente (2015).

Note that while the simplified system above was used to sketch how pronoun resolution may proceed for object and subject antecedents with *pro* and overt pronouns, the same reasoning holds for any type of non-subject antecedent that is less prominent in the discourse than a subject.

Because of the inference that a null form in BP should not be associated with a non-subject antecedent, the reported judgment we get is that only a subject antecedent is possible. As a result, BP has frequently been mischaracterized as a language where *pro* is syntactically licensed only with a c-commanding antecedent. However, I have derived this without c-command in pragmatic terms. Since my account relies on a pragmatic inference, the subject antecedent preference should be cancelable if we provide the right pragmatic conditions for a non-subject antecedent to be associated with *pro*, as in (48), repeated from chapter 1.

(48) [O pai d-ela$_i$]$_j$ disse que *pro* só vai namorar depois dos trinta. (BrPt)
the father of-her says that *pro* only goes-to date-INF after of-the thirty
"[Her$_i$ father]$_j$ says that she$_i$ can only date after she's thirty."

Indeed, we find that in this example (where the intended referent is clear), a null subject with a non-c-commanding antecedent is perfectly acceptable.

4.5 Summary of Chapter 4

In this chapter, we began with the observation that *Minimize Restrictors!* could be used to explain varying distributions of *pro* and overt subject pronouns in Brazilian Portuguese and Spanish, as the overt pronouns in each language carry different numbers of semantic restrictions. However, I first needed to adopt a non-presuppositional account of pronominal gender, for which I turned to Esipova (2019). I then modified my denotations of overt pronouns in Spanish to make use of Esipova's notion of contextually-dependent form indexicals, then justified my choice not to include these indexicals for my analysis of BP overt pronouns. Then, I returned to *Minimize Restrictors!* and used it to account for why overt pronouns are omitted in favor of *pro* more often in Spanish than in Brazilian Portuguese. I also accounted for interpretative contrasts associated with overt pronouns in Brazilian Portuguese and Spanish under *Minimize Restrictors!*. Finally, I sketched a model of pronoun resolution under which speakers of Brazilian Portuguese make an inference that

pro is unlikely to be uttered with an intended non-subject antecedent, thereby eliminating the need to posit a c-command requirement in the syntax.

CHAPTER 5

BROADENING THE EMPIRICAL SCOPE

5.1 Overview of Chapter 5

In this chapter I will extend the empirical generalizations made so far for Brazilian Portuguese and Spanish[1] to other Romance languages that pattern together in similar ways, namely French and Dominican Spanish (which pattern with Brazilian Portuguese) and Italian (which patterns with Standard Spanish). I will refer to French, Dominican Spanish, and Brazilian Portuguese as the 'weak pronoun' languages, as these languages all have a weak (i.e. semantically minimally restricted) overt subject pronoun in their pronominal inventory. I will refer to Italian and Standard Spanish as the 'strong pronoun' languages, as they only have strong overt subject pronouns.

5.2 Animacy and biosocial gender by language group

I will begin by showing that adding additional languages to our sample strengthens the following correlation: if a language permits its overt subject[2] pronouns to refer to inanimate entities, then it also permits them to realize a gender feature that does not correspond to the real-world biosocial gender of its referent in certain circumstances. We see this is true for both French (49-50), Brazilian Portuguese (51-52, (repeated from above)), whose subject pronouns may be inanimate and may also take a value for gender other than the value corresponding to the biosocial gender of the person being discussed. A similar generalization possibly holds for Dominican Spanish (see Bullock and Toribio 2009 regarding (in)animacy).

(49) *French subject pronouns may be inanimate*[3]:

[1]Up to now, the judgments referred to as 'Spanish' apply to at least standard Rioplatense Spanish and Castilian Spanish, and likely also Mexican Spanish and other dialects. My consultants have been speakers from Argentina and Spain.

[2]Some of the generalizations made here hold for pronouns in other syntactic positions, but because the focus of the thesis is on null subjects and their overt counterparts, I will restrict the discussion to subjects only.

[3]Example obtained from http://opus.nlpl.eu/OpenSubtitles2016.php. Note that the French pronoun *il* may also be used to refer to animate referents, like English *he*.

Context: Talking about an atom (*un atome*), which is syntactically masculine in French.

a. Quand il se décompose il libère une toxine indétectable. (French)
When 3-MASC SE decomposes 3-MASC releases a toxin undetectable.

"When it decomposes, it releases an undetectable toxin."

(50) *French subject pronouns can go against biosocial gender*:

Context: A male sentinel (*sentinelle*, invariable syntactically feminine) is being discussed.

(adapted from Ihsane & Sleeman 2016).

a. La sentinelle arriva. Elle avait une longue barbe. (French)
The.FEM sentinelle.FEM arrived 3S-FEM had a long beard.
"The sentinelle arrived. He (lit. *she*) has a long beard."

(51) *Brazilian Portuguese subject pronouns may be inanimate*:

Context: A speaker comments on the moon (*a lua*, feminine).

a. Ela é bonita. (Brazilian Portuguese)
3S-FEM is beautiful
"It is beautiful".

(52) *Brazilian Portuguese subject pronouns may go against biosocial gender*:

a. Uma pessoa que realmente acha que ela mesma$_i$ pode fazer a
a.FEM person.FEM that really believes that 3S-FEM same can make-INF a
diferença na sociedade, é o João$_i$. (Brazilian Portuguese)
difference in-the society, is the João
"A person x who really thinks that x can make a difference in society, that person is John."

(53) *Dominican Spanish subject pronouns may be inanimate*:

Context: A speaker comments on the moon (*la luna*, feminine).

a. Ella es bonita. (Dominican Spanish)
3S-FEM is beautiful
"It is beautiful".

Although I don't have data from any Dominican Spanish-speaking consultants regarding gender mismatch scenarios like in Brazilian Portuguese, the fact that gendered Dominican Spanish overt

subject pronouns are compatible with inanimates tells us that at least in some cases gender on overt pronouns does not correspond to biosocial gender. This is because there is no correlate of biosocial gender for inanimate objects.

Moving onto the 'strong pronoun' pronoun languages, we see that in Standard Spanish (54, repeated from above) & Italian (56), overt pronominal subjects referring to inanimate referents are categorically ungrammatical and unattested (see Enríquez 1984 for Spanish[4] and Carminati 2002 for Italian). Additionally, in both Italian (57) and Standard Spanish (55, repeated) the gender realized on subject pronouns may *not* conflict with the biosocial gender of the person being discussed, regardless of the presence of a potential syntactic antecedent of the opposite syntactic gender.

(54) *Spanish overt subject pronouns may not be inanimate*:

Context: A speaker comments on the moon (*la luna*, feminine).

a. *Ella es bonita. ((non-Dominican) Spanish)
3S-FEM is beautiful
Intended: "It is beautiful".

(55) *Spanish overt subject pronouns must match biosocial gender*:

a. Una persona que realmente cree que {$^{??}$ella misma$_i$ / ^{OK}él mismo$_i$}
a.FEM person.FEM that really thinks that 3S-FEM same / 3S-MASC same
va a marcar una diferencia en la sociedad, es Juan$_i$. (Spanish)
goes to marc-INF the difference in the world is John
"A person x who really believes that x can make a difference in the world, that person is John."

(56) *Italian overt subject pronouns may not be inanimate*:

Context: A speaker comments on the moon (*la luna*, feminine).

a. *Lei è bella. (Italian)
3S-FEM is beautiful
Intended: "It is beautiful".

[4] Actually, Enríquez reports that five cases making up a total of 0.15% of overt subject pronouns in Madrid Spanish had inanimate referents. However, these few cases were not true exceptions as according to her they were used to 'personify' a typically inanimate object.

(57) *Italian overt subjet pronouns must match biosocial gender*[5]:

a. Una persona che crede davvero che {$^{??}$lei stessa$_i$ / ^{OK}lui stesso$_i$}
a.FEM person.FEM that believes of-truth that 3S-FEM same / 3S-MASC same
possa fare la differenza nel mondo, è Giovanni$_i$. (Italian)
can make the difference in-the world, is John.
"A person x who really believes that x can make a difference in the world, that person is John."

We see from this collection of data that languages in this sample either both allow inanimates and gender on animate-referring pronouns that goes against the biosocial gender of the person being discussed, *or* both disallow inanimates and require that animate-referring pronouns have a value for gender that matches the biosocial gender of the person being discussed. This is a nontrivial result; we could easily imagine a configuration in which French and Brazilian Portuguese allow their subject pronouns to have inanimate referents, but nonetheless determine gender on animate-referring pronouns from the discourse context. Likewise, we could imagine a configuration in which Spanish and Italian disallow inanimate-referring subject pronouns but are agnostic about whether the gender realized by these pronouns corresponds to the gender of an antecedent like *la persona* ('the person', feminine) or the biosocial gender of a male referent who typically uses masculine pronouns. Because biosocial gender and animacy restrictions do in fact pattern tightly together, this provides support for my hypothesis that they stem from a common source, namely the presence of PERSON. I suggest extending my proposal from chapter 2.1 to French, Dominican Spanish, and Italian, such that Standard Spanish and Italian subject pronouns host PERSON while French, Brazilian Portuguese, and Dominican Spanish do not.

5.3 Distributions of null subjects by language group

In both 'strong pronoun' languages in my sample, Italian and Spanish, null subjects are freely permitted and deployed instead of overt pronouns the vast majority of the time. As previously cited, Enríquez (1984) reports a subject omission rate of about 80% for Madrid Spanish. Italian

[5]I again report the majority judgment here; two of my Italian-speaking consultants firmly reject this sentence, while one reports the sentence to be marginally acceptable. As with Spanish, there may be some interspeaker variation, but in any case there is a still a contrast with Brazilian Portuguese.

patterns with Spanish in having a similarly high rate: Lorusso, Caprin, & Guasti (2005) report a subject omission rate of 74% for Italian.

The other three Romance varieties, French, Brazilian Portuguese, and Dominican Spanish, omit subject pronouns substantially less frequently than varieties in the prototypical group, or not at all. As cited in chapter 4, Brazilian Portuguese has a subject pronoun omission rate of just 26% (Duarte 1993, 2000). Dominican Spanish also has a low rate of 39%. Finally, French used to allow null subjects, completely lost null subjects by the end of the Middle French period (Kaiser 2009). Interestingly, the loss of *pro* the diachrony of French appears to have been preceded by a change in the inventory of overt pronominal forms. To quote Roberts (2014), "there was an enrichment of the overt pronoun inventory at the expense of *pro*,". More specifically, in the Early Middle French period a series of strong/tonic pronouns (e.g. 1st-person *moi*, 2nd-person *toi*, and 3rd-person masculine *lui*) emerged and appeared in complementary distribution with a now-weak/atonic series of subject pronouns (e.g. 1st-person *je*, 2nd-person *tu*, 3rd-person masculine *il*). In Roberts' analysis, the weakening of this latter series of pronouns through the emergence of the strong series encroached on the distribution of *pro*, eventually rendering *pro* obsolete.

Since all three of the 'weak pronoun' languages in this sample share low or non-existent rates of subject pronoun omission, I take this as empirical support for my proposal. This correlation reinforces my hypothesis that having a PERSON-less weak subject pronoun inventory leads to higher rates of overt subject expression/lower rates of *pro*-drop, which I related to *Minimize Restrictors!* in the previous chapter. I similarly take the fact that the 'strong pronoun' languages have high rates of *pro*-drop as empirical support for my analysis.

The interpretative differences discussed for overt subjects in Brazilian Portuguese vs. Spanish appear to extend to the other languages as well. For example, just as in Brazilian Portuguese (58), repeated from chapter 4), but unlike in Standard Spanish, in Dominican Spanish it is perfectly felicitous to realize an embedded overt subject with a matrix subject antecedent (59, adapted from Martínez-Sanz 2011):

(58) O atleta$_i$ consultou o médico depois que ele$_i$ regressou da viagem à Itália.
the athlete consulted the doctor after that he returned from-the trip to-the Italy
(Brazilian Portuguese)

"The athlete consulted the doctor after he returned from the trip to Italy."

(59) Tío Papi$_i$ murió tan feliz que *pro*$_i$ no sabe que él$_i$ murió. (Dominican
Tío Papi died so happy that *pro* NEG knows that 3-MASC died
Spanish)

"Tío Papi died so peacefully that he doesnt even know that he's dead."

Finally, French clearly patterns with Dominican Spanish and Brazilian Portuguese in freely permitting embedded overt subject pronouns with matrix subject antecedents, as there is no null counterpart to the overt subject pronoun in French.

5.4 Summary of Chapter 5

In this chapter I demonstrated that the previously discussed observations surrounding pronouns in Brazilian Portuguese and Spanish can be extended to French, Italian, and Dominican Spanish as well. I showed that, like Brazilian Portuguese overt subject pronouns, Dominican Spanish and French subject pronouns may have inanimate referents and may be valued for gender in a way that does not match the gender of the real-life individual being discussed. Italian on the other hand, patterns with (non-Dominican) Spanish in requiring animate referents for overt subject pronouns and requiring that the gender of an overt subject pronoun correspond to the real-life gender of the individual being discussed. Finally, I showed that languages which group together in terms of the semantic properties of their overt subject pronouns also pattern together in terms of how/to what extent they license null subjects (or did so historically, in the case of French).

CHAPTER 6

ACQUIRING NULL AND OVERT SUBJECTS

6.1 Overview of Chapter 6

An advantage of the present proposal for an overt pronoun-centered approach to variation between NSLs is that it simplifies the *pro*-drop-acquiring child's learning path in a number of ways. My account reduces the syntactic difference between 'consistent' null subject languages like Standard Spanish and 'partial' null subject languages like Brazilian Portuguese to a difference in these language' overt subject pronouns. As a result, much of the work of learning the target distributions of null and overt subjects in a given null subject language is reduced to learning whether PERSON is present or absent on overt subject pronouns.

6.2 Acquiring conditions on *pro*-drop

Let us consider the task that a child acquiring a *pro*-drop language faces in terms of learning the appropriate syntactic and/or pragmatic contexts for when *pro*-drop is licit. Per the D(efiniteness)-feature approach of Holmberg, Nayudu, and Sheehan (2009) discussed in chapter 1, the lack of a D-feature on T in Brazilian Portuguese means that definite null subjects are generally not licensed, and only licit if some alternative mechanism is present (e.g. control by a c-commanding matrix antecedent). Under this view, the child must learn that (i) subject pronouns may be omitted in the target language and (ii) *pro*-drop is generally *not* permissible and limited only to certain contexts. This requires the child to rely on the negative evidence that *pro*-drop does not occur in other contexts to learn that it is restricted to only the observed contexts. Alternatively, one could posit that the child posits the most conservative hypothesis consistent with the instances of *pro*-drop they observe and only extends the hypothesis if *pro*-drop occurs in other contexts (as would occur in a consistent null subject language like Spanish). Given that *pro*-drop is not categorically restricted to cases in which it has a c-commanding subject antecedent, though, this is empirically problematic for

Holmberg et al.'s account. By contrast, my account only requires the child to learn whether or not the overt subject pronouns in their target language host PERSON, then incorporate this knowledge with broader pragmatic knowledge utilized in other linguistic domains (for discussion of the extent of children's pragmatic knowledge, see Pérez-Leroux and Roeper 1999). Once the child converges on the target representation of overt pronouns (either with or without PERSON), they can make use of general Gricean pragmatic principles like the Maxim of Manner (around which Schlenker bases *Minimize Restrictors!*) to establish the discourse conditions under which null and overt pronouns are appropriate.

6.3 Acquiring the Overt Pronoun Constraint

The Overt Pronoun Constraint, discussed at length in chapter 3, is repeated below:

(60) Overt pronouns cannot occur under the scope of a quantifier iff the alternation overt/empty obtains.

Stated as above, this constraint requires the Spanish-acquiring child to manage the conflicting knowledge that in bound-variable contexts, a subject pronoun should in principle be able to be either covert or overt, but precisely because of this possible alternation only the covert form is possible. Furthermore, the only evidence available to the child in the input for the impossibility of an overt form is the indirect negative evidence that overt pronouns do not occur in this context. However, if the OPC follows from general constraints on pronominal coreference as Ive suggested in this thesis, Spanish-acquiring children would only need to acquire the target representations of null and overt pronouns to learn the OPC; once they learn that overt 3rd-person pronouns host PERSON and are inherently referential, they have enough information to infer that this argument makes a poor bound variable.

6.4 Acquiring PERSON

The question that remains is how children converge on the target representations of overt pronouns in their language (with or without PERSON) in the first place. If there is a robust relationship between

animacy and PERSON as suggested by Rooryck (2000), Lochbihler, Oxford and Welch (2015), and others, children may be able to track the animacy of subject referents to acquire PERSON (or a lack thereof). For children acquiring Brazilian Portuguese, Dominican Spanish, or French, the learning path is straightforward: once a significant number of instances of inanimate overt subject pronouns have been observed, the child may conclude that these forms lack PERSON. However, the question of how children acquiring a language like Spanish or Italian learn that overt subject pronouns may only refer to animate entities is slightly more complex. Since inanimate overt subject pronouns are unattested in the input for these languages, children lack direct negative evidence for the impossibility of inanimate overt subject pronouns. Further complicating matters, a Spanish-acquiring child's overt subject pronoun input that consists exclusively of animate referents is still largely compatible with a PERSON-less representation of overt pronouns. This is because, even in languages like French or Brazilian Portuguese, the default interpretation of 3rd-person subject pronouns in out-of-the-blue is an animate interpretation (Rooryck 2000 for French, Cristina Schmitt, p.c. for Brazilian Portuguese). Barbosa, Duarte, and Kato (2005) also document that despite attested cases of inanimate overt subject pronouns in Brazilian Portuguese, overt pronominal form is strongly positively correlated with animacy. At best, then, a Spanish-acquiring child has very weak indirect negative evidence that overt pronouns host PERSON and are categorically restricted to animate referents.

One promising way to avoid this learning problem is to posit that there is a universal bias for gendered pronominal forms to be animate. This appears to be supported by the cross-linguistic generalization that, while there are languages in which grammatical gender sometimes has a real-world correlate (as in Spanish) and languages in which grammatical gender always corresponds to a real-world correlate (as in English), there are no languages in which there is a grammatical gender system that is never associated with properties of animate individuals (Kramer 2020). If the unmarked option is for grammatical gender to be a property of animate-denoting forms, then perhaps this is the hypothesis that children begin with. In slightly more formal terms, perhaps the universal default is for gendered pronominals to host PERSON, in which case children acquiring

Spanish and Italian begin with the target representation of overt subject pronouns and are never presented with evidence that forces them to abandon this representation. For French, Brazilian Portuguese, and Dominican Spanish, children abandon their initial representation of a PERSON-hosting overt subject pronoun when present with a significant number of inanimate overt subject pronoun.

6.5 Summary of Chapter 6

In this chapter I outlined how the present approach simplifies the learning task faced by the *pro*-drop acquiring child in a number of ways. I argued that, under my account, knowing whether a given *pro*-drop language is partial (like Brazilian Portuguese) or consistent (like Italian) may be reduced to knowing whether the language's overt subject pronouns host PERSON.

CHAPTER 7

REMAINING QUESTIONS AND CONCLUSIONS

7.1 Overview of Chapter 7

In this chapter I will highlight a number of promising areas for future research, address potential issues with the analysis I've presented so far, then conclude the thesis.

7.2 Diachronic change

An important question relates to what predictions the theory of null subject languages presented in this thesis makes for diachronic change within null subject languages. This type of discussion in the Romance null subjects literature is usually framed in terms of loss of the null subject parameter; for example, the loss of null subjects in French (Roberts 2014), the apparent loss of Avoid Pronoun in Brazilian Portuguese (Duarte 2000), and a similar loss of widespread *pro*-drop in Dominican Spanish (Camacho 2008).

As mentioned in chapter 5, Roberts (2014) highlights that in the case of French, the emergence of weak pronouns in the Middle French period appears to have triggered a reduction in *pro*-drop. Given that I have shown Brazilian Portuguese and Dominican Spanish pronouns to be 'weaker' than their counterparts in languages like Standard Spanish and Italian, the same trajectory may be possible for these languages as well. However, the opposite path may be possible as well: the rate of *pro*-drop decreases in favor of overt subject expression, causing learners in subsequent generations to posit a semantically less restricted representation of the overt pronoun in line with their knowledge of *Minimize Restrictors!*. A version of this is put forth by Camacho (2008), who suggests that a statistical increase in the frequency of overt pronouns in Dominican Spanish and Brazilian Portuguese triggered a change in their statuses as null subject languages. He argues that this initial change may have been arisen due to a change in the morphological paradigm or at random; in any case, the result is that the division of labor between null and overt subjects in terms

of discourse function is blurred as overt subjects become more frequent.

Camacho's account for Dominican Spanish and Brazilian Portuguese is certainly compatible with my analysis. However, there is reason to be somewhat skeptical of his account for 'change' in the status of Brazilian Portuguese overt subject pronouns, namely because it's not entirely clear that Brazilian Portuguese overt subject pronouns are distinct from European Portuguese overt subject pronouns as is standardly assumed. Camacho frames his analysis in terms of the eventual divergence of Brazilian Portuguese from the consistent NSL status of European Portuguese, but overt pronouns in European Portuguese display the same properties as described here for Brazilian Portuguese. Posio (2012) and others have documented many cases of overt subject pronouns with inanimate referents in European Portuguese. Additionally, the European Portuguese speakers I have consulted report the same judgments as Brazilian Portuguese speakers for the gender mismatch data presented in chapter 2; that is, when a conflict arises between the arbitrary syntactic gender of an antecedent like *uma pessoa* ('a person') and the biosocial gender of the individual being discussed (e.g. a man named John), European Portuguese speakers allow this to be resolved in favor of the syntactic gender on *uma pessoa*, unlike Standard Spanish and Italian speakers. It is also unlikely that this is a recent innovation in the history of Portuguese, as there are examples of inanimate overt subject pronouns from at least as early as the 18th century:

(61) Mas se assi é, ela é a milhor invenção que eu vi. (16th-century European Portuguese)
but if so is, 3-FEM is the.FEM best invention.FEM that I saw

"But if it is so, it is the best invention that I've seen." (Camões, *El-Rei Seleuco*, 1545)

In light of this data from contemporary and early Modern Portuguese, it may not be necessary to analyze Brazilian Portuguese as a particularly innovative null subject variety. To the extent that Brazilian Portuguese and European Portuguese do differ from each other in terms of null subjects (see Barbosa, Duarte and Kato (2005) for discussion), we should at least refrain from positing this difference in terms of the makeup of their overt pronouns.

In any case, my analysis provides several possible pathways for diachronic change: a change in

the distribution of null subjects may lead to a change in the overt pronoun inventory of given null subject language, or a change in the overt pronoun inventory may lead to changes in the distributions of null subjects.

7.3 Exceptions to the animacy restriction

Although the animacy restriction on overt pronouns holds categorically for standalone subject pronouns in Spanish and Italian, overt pronouns (at least in Spanish) appear not to obey this restriction in a few limited contexts. In particular, the overt forms *él* and *ella*, which are restricted to animate referents in subject position, may refer to inanimates when they appear as the objects of prepositions (64a). Additionally, there are limited cases of overt pronouns in subject position that may refer to inanimate referents, namely when the pronoun is quantified (62a), or when the pronoun is modified by an adjective (63).

(62) *Quantified overt subject pronouns may be in inanimate in Spanish.*

Context[1]: a description about a website (*un sitio*, syntactically masculine).

a. Todo él es bilingüe (Castellano-Inglés). (Spanish)
all 3S-MASC is bilingual (CastilianEnglish)
"All of it (lit. 'him') is bilingual (SpanishEnglish)."

(63) *Adjectivally modified overt subject pronouns may be inanimate in Spanish.*

Context: discussing a clock. (Adapted from Bosque et al 2009)

a. Ocupaba él solito toda la habitación. (Spanish)
Occupied 3S-MASC alone all the room
It (lit. *he*) alone took up the entire room."

(64) *Spanish* (Adapted from Camacho 2018, p. 351)

a. Aunque no hay otros muebles sino la mesa, les pido que no se
Although NEG are other furnitures besides the table, 3PL.DAT ask.1S that no SE
suban sobre ella. (Spanish)
mount.3PL on 3S-FEM
Even though there's no other furniture besides the table, I ask that you don't sit on it (lit. 'her').

[1]Sentence obtained from https://skyandtelescope.org/clubs-organizations/observatorio-arval/.

The common factor uniting all of these cases of inanimate *él* and *ella* is the impossibility of *pro*. Spanish does not license null prepositional objects, and it is likely also illicit to overtly quantify over a silent argument or adjoin an overt adjectival to one. I therefore suggest that what occurs here is that the only overt 3rd-person pronominal forms available, namely *él* or *ella*, are inserted, and as such these forms are coerced into referring to inanimates, perhaps via Impoverishment (Bonet 1991, Halle and Marantz 1993). A separate logical possibility would be to posit that there is a homophonous *él*/*ella* that may refer to inanimates. However, the fact that *él* and *ella* must be coerced into referring to inanimates under my analysis is supported by the observation that several of my consultants find inanimate *él*/*ella* unnatural even in (62a)(64a), despite this form being the only pronoun available.

7.4 Null impersonal subjects

One aspect of Holmberg et al.'s D-feature hypothesis regarding the difference between Brazilian Portuguese and Spanish that I have not discussed is the availability of null impersonal null subjects. According to Holmberg (2005), Holmberg, Nayudu and Sheehan (2009) and subsequent work, the absence of a D-feature on T in Brazilian Portuguese not only means that definite null subjects are severely restricted, it also has the consequence of making available indefinite, impersonal null subjects akin to English 'one'. Brazilian Portuguese forms a minimal pair with Spanish in this respect, as shown in the contrast between BP (65) and Spanish (66) below:

(65) É assim que {*pro* / se} faz o bolo. (Brazilian Portuguese)
is so that *pro* / SE makes the cake
"This is how one makes the cake."

(66) Es así que {**pro* / se} hace el pastel. (Spanish)
is so that *pro* / SE makes the cake.
"This is how one makes the cake."

Within Holmberg et al.'s analysis, the presence of the D-feature on T in Spanish and other consistent null subject languages blocks the availability of an indefinite, impersonal null subject. To the extent that the empirical contrast captured by Holmberg et al. is real, my account has

nothing to say regarding the (non)availability of impersonal null subjects in null subject languages of different types. However, I believe that in the case of Spanish and Brazilian Portuguese, there is reason to doubt that the minimal pair above is definitively related to the presence/absence of a D-feature or any other null subject parameter. This is because the attested cases of null impersonals in Brazilian Portuguese are restricted to cases such as (65) in which the null subject alternates with *se*. I contend that this particular contrast may be an independent fact about the realization of SE in Brazilian Portuguese and not a direct empirical consequence of how null subjects are licensed. This contention is supported by the fact that Dominican Spanish, which patterns with Brazilian Portuguese as a partial null subject language in (i) having low rates of *pro*-drop, (ii) not requiring special pragmatic conditions for deploying overt subject pronouns, and (iii) having semantically weak overt pronouns, patterns with Spanish and against BP in disallowing a null subject in sentences like (66). I thus suggest that although Brazilian Portuguese allows impersonal null subjects, this fact may be unrelated to the licensing of null subjects at large. In any case, the presence of null impersonals in BP does not intersect with my analysis in any meaningful way.

7.5 Conclusion

I began this thesis by highlighting empirical shortcomings of accounts that posit categorical differences between NSLs in how/whether null subjects are licensed in the inflectional domain: they predict cross-linguistic contrasts in availability of null subjects in contexts where such a contrast does not reliably obtain categorically. I then suggested an alternative wherein the differences in distributions of null subjects between NSLs are due to variation in the makeup of overt pronouns, rather than variation in the inflectional domain.

To this end, I argued for a novel treatment of the Overt Pronoun Constraint, a property typically associated with the licensing of null subjects, but which I showed can be reduced to a side effect of structural differences in pronouns more generally. In order to do so, I used contrasting data regarding animacy and gender on pronouns in Brazilian Portuguese and Spanish to propose fine-grained distinctions between the semantic and syntactic makeup of overt subject pronouns in both.

In addition, I showed that a wide range of quantitative and qualitative differences in the respective distributions of null and overt subjects, (e.g. frequency of *pro*-drop, pronoun resolution biases, pragmatic effects) can be accounted for using *Minimize Restrictors!* if the semantic differences I argue for are taken into account. Beyond this, I showed that the generalizations I made regarding Spanish and Brazilian Portuguese can also be extended to French, Dominican Spanish, and Italian. Finally, I considered the implications of my theory for the acquisition of null subjects.

Future experimental and theoretical research in a number of domains will be necessary to test the predictions made by the hypotheses laid forth in this work. This thesis has been a first step in integrating an account of diverse NSLs with a fine-grained consideration of pronominal typology across languages. However, there are a number of empirical and theoretical questions that still need to be addressed, such as how first- and second-person pronouns fit into the picture, the extent of the interaction between the inflectional domain and the pronominal domain, and the consequences of my pronoun-centered analysis for NSLs unrelated to those considered here.

BIBLIOGRAPHY

BIBLIOGRAPHY

Ackerman, Lauren (2019). "Syntactic and cognitive issues in investigating gendered coreference". In: *Glossa: A Journal of General Linguistics* 4.1.

Ahn, Dorothy (2019). "THAT thesis: A competition mechanism for anaphoric expressions". PhD thesis. Harvard University.

Alonso-Ovalle, Luis et al. (2002). "Null vs. overt pronouns and the topic-focus articulation in Spanish". In: *Italian Journal of Linguistics* 14, pp. 151–170.

Barbosa, Pilar, Mary Kato, and Eugênia Duarte (2005). "Null Subjects in European and Brazilian Portuguese". In: *Journal of Portuguese Linguistics*.

Barbosa, Pilar P (2019). "*pro* as a Minimal nP: Toward a Unified Approach to *Pro*-Drop". In: *Linguistic Inquiry* 50.3, pp. 487–526.

Bjorkman, Bronwyn M (2017). "Singular *they* and the syntactic representation of gender in English". In: *Glossa: A Journal of General Linguistics* 2.1, p. 80.

Bonet i Alsina, M Eulalia (1991). "Morphology after syntax–pronominal clitics in romance". PhD thesis. Massachusetts Institute of Technology.

Bosque, Ignacio Académico Ponente, Victor Dir Presidente Garca de la Concha, and Humberto Srio López Morales (2009). *Nueva gramática de la lengua española: sintaxis II.*. Real Academia Española.

Bullock, Barbara E and Almeida Jacqueline Toribio (2009). "Reconsidering Dominican Spanish: data from the rural Cibao". In: *Revista Internacional de Lingüstica Iberoamericana*, pp. 49–73.

Camacho, José (2008). "Syntactic variation: The case of Spanish and Portuguese subjects". In: *Studies in Hispanic and Lusophone Linguistics* 1.2, pp. 415–434.

— (2018). *Introducción a la sintaxis del español*. Cambridge University Press.

Cardinaletti, Anna and Michal Starke (1999). "The typology of structural deficiency: A case study of the three classes of pronouns". In: *Clitics in the languages of Europe* 8.

Carminati, Maria Nella (2002). "The processing of Italian subject pronouns." PhD thesis. University of Massachusetts Amherst.

Chomsky, Noam (1981). "Lectures on Government and Binding (Dordrecht: Foris)". In: *Studies in Generative Grammar* 9.

Çnar, Oktay and Sinan Çakr (2019). "The Universality of the Overt Pronoun Constraint: The Re-analysis of the Turkish Case". In: *Australian Journal of Linguistics* 39.4, pp. 463–484.

Conrod, Kirby (2019). "Pronouns raising and emerging". PhD thesis. University of Washington.

Davies, Mark (2016). "Corpus del Español: Web/Dialects". In: *BYU Digital Corpora*.

Davies, Mark and MJ Ferreira (2016). "Corpus do Português: Web/Dialects". In: *BYU Digital Corpora*.

De la Fuente, Israel (2015). "Putting pronoun resolution in context: The role of syntax, semantics, and pragmatics in pronoun interpretation". PhD thesis. Université Paris Diderot.

Duarte, Maria Eugênia Lamoglia (1993). "Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil". In: *Português brasileiro: uma viagem diacrônica. Campinas: Editora da Unicamp*, pp. 107–128.

— (2000). "The loss of the 'Avoid Pronoun' principle in Brazilian Portuguese". In: *Brazilian Portuguese and the null subject parameter*, pp. 17–36.

Enrquez, Emilia V (1984). *El pronombre personal sujeto en la lengua española hablada en Madrid*. Vol. 4. Editorial CSIC-CSIC Press.

Esipova, Maria (2019). "Composition and Projection in Speech and Gesture." PhD thesis. New York University.

Fernandes, Eunice G et al. (2018). "Adaptation in pronoun resolution: Evidence from Brazilian and European Portuguese." In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 44.12, p. 1986.

Filiaci, Francesca, Antonella Sorace, and Manuel Carreiras (2013). "Anaphoric biases of null and overt subjects in Italian and Spanish: A cross-linguistic comparison". In: *Language, Cognition and Neuroscience* 29.7, pp. 825–843.

Frascarelli, Mara and Ángel L Jiménez-Fernández (2019). "Understanding Partiality in *pro*-Drop Languages: An Information-Structure Approach". In: *Syntax* 22.2-3, pp. 162–198.

Halle, M. and A. Marantz (1993). "Distributed Morphology and the Pieces of Inflection". In: *K. Hale and S. Keyser, eds., The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. MIT Press, pp. 111–176.

Heim, Irene (1991). "Artikel und Definitheit [Articles and Definiteness]". In: *A. von Stechow D. Wunderlich (eds.), Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*. de Gruyter, pp. 487–535.

Holmberg, Anders (2005). "Is there a little *pro*? Evidence from Finnish". In: *Linguistic Inquiry* 36.4, pp. 533–564.

Holmberg, Anders, Aarti Nayudu, and Michelle Sheehan (2009). "Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi". In: *Studia Linguistica* 63.1, pp. 59–97.

Ihsane, Tabea and Petra Sleeman (2016). "Gender agreement with animate nouns in French". In: *Selected proceedings of the 43rd linguistic symposium on Romance languages*, pp. 159–176.

Jaeggli, Osvaldo A (1986). "Arbitrary plural pronominals". In: *Natural Language & Linguistic Theory* 4.1, pp. 43–76.

Kaiser, Georg A (2009). "Losing the null subject: A contrastive study of (Brazilian) Portuguese and (Medieval) French". In: *Proceedings of the Workshop "Null-subjects, expletives, and locatives in Romance"*, pp. 131–156.

Konnelly, Lex and Elizabeth Cowper (2020). "Gender diversity and morphosyntax: An account of singular *they*". In: *Glossa: A Journal of General Linguistics* 5.1.

Kratzer, Angelika and Irene Heim (1998). *Semantics in generative grammar*. Vol. 1185. Blackwell Oxford.

Kuerová, Ivona (2018). "-Features at the Syntax-Semantics Interface: Evidence from Nominal Inflection". In: *Linguistic Inquiry* 49.4, pp. 813–845.

Landau, Idan (2010). *The locative syntax of experiencers*. Vol. 53. MIT press.

— (2016). "DP-internal semantic agreement: A configurational analysis". In: *Natural Language & Linguistic theory* 34.3, pp. 975–1020.

Levinson, Stephen C (1998). "Minimization and conversational inference". In: *Pragmatics: Vol. 4 Presupposition, implicature and indirect speech acts*. Routledge, pp. 545–612.

Lochbihler, Bethany, Will Oxford, and Nicholas Welch (2015). "The person-animacy connection: Evidence from Algonquian and Dene". Talk presented at Gender, Class, and Determination Workshop at the University of Ottawa. URL: `https://uocal.uottawa.ca/en/node/12202`.

Longobardi, Giuseppe (2008). "Reference to individuals, person, and the variety of mapping parameters". In: *Essays on nominal determination: From morphology to discourse management* 189, p. 211.

Lorusso, Paolo, Claudia Caprin, and Maria Teresa Guasti (2005). "Overt subject distribution in early Italian children". In: *Proceedings of the 29th Annual Boston University Conference on Language Development*.

Martnez-Sanz, Cristina (2011). *Null and overt subjects in a variable system: The case of Dominican Spanish*. University of Ottawa (Canada).

Matushansky, Ora (2013). "Gender confusion". In: Oxford University Press, pp. 1–60.

Mayol, Laia (2011). "An account of the variation in the rates of overt subject pronouns in Romance". In: *Spanish in Context* 9.3, pp. 420–442.

Montalbetti, Mario M (1984). "After binding: On the interpretation of pronouns". PhD thesis. Massachusetts Institute of Technology.

Negrão, Esmeralda (1999). "O português brasileiro: uma lngua voltada para o discurso". In: *USP: Tese de Livre-docência*.

Nunes, Jairo (2008). "Inherent case as a licensing condition for A-movement: The case of hyper-raising constructions in Brazilian Portuguese". In: *Journal of Portuguese Linguistics* 7.2.

— (2019). "Remarks on Finite Control and Hyper-Raising in Brazilian Portuguese". In: *Journal of Portuguese Linguistics* 18.1.

Okuma, Tokiko (2015). "Overt Pronoun Constraint effects in second language Japanese". PhD thesis. McGill University Libraries.

Pérez-Leroux, Ana T and Thomas Roeper (1999). "Scope and the structure of bare nominals: Evidence from child language". In: *Linguistics* 37.5, pp. 927–960.

Perlmutter, David (1971). "Deep and surface structure constraints". In: *Syntax*.

Pesetsky, David (2013). *Russian case morphology and the syntactic categories*. Vol. 66. MIT Press.

Posio, Pekka et al. (2012). "Pronominal subjects in Peninsular Spanish and European Portuguese: Semantics, pragmatics, and formulaic sequences". PhD thesis. University of Helsinki.

Ritter, Elizabeth (1995). "On the syntactic category of pronouns and agreement". In: *Natural Language & Linguistic Theory* 13.3, pp. 405–443.

Rizzi, Luigi (1986). "Null objects in Italian and the theory of pro". In: *Linguistic Inquiry* 17.3, pp. 501–557.

Roberts, Ian (2014). "Taraldsens Generalization and Language Change". In: *Functional structure from Top to Toe* 9.

— (2019). *Parameter Hierarchies and Universal Grammar*. Oxford University Press, USA.

Rodrigues, Cilene (2002). "Loss of Verbal Morphology and Referential Subjects in Brazilian Portuguese." In: *Syntactic Effects of Morphological Change.*, pp. 160–178.

Rodrigues, Cilene Aparecida Nunes (2004). "Impoverished morphology and A-movement out of Case domains". PhD thesis.

Rooryck, Johan (2000). *Configurations of sentential complementation: Perspectives from Romance languages*. Routledge.

Schlenker, Philippe (2005). "*Minimize restrictors!* (Notes on definite descriptions, Condition C and epithets)". In: *Proceedings of Sinn und Bedeutung*. Vol. 9, pp. 385–416.

Sheehan, Michelle (2006). "The EPP and null subjects in Romance". PhD thesis. Newcastle University.

Sigurðsson, Halldór Ármann (2019). "Gender at the edge". In: *Linguistic Inquiry* 50.4, pp. 723–750.

Steriopolo, Olga Martina Wiltschko (2010). *Distributed GENDER hypothesis*.

Sudo, Yasutada et al. (2012). "On the semantics of phi features on pronouns". PhD thesis. Massachusetts Institute of Technology.